

Epidemiological Studies

Second Edition

Following on in the footsteps of its acclaimed and popular predecessor, this new edition builds on the successful features that engaged readers of the first edition: it explains the nuts and bolts of epidemiology and serves as a handbook for those who wish to do epidemiology; it uses relevant exercises and examples, taken from real life, to illustrate how to set up a study; it aims to help produce valid results that will satisfy grant bodies, ethical committees and journal editors; ultimately it bridges the gap between theory and practice. By making the subject so easily accessible, it will be an excellent introduction for anyone who is training in epidemiology and public health, and for all those involved in medical research. This edition includes numerous improvements and several new chapters which will further enhance its appeal and usefulness.

Epidemiological Studies

A Practical Guide

Second Edition

Alan J. Silman

The Medical School, University of Manchester

and

Gary J. Macfarlane

The Medical School, University of Manchester



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
 103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

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/9780521810975

© Cambridge University Press 1995, 2002

This publication is in 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 1995

Second edition 2002

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

Library of Congress Cataloging in Publication data

Silman, Alan J.

Epidemiological studies : a practical guide / Alan J. Silman, Gary J. Macfarlane. – 2nd ed.
 p. cm.

Includes index.

ISBN 0 521 81097 3 (hbk.) – ISBN 0 521 00939 1 (pbk.)

1. Epidemiology – Methodology. I. Macfarlane, Gary J. II. Title.

[DNLM: 1. Epidemiologic Methods. WA 950 S584e 2002]

RA652.4.S55 2002

614.4–dc21 2001043361

ISBN 978-0-521-81097-5 Hardback

ISBN 978-0-521-00939-3 Paperback

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 which 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

<i>Scope of this volume</i>	<i>page xi</i>
<i>Acknowledgements</i>	<i>xiv</i>

PART I INTRODUCTION

1	Scope of epidemiological enquiry and overview of main problem areas	3
1.1	What questions can epidemiology answer?	3
1.2	What are the major issues in conducting epidemiological research?	7

PART II MEASURING THE OCCURRENCE OF DISEASE

2	Which measure of disease occurrence?	13
2.1	Incidence	13
2.2	Prevalence	16
2.3	Choice of measure	17
3	Comparing rates: between and within populations	20
3.1	Introduction	20
3.2	Standardisation	21
3.3	Comparison of rates over time	24

PART III STUDYING ASSOCIATIONS BETWEEN RISK FACTORS AND DISEASE

4	Which type of study?	31
4.1	The ecologic study	31
4.2	The migrant study	33
4.3	The cross-sectional study	35
4.4	The case-control study	37
4.5	The cohort study	39
4.6	Choice of study design	41
5	Which measure of association?	45
5.1	Relative risks	45
5.2	Odds ratios	47
5.3	Attributable risks	48
5.4	Precision of measures in association	49
5.5	Categorisation of exposures	49

PART IV SELECTION OF POPULATIONS AND SAMPLES TO STUDY

6	Studies of disease occurrence. I: Identification of the population	53
6.1	Representativeness	54
6.2	Access	55
6.3	Population data accuracy	56
6.4	Study size	57
7	Studies of disease occurrence. II: Assessing disease status in study populations	60
7.1	Approaches to measuring incidence	60
7.2	Use of diagnosed cases: retrospective review or prospective notification?	63
7.3	Defining cases with the catchment population approach	65
7.4	Use of cross-sectional population surveys to assess incidence	67
7.5	Approaches to measuring prevalence	68
7.6	Catchment population methods for measuring prevalence	69
7.7	Population surveys	70
7.8	Other (indirect) measures	71

vii **Contents**

8	Studies of disease causation. I: Selection of subjects for case-control studies	74
8.1	Recruitment of cases	74
8.2	Recruitment of controls	82
8.3	One or two control groups?	87
8.4	Matching	88
8.5	Study size	90
9	Studies of disease causation. II: Selection of subjects for cohort (longitudinal) studies	93
9.1	Retrospective or prospective study cohorts?	93
9.2	How should exposure be categorised?	95
9.3	Study size	97

PART V INFORMATION FROM EPIDEMIOLOGICAL SURVEYS

10	Collecting information	103
10.1	Interview or subject completing questionnaire?	103
10.2	How to formulate a questionnaire	106
11	Obtaining valid information	111
11.1	Introduction	111
11.2	Sensitivity and specificity	112
11.3	Validity for variables that are not dichotomous	115
11.4	Possible approaches for independent validation	116
11.5	Misclassification	118
12	Repeatability	120
12.1	Introduction	120
12.2	Study designs to measure repeatability	122
13	Maximising participation	128
13.1	Introduction	128
13.2	Reasons for non-participation	128
13.3	Maximising participation in follow-up	132

viii **Contents**

14	Conducting a pilot study	138
	14.1 Aims	138

**PART VI ANALYSIS AND INTERPRETATION OF
EPIDEMIOLOGICAL DATA**

15	Preparation of survey data for statistical analysis	145
	15.1 Introduction	145
	15.2 Initial checking for completeness and accuracy	147
	15.3 Linkage by subject of data from multiple sources	149
	15.4 Development of a data coding schedule	149
	15.5 Development of a computer database	151
	15.6 Procedure for data entry	153
	15.7 Checking for errors in entered data	154
	15.8 Missing data	155
	15.9 Recoding of entered data	156
	15.10 Storage of data and data set	157
16	Introductory data analysis: descriptive epidemiology	158
	16.1 Introduction	158
	16.2 Incidence rates	158
	16.3 Prevalence (proportions)	160
	16.4 Crude, age-specific and standardised rates	163
17	Introductory data analysis: analytical epidemiology	168
	17.1 Introduction	168
	17.2 Effect measurement, interval assessment and significance testing	170
	17.3 Analysis of case-control studies	171
	17.4 Analysis of cohort studies	179
	17.5 Conclusion	187
18	Confounding	188
	18.1 Introduction	188
	18.2 Minimising confounding in study design	190
	18.3 Conduct of study	191
	18.4 Analysis	192

ix Contents

19	Bias	201
	19.1 Introduction	201
	19.2 Major sources of bias	203
	19.3 Selection bias	204
	19.4 Information bias	209
	19.5 Is an unbiased study ever possible?	212

PART VII OTHER PRACTICAL ISSUES

20	Ethical issues in epidemiology	215
	20.1 Introduction	215
	20.2 Ethical approval	215
	20.3 Ethical constraints in maximising response rate	216
	20.4 Confidentiality and data protection	218
	20.5 Detection of previously unrecognised disease	221
	20.6 Consent	225
21	The costs of epidemiological studies	228
	21.1 Costs versus design	228
	21.2 Costing an epidemiological study	229
	21.3 Possibilities for cost containment	233
	21.4 Wasting resources	235
	<i>Index</i>	237

Scope of this volume

A quick browse through the epidemiology section of medical libraries and bookshops reveals that most volumes with ‘epidemiology’ in their title are either theoretical texts covering the concepts of the discipline, with varying degrees of biostatistical complexity, or are reviews of the available ‘epidemiological’ data on one disease or a group of related diseases. The problem with the latter is that such reviews, often out of necessity, present the available published data without critically reviewing the studies that lead to their generation. In this volume an attempt has been made to bridge the gap between concept and result and to provide a practical guide to epidemiological practice. It is therefore written as a handbook to those who wish to do epidemiology, rather than restricting its aim to those who wish to understand the discipline. The underlying theme, however, has been how to complete a study without compromising validity (in its broadest sense). The hope therefore is that sufficient guidance is provided to produce work that will satisfy grant bodies, ethical committees and journal editors and reviewers.

This volume has taken a deliberate ‘holistic’ view of epidemiological method. The practical issues, for example, of approaching potential study populations, maximising response and questionnaire design, are similarly independent of the epidemiological study design chosen.

The volume is divided into a number of parts. The first part provides the uninitiated with a list of subjects that are relevant for the application of epidemiological method and follows this with an outline of the main problem areas in undertaking epidemiological research. The remainder of the book aims to cover these issues systematically in completing an epidemiological study.

The second part addresses the options available for measuring the occurrence of disease and methodological considerations in comparing occurrence between populations and over time.

xii **Scope of this volume**

The third part focuses on the choice of an appropriate design to address the questions posed. As such, this part acts also as an introduction to the underlying concepts behind the varying options for choice of study. The objective is that this part provides sufficient material to enable the investigator to choose and justify the optimal method(s) for study.

The fourth part covers the problems inherent in selecting populations for study. The relative merits of the different choices available are discussed together with the practical issues of how to recruit subjects, including, for example, suggestions for letters of invitation. In this part as in the sections on analysis in Part VI, the separate requirements of different study approaches, for example case control and cohort studies, are considered individually.

The fifth part addresses the problems of data collection within epidemiological studies with the objective of selecting the most effective survey method to achieve the goals targeted. Practical issues covered include the assessment of validity and reproducibility of survey methods, suggestions for design of survey instruments and a detailed account of how to maximise participation, possibly the greatest practical problem in conducting a survey. These issues are also considered collectively in the section on conducting a pilot study.

The sixth part covers analysis and interpretation of the data collected. The first focus is on the preparation of data for analysis, including the important but often ignored steps of checking for data errors and ensuring a clean set of data for subsequent analysis. Chapters 16 and 17, on introductory epidemiological data analysis, are not meant to be a substitute for the necessary detailed information on this topic that is required to be obtained from statistics textbooks and the manuals accompanying analytical software packages. Their aim is to provide the necessary formulae with simple worked examples to permit the investigator to understand how measures of occurrence and effect are calculated and to undertake these on a hand calculator. These chapters assume only a limited knowledge of statistics such as an understanding of summary measures (e.g. the mean), and measures of dispersion (e.g. standard deviation). It will also be necessary to understand the concept behind statistical significance testing and particularly the use and limitations of the confidence interval. The subsequent chapters cover the topics of confounding and bias, with the goal that the investigator

xiii **Scope of this volume**

will be able to assess the impact, if any, of these phenomena on the results obtained.

The final part covers two important practical areas: how to ensure that a proposed study is ethically sound and how to minimise the costs of undertaking studies.

The text is liberally illustrated with examples from many areas of medicine in an attempt to clarify the points made. Many of the examples are drawn from the real life experience of the authors and their colleagues, others are based on reports of research, both published and unpublished, with the remainder unapologetically invented to illustrate a point where necessary.

Finally, where appropriate, exercises are set at the end of a chapter, with model solutions provided at the end of the book.

Acknowledgements

From first edition

The genesis of this volume lies in the organisation of a teaching programme in epidemiology to those entering public health in the southeast of England. There was a clear need for those students to receive sufficient training to undertake simple epidemiological studies themselves. With no appropriate practical guide available, a course was developed to achieve this goal. With suitable refinements resulting from continuing feedback from the participants, the material developed for this course formed the basis of the current volume. I therefore acknowledge the contribution made by successive cohorts of attenders. The material has grown and altered in further discussions with clinical colleagues in both primary and hospital care who have themselves desired to dip their toes into the fascinating waters of epidemiological science.

I am, however, particularly grateful to my close colleagues Peter Croft, Deborah Symmons and Paul Brennan for their frank and constructive criticisms as the work has progressed; any errors or confusions that remain are mine. My secretary Cath Barrow has worked tirelessly and cheerfully and has coped with substantial rewrites without complaint.

Second edition

The second edition has further benefited from feedback of both staff and postgraduate students at the School of Epidemiology and Health Sciences, University of Manchester. In addition, Lesley Jordan has ably and willingly co-ordinated the administrative work involved in producing this revised edition.