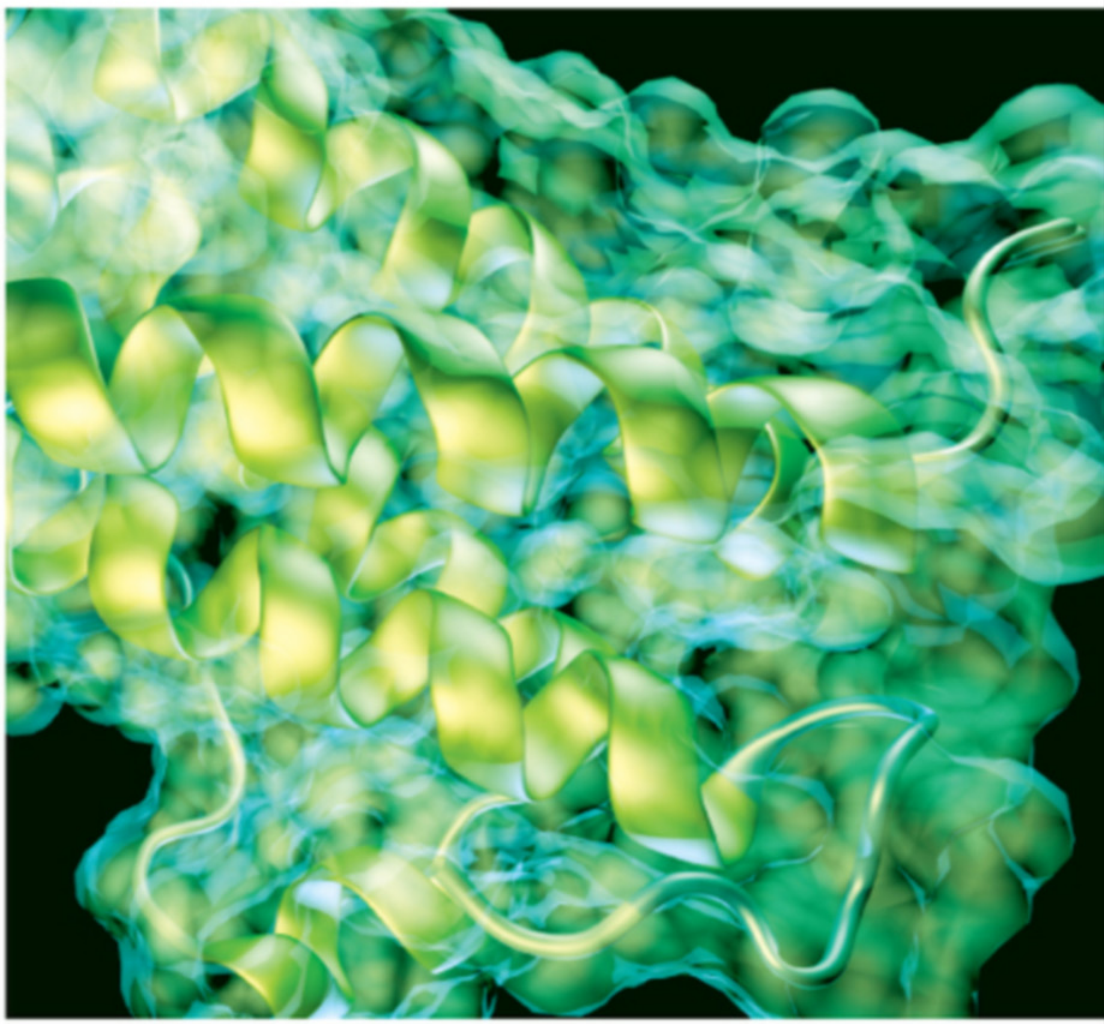
LECTURE NOTES

Clinical Biochemistry

GEOFFREY BECKETT
SIMON WALKER
PETER RAE
PETER ASHBY

8th edition



WILEY-BLACKWELL

Lecture Notes: Clinical Biochemistry

Lecture Notes

Clinical Biochemistry

Geoffrey Beckett

BSc PhD FRCPath
Consultant Clinical Scientist
Honorary Reader in Clinical Biochemistry
Department of Clinical Biochemistry
The Royal Infirmary of Edinburgh
Edinburgh

Simon Walker

MA MB BS DM FRCPath
Senior Lecturer in Clinical Biochemistry
Honorary Consultant Clinical Biochemist
Department of Clinical Biochemistry
The Royal Infirmary of Edinburgh
Edinburgh

Peter Rae

BA PhD MBChB FRCPE FRCPath
Consultant Clinical Biochemist
Honorary Senior Lecturer in Clinical Biochemistry
Department of Clinical Biochemistry
The Royal Infirmary of Edinburgh
Edinburgh

Peter Ashby

BA PhD FRCPath
Consultant Clinical Scientist
Honorary Senior Lecturer in Clinical Biochemistry
Department of Clinical Biochemistry
The Western General Hospital
Edinburgh

Eighth Edition

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2010, © 2010 by Geoffrey Beckett, Simon Walker, Peter Rae, Peter Ashby
Previous editions 1975, 1980, 1984, 1988, 1993, 1998, 2005

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing programme has been merged with Wiley's global Scientific, Technical and Medical business to form Wiley-Blackwell.

Registered office: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

Library of Congress Cataloging-in-Publication Data

Lecture notes. Clinical biochemistry. – 8th ed. / edited by Geoffrey Beckett ... [et al.].
p. ; cm.
Includes bibliographical references and index.
Includes bibliographical references and index.
ISBN 978-1-4051-9305-4 (pbk. : alk. paper)
1. Clinical chemistry. 2. Clinical biochemistry. I. Beckett, G. J.
[DNLM: 1. Biochemistry. 2. Clinical Chemistry Tests. 3. Clinical Laboratory Techniques. 4. Pathology, Clinical–methods. QU 4 L4705 2010]
RB40.L4 2010
616.07'56–dc22
2009031374

A catalogue record for this book is available from the British Library.

Set in 8 on 12pt Stone Serif by Toppan Best-set Premedia Limited
Printed in Malaysia

1 2010

Contents

Preface

This is the eighth edition of the book first conceived by Professor Gordon Whitby, Dr Alistair Smith and Professor Iain Percy-Robb in 1975. As with the first edition, the book has been written primarily for medical students and junior doctors. The changes that have been introduced into the undergraduate medical teaching curriculum, including systems-based medicine, means that the book is now of relevance to each of the years of the course. In addition, we believe that the book is of value to specialist registrars, clinical scientists and biomedical scientists who are studying for higher qualifications to pursue a career in clinical biochemistry and metabolic medicine.

As with previous editions, the book has been modified in response to a number of suggestions made by a group of students commissioned by Wiley to review the previous edition. In response to their comments, Wiley have introduced a further colour into this edition and improved the format of the text and tables to make the content clearer. The order of the chapters has also been revised to allow all endocrine chapters to run consecutively. On our part, we have reviewed and updated all chapters to reflect current clinical practice and national guidelines, and we have included additional case studies since these have proved to be a very popular component of the book. It is interesting to note how lack of funding can prevent a useful test being introduced into clinical practice. In the 7th edition we were predicting that brain natriuretic peptide (BNP) measurements would play a pivotal role in the diagnosis of suspected heart failure in primary care. In the event, the introduction of BNP measurement into primary care in the UK has been slow due to lack of availability of adequate funding.

We have retained the objectives that have characterised the book throughout each of its previous editions. As a consequence, the reader should gain a knowledge and understanding of the value, limitations and interpretation of the many biochemical tests that are in common use in modern medicine. In addition we hope that this book will allow the reader to answer the questions first raised in Asher's Catechism published in the *British Medical Journal* in 1954 under the title 'Straight and Crooked Thinking in Medicine', namely:

Why do I request this test?

What will I look for in the result?

If I find what I am looking for, will it affect my diagnosis?

How will this investigation affect my management of the patient?

We would like to thank Dr Jean Kirk for her help with the paediatric biochemistry section; Dr Allan Deacon for his views regarding the investigation of porphyria; and Dr Gordon Brydon for helpful comments concerning tests of gastrointestinal function. We also wish to thank the staff at Wiley for their continued interest and support towards this title since its conception in 1975, and for this edition particularly Laura Murphy and Ben Townsend.

Geoff Beckett
Simon Walker
Peter Rae
Peter Ashby

List of abbreviations

ABP androgen-binding protein
A&E Accident and (&) Emergency
ACE angiotensin-converting enzyme
ACTH adrenocorticotrophic hormone
ADH antidiuretic hormone
AFP α-fetoprotein
AI angiotensin I
AII angiotensin II
AIP acute intermittent porphyria
ALA aminolaevulinic acid
ALP alkaline phosphatase
ALT alanine aminotransferase
AMP adenosine 5-monophosphate
ANP atrial natriuretic peptide
API α_1-protease inhibitor
AST aspartate aminotransferase
ATP adenosine triphosphate
ATPase adenosine triphosphatase
BChE butylcholinesterase
BMI body mass index
BMR basal metabolic rate
BNP B-type natriuretic peptide
CAH congenital adrenal hyperplasia
cAMP cyclic adenosine monophosphate
CBG cortisol-binding globulin
CCK-PZ cholecystokinin-pancreozymin
CDT carbohydrate-deficient transferrin
CEA carcinoembryoinic antigen
ChE cholinesterase
CK creatine kinase
CKD chronic kidney disease
CNS central nervous system
CoA coenzyme A
COC combined oral contraceptive
COHb carboxyhaemoglobin
CRH corticotrophin-releasing hormone
CRP C-reactive protein
CSF cerebrospinal fluid
CT computed tomography
DDAVP 1-deamino,8-D-arginine vasopressin
DHEA dehydroepiandrosterone
DHEAS dehydroepiandrosterone sulphate
DHCC dihydrocholecalciferol
DHT dihydrotestosterone
DIT di-iodotyrosine
DKA diabetic ketoacidosis
DPP-4 dipeptidyl peptidase-4
DVT deep venous thrombosis
ECF extracellular fluid
ECG electrocardiogram/electrocardiography
EDTA ethylenediamine tetraacetic acid
eGFR estimated glomerular filtration rate
ERCP endoscopic retrograde cholangiopancreatography
ESR erythrocyte sedimentation rate
FAD flavin–adenine dinucleotide
FAI free androgen index
FBHH familial benign hypocalciuric hypercalcaemia
FOB faecal occult blood
FSH follicle-stimulating hormone
FT3 free tri-iodothyronine
FT4 free thyroxine
GAD glutamic acid decarboxylase
Gal-1-PUT galactose-1-phosphate uridylyl-transferase
GC–MS gas chromatography–mass spectrometry
GFR glomerular filtration rate
GGT γ-glutamyltransferase
GH growth hormone
GHD growth hormone deficiency
GHRH growth hormone-releasing hormone
GI gastrointestinal
GIP glucose-dependent insulinotrophic peptide
GLP-1 glucagon-like polypeptide-1
GnRH gonadotrophin-releasing hormone
GP general practitioner
GSA glucocorticoid-suppressible hyperaldosteronism
GTT glucose tolerance test

List of abbreviations

Hb haemoglobin
HC hereditary coproporphyria
HCC hydroxycholecalciferol
hCG human chorionic gonadotrophin
HDL high-density lipoprotein
HDU high dependency unit
HGPRT hypoxanthine-guanine phosphoribosyltransferase
5-HIAA 5-hydroxyindoleacetic acid
HIV human immunodeficiency virus
HLA human leucocyte antigen
HMG-CoA β-hydroxy-β-methylglutaryl-coenzyme A
HNF hepatic nuclear factor
HPA hypothalamic–pituitary–adrenal
HPLC high-performance liquid chromatography
HRT hormone replacement therapy
hsCRP high sensitive C-reactive protein
5-HT 5-hydroxytryptamine
5-HTP 5-hydroxytryptophan
ICF intracellular fluid
ICU intensive care unit
IDL intermediate-density lipoprotein
IFCC International Federation for Clinical Chemistry
IFG impaired fasting glycaemia
Ig immunoglobulin
IGF insulin-like growth factor
IGFBP insulin-like growth factor-binding protein
IGT impaired glucose tolerance
IM intramuscular
INR international normalised ratio
IV intravenous
LCAT lecithin cholesterol acyltransferase
LDH lactate dehydrogenase
LDL low-density lipoprotein
LH luteinising hormone
LHRH luteinising hormone-releasing hormone
Lp (a) lipoprotein (a)
LSD lysergic acid diethylamide
MCAD medium chain acyl-CoA dehydrogenase
MCV mean cell volume
MDRD Modification of Diet in Renal Disease
MEGX monoethylglycinexylidide
MEN multiple endocrine neoplasia
MGUS monoclonal gammopathy of unknown significance
MIH Mullerian inhibitory hormone
MIT mono-iodotyrosine
MODY maturity onset diabetes of the young
MOM multiples of the median
MRI magnetic resonance imaging
MSAFP maternal serum α-fetoprotein
NAD nicotinamide–adenine dinucleotide
NADP NAD phosphate
NAFLD non-alcoholic fatty liver disase
NASH non-alcoholic steatohepatitis
NICE National Institute for Health and Clinical Excellence
NTD neural tube defect
NTI non-thyroidal illness
OGTT oral glucose tolerance test
PAPP-A pregnancy-associated plasma protein A
PBG porphobilinogen
PCOS polycystic ovarian syndrome
PCT porphyria cutanea tarda
PE pulmonary embolism
PEM protein-energy malnutrition
PKU phenylketonuria
POCT point of care testing
POP progestogen-only pill
PP pyridoxal phosphate
PRA plasma renin activity
PRPP 5-phosphoribosyl-1-pyrophosphate
PSA prostate-specific antigen
PT prothrombin time
PTC percutaneous transhepatic cholangiography
PTH parathyroid hormone
PTHrP PTH-related protein
RDA recommended dietary allowance
RF rheumatoid factor
ROC receiver operating characteristic
SAH subarachnoid haemorrhage
SD standard deviation
SHBG sex hormone-binding globulin
SI Système International
SIADH inappropriate secretion of ADH
SUR sulphonylurea receptor
T3 tri-iodothyronine
T4 thyroxine
TBG thyroxine-binding globulin

TDM therapeutic drug monitoring
TIBC total iron-binding capacity
TPMT thiopurine methyltransferase
TPN total parenteral nutrition
TPOAb thyroid peroxidase antibody
TPP thiamin pyrophosphate
TRAb thyrotrophin receptor antibody
TRH thyrotrophin-releasing hormone
TSH thyroid-stimulating hormone
TSI thyroid-stimulating immunoglobulin
tTG tissue transglutaminase
U&Es urea and electrolytes
UFC urinary free cortisol
VIP vasoactive intestinal peptide
VLDL very low density lipoprotein
VMA vanillylmandelic acid
VP variegate porphyria
WHO World Health Organization

Chapter 1

Requesting and interpreting tests

Introduction

Biochemical tests are crucial to many areas of modern medicine. Most biochemical tests are carried out on blood using plasma or serum, but urine, cerebrospinal fluid (CSF), faeces, kidney stones, pleural fluid, etc. are sometimes required. Plasma is obtained by taking whole blood into an anti-coagulant and represents the aqueous supernatant obtained when all the cellular elements have been separated by centrifugation. Serum is the corresponing aqueous phase when blood is allowed to clot. For many (but not all) biochemical tests on blood, it makes little difference whether plasma or serum is used and the terms are often used interchangeably.

There are many hundreds of tests available in clinical biochemistry that include many specialist tests. However, a core of common tests makes up the majority of test requesting in clinical biochemistry. These core tests will be offered by almost all clinical biochemistry laboratories and will be available 24 h daily for more urgent situations. It is also sometimes appropriate to bring tests together in profiles, especially where a group of tests can provide better understanding of a problem than a single test (e.g. the liver function test profile). Many of the other more specialist tests are restricted to larger laboratories or, in some cases, to a very small number of centres offering a regional or national service.

In dealing with the large number of routine test requests, the modern clinical biochemistry laboratory depends heavily on automated instrumentation. This is most often linked to a laboratory computing system which assigns test requests to electronic patient files, maintains a cumulative patient record and regulates the printing of reports. Increasingly, test requests can be electronically booked at the ward, clinic or even general practitioner (GP) surgery via a terminal linked to the main laboratory computer. Equally, the test results can be displayed on computer screens at distant locations, even negating the need for issuing printed reports.

In this first chapter, we set out some of the principles of requesting tests and of the interpretation of results. The effects of analytical errors and of physiological factors, as well as of disease, on test results are stressed. Biochemical testing in differential diagnosis and in screening is discussed.

Collection of specimens

Test requests require unambiguous identification of the patient (patient's name, sex, date of birth and, increasingly, a unique patient identification number), together with the location, the name of

Lecture Notes: Clinical Biochemistry, 8e. By G. Beckett, S. Walker, P. Rae & P. Ashby. Published 2010 by Blackwell Publishing.

the requesting doctor and the date and time of sampling. Each test request must specify which analyses are requested and provide details of the nature of the specimen itself and relevant clinical diagnostic information. Traditionally, this information is provided through the request form with appropriate parallel labelling of the specimen itself. Increasingly, this information is provided electronically so that only the sample itself need be sent to the laboratory with its own unique identifier (e.g. a bar code which links it to the electronic request).

Because of the large number of samples which are processed by most clinical biochemistry laboratories, every step needs to be taken to avoid errors. Regrettably, errors do rarely occur and can be dividied according to the error source:

- Pre-analytical. For example, assigning a specimen to the wrong patient at the ward end or taking a sample at the wrong time (e.g. digoxin level is requested on a sample shortly after digoxin has been administered (pp. 279)) or mislabelling of an aliqout of serum taken at specimen reception. Most errors fall into this category (see Table 1.1).
- Analytical. For example, a small sample volume may lead to a pipetting error where insufficient sample is used for the assay. Again, developments in automated sample detection and pipetting mean these problems are very unusual.
- Post-analytical. These are increasingly rare because of electronic download of results from the analyser but might include transcription errors when entering results into the lab computer manually.

On the scale of the requesting of biochemical tests, errors are fortunately rare. However, occasional blunders do arise and, if very unexpected results are obtained, it is incumbent on the requesting doctor to contact the laboratory immediately to look into the possibility that a blunder may have occurred.

Table 1.1 Some more common causes of pre-analytical errors arising from use of the laboratory

Error	Consequence
Crossover of addressograph labels between patients	This can lead to two patients each with the other's set of results. Where the patient is assigned a completely wrong set of results, it is important to investigate the problem in case there is a second patient with a corresponding wrong set of results.
Timing error	There are many examples where timing is important but not considered. Sending in a blood sample too early after the administration of a drug can lead to misleadingly high values in therapeutic monitoring. Interpretation of some tests (e.g. cortisol) is critically dependent on the time of day when the blood was sampled.
Sample collection tube error	For some tests the nature of the collection tube is critical, which is why the Biochemistry Laboratory specifies this detail. For example, using a plasma tube with lithium–heparin as the anti-coagulant invalidates this sample tube for measurement of a therapeutic lithium level! Electrophoresis requires a serum sample; otherwise, the fibrinogen interferes with the detection of any monoclonal bands. Topping up a biochemistry tube with a haematology (potassium ethylenediamine tetraacetic acid (EDTA) sample) will lead to high potassium and low calcium values in the biochemistry sample.
Sample taken from close to the site of an intravenous (IV) infusion	The blood sample will be diluted so that all the tests will be correspondingly low with the exception of those tests which might be affected by the composition of the infusion fluid itself. For example, using normal saline as the infusing fluid would lead to a lowering of all test results, but with sodium and chloride results which are likely to be raised.

The use of clinical biochemistry tests

Biochemical tests are most often *discretionary*, meaning that the test is requested for defined diagnostic purposes, as distinct from screening, where a disease is sought without there being any specific indication of its presence in the individual. The justification for discretionary testing is well summarised by Asher (1954):

1 Why do I request this test?
2 What will I look for in the result?
3 If I find what I am looking for, will it affect my diagnosis?
4 How will this investigation affect my management of the patient?
5 Will this investigation ultimately benefit the patient?

The principal reasons for requesting biochemical tests are as follows (where the first two categories would be defined as discretionary):

- To assist in diagnosis. For example, the diagnosis of diabetes mellitus is crucially dependent on the measurement and interpretation of plasma [glucose]. Biochemical tests may also aid the differential diagnosis or indicate the severity of a disease (see also Table 1.2).
- In disease monitoring. A good example is the use of arterial blood gases to follow the progress of someone admitted with a severe pneumonia or creatinine in an individual with chronic renal failure (see also Table 1.2).
- In prognosis or disease risk assessment. Serum cholesterol (pp. 192) or high-sensitive C-reactive protein (hsCRP) (pp. 192) are used in the assessment of cardiovascular risk, for example.
- In screening for disease. An example here would be measurement of thyroid-stimulating hormone (TSH) to screen for neonatal hypothyroidism.
- Miscellaneous, for example for forensic purposes or ethically approved research.

Screening may take two forms:

- In well-population screening a spectrum of tests is carried out on individuals from an apparently healthy population in an attempt to detect presymptomatic or early disease. It is easy to miss significant abnormalities in the 'flood' of data coming from the laboratory, even when the abnormalities are 'flagged' in some way. Most of the abnormalities detected will be of little or no significance, yet may need additional time-consuming and often expensive tests to clarify their importance (or lack of it). For these and other reasons, the value of well-population screening has been called into question and certainly should only be initiated under certain specific circumstances which are listed in Table 1.3.

Table 1.2 Test selection for the purposes of discretionary testing

Category	Example
To confirm a diagnosis	Serum [free T4] and [thyroid-stimulating hormone, (TSH)] in suspected hyperthyroidism
To aid differential diagnosis	To distinguish between different forms of jaundice
To refine a diagnosis	Use of adrenocorticotrophic hormone (ACTH) to localise Cushing's syndrome
To assess the severity of disease	Serum [creatinine] or [urea] in renal disease
To monitor progress	Plasma [glucose] and serum [K^+] to follow treatment of patients with diabetic ketoacidosis (DKA)
To detect complications or side effects	Alanine aminotransferase (ALT) measurements in patients treated with hepatotoxic drugs
To monitor therapy	Serum drug concentrations in patients treated with anti-epileptic drugs

Table 1.3 Requirements for well-population screening

The disease is common or life-threatening
The tests are sensitive and specific
The tests are readily applied and acceptable to the population to be screened
Clinical, laboratory and other facilities are available for follow-up
Economics of screening have been clarified and the implications accepted

Table 1.4 Examples of tests used in case-finding programmes

Programmes to detect diseases in	Chemical investigations
Neonates	
PKU	Serum [phenylalanine]
Hypothyroidism	Serum [TSH] and/or [thyroxine]
Adolescents and young adults	
Substance abuse	Drug screen
Pregnancy	
Diabetes mellitus in the mother	Plasma and urine [glucose]
Open neural tube defect (NTD) in the foetus	Maternal serum [α-fetoprotein]
Industry	
Industrial exposure to lead	Blood [lead]
Industrial exposure to pesticides	Serum cholinesterase activity
Elderly	
Malnutrition	Serum vitamin D levels
Thyroid dysfunction	Serum [TSH] and/or [thyroxine]

- In case-finding screening programmes appropriate tests are carried out on a population sample known to be at high risk of a particular disease. These are inherently more selective and yield a higher proportion of useful results (Table 1.4).

Point of care testing (POCT) (Table 1.5)

There are occasions when the urgency of the clinical situation requires that blood testing on patient samples is performed near the patient (point of care testing). Furthermore in the UK the government, in outlining the future of the National Health Service, has indicated a desire to move laboratory testing from the hospital laboratory into the community setting. High street pharmacies have taken up these opportunities and can, for example, provide cholesterol and glucose testing while you wait. In addition, there is an increasing number of urine test sticks that are sold for home use (e.g. pregnancy and ovulation testing by measuring human chorionic gonadotrophin (hCG) and luteinising hormone (LH), respectively).

POCT eliminates the need to send the specimen to the laboratory, and will usually allow a more rapid turnaround time. POCT is particularly suitable for use in intensive care units (ICUs), high-dependency units (HDUs), Accident and Emergency (A&E) departments and specialist clinics. Small dedicated analysers are often introduced into these centres.

Since decisions to initiate treatment are often made on the basis of POCT it is vital that confidence can be placed in the results obtained by such methods, such that the risk to the patient is minimised. It is thus essential that POCT is carried out by staff who are suitably trained and that the reliability of the tests is monitored on a regular basis using appropriate quality control measures.

If POCT is to be introduced into a ward or outpatient department it is essential that:

- The laboratory is consulted to advise on the choice of method, staff training and quality control issues.
- Only properly trained staff should be permitted to use the equipment.
- There are simple sets of written instructions, which must include simple quality control procedures and what to do when the instrument seems to be performing unreliably.
- Health and safety issues must be considered, e.g. risk of exposure to hepatitis or human immunodeficiency virus (HIV).
- Quality control is monitored on a regular basis, preferably by the main laboratory.
- Records are kept of patient results, quality control and the personnel that performed the test.

Advantages include:

- Rapid access to results on acutely ill patients.
- Closer (more frequent) monitoring, whether acutely ill (e.g. blood gases in ICU) or in the home (e.g. glucose meters).
- 24 h availability.

Disadvantages include:

- Typically more expensive than a main analytical laboratory test.

Table 1.5 Examples of POCT that are in common use

Common POCT in blood	Common POCT in urine
Blood gases	Glucose
Glucose	Ketones
Urea and creatinine	Red cells/haemoglobin
Na, K and Ca	Bilirubin
Bilirubin	Urobilinogen
Salicylate	pH
Paracetamol	Protein
Alcohol	hCG
Troponin	Drugs of abuse

- Requires wider staff training with less ability to regulate access by untrained individuals.
- Calibration and quality control requirements are potentially less robust.
- Results not often integrated into the full electronic patient record.

As outlined above, where it is appropriate to introduce POCT, it is critical that matters of training, analytical performance, quality control, and health and safety are properly addressed. 'Smart' systems are also available which only allow password access to trained individuals and can also prevent issue of results if calibration is unsatisfactory or quality control failure occurs. POCT instruments can also be networked and performance monitored from the central laboratory.

Interpretation of clinical biochemistry tests

Most reports issued by clinical biochemistry laboratories contain numerical measures of concentration or activity, expressed in the appropriate units. Typically, the result is interpreted in relation to a reference range (see p. 7) for the analyte in question.

The following questions should be considered when interpreting the results:

1 Is each result normal or abnormal? Reference ranges (often incorrectly called normal ranges) are needed in order to answer questions about quantitative data.

2 Does each result fit in with my previous assessment of this patient? If not, can I explain the discrepancy?

3 Has a significant change occurred in any of the results from those previously reported?

4 Do any of the results alter my diagnosis of this patient's illness or influence the way in which the illness should be managed?

5 If I cannot explain a result, what do I propose to do about it?

This section discusses the interpretation of laboratory results and the factors that may cause them to vary, under the following main headings:

1 *Analytical factors* These cause errors in measurement.

2 *Biological and pathological factors* Both these sets of factors affect the concentrations of analytes in blood, urine and other fluids sent for analysis.

Sources of variation in test results

Analytical sources of variation

Systematic and random variation

Analytical results are subject to error, no matter how good the laboratory and no matter how skilled the analyst. These errors may be due to lack of accuracy, that is, always tend to be either high or low, or may be due to random effects and lack precision, that is, may be unpredictably high or low.

Accuracy

An accurate method will, on average, yield results close to the true value of what is being measured. It has no systematic bias.

Precision

A precise method yields results that are close to one another (but not necessarily close to the true value) on repeated analysis. If multiple measurements are made on one specimen, the spread of results will be small for a precise method and large for an imprecise one.

The 'dartboard' analogy is often used to illustrate the different meanings of the terms accuracy and precision, and this is illustrated in Figure 1.1.

The standard deviation (SD) is the usual measure of scatter around a mean value. If the spread of results is wide, the SD is large, whereas if the spread is narrow, the SD is small. For data that have a Gaussian distribution, as is nearly always the

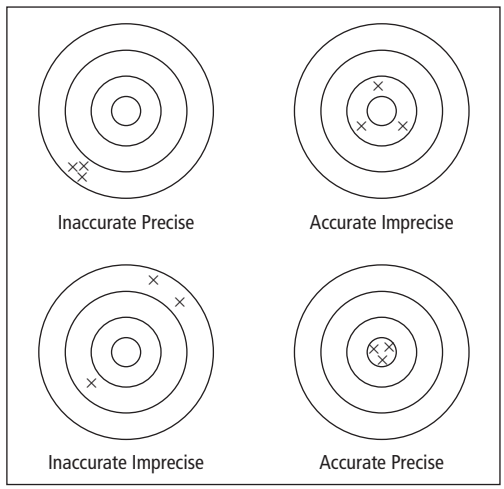


Figure 1.1 The 'dartboard' analogy can be used to illustrate accuracy and precision.

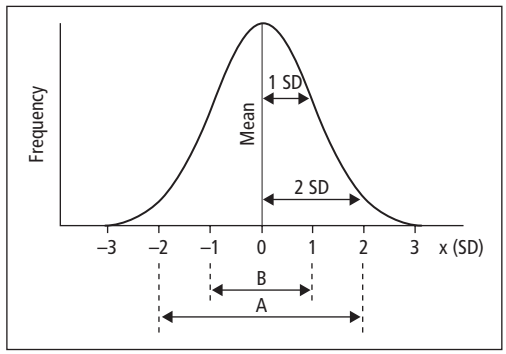


Figure 1.2 Diagram of a Gaussian (normal or symmetrical) distribution curve. The span (A) of the curve, the distance between the mean ± 2 SD, includes about 95% of the 'population'. The narrower span (B), the distance between the mean ± 1 SD, includes about 67% of the 'population'.

case for analytical errors, the shape of the curve (Figure 1.2) is completely defined by the mean and the SD, and these characteristics are such that:

- About 67% of results lie in the range mean ± 1 SD.
- About 95% of results lie in the range mean ± 2 SD.
- Over 99% of results lie in the range mean ± 3 SD.

Blunders

These are grossly inaccurate results that bear no constant or predictable relationship to the true value. They arise, for instance, from mislabelling of specimens at the time of collection, or transcription errors when preparing or issuing reports (see Table 1.1).

Serial results in the same patient

Doctors often have to interpret two or more sets of results for the same analysis or group of analyses performed on different occasions on the same patient. An important question is whether an analytical change is due mainly to laboratory imprecision or a to true change in the patients's clinical condition. Without elaborating on the statistical aspects of this, the following rule may be applied: if the results for analyses performed on specimens collected on different occasions, but under otherwise identical conditions, differ by more than 2.8 times the analytical SD then there is a chance of over 95% that a genuine change in concentration of the substance has occurred.

Biological causes of variation

As well as analytical variation, test results also show biological variation in both health and disease. Key questions are:

- How do results vary in health?
- How do results vary in disease?

How do results vary in health?

The concentrations of all analytes in blood vary with time due to diverse physiological factors *within* the individual. There are also differences *between* individuals.

Within-individual variation

The following may be important causes of within-individual variation:

1 *Diet* Variations in diet can affect the results of many tests, including serum [triglyceride], the response to glucose tolerance tests and urinary calcium excretion.

2 *Time of day* Several plasma constituents show diurnal variation (variation with the time of day), or a sleep/wake cycle. Examples include iron, adrenocorticotrophic hormone (ACTH) and cortisol concentrations.

3 *Posture* Proteins and all protein-bound constituents of plasma show significant differences in concentration between blood collected from upright individuals and blood from recumbent individuals. Examples include serum calcium, cholesterol, cortisol and total thyroxine concentrations.
4 *Muscular exercise* Recent exercise, especially if vigorous or unaccustomed, may increase serum creatine kinase (CK) activity and blood [lactate], and lower blood [pyruvate].
5 *Menstrual cycle* Several substances show variation with the phase of the cycle. Examples include serum [iron], and the serum concentrations of the pituitary gonadotrophins, ovarian steroids and their metabolites, as well as the amounts of these hormones and their metabolites excreted in the urine.
6 *Drugs* These can have marked effects on chemical results. Attention should be drawn particularly to the many effects of oestrogen-containing oral contraceptives on serum constituents (p. 166).
Even after allowing for known physiological factors that may affect plasma constituents and for analytical imprecision, there is still considerable residual individual variation (Table 1.6). The magnitude of this variation depends on the analyte, but it may be large and must be taken into account when interpreting successive values from a patient.

Between-individual variation

Differences between individuals can affect the concentrations of analytes in the blood. The following are the main examples:

Table 1.6 Residual individual variation of some serum constituents (expressed as the approximated day-to-day, within-individual coefficient of variation)

Serum constituent	CV (%)	Serum constituent	CV (%)
Sodium	1	ALT activity	25
Calcium	1–2	AST activity	25
Potassium	5	Iron	25
Urea	10		

CV = coefficient of variation.

1 *Age* Examples include serum [phosphate] and alkaline phosphatase (ALP) activity, and serum and urinary concentrations of the gonadotrophins and sex hormones.
2 *Sex* Examples include serum creatinine, iron, urate and urea concentrations and γ-glutamyltransferase (GGT) activity, and serum and urinary concentrations of the sex hormones.
3 *Race* Racial differences have been described for serum [cholesterol] and [protein]. It may be difficult to distinguish racial from environmental factors, such as diet.

Reference ranges

When looking at results, we need to compare each result with a set of results from a particular defined (or reference) population. This reference range is determined, in practice, by measuring a set of reference values from a sample of that population, usually of healthy individuals. The nature of the reference population should be given whenever reference ranges are quoted, although a healthy population is usually assumed. Even age-matched and sex-matched reference ranges are often difficult to obtain, since fairly large numbers of individuals are needed. In practice, blood donors are very often selected as the most readily available reference population.

Distribution of results in a reference population

When results of analyses for a reference population are analysed, they are invariably found to cluster around a central value, with a distribution that may be symmetrical (often Gaussian, Figure 1.3a) or asymmetrical (often log-Gaussian, Figure 1.3b). However, reference ranges can be calculated from these data without making any assumptions about the distribution of the data, using non-parametric methods.

Because of geographical, racial and other biological sources of variation between individuals, as well as differences in analytical methods, each laboratory should ideally define and publish its own reference ranges. Conventionally, these include the central 95% of the results obtained for each analysis from the reference population. This 95%

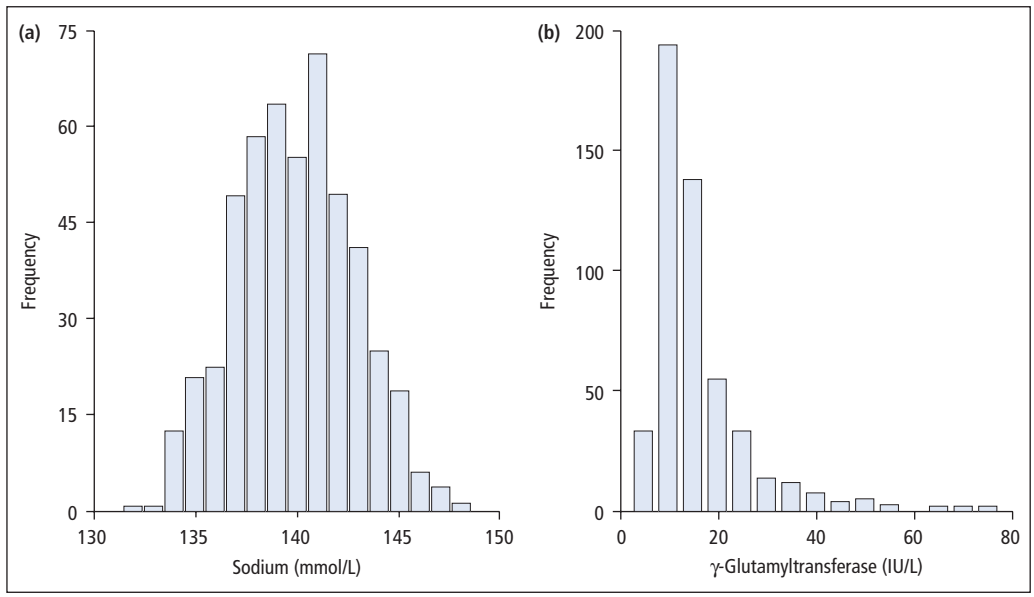


Figure 1.3 Histograms showing the relative frequency with which results with the values indicated were obtained when serum [Na^+] and γ-glutamyltransferase (GGT) activities were measured in a reference population of healthy adult women. (a) The sodium data are symmetrically distributed about the mean whereas (b) the GGT data show a log-Gaussian distribution.

figure is arbitrary, selected in order to minimise the overlap between results from diseased populations and from healthy individuals.

Analytical factors can affect the reference ranges for individual laboratories. If an *inaccurate* method is used, the reference range will reflect the method bias. If an *imprecise* method is used, the reference range will be widened, that is, the observed span of results (reflected in the SD) will be greater. In statistical terms, the observed variance (i.e. the square of the SD) of the population results will equal the sum of the true or biological variance of the population plus the analytical variance of the method.

How do results vary in disease?

Biochemical test results do not exist in isolation, since, by the time tests are requested, the doctor will often have made a provisional diagnosis and a list of differential diagnoses based on each patient's symptoms and signs.

For example, in a patient with severe abdominal pain, tenderness and rigidity, there may be several differential diagnoses to consider – including, for example, acute pancreatitis, perforated peptic ulcer and acute cholecystitis. In all three conditions, the serum amylase activity may be raised, that is, above the upper reference value for healthy adults. So healthy adult reference ranges (in this instance) are irrelevant, since healthy adults do not have abdominal pain, tenderness and rigidity! Instead, we need to know how the serum amylase activity might vary in the clinically likely differential diagnoses. It would be useful to know, for instance, whether very high serum amylase activities are associated with one of these diagnostic possibilities, but not with the other two.

To summarise, to interpret results on patients adequately, we need to know:

- the reference range for healthy individuals of the appropriate age range and of the same sex;
- the values to be expected for patients with the disease, or diseases, under consideration;

- the prevalence of the disease, or diseases, in the population to which the patient belongs.

The assessment of diagnostic tests

In evaluating and interpreting a test, it is necessary to know how it behaves in health and disease. Central to understanding here are the terms sensitivity and specificity.

- Test sensitivity refers to how effective the test is in detecting individuals who have the disease in question. It is expressed as the percentage of true positives in all the individuals who have disease (all the individuals with disease will encompass the true positives (TP) and false negatives (FN)). So:

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \times 100\%$$

- Test specificity is a measure of how good the test is at providing a negative result in the absence of disease. It is expressed as the percentage of true negatives in all those without the disease (all the individuals without disease will encompass the true negatives (TN) and the false positives (FP). So:

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \times 100\%$$

The ideal test is 100% sensitive (positive in all patients with the disease) and 100% specific (negative in all patients without the disease), shown diagrammatically in Figure 1.4a. This ideal is rarely achieved; there is usually overlap between the healthy and diseased populations (Figure 1.4b). In practice, we have to decide where to draw dividing lines that most effectively separate 'healthy' from 'diseased' groups, or disease A from disease B. We can illustrate this by means of the following hypothetical example.

The effectiveness of a test can also be defined in terms of the predictive value of a positive result and the predictive value of a negative result. The positive predictive value is:

$$\text{TP}/(\text{TP} + \text{FP}) \times 100\%$$

A test with a high positive predictive value will, by definition, have few false positives. This would be important in a situation where a high number of false positives would otherwise lead to extensive and costly further investigation.

The negative predictive value is defined as follows:

$$\text{TN}/(\text{TN} + \text{FN}) \times 100\%$$

A test with a high negative predictive value would, by definition, have few false negatives. This would be particularly important, for example, in a test which was used for a screening programme where it is essential not to miss a case of the disease in question.

In defining the presence or absence of a disease, a cut-off may be assigned to a test. Consider the situation where a high value for a particular test equates with the presence of a particular disease. A value above the cut-off would then define the presence of the disease and a value below the cut-off, the absence of disease. A cut-off which is set at a

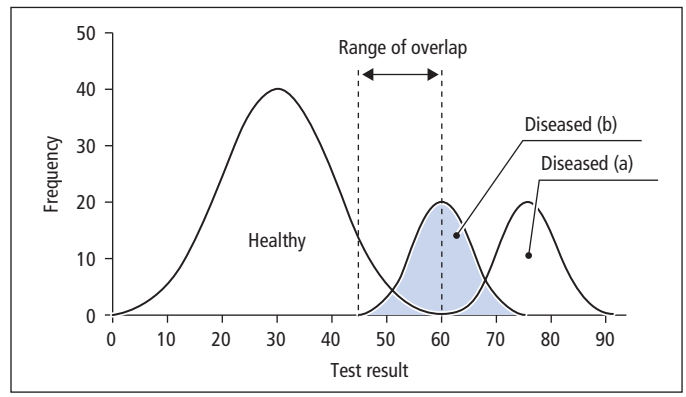


Figure 1.4 Diagrammatic representations of the distributions of results obtained with a test (a) that completely separates healthy people from people with a disease without any overlap between the distribution curves (i.e. an ideal test with 100% sensitivity and 100% specificity), and a test (b) that is less sensitive and less specific, in which there is an area of overlap between the distribution curves for healthy people and people with disease.

higher level will increase the test specificity at the expense of test sensitivity (more false negatives), whilst a cut-off set at a lower value will increase test sensitivity at the expenses of test specificity (more false positives).

In evaluating tests for decision making, it is clearly important to decide on the relative importance of sensitivity versus specificity in the context for which a test is used. To that end, it is helpful to be able to make comparisons of different tests with respect to sensitivity and specificity. This is often best carried out by plotting the test sensitivity against specificity and constructing a so-called receiver operating characteristic (ROC) curve. These curves will highlight which test is best suited to which requirement and will also help to define which cut-off to select in order to balance specificity versus sensitivity. This is illustrated in Figure 1.5.

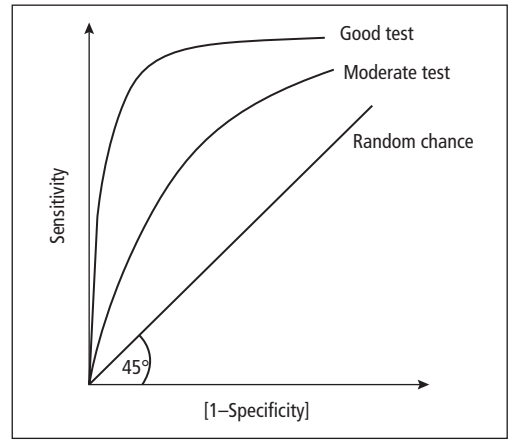


Figure 1.5 Schematic representation of a receiver operating characteristic (ROC) plot. A random test produces a straight line set at 45° to the axes. A discriminatory, good test produces a graph with a steep slope from the origin, displaying high sensitivity at high specificity. Less discriminatory tests produce curves at intermediate positions, as shown.
(Adapted with the authors' permission from: Roulston, J.E. and Leonard, R.F.C. 'Serological tumour markes: an Introduction' Publ. Churchill Livingstone 1993)

Screening for rare diseases

For diseases that are rare, tests of extremely high sensitivity and specificity are required. To illustrate this, consider an inherited metabolic disorder with an incidence of 1 : 5000; this is similar to that of some of the more common, treatable, inherited metabolic diseases such as phenylketonuria (PKU) or congenital hypothyroidism. Assume that we have a test with a good performance, that is, a sensitivity and specificity of 99.5% (Table 1.7).

Table 1.7 shows that for every neonate affected by the disorder who has a positive test result, there will be about 25 (4999/199) neonates who also have a positive test but who do not have the disease. Two important points emerge:

1 Tests with very high sensitivity and with very low false-positive rates are required when screening for rare disorders.

2 A heavy investigative load will result from the screening programme, since all the false positives will have to be followed up to determine whether or not they indicate the presence of disease.

The traditional 95% reference range (see above) is not relevant to screening for rare conditions, since the rate of false positives would be far too high. The cut-off value has to be altered to decrease the false-positive rate, at the probable expense of missing some patients who have the condition for which screening is being carried out.

Table 1.7 A hypothetical set of results of a screening test for a relatively common inherited metabolic disorder in neonates

Diagnostic category	Positive results	Negative results	Total
Disease present	199	1	200
Disease absent	4999	994 801	999 800
Total	5198	994 802	1 000 000
Predictive value	3.8%	100%	

Assumptions: sensitivity of the test 99.5%, false-positive rate 0.5% (specificity 99.5%), prevalence of the disorder, 1 : 5000; 1 000 000 neonates screened.
Note that the prevalence of PKA and of hypothyroidism in the UK is about 1 : 5000 live births, and that about 800 000 neonates in the UK are screened annually.

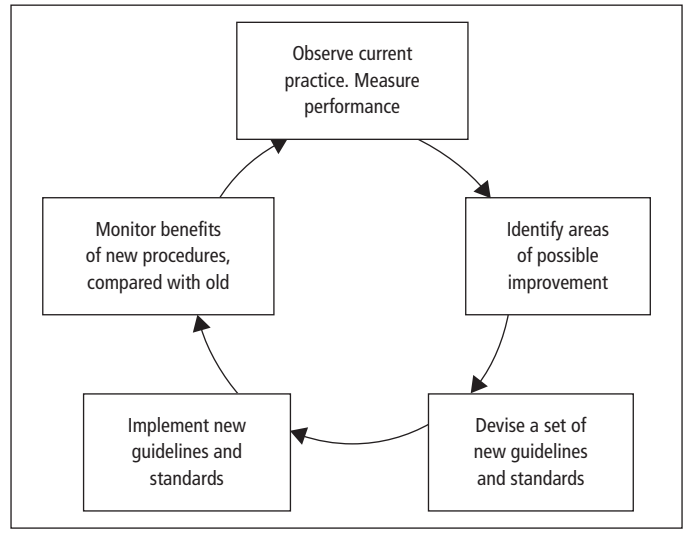


Figure 1.6 The audit cycle.

Audit in clinical biochemistry

Audit is the process whereby the procedures involved in patient care are monitored in order to give high priority to the delivery of an efficient and cost-effective service. The measure of health outcome is benefit to the patient.

The value of audit can most readily be seen in those specialties concerned directly with patient care, but the principles are applicable to all clinical and investigational specialties (e.g. radiology), as well as laboratory-based specialties such as clinical biochemistry. For example, the monitoring of laboratory performance may identifiy that reports are arriving too late and too often at the wrong location. This would precipitate a review of the form printing and delivery process, implementation of a change in the arrangements and a re-monitoring of the delivery process to ensure that the original problem had been overcome.

The audit process

There is an essential sequence to auditing activities (Figure 1.6):

1 Identify an area of concern or interest, particularly if it is felt that there is room for improvement in the service, or if the same quality of service can be provided more economically.
2 Review and analyse the present procedures.
3 Identify specific aspects that might be capable of improvement.
4 Identify alternative procedures or standards that might lead to improvement.
5 Take the practical steps necessary to implement any changes proposed.
6 Compare the performance after the changes with those before them.
7 It must be emphasised that the final stage of analysis of the effects of any change is an integral part of the audit process; it is essential to know whether the measures taken have improved the service or made it more cost-effective. Sometimes, changes have no effect, or even adverse effects.

Reference

Asher, R. (1954) Straight and crooked thinking in medicine. *British Medical Journal*, **2**, 460–2.

Keypoints

- As a rule, only those tests that may contribute to diagnosis and management of patients should be requested.
- Population screening is indicated where reliable tests are able to detect important and treatable disease; case-finding screening is valuable in groups at high risk of a specific disorder.
- There is an overlap between test results obtained in health and those obtained in disease or (where the test is being used for screening) results from affected individuals.
- Effective tests reduce this overlap to a minimum, but there will always be a trade-off between maximising the sensitivity of a test, that is, detecting as many affected individuals as possible, and maximising its specificity, that is, minimising the number of non-affected individuals classified as abnormal by the test.
- Strategies to maximise a test's value depend both on the test itself and on the prevalence of the disease in the population being studied.

Case 1.1

A new test is marketed which claims to diagnose heart failure. The test has a specificity of 70% and a sensitivity of 95% at the manufacturer's recommended cut-off for diagnosis. The Admissions Unit decides to use the test as part of an admission profile on breathless patients admitted for further assessment over the age of 65 years in order to exclude heart failure. Assuming a prevalence of 20% for heart failure in this population, calculate how many false negatives would be recorded after the first 1000 patients meeting the testing criteria had passed through the unit. Given that other tests can be used to establish a diagnosis of heart failure, do you think that the cut-off selected is sensible? (prevalence figures are for illustrative purposes only).

Comment: This is best examined by constructing a table as follows:

	Positive results	Negative results	Totals
Heart failure present	190 TP	10 FN	200
Heart failure absent	240 FP	560 TN	800
Total	430	570	1000

Because the test has a relatively high sensitivity, the table shows that it identifies the majority of patients with heart failure which is what is required in a test to rule out heart failure. Because the test lacks specificity, it can also be seen from the table that it identifies a considerable number of patients with positive results who do not have heart failure. In fact, the test is positive on more occasions in patients who do not have heart failure than in those with heart failure. Because other tests are available to the clinician, the false-positive patients can be separated from the true-positive patients on the basis of these further investigations. The 560 patients where the result is a true negative would then not need to go through more expensive further investigations. In this example, the test has been valuable in ruling out patients who would not require further investigation but ruling in those who would benefit. Clearly, it is not a perfect test but would potentially prevent costly further investigations in a significant number of patients and, provided that the test itself is not too expensive, ultimately be worthy of consideration in terms of health economics.

Case 1.2

A 72-year-old man is admitted vaguely unwell with some nausea and associated vomiting, though not severe. He appears rather pale and wasted with a low blood pressure. He is on treatment with digoxin for his atrial fibrillation and the suspicion arises that his symptoms may arise from digoxin toxicity. This would also help explain the raised potassium result for which there is no other clear cause. The most recent digoxin dose had been taken just before his admission to the hospital. The house officer telephones to request an additional digoxin measurement on the admission sample and this is reported back as raised. On this basis, the digoxin is withheld and his condition monitored. Little improvement is noted and the nausea becomes worse, accompanied by a worsening of his atrial fibrillation. Further advice is sought.

Comment on this case with particular reference to the raised digoxin and the worsening of his atrial fibrillation.

Comment: The timing of a blood test is crucial to the interpretation of a number of drugs whose concentration in blood is monitored for therapeutic purposes. This is most certainly the case with digoxin where the blood sample should *not* be taken within 6 h of the most recent digoxin dose. The House Officer has requested digoxin as an additional test on the patient's admission sample, without reference to the exact time when the patient took his dose of digoxin prior to admission. In fact, the time elapsed between taking the drug and the blood sample was about 1 h. The raised digoxin concentration is uninterpretable and it may well be that the patient has digoxin levels within the therapeutic range or even on the low side. This turned out to be the case, explaining the worsening in his condition when the drug was inappropriately withheld.

An isolated raised potassium result can be a very important finding which reflects underlying pathology such as renal disease, DKA, etc. Although there was no immediate explanation for this man's raised potassium, it became evident what the problem was when the full blood count report was received. This showed a very high lymphocyte count consistent with chronic lymphocytic leukaemia. In this condition, the white cells are fragile and can lyse on blood sampling. With the high white cell count, it is then possible to measure a spuriously high potassium level in the corresponding biochemistry sample.

Case 1.3

The following results were obtained on a 54-year-old woman after surgery for ovarian cancer. Can you account for the abnormalities found?

Serum	Result	Reference range
Urea	2.0	2.5–6.6 mmol/L
Sodium	147	135–145 mmol/L
Potassium	2.0	3.6–5.0 mmol/L
Total CO_2	10.0	22–30 mmol/L
Bilirubin	7.0	3–16 μmol/L
ALT	11.0	10–50 U/L
ALP	35.0	40–125 U/L
Total protein	42.0	60–80 g/L
Calcium	1.6	2.1–2.6 mmol/L

Comments: Many of these results are abnormal and, with the exception of the sodium result, are abnormally low. In a post-operative patient, a set of results like this should immediately raise the suspicion that the blood sample was taken close to the site of an IV infusion. The fluid infused would dilute the blood at the site of sampling, leading to a consequent lowering of the concentration of all the analytes measured. If the IV infusion was normal saline, this would then account for the fact that only the sodium value is high while all the other values are low. When the Duty Biochemist contacted the House Officer on the ward, he did admit that he had had difficulty taking a blood sample from the patient and did recollect that he sampled from close to the site of the IV infusion. A repeat blood sample was requested from a site away from the infusion and confirmed the original error since all the results were within the reference range, apart from the sodium which was slightly low at 132 mmol/L.

Case 1.4

The following set of results was obtained on a young man admitted with a fractured femur after a motorcycle accident. He appeared stable and had no previous past medical history of note. The houseman was at a loss to explain the results but remembered that he had topped up the sample shortfall in the Biochemistry tube from the haematology full blood count tube. Can you account for the results?

Serum	Result	Reference range
Urea	6.4	2.5–6.6 mmol/L
Sodium	138	135–145 mmol/L
Potassium	16.1	3.6–5.0 mmol/L
Total CO_2	32	22–30 mmol/L
Bilirubin	14	3–16 μmol/L
ALT	40	10–50 U/L
ALP	38	40–125 U/L
Total protein	75	60–80 g/L
Calcium	0.6	2.1–2.6 mmol/L
Albumin	32	35–50 g/L

Comments: This particular case illustrates the importance of using the correct blood sample tube. In transferring some of the blood from the Haematology tube to the Biochemistry tube, the doctor had not appreciated that the anti-coagulant in the Haematology (pink) tube was potassium EDTA. This explains the high potassium and the low calcium since the EDTA chelates the calcium. leading to a low result on analysis.

Chapter 2

Disturbances of water, sodium and potassium balance

Fluid loss, retention or redistribution are common clinical problems in many areas of clinical practice. The management of these conditions is often urgent, and requires a rapid assessment of the history and examination, and of biochemical and other investigations.

In this chapter we consider:

- The distribution of water, Na^+ and K^+ in the different fluid compartments of the body, and their control by hormonal and other factors.
- Clinical effects and management of different types of loss, retention or redistribution of fluid.
- Causes and investigation of hypernatraemia and hyponatraemia.
- Causes and investigation of hyperkalaemia and hypokalaemia.
- Fluid and electrolyte problems in surgical patients, and the metabolic response to trauma.

Water and sodium balance

Both the internal and external balance of any substance must be considered. The internal balance is the distribution between different body compartments, while the external balance matches input with output. The movements of Na^+ and water that occur all the time between plasma and glomerular filtrate, or between plasma and gastrointestinal (GI) secretions, provide the potential for large losses, with consequent serious and rapid alterations in internal balance. For example, about 25 000 mmol of Na^+ are filtered at the glomerulus over 24 h, normally with subsequent reabsorption of more than 99%. Likewise, 1000 of mmol Na^+ enter the GI tract in various secretions each day, but less than 0.5% (5 mmol) is normally lost in the faeces.

Internal distribution of water and sodium

In a 70 kg adult, the total body water is about 42 L comprising about 28 L of intracellular fluid (ICF) and 14 L of extracellular fluid (ECF) water. The ECF water is distributed as 3 L of plasma water and 11 L of interstitial water. The total body Na^+ is about 4200 mmol and is mainly extracellular – about 50% is in the ECF, 40% in bone and 10% in the ICF.

Two important factors influence the distribution of fluid between the ICF and the intravascular and extravascular compartments of the ECF:

1 *Osmolality* This affects the movement of water across cell membranes.

2 *Colloid osmotic pressure* Together with hydrodynamic factors, this affects the movement of water and low molecular mass solutes (predominantly NaCl) between the intravascular and extravascular compartments.

Lecture Notes: Clinical Biochemistry, 8e. By G. Beckett, S. Walker, P. Rae & P. Ashby. Published 2010 by Blackwell Publishing.

Osmolality, osmolarity and tonicity

The *osmolality* is the number of solute particles per unit weight of water, irrespective of the size or nature of the particles. Therefore, a given weight of low molecular weight solutes contributes much more to the osmolality than the same weight of high molecular weight solutes. The units are mmol/kg of water. This determines the osmotic pressure exerted by a solution across a membrane. Most laboratories can measure plasma osmolality, but it is also possible to calculate the approximate osmolality of plasma using a number of formulae of varying complexity. The following formula has the benefit of being easy to calculate and performs as well as more complex versions (all concentrations must be in mmol/L):

$$\text{Calculated osmolality} = 2[Na^+] + 2[K^+] + [\text{glucose}] + [\text{urea}]$$

This formula includes all the low molecular weight solutes contributing to plasma osmolality. Values for Na^+ and K^+ are doubled so as to allow for their associated anions, such as chloride. The formula is approximate and is not a complete substitute for direct measurement. Calculated osmolality is usually close to measured osmolality, but they may differ considerably for two different types of reason. First, there may be large amounts of unmeasured low molecular mass solutes (e.g. ethanol) present in plasma. These will contribute to the measured osmolality, but will obviously not be taken into account in the osmolality calculated from this formula. This will cause an 'osmole gap', with measured osmolality being greater than calculated osmolality. The other cause of a discrepancy is when there is a gross increase in plasma protein or lipid concentration, both of which decrease the plasma water per unit volume. This affects some methods of measurement of $[Na^+]$, giving an artefactually low result ('pseudohyponatraemia', see p. 23). This will result in an erroneously low calculated osmolality.

The osmolality of urine is usually measured directly, but is also linearly related to its specific gravity (which can be measured using urine dipsticks), unless there are significant amounts of glucose, protein or X-ray contrast media present.

The *osmolarity* is the number of particles of solute per litre of solution. Its units are mmol/L. Its measurement or calculation has been largely replaced by osmolality.

Tonicity is a term often confused with osmolality. However, it should only be used in relation to the osmotic pressure due to those solutes (e.g. Na^+) that exert their effects across cell membranes, thereby causing movement of water into or out of the cells. Substances that can readily diffuse into cells down their concentration gradients (e.g. urea, alcohol) contribute to plasma osmolality but not to plasma tonicity, since after equilibration their concentration will be equal on both sides of the cell membrane. Tonicity is not readily measurable.

The tonicity of ICF and ECF equilibrate with one another by movement of water across cell membranes. An increase in ECF tonicity causes a reduction in ICF volume as water moves from the ICF to the ECF to equalise the tonicity of the two compartments, whereas a decrease in ECF tonicity causes an increase in ICF volume as water moves from the ECF to the ICF.

Colloid osmotic pressure (oncotic pressure)

The osmotic pressure exerted by plasma proteins across cell membranes is negligible compared with the osmotic pressure of a solution containing NaCl and other small molecules, since they are present in much lower molar concentrations. In contrast, small molecules diffuse freely across the capillary wall, and so are not osmotically active at this site, but plasma proteins do not readily do so. This means that plasma [protein] and hydrodynamic factors together determine the distribution of water and solutes across the capillary wall, and hence between the intravascular and interstitial compartments (Figure 2.1).

Regulation of external water balance

Typical daily intakes and outputs of water are given in Table 2.1. Water intake is largely a consequence

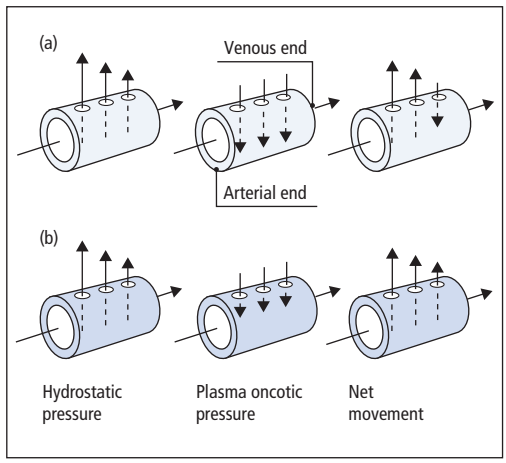


Figure 2.1 Movements of water and low molecular mass solutes across the capillary wall when the plasma [protein] is (a) normal and (b) low. The effects shown are: hydrostatic pressure, which drives water and low molecular mass solutes *outwards* and decreases along the length of the capillary; and plasma oncotic pressure, which attracts water and low molecular mass solutes *inwards* and is constant along the length of the capillary. The net movement of water and low molecular mass solutes across the capillary wall is governed by the net effect of hydrostatic and plasma oncotic pressures.

Table 2.1 Average daily water intake and output of a normal adult in the UK

Intake of water	mL	Output of water	mL
Water drunk	1500	Urine volume	1500
Water in food	750	Water content of faeces	50
Water from metabolism of food	250	Losses in expired air and insensible perspiration	950
Total	2500	Total	2500

of social habit and is very variable, but is also controlled by the sensation of thirst. Its output is controlled by the action of vasopressin, also known as antidiuretic hormone (ADH). In states of pure water deficiency, plasma tonicity increases, causing a sensation of thirst and stimulating vasopressin secretion, both mediated by hypothalamic osmoreceptors. Vasopressin then promotes water reabsorption in the distal nephron, with consequent production of small volumes of concentrated urine. Conversely, a large intake of water causes a fall in tonicity, suppresses thirst and reduces vasopressin secretion, leading to a diuresis, producing large volumes of dilute urine.

Secretion of vasopressin is normally controlled by small changes in ECF tonicity, but it is also under tonic inhibitory control from baroreceptors in the left atrium and great vessels on the left side of the heart. Where haemodynamic factors (e.g. excessive blood loss, heart failure) reduce the stretch on these receptors, often without an accompanying change in ECF tonicity, a reduction in tonic inhibitory control stimulates vasopressin secretion. The resulting water retention causes hyponatraemia and is relatively ineffective in expanding the intravascular compartment, since water diffuses freely throughout all compartments (Figure 2.2).

Regulation of external sodium balance

Dietary intakes of Na^+ (and Cl^-) are very variable worldwide. A typical 'Western' diet provides 100–200 mmol of both Na^+ and Cl^- daily, but the total body Na^+ can be maintained even if intake is less than 5 mmol or greater than 750 mmol daily. Urinary losses of Na^+ normally closely match intake. There is normally little loss of these ions through the skin or in the faeces, but in disease the GI tract can become a major source of Na^+ loss.

The amount of Na^+ excreted in the urine controls the ECF volume since, when osmoregulation is normal, the amount of extracellular water is controlled to maintain a constant concentration of extracellular Na^+. A number of mechanisms are important regulators of Na^+ excretion:

- *The renin–angiotensin–aldosterone system* Renin is secreted in response to a fall in renal afferent arteriolar pressure or to a reduction in supply of Na^+ to the distal tubule. It converts angiotensinogen in plasma to angiotensin I (AI), which in turn is converted to angiotensin II (AII) by angiotensin-converting enzyme (ACE). Both AII and its metabolic product angiotensin III (AIII) are pharmacologi-

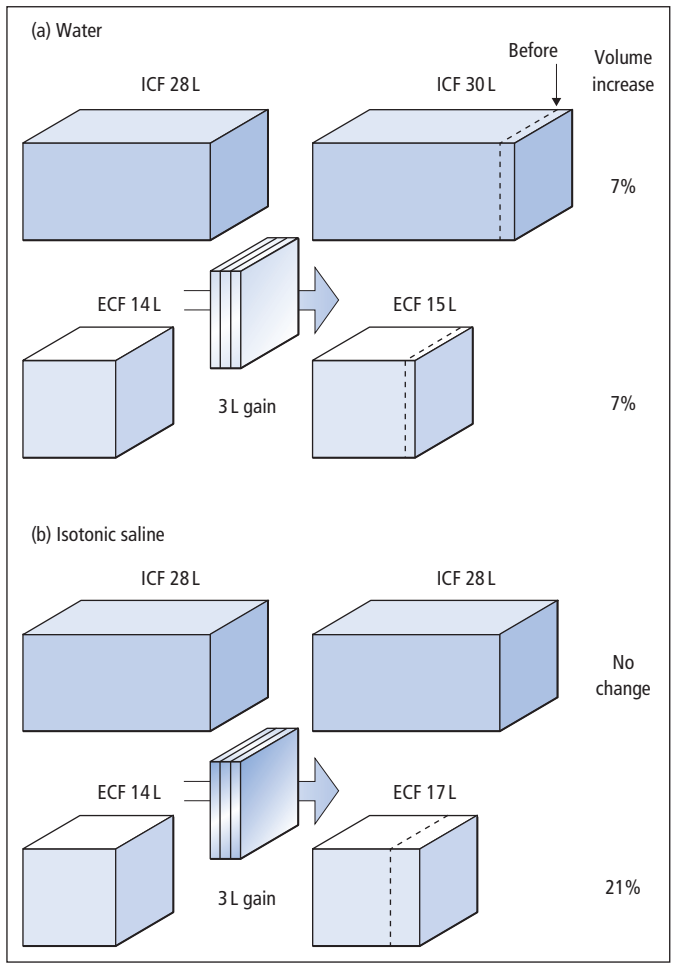


Figure 2.2 Different effects on the body's fluid compartments of fluid gains of 3 L of (a) water and (b) isotonic saline. The volumes shown relate to a 70 kg adult.

cally active, and stimulate the release of aldosterone from the adrenal cortex. Aldosterone acts on the distal tubule to promote Na^+ reabsorption in exchange for urinary loss of H^+ or K^+. Since Na^+ cannot enter cells freely, its retention (with isosmotically associated water) contributes solely to ECF volume expansion, unlike pure water retention (Figures 2.2 and 2.3). Although the renin–angiotensin–aldosterone system causes relatively slow responses to Na^+ deprivation or Na^+ loading, evidence suggests that this is the main regulatory mechanism for Na^+ excretion.

- *The glomerular filtration rate (GFR)* The rate of Na^+ excretion is often related to the GFR. When the GFR falls acutely, less Na^+ is filtered and excreted, and vice versa. However, this only becomes a limiting factor in Na^+ excretion at very low levels of GFR.
- *Atrial natriuretic peptide (ANP)* This peptide secreted by cardiocytes of the right atrium of the heart promotes Na^+ excretion by the kidney, apparently by causing a marked increase in GFR. The importance of the ANP regulatory mechanism is not yet clear, but it probably only plays a minor role. Other structurally similar peptides have been identified, including brain or B-type natriuretic peptide (BNP), secreted by the cardiac ventricles and with similar properties to ANP. BNP is increas-