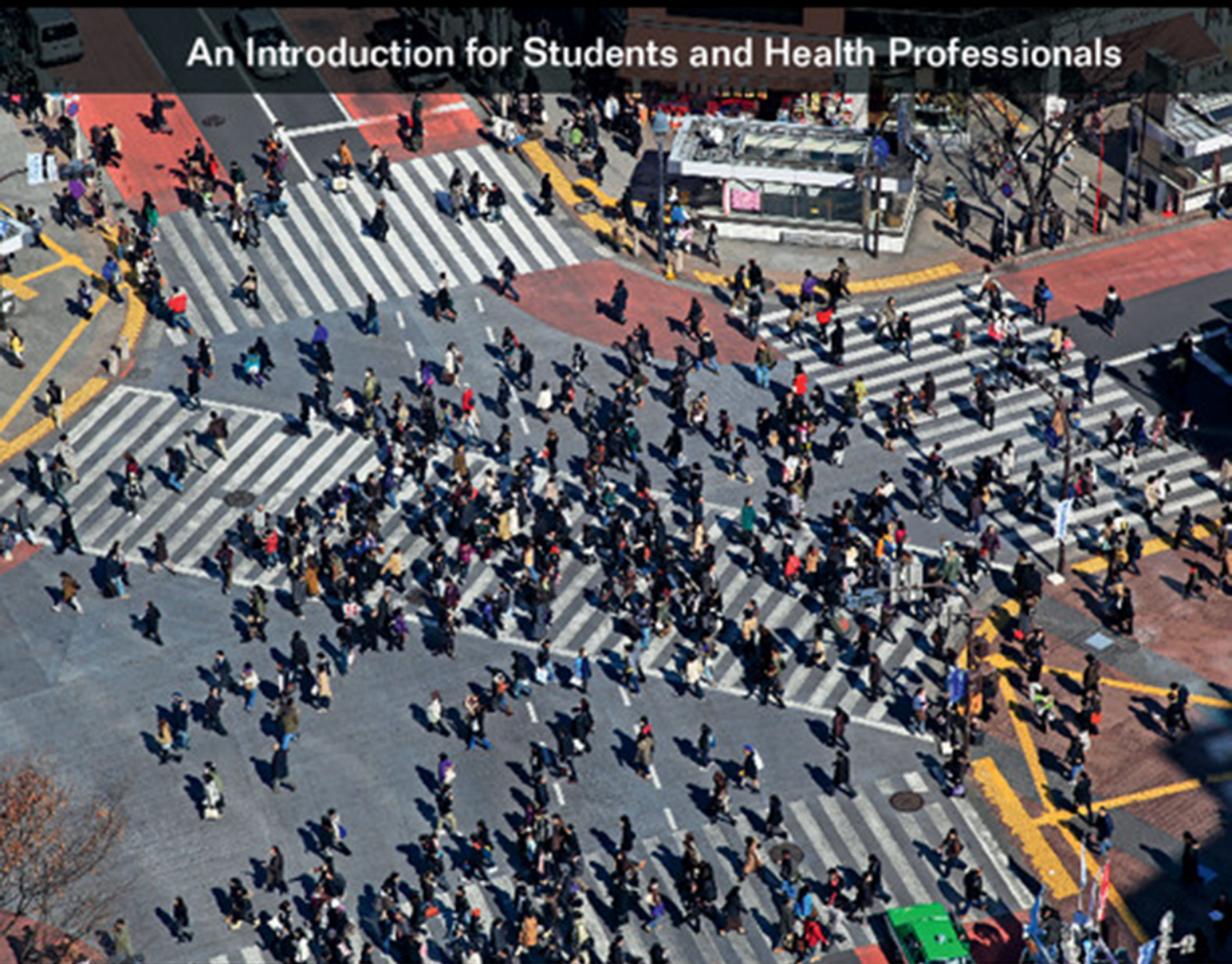


THIRD EDITION

Penny Webb | Chris Bain | Andrew Page

ESSENTIAL Epidemiology

An Introduction for Students and Health Professionals



CAMBRIDGE

Medicine

Essential Epidemiology

An Introduction for Students and Health Professionals

Third Edition

Penny Webb, MA (Cambridge), DPhil (Oxford), is a Principal Research Fellow and Group Leader at the QIMR Berghofer Medical Research Institute, Brisbane, and Honorary Professor in the School of Public Health, University of Queensland. She taught basic and intermediate epidemiology to public health students across Australia for 5 years and has worked as a visiting scientist at the International Agency for Research on Cancer, France, and Harvard University, USA. She has published more than 250 original research papers in the field of cancer epidemiology.

Chris Bain, MB BS (UQ), MPH, MSc (Harvard), formerly Reader in Epidemiology, University of Queensland, is currently a Visiting Scientist at the QIMR Berghofer Medical Research Institute and National Centre for Epidemiology and Population Health, Australian National University. He has taught epidemiology to public health and medical students for over 3 decades, and has co-authored a book on systematic reviews, as well as many research papers. He has had wide exposure to international epidemiological practice and teaching in the UK and US.

Andrew Page BA(Psych) Hons (Newcastle), PhD (Sydney) is Professor of Epidemiology in the School of Medicine at Western Sydney University, Australia. He has been teaching basic and intermediate epidemiology and population health courses to health sciences students for 10 years, and has published over 140 research articles and reports across a diverse range of population health topics. He has been a Research Associate at the University of Bristol and has also worked at the University of Queensland and University of Sydney in Australia.

Essential Epidemiology

An Introduction for Students
and Health Professionals

Third Edition

Penny Webb

Principal Research Fellow,
QIMR Berghofer Medical Research Institute and
Honorary Professor,
University of Queensland, Brisbane, Australia.

Chris Bain

Visiting Scientist,
QIMR Berghofer Medical Research Institute,
Brisbane, Australia.

Andrew Page

Professor of Epidemiology,
School of Medicine,
Western Sydney University, Australia.



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi - 110025, India
79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107529151

© P. Webb, C. Bain and S. Pirozzo 2005

© P. Webb and C. Bain 2011

© Penny Webb, Chris Bain and Andrew Page 2017

This publication is copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in 2005

Second edition 2011

Third edition 2017 (version 2, December 2017)

Typeset by SPi Global

Printed in Singapore by Markono Print Media Pte Ltd, November 2017

A catalogue record for this publication is available from the British Library

A Cataloguing-in-Publication entry is available from the catalogue of the National Library of Australia at www.nla.gov.au

ISBN 978-1-107-52915-1 Paperback

Additional resources for this publication at www.cambridge.org/9781107529151

Reproduction and communication for educational purposes

The Australian *Copyright Act 1968* (the Act) allows a maximum of one chapter or 10% of the pages of this work, whichever is the greater, to be reproduced and/or communicated by any educational institution for its educational purposes provided that the educational institution (or the body that administers it) has given a remuneration notice to Copyright Agency Limited (CAL) under the Act.

For details of the CAL licence for educational institutions contact:

Copyright Agency Limited
Level 15, 233 Castlereagh Street
Sydney NSW 2000
Telephone: (02) 9394 7600
Facsimile: (02) 9394 7601
E-mail: info@copyright.com.au

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

.....
Every effort has been made in preparing this book to provide accurate and up-to-date information that is in accord with accepted standards and practice at the time of publication. Although case histories are drawn from actual cases, every effort has been made to disguise the identities of the individuals involved. Nevertheless, the authors, editors and publishers can make no warranties that the information contained herein is totally free from error, not least because clinical standards are constantly changing through research and regulation. The authors, editors and publishers therefore disclaim all liability for direct or consequential damages resulting from the use of material contained in this book. Readers are strongly advised to pay careful attention to information provided by the manufacturer of any drugs or equipment that they plan to use.

Contents

Foreword	<i>page</i> ix
Preface	xi
List of Contributors	xv
1 Epidemiology is ...	1
A case of food poisoning	3
Subdisciplines of epidemiology	4
On epidemics	6
An historical epidemic	8
The beginnings	11
What does epidemiology offer?	17
What do epidemiologists do?	19
A natural experiment	24
Conclusions	27
2 How long is a piece of string? Measuring disease frequency	31
What are we measuring?	33
The concepts: prevalence and incidence	34
Measuring disease occurrence in practice: epidemiological studies	41
Measuring disease occurrence in practice: using routine data	48
Other measures commonly used in public health	55
Global health indicators	59
Summary	72
Questions	72
3 Who, what, where and when? Descriptive epidemiology	75
Case reports and case series	76
Vital statistics and mortality data	77
Morbidity data	86
Creative use of existing data	93
Confidentiality	100
Summary	100
Questions	100
4 Healthy research: study designs for public health	104
The ideal study	106
Intervention studies or trials	108

Observational studies	114
A word about ethics	130
Summary	133
Questions	135
5 Why? Linking exposure and disease	139
Looking for associations	140
Ratio measures (relative risk)	142
Difference measures (attributable risk)	147
Relative risk versus attributable risk: an example	155
Case-control studies	156
Looking for associations when the measures are continuous	163
Summary	163
Questions	165
6 Heads or tails: the role of chance	169
Random sampling error	170
Statistical significance: could an apparent association have arisen by chance?	171
Confidence intervals	173
Power: could we have missed a true association?	175
Interpreting p -values and confidence intervals	177
Statistical versus clinical significance	181
Summary	183
Questions	183
7 All that glitters is not gold: the problem of error	185
Sources of error in epidemiological studies	187
Selection bias	187
Measurement or information error	200
Summary	213
Questions	214
8 Muddied waters: the challenge of confounding	217
An example of confounding: is alcohol a risk factor for lung cancer?	219
Characteristics of a confounder	221
The Effects of confounding	223
Control of confounding	228
Confounding: the bottom line	243
Questions	244
9 Reading between the lines: reading and writing epidemiological papers	248
The research question and study design	250

Internal validity	251
So what? are the results important?	260
Generalisability (external validity)	262
Descriptive studies	263
Writing papers	264
Summary: one swallow doesn't make a summer	265
Questions	266
10 Who sank the boat? Association and causation	269
What do we mean by a cause?	271
Association versus causation	275
Evaluating causation	276
An example: does <i>H. pylori</i> cause stomach cancer?	285
Conclusion	286
Questions	286
11 Assembling the building blocks: reviews and their uses	288
What is a systematic review?	289
Identifying the literature	291
Appraising the literature	292
Summarising the data	294
Drawing conclusions	300
Assessing the quality of a systematic review	303
Making judgements in practice	303
The end result	307
Conclusion	309
Questions	310
12 Surveillance: collecting health-related data for epidemiological intelligence and public health action	313
<i>Martyn Kirk and Adrian Sleigh</i>	
The scope of surveillance	314
Why conduct surveillance?	315
Surveillance essentials	316
Types of surveillance	322
Summary	329
Questions	329
13 Outbreaks, epidemics and clusters	333
<i>Martyn Kirk and Adrian Sleigh</i>	
Outbreaks, epidemics, and clusters	335
Epidemiology of infectious diseases	336
Non-infectious clusters and outbreaks	343
Outbreak management and investigation	345

Evidence for causation	350
Summary	352
Questions	353
14 Prevention: better than cure?	356
Disease prevention in public health	357
The scope for preventive medicine	362
Strategies for prevention	365
The population attributable fraction as a guide to prevention	370
Prevention in practice	373
Evaluation of preventive interventions in practice	376
A final (cautionary) word	378
Questions	378
15 Early detection: what benefits at what cost?	381
Why screen?	382
The requirements of a screening programme	385
Evaluation of a screening programme	401
Summary	412
Questions	413
16 Epidemiology and the public's health	417
Translating epidemiological research into practice	418
Challenges	420
Synthesis and integration	422
Limiting error	424
Improving measurement	424
A final word	427
Answers to questions	430
Appendix 1: Direct standardisation	451
Appendix 2: Standard populations	453
Appendix 3: Calculating risk and lifetime risk from routine data	455
Appendix 4: Indirect standardisation	457
Appendix 5: Calculating life expectancy from a life table	459
Appendix 6: Why the odds ratio approximates the relative risk for a rare disease	461
Appendix 7: Formulae for calculating confidence intervals for common epidemiological measures	462
Appendix 8: The Mantel-Haenszel method for calculating pooled odds ratios	464
Glossary	466
Index	483

I am delighted to write the new foreword for the third edition of *Essential Epidemiology*.

This well known, highly respected and engaging book, aimed at the Masters level in Epidemiology, is a timely evolution from the earlier editions, bringing in new material and educational approaches (fully described in the authors' preface).

As someone who has been heavily involved in my own institution's undergraduate, masters, and doctoral level research and service training over many decades, and been external examiner for many other institutions, I see the value of this book to many international constituencies. The key audience will, of course, be those involved in Masters studies of Epidemiology, Public Health and related disciplines. However I would also recommend this book to those involved in undergraduate teaching of epidemiology, students and teachers, those in post-Master's training or practice, as well as those in the biomedical and social sciences who wish to understand and utilise the perspectives and principles of epidemiology.

The value of the book is that the authors have based it on many years of teaching students on the ground and the latest revision and refreshment ensures that it maintains relevance. Thus this book retains the major content of the earlier versions, with sound grounding in the core principles and practice of epidemiology, as well as incorporating new areas. It is vital that future epidemiological research is relevant to the challenges we face globally. The book continues to provide this wider perspective, as well as the more technical approaches that are used when merged with other fields such as genetics. This new edition also provides on-line further materials (including expanding on some trickier methodological topics) and full teaching materials (more on questions and answers, lecture slides), which allow students to engage in more active learning and teachers to draw on presentations which they can use and adapt.

As is clear I recommend this book strongly to those in relevant training and those involved in their education as an up to date, highly accessible and excellent resource.

Carol Brayne
Professor of Public Health Medicine
University of Cambridge

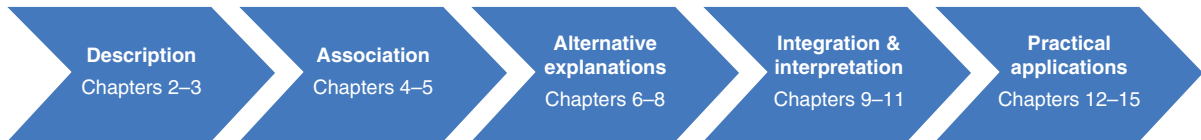
This book grew out of our collective experience of teaching introductory epidemiology both in the classroom and to distance students enrolled in public health and health studies programmes in the School of Public Health (formerly the Department of Social and Preventive Medicine and then School of Population Health), University of Queensland. It began life as a detailed set of course notes that we wrote because we could not find a single epidemiology text that covered all of the areas we felt were important in sufficient detail. As the notes were to be used primarily by distance students, we tried hard to make them accessible with lots of examples, minimal jargon and equations, and by engaging readers in ‘doing’ epidemiology along the way. Feedback from students and colleagues convinced us that the notes were both approachable and practical and the result is this text, which we offer as a practical introduction to epidemiology for those who need an understanding of health data they meet in their everyday working lives, as well as for those who wish to pursue a career in epidemiology.

The first revision of the text reflected evolution, not revolution. We listened to the feedback we received from instructors and students and tried to simplify and clarify some of the trickier bits of the original text while maintaining a very ‘hands-on’ approach. We added new material to reflect contemporary epidemiological practice in public health and re-ordered some of the existing elements to improve the flow and enhance the continuity between chapters. New and expanded topics included a look at how we measure the burden of disease, greater discussion of issues relevant to ethics and privacy, appendices covering life tables and calculation of confidence intervals for common epidemiological measures, and a glossary.

This, the third edition, reflects further evolution. With our new co-author Professor Andrew Page, and inspired by colleagues at a workshop on methods of teaching modern epidemiology convened by Professors Diana Safarti and John Lynch at the University of Otago, New Zealand in 2014, we have injected some more modern approaches to causal thinking, bias and confounding. These changes are most obvious in [Chapter 4](#) (Study Design), which we have restructured to show more clearly how each design contrasts with the ‘ideal’ (counterfactual) experiment, [Chapters 7](#) (Bias) and [8](#) (Confounding), and [Chapter 10](#) (Causation). A series of constructive reviews from teachers using the book helped us to identify and correct some faults, convinced us to retain the infectious disease elements of the text but in a more focussed and practical

form, and led us to add recommendations for ‘further reading’ for those who want deeper insights into some of the issues discussed. We have also refined the final chapter, which now builds on the experience of the earlier material to consider the role and value of epidemiology in translational research. Another major enhancement is the greatly expanded website, which provides the reader with access to additional examples and useful links, many of the references cited in the book (subject to copyright), additional questions with comprehensive worked answers and a ‘Test Yourself’ set of interactive multiple choice questions (and answers) for each of the main content chapters. For lecturers there are also more detailed sets of teaching slides for each chapter.

Our overall aims are, however, unchanged. Firstly, to give students a good understanding of the fundamental principles common to all areas of epidemiology, including the study of both infectious and chronic diseases as well as public health and clinical epidemiology, and to show the essential role of epidemiology in a broad range of health monitoring and research activities. Secondly, and perhaps more importantly, we have endeavoured to do this in a way that is both approachable and engaging, that minimises mathematical jargon and complex language without sacrificing accuracy, and that encourages study and stimulates epidemiological thought.



As previously, [Chapter 1](#) is a general introduction that both answers the question ‘what is epidemiology and what can it do?’ and presents the main concepts that are the focus of the rest of the book. The next chapters are divided into five separate sections. The first covers the basic principles and underlying theory of epidemiology in a very ‘hands-on’ way. We start by looking at how we can measure disease and, new to this edition, the overall burden of disease in a population ([Chapter 2](#)), followed by a look at the role of descriptive epidemiology in describing health patterns ([Chapter 3](#)). We move on to look at the types of study that we use to identify potential causes of disease, including an expanded discussion of the potential of record linkage ([Chapter 4](#)) and how we quantify the associations between cause and outcome ([Chapter 5](#)). In the next section we look at the role of chance in epidemiology ([Chapter 6](#)), consider the thorny issues of error and bias ([Chapter 7](#)) and give a practical overview of the problem of confounding ([Chapter 8](#)). This leads to

the third section, where we integrate this information in a practical look at how we read and interpret epidemiological reports (Chapter 9), think about assessing causation (Chapter 10), and finally synthesise a mass of information in to a single review to make practical judgements regarding the likelihood that a relation is causal (Chapter 11). In the final section we look at some specific applications of epidemiology, including its role in surveillance (Chapter 12), outbreak control (Chapter 13), prevention – including a discussion of how we can assess the impact of different preventive interventions on the health of a population (Chapter 14) and screening (Chapter 15). The greatly revised Chapter 16 then concludes by reviewing core concepts of the earlier material to address some of the challenges that face a modern epidemiologist who desires to improve health through ‘translation’ of research into practice.

Symbols

Throughout the book we have used **bold** typeface to indicate terms included in the glossary and the following symbols are used to define key elements within the text.

We strongly believe that the best way to learn anything is by actually doing it and so have included questions within the text for those who like to test their understanding as they go. Because we also know how frustrating it is to have to search for answers, we have provided these immediately following the questions for those in a hurry to proceed: The questions at the end of the chapters also have full worked answers at the end of the book.

We have used numerous real-life examples from all around the world to illustrate the key points and to provide additional insights in some areas. Extra examples that provide added interest and complement the main message in the text are given in boxes featuring this symbol.

Many books present clinical epidemiology as a separate discipline from public health epidemiology – a distinction that is strengthened by the fact that clinical epidemiologists have developed their own names for many standard epidemiological terms. In practice all epidemiology is based on the same underlying principles, so we have integrated the two approaches throughout the book but have also highlighted specific examples more relevant to the clinical situation. (Please note that this book does not offer a comprehensive coverage of clinical epidemiology; rather, we aim to show the similarity of the two areas where they overlap.)

We have deliberately tried to keep the main text free of unnecessary detail and equations, but have included some epidemiological ‘extras’. This material is not essential to the continuity of the core text but provides some additional information for those who like to see where things have come from or want a more detailed perspective.





New to this third edition, we have identified areas where additional material is available online; www.cambridge.org/9781107529151. This material includes additional reading, links to the papers that we have cited and additional questions and answers.

Acknowledgements

If we were to name everyone who had contributed in some way to this book, the list would be endless. We would, however, like to acknowledge some of the great teachers (and their books) from whom we have learned most of what we know, and the books we have relied on heavily for our teaching. These include Brian MacMahon (*Epidemiology: Principles and Methods*, MacMahon and Pugh, 1970), Olli Miettinen, Charlie Hennekens (*Epidemiology in Medicine*, Hennekens and Buring, 1987), Ken Rothman (*Modern Epidemiology*, 1986), *Foundations of Epidemiology* (Lillienfeld and Lilienfeld, 1980), and *Epidemiology* (Gordis, 1996). We would also like to thank our colleagues and friends, especially the Fellows from the then School of Population Health, University of Queensland, and the staff and students from the then Cancer and Population Studies Group at the QIMR Berghofer Medical Research Institute, whose constructive feedback helped shape the first edition back in 1995.

Particular thanks go to our former colleague and co-author of the first edition, Sandi Pirozzo, who has since moved on to a rewarding new career post-epidemiology; we remain grateful for her prior contributions. Also to Adrian Sleight (Australian National University) who wrote the original chapters on Outbreak Investigation and Surveillance and to Martyn Kirk (Australian National University) who extensively revised these chapters for this third edition. We are grateful to members of the former Burden of Disease group at the School of Public Health, University of Queensland, especially Theo Vos, Stephen Begg and Alan Lopez for their suggestions regarding our consideration of the 'Burden of Disease' for the second edition, and to Chalapati Rao (Australian National University) whose constructive feedback helped us to update [Chapters 2 and 3](#) for this edition. The excellent critiques and suggestions we received from Michael O'Brien and Kate Van Dooren regarding the first edition helped improve the cohesion and internal 'sign-posting' of the book. We also thank Susan Jordan, Kate Van Dooren and Keren Papier, who helped pull everything together for the first, second and third editions, respectively. Finally, we would like to acknowledge the School of Public Health, University of Queensland, which provided the intellectual environment that led to us writing this book in the first place, the team from Otago University in New Zealand who provided very helpful feedback on the first edition, and the many users of the first two editions who provided the critical feedback that has directly led to this new and hopefully improved third edition.

Contributors

Martyn Kirk

Associate Professor, National Centre for Epidemiology and Population Health, Research School of Population Health, Australian National University, Canberra, Australia

Adrian Sleigh

Emeritus Professor, National Centre for Epidemiology and Population Health, Research School of Population Health, Australian National University, Canberra, Australia

Epidemiology is ...

A case of food poisoning	3
Subdisciplines of epidemiology	4
On epidemics	6
An historical epidemic	8
The beginnings	11
What does epidemiology offer?	17
Description of health status of populations	18
Causation	18
Evaluation of interventions	18
Natural history and prognosis	18
What do epidemiologists do?	19
Descriptive studies: person, place and time	19
By 'person'	19
By 'place'	20
By 'time'	20
Analytic studies	23
Intervention Studies	24
A natural experiment	24
Conclusions	27

Box 1.1 Epidemiology is ...

'The science of epidemics' (*Concise Oxford Dictionary*, 1964)

'The science of the occurrence of illness' (Miettinen, 1978)

'The study of the **distribution** and **determinants** of disease in humans'
(MacMahon and Pugh, 1970)

'The study of the distribution and determinants of **health-related states or events** in specified populations, and the **application of this study to control of health problems**' (Porta, 2008)

'The study of the occurrence and distribution of health-related events, states and processes in specified populations, including the study of the determinants influencing such processes, and the application of this knowledge to control relevant health problems' (Porta, 2014)

So what is epidemiology anyway? As shown in [Box 1.1](#), the *Concise Oxford Dictionary* (1964) defined it accurately, but not very helpfully, as ‘the science of epidemics’. In 1970, MacMahon and Pugh came up with something a bit more concrete: ‘the study of the *distribution* and *determinants* of disease’. Their definition succinctly identifies the two core strands of traditional epidemiology: *who* is developing disease (and *where* and *when*), and *why* are they developing it? The next definition, from the 2008 edition of the *Dictionary of Epidemiology* (Porta, 2008), takes things two steps further by broadening the scope to include health in general, not just disease, as well as highlighting the essential role of epidemiology in translating research findings into health policy and medical practice to control disease. The most recent definition (Porta, 2014) elaborates further still but, in doing so, loses some of the elegance of the earlier versions.

Epidemiology, therefore, is about measuring disease or other aspects of health, identifying the causes of ill-health and intervening to improve health; but what do we mean by ‘health’? Back in 1948, the World Health Organization (WHO, 1948) defined it as ‘... a state of physical, mental and social well-being’. In practice, what we usually measure is physical health, and this focus is reflected in the content of most routine reports of health data and in many of the health measures that we will consider here. However, methods that attempt to capture the more elusive components of mental and social well-being are now emerging. Instead of simply measuring ‘life expectancy’, WHO introduced the concepts of ‘health-adjusted life expectancy’ (HALE) and subsequently ‘disability-adjusted life years’ (DALYs) to allow better international comparisons of the effectiveness of health systems. In doing so, they recognised that it is not longevity per se that we seek, but a long and healthy life. We will discuss these and other measures in more detail in [Chapter 2](#).

Perhaps epidemiology’s most fundamental role is to provide a logic and structure for the analysis of health problems both large and small or, as described by Wade Hampton Frost, epidemiology involves the ‘orderly arrangement of [established facts] into chains of inference which extend more or less beyond the bounds of direct observation’ (Frost, 1927). It emphasises the sound use of numbers – we have to count and we have to think. We have to think about what is worth counting and how best to count it, about what is practical and, importantly, about how well we (or others) finally measured whatever it was we set out to measure, and what it all means. Accurate measurement of health is clearly the cornerstone of the discipline, but we believe the special value of epidemiology flows from a way of thought that is open, alert to the potential for error, willing to consider alternative explanations and, finally, constructively critical and pragmatic.

We offer this book as an aid to such thought. It does not aim to turn you into a practising epidemiologist overnight, but will give clear directions if that is where you decide to go. Its primary goal is to help you interpret the mass of

Table 1.1 Numbers of people who became ill after eating various foods at a youth camp.

Food	People who ate the food		People who didn't eat the food	
	Total	Number ill	Total	Number ill
Friday dinner:				
Hot chicken	343	156	231	74
Peas	390	175	184	55
Potato fries	422	184	152	46
Saturday lunch:				
Cold chicken	202	155	372	75
Salad	385	171	189	59
Saturday dinner:				
Fruit salad	324	146	250	84

(Adapted from Hook *et al.*, 1996, with permission from John Wiley and Sons. © 1996 The Public Health Association of Australia Inc.)

epidemiological literature and the various types of health data that you may come across. We hope that you will see, by reading and by doing, that the fundamental concepts and tools of epidemiology are relatively simple, although the tasks of integrating, synthesising and interpreting health information are more challenging. But before we go any further, let us do some public health epidemiology.

A case of food poisoning

Epidemiology is a bit like detective work in that we try to find out why and how disease occurs. Our first example illustrates this. After an outbreak of food poisoning at a youth camp, the local public health unit was called in to identify the cause (Hook *et al.*, 1996). They first asked everyone at the camp what they had eaten prior to the outbreak and some results of this investigation are shown in Table 1.1.

Looking at the numbers in Table 1.1, it is difficult to see which of the foods might have been responsible for the outbreak. (Note that everyone is recorded as either having eaten or not eaten each food; and that most people will have eaten more than one of the foods.) More people became ill after eating potato fries than after eating cold chicken (184 versus 155) – but then more people ate the fries (422 versus 202). How then can we best compare the two foods? One simple way to do this is to calculate the *percentage* of people who became ill among those who ate (or did not eat) each type of food. For example, 156 out of 343 people who ate hot chicken became ill and

$$156 \div 343 = 0.45 = 45\%$$

So 45% of people who ate hot chicken became sick. This is known as the **attack rate** for hot chicken, i.e. 45% of hot chicken eaters were 'attacked' by food poisoning.



Calculate the attack rates for the other foods. Which food has the highest attack rate?

Although cold chicken has the highest attack rate (77%), not everyone who ate it (or, more precisely, who *reported* eating it) became ill and 20% or one in five people who did *not* eat cold chicken still became ill. This is to be expected; no matter what the cause of concern, it is rare that everyone who is exposed to it will show the effects (in this case, become ill). What can help here is to work out how much *more likely* people who ate a particular food were to become ill than those who did not eat it. For example, 45% of people who ate hot chicken became ill, compared with 32% of people who did not eat hot chicken. Hot-chicken eaters were therefore 1.4 times ($45\% \div 32\% = 1.4$) more likely to become ill than people who did not eat hot chicken. This measure gives us the risk of sickness in hot-chicken eaters *relative* to non-eaters, hence its name – **relative risk**.



Calculate the relative risk of developing food poisoning associated with each of the other food items. Which food is associated with the highest relative risk of sickness?

We can now conclude that the food item most likely to have been responsible for the outbreak was the cold chicken – people who ate this were almost four times as likely to become ill as those who did not. This is quite a strong relative risk; in comparison, eating any of the other foods was associated with no more than one and a half times the risk of disease. The relevant data, including the attack rates and relative risks, are summarised in [Table 1.2](#), which is much more informative than the raw numbers of [Table 1.1](#).

In identifying the cause of the outbreak you have just solved an epidemiological problem. The 'attack rates' and 'relative risks' that you used are simple to calculate and are two very useful epidemiological measures. We will discuss them further in [Chapters 2 and 5](#) and they will appear throughout the book.

Subdisciplines of epidemiology

The outbreak investigation above is an example of what might be called *public health epidemiology*, or *infectious disease epidemiology*, with the first name reflecting the broad field of application and the second the nature of both the aetiological (causal) agent and the disease. It is quite common now to specify such subfields of epidemiology, which range on the one hand from *nutritional* through *social* to *environmental* and *eco-epidemiology*, and on the other from

Table 1.2 Numbers of people who became ill after eating various foods at a youth camp and attack rates and relative risks for each food.

Food	People who ate the food			People who didn't eat the food			Relative risk ^a
	Total	Number ill	Attack rate	Total	Number ill	Attack rate	
Friday dinner:							
Hot chicken	343	156	45%	231	74	32%	1.4
Peas	390	175	45%	184	55	30%	1.5
Potato fries	422	184	44%	152	46	30%	1.4
Saturday lunch:							
Cold chicken	202	155	77%	372	75	20%	3.8
Salad	385	171	44%	189	59	31%	1.4
Saturday dinner:							
Fruit salad	324	146	45%	250	84	34%	1.3

^a Note, relative risks are calculated using the exact percentages and not the rounded values shown.

(Adapted from Hook *et al.*, 1996, with permission from John Wiley and Sons. © 1996 The Public Health Association of Australia Inc.)

cancer to injury or *perinatal epidemiology*: the former grouping being exposure-oriented and the latter focused on the particular disease or outcome. Nonetheless, the core methods and techniques of epidemiology remain common to all subdisciplines, so the contents of this book are relevant to all. Setting subspeciality boundaries largely reflects the explosion of knowledge in these areas, although some areas do present special challenges. For example, capturing a person's usual diet is remarkably challenging and the subsequent data analysis equally so; epidemiologists coming fresh to the field of nutritional epidemiology will need to develop experience and expertise in that specific area. You will meet examples from a wide cross-section of health research as you read on, and the common threads of logic, study design and interpretation will, we trust, become apparent.

It is of some interest to know a bit more about a few of the special epidemiologies. *Occupational epidemiology* has the longest history of all, with influential early observations of diseases linked to occupations such as mining appearing in the sixteenth century, and a systematic treatise on occupational diseases was published by Ramazzini back in 1700 (Rosen, 1958). Occupational health research in general, and epidemiology in particular, continue to contribute to enhancing workplace health today. Seminal contributions in the field include identification of the pulmonary (lung) hazards of asbestos for miners and construction workers (Selikoff *et al.*, 1965) and the work practices that led to an epidemic of a rare fatal cancer in workers in the polyvinyl chloride industry (Makk *et al.*, 1974). Company records of job tasks can

provide measures of past exposure among employees, allowing researchers to look back in time and link, for example, past asbestos exposure to subsequent deaths in the workforce. (This type of study is a *historical cohort design* – see [Chapter 4](#). It is only possible when there are good records of both exposure and outcome, usually death, and for this reason has proved particularly useful in occupational studies where such records often do exist.)

Far more modern are the subdisciplines of *molecular* and *clinical epidemiology*. The former aims to weld the population perspective of epidemiology with our rapidly increasing understanding of how variations in genes and their products affect the growth, form and function of cells and tissues. It thus has the potential to define genetic contributions to disease risk and can also provide biological markers of some exposures (e.g. changes to DNA following exposure to tobacco smoke). In contrast, clinical epidemiology differs from other branches of epidemiology in its focus on enhancing clinical decisions to benefit *individual patients*, rather than improving the health of *populations*. For this reason, clinical epidemiology is sometimes regarded as a separate discipline, a view encouraged by the fact that it has developed its own names for many standard epidemiological measures. The foundations are, however, identical to those of public health epidemiology and when appropriate we will discuss the two in parallel, highlighting any differences in language or approach along the way. There is also increasing interest in *lifecourse epidemiology*, which attempts to integrate events across the lifetime, often going right back to conception and sometimes to previous generations, to understand disease risk.

On epidemics

If we take the word ‘epidemiology’ itself, its origins from ‘*epidemic*’ are clear. If we talk about an epidemic we immediately conjure up pictures of an acute outbreak of infectious disease but, both for practical and for etymological reasons, it seems reasonable to use the term to describe a notable excess of any disease over time. Many developed countries could, for example, be described as undergoing an epidemic of lung cancer over the last few decades ([Figure 1.1](#)). Notably the pattern of lung cancer over time differs for men and women; rates in men rose sharply between 1950 and 1980 but have been falling for some years now, while those in women rose later and started to fall more recently – a consequence of the fact that, as a group, women took up smoking more recently than men. To describe this excessive occurrence of disease (or death) as an ‘epidemic’ captures some of the urgency the numbers demand.

The derivation of the word ‘epidemiology’ itself is from the Greek *epi*, upon, *demos*, the people, and *logia*, study. Literally, therefore, it means the ‘study

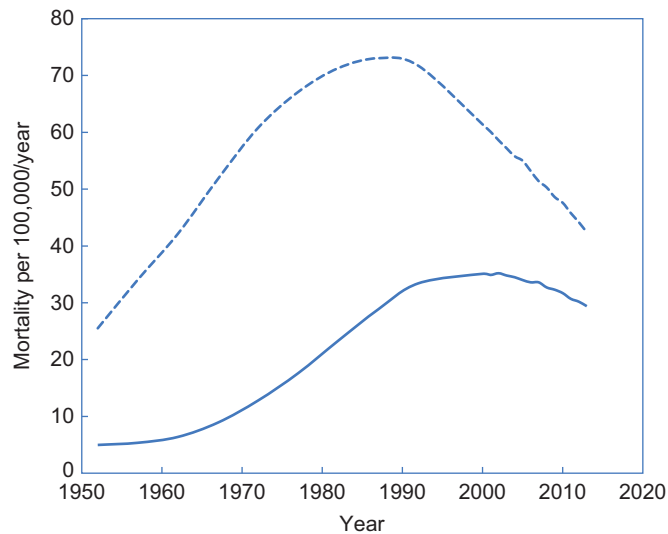


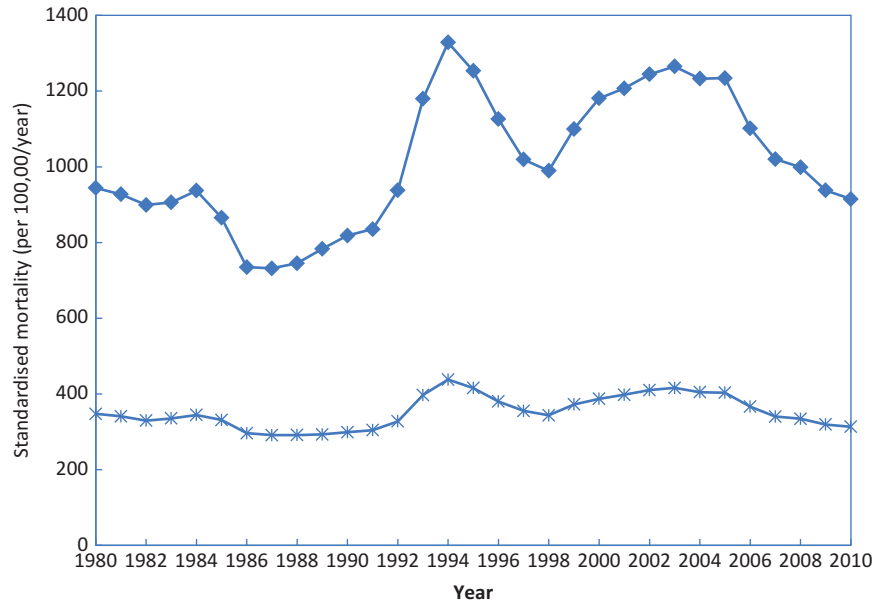
Figure 1.1 Time trends in lung cancer mortality rates in the USA (age-standardised to the 1970 US population) for white men (---) and women (—). (Drawn from: CDC Wonder Database (CDC), accessed 26 February 2015.)

(of what is) upon the people’. Such study suggests a simple set of questions that have long lain at the heart of epidemiology.

- **What** disease/condition is present in excess?
- **Who** is ill?
- **Where** do they live?
- **When** did they become ill?
- **Why** did they become ill?

The first question reflects the need for a sound, common definition of a disease so that like is compared with like. Epidemiology is all about comparison – without some reference to what is usual, how can we identify excess? The next three questions form the mantra of descriptive epidemiology: ‘*person, place and time*’. As Figure 1.2 shows, an ‘epidemic of premature mortality’ occurred among young and middle-aged men in Russia in the mid-1990s and again in the early 2000s. This description captures the essence of the problem and prompts the next questions: what caused these epidemics? What changed in the circumstances of younger Russian men to reverse the pattern of falling mortality in the early 1980s and then cause it to almost double in less than 10 years? And why did this happen again in the late 1990s? Other data show that there were no such mortality changes in Western Europe, or among older Russian men or infants, or (to the same extent) in Russian women. This simple graph captures a public health disaster for Russia and prompts urgent causal speculation: *Why did this happen?* Solving and responding to this final question is critical for public health progress, but there is clearly no simple solution. In this case, a high proportion of the deaths were

Figure 1.2 Changes in all cause mortality at ages 0–64 years in the Russian Federation from 1980 to 2010 (◆ men, ✕ women). (Data from: the European Health for All Database. WHO Regional Office for Europe, Copenhagen, Denmark, <http://data.euro.who.int/hfad/>, accessed 27 February 2015.)



linked to excess consumption of alcohol during the 1990s: increases in mortality coincided with periods of economic and societal crisis, and rates fell when the economic situation improved (Zaridze *et al.*, 2009). The earlier decline during the 1980s coincided with an anti-alcohol campaign involving higher taxes and reduced production which led to sharp decreases in alcohol consumption in the short term, and lower rates of alcohol-related mortality and suicide (Pridemore and Spivak, 2003). This example highlights the importance of paying close attention to descriptive data that provide a ‘community diagnosis’ or take the public health ‘pulse’ of a nation. Much can be gleaned from apparently simple data to give a quite precise description of the overall health of a population or a more specific health event, as the following exercise shows.

An historical epidemic

Table 1.3 shows some data that relate to an actual human experience. It tells you how many people there were in various age, sex and socioeconomic groups and what percentage of these people died during the ‘epidemic’. The challenge is to use these data to describe the event systematically in terms of **whom** this happened to (we have no data on place or time) and then to think about the sort of event that might have induced such a pattern.

Table 1.3 An historical event.

SES ^a	Adult males		Adult females		Children (both sexes)		Total population	
	Total	% Dead	Total	% Dead	Total	% Dead	Total	% Dead
High	175	67.4	144	2.8	6	–	325	37.5
Medium	168	91.7	93	14.0	24	–	285	58.6
Low	462	83.8	165	53.9	79	65.8	706	74.8
Other	885	78.3	23	13.0	0	–	908	76.7
Total	1690	80.0	425	25.6	109	47.7	2224	68.0

^a SES, socioeconomic status.

(Source: www.anesi.com/titanic.htm, The *Titanic* casualty figures (and what they mean), accessed 29 April 2015.)

The following questions are designed to help you identify key features of the data.

1. What is distinctive about this isolated population with regard to:
 - the numbers of men and women (sex distribution),
 - the numbers of adults and children (age distribution), and
 - the numbers in each socioeconomic group (socioeconomic distribution)?
2. What strikes you about the percentage of people who died (the ‘death rate’)? Is this different for (a) adults and children, (b) men and women, (c) high and low socioeconomic status (SES) and (d) any particular combinations of the above?
3. How many times more likely were:
 - men to die than women, and
 - those of low SES to die than those of high SES?
4. To what historical event might these data refer?



Table 1.3 displays more complicated data than Table 1.2 because you had to consider the joint effects of three factors (sex, SES, and age) on mortality. The sequence of questions above underlines a general principle in describing such tables – i.e. to look at overall patterns first, then move on to more detail. We all see things in different ways, but until you develop your own style the approach shown in Box 1.2 may help you avoid getting lost in the array of possible relationships. You need first to grasp the size of the *whole group* under study and how many died; then check the overall patterns (the numbers and death rates¹) across each ‘exposure’ separately (sex, SES, age).

¹ As you will see in Chapter 2, these are not technically ‘rates’ in the true sense of the word but it is convenient to call them rates as they are essentially identical in form to the attack rates in Table 1.2.