METHODS IN EPIDEMIOLOGIC RESEARCH

‘*A comprehensive text*
for the discipline’

www.upei.ca/mer

METHODS IN EPIDEMIOLOGIC RESEARCH

Ian Dohoo
Professor Emeritus of Epidemiology
Department of Health Management
University of Prince Edward Island
Charlottetown PEI Canada

Wayne Martin
Professor Emeritus of Epidemiology
Department of Population Medicine
University of Guelph
Guelph Ontario Canada

Henrik Stryhn
Professor of Biostatistics
Department of Health Management
University of Prince Edward Island
Charlottetown PEI Canada

VER inc. Published by VER Inc. · Charlottetown · Prince Edward Island · Canada

This book was prepared using the open source office software—OpenOffice (http://www.openoffice.org/)

Editor/Proofreader/Compositor/Publication Coordinator: S Margaret McPike
Cover designer: Gregory Mercier

VER Inc, PO Box 491, Charlottetown, Prince Edward Island, Canada, C1A 7L1

Printed in Canada

10 9 8 7 6 5 4 3 2 1

National Library of Canada Cataloguing in Publication
Methods in epidemiologic research / Ian Dohoo ... [*et al*].
Includes index.
ISBN 978-0-919013-73-5
1. Human Epidemiology – Research – Textbooks
I. Dohoo, Ian Robert

Library catalogue information available from www.upei.ca/mer

Dedication

This text is dedicated to all of the graduate students who have challenged and inspired us throughout our careers, and to our families who have supported us, especially during the writing of this text.

CONTENTS

FOREWORD – DR. JOSEPH HILBE

Methods in Epidemiologic Research is a revision of the authors' very well-received 2009 volume, *Veterinary Epidemiologic Research*, 2nd edition; rather than examples from veterinary epidemiology, this new text now employs examples from across the spectrum of human epidemiology. References have also been updated.

The authors of this text have produced what comes close to being called an encyclopedia of applied epidemiologic analysis. In fact, the text is best characterized as a well-written comprehensive presentation of all foremost areas of epidemiological research. Of the some 850 pages of text, nearly 350 are devoted to descriptive statistics, varieties of study design, questionnaire and screening techniques, sampling, measures of association, and confounding. Epidemiologists are presented with clear guidelines for nearly everything they need to know about gathering data and study design for their individual research projects. Four hundred plus pages are then devoted to nearly all areas of epidemiologic modeling. Major chapters are given for linear modeling, all of the major categorical response models, random and mixed effects modeling, longitudinal and clustered models, survival models, meta-analysis, and spatial analysis, chapters are also provided on infectious disease epidemiology and Bayesian methodology. Bayesian methods are increasing in popularity due to ever-faster and more powerful personal computers, as well as enhanced Monte Carlo algorithms. The authors have presented a fine overview of the subject, and have given readers several excellent worked-out examples to help clarify the basics of the methodology.

Numerous examples are given throughout the text, accompanied by modeling results and thorough interpretation. Guidelines are presented together with caveats which the researcher should keep in mind when modeling. It is a text that researchers from most disciplines will find to be useful and informative.

The earlier version of this text sold extremely well to veterinarians as well as to those in biostatistics and human epidemiology. The enthusiasm with which the earlier text was accepted motivated the creation of this volume, which promises to be widely used as the standard single-volume resource on epidemiologic research throughout the world.

Joseph M. Hilbe, J.D., Ph.D.

Emeritus Professor, University of Hawaii,

Adjunct Professor of Statistics, Arizona State University,

Solar System Ambassador, Jet Propulsion Laboratory, California Institute of Technology

FOREWORD – DR. JAMES ANTHONY

This brief note is not enough praise for what the authors have accomplished in a revision of their excellent text on *Veterinary Epidemiologic Research* (VER) for a human medicine audience. The generally superior coverage of epidemiological research at an intermediate level, coupled with exercises and problem sets with Stata software code, prompted us to make VER the required textbook for the third course in our epidemiology graduate sequence, which draws enrollees both from the College of Veterinary Medicine and the College of Human Medicine. The strengths of VER and its examples and exercises drawn from veterinary medicine have served us well at this intermediate stage in the course sequence. It was a delight to learn that the textbook was to be converted into one more accessible to people working in 'human epidemiology'—*Methods in Epidemiologic Research*. The result will be an increase in the breadth of appreciation for what the authors have accomplished.

James C. Anthony, M.Sc., Ph.D.

Professor of Epidemiology & Biostatistics, College of Human Medicine, Michigan State University East Lansing, Michigan

Adjunct Professor, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, and

Profesor Honorario, Universidad Peruana Cayetano Heredia, Lima, Peru

PREFACE

This book started its life as a veterinary text: *Veterinary Epidemiologic Research*. The first edition of that text was published in 2003 and the 2nd edition in late 2009. Our goal in the publication of this veterinary text was to produce a text that was comprehensive, but accessible to both researchers and graduate students. Since its initial publication, *Veterinary Epidemiologic Research* has become the standard textbook used in most graduate programs in veterinary epidemiology around the world; we are most appreciative of the strong support for the text that we have received from the veterinary epidemiological community.

There has been considerable convergence of veterinary and human health worlds with the development of the 'One Medicine/One Health' movement (discussed in Chapter 1). In the population health arena, this is facilitated by the fact that methods used in veterinary epidemiology are virtually identical to those used in 'human' (medical) epidemiology. However, given the title and focus of our original book (veterinary epidemiology), it is not surprising that only a few people in the medical epidemiology community were aware of the text, or had a chance to review it. Fortunately, a number of those that did encounter the book were complimentary and encouraged us to consider preparation of a version suitable for use by medical epidemiologists and public health practitioners. Most notable was the encouragement we received from Drs. Joseph Hilbe and James Anthony who kindly agreed to write forewords for this text.

Methods in Epidemiologic Research is based heavily on the material covered in *Veterinary Epidemiologic Research,* with 2 major changes. Firstly, all of the veterinary examples have been replaced with human medical examples. We are indebted to those in the medical epidemiology and public health communities who have kindly shared datasets with us for the examples used in the text (please see the Acknowledgements section for details). Secondly, nearly all references have been switched from veterinary to human examples and references have been updated (particularly for the chapters dealing with study design). It is our sincere hope that we have produced a text that will serve the medical epidemiological community as well as our previous text has served the veterinary community.

While the text has been fully converted to be understandable by medical epidemiologists and public health practitioners, it retains a few characteristics from its veterinary roots. For example, the chapters on the handling and modelling of spatial data (Chapters 26 and 27) have remained unchanged from the veterinary text, because they dealt with an issue relevant to both veterinary and human health (avian influenza in both poultry and human populations). In a few select instances, methods which have been more widely adopted in veterinary medicine, but which have relevance in human epidemiology (*eg* targeted surveillance in Section 2.9) have been retained. The problem of 'clustered data' is of particular importance in animal health research (because animals are frequently kept in clusters (*eg* herds)), but this problem is also encountered in medical epidemiology, so we have retained our very thorough coverage of this topic.

Before reviewing the content of the text we thought we should address the two most common questions we received about the veterinary version of this text.

- Why are the 2X2 tables oriented the way they are (disease in rows, exposure in columns)? The answer to this is that we feel that the text *Modern Epidemiology* (Rothman *et al*, 2008) is a key reference text in the field of epidemiology and have

chosen to be consistent with their format.

- Why does the title use the word 'epidemiologic' instead of 'epidemiological?' According to "Scientific Style and Format—The CSE Manual for Authors, Editors and Publishers" (Council of Science Editors—Style Manual Committee, 2006), either is acceptable. Once again, we deferred to a text which we felt was seminal in the development of epidemiologic methods *Epidemiologic Research: Principles and Quantitative Methods* (Kleinbaum *et al*, 1982).

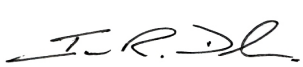
This text focuses on both design and analytic issues. Chapters 1 through 6 focus on basic epidemiologic principles. Chapters 7–11 focus on study design issues for observational studies and controlled trials. There has been much discussion over the past decade about the need for epidemiologists to thoroughly report their research findings (and by doing so this will help ensure high-quality study designs in the future) and we have cited the summary recommendations in these chapters.

Chapters 14–19 cover a range of multivariable models. Chapter 19 (Modelling Survival Data) attempts to provide a comprehensive coverage of the most commonly used methods in the analysis of time-to-event data.

Chapters 20–23 deal with the issue of clustered data, including a thorough description of methods for analysing repeated measures data. Chapters 24–30 cover a range of specialised topics including: Bayesian methods (Chapter 24—contributed by Henrik Stryhn in collaboration with William Browne), two chapters on presenting and analysing spatial data (Chapter 25 and 26—contributed by Javier Sanchez and Dirk Pfeiffer), an introduction to infectious disease epidemiology (Chapter 27—contributed by Graham Medley in collaboration with Ian Dohoo), and meta-analysis (Chapter 28).

Supplementary materials for this text will all be made available at upei.ca/mer. These will include datasets, computer programs for all examples presented (initially Stata "do files" with the expectation that programs for other statistical packages will be added later).

All of the datasets used in these examples are described in the text (Chapter 31) and are available through upei.ca/mer. Virtually all of the examples have been analysed using the statistical program Stata™—a program which provides a unique combination of statistical and epidemiological tools and which we use extensively in our teaching. Version 12 of Stata was used throughout. In the future, we hope to add additional sets of sample problems, program code in other languages and additional supplemental material to the website.



We hope that you find *Methods in Epidemiologic Research* useful in your studies and your research.

REFERENCES

Council of Science Editors - Style Manual Committee. Scientific style and format: The CSE manual for authors, editors and publishers. 7th ed. Reston (VA): The Council; 2006.

Kleinbaum D, Kupper L, Morgenstern H. Epidemiologic Research: Principles and Quantitative Methods. London: Lifetime Learning Publications; 1982.

Rothman K, Greenland S, Lash T. Modern Epidemiology, 3rd Ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

ACKNOWLEDGEMENTS

We are indebted to the many people who have provided useful feedback on *Veterinary Epidemiologic Research*. Their input has been appreciated and incorporated into the writing of *Methods in Epidemiologic Research*. However, we would like to highlight and acknowledge the contribution of 3 particular individuals.

- Dr. Garry Anderson, University of Melbourne, provided extensive constructive feedback on *Veterinary Epidemiologic Research* and was a primary reviewer of the 2nd edition of that text.
- Carolyn Dohoo read many of the chapters of *Methods in Epidemiologic Research* and provided much useful feedback on the clarity and validity of the material presented.
- Craig Jones contributed greatly to the clarity of presentation of material in a number of chapters in *Methods in Epidemiolgic Research*.

We would like to express our sincere appreciation to all of these individuals.

We believe the value of this book has been greatly enhanced by the provision of a substantial number of 'real-life' datasets. The details of these datasets (and who contributed them) is described in Chapter 31, but at this point we would like to highlight the major contributors.

- Pasha Marcynuk - study on rainwater cisterns and acute gastrointestinal illness
- Dr. Kate Thomas - Canadian survey data on acute gastrointestinal illness
- Dr. Robert Goldberg and Darlene Lessard - myocardial infarction data from the Worcester Heart Attack Study database
- Dr. David Fisman - norovirus diagnostic test data
- Dr. Oliver Bucher - meta-analysis data for *Salmonella* control in poultry products
- Dr. Dirk Pfeiffer - spatial data on human and poultry cases of avian influenza in Thailand.

Other datasets were obtained from public domain sources and we appreciate the effort that the creators of those datasets went to to make them publicly available.

As we did with *Veterinary Epidemiologic Research*, we prepared this book using open source software (OpenOffice—www.openoffice.org). We are deeply indebted to Margaret McPike who has done all of the editing, proofreading, and formatting of this text. As with *Veterinary Epidemiologic Research*, we published this book ourselves, which entailed taking complete responsibility for these activities. Margaret dedicated herself to this task. All of the credit for layout of the book, and the clarity of the format, goes to Margaret.

We would like to thank Gregory Mercier, who did the graphic design work for the cover. We would also like to thank Bill Rising of Stata Corp. who reviewed all of the analytical methods chapters and provided some very constructive feedback, particularly in terms of the program files (Stata -do- files) which are available on the book's website.