

# **METHODS IN EPIDEMIOLOGIC RESEARCH**

***‘A comprehensive text  
for the discipline’***

**[www.upei.ca/mer](http://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



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

© 2012, by VER Inc. All rights reserved. This book is protected by copyright. No part of this book may be reproduced, stored in a retrieval system or transmitted, in any form or by any means—electronic, mechanical, photocopy, recording or otherwise—without the written permission of the publisher. For information write:

**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](http://www.upei.ca/mer)

---

Care has been taken to confirm the accuracy of the information presented. Nevertheless, the authors, editor and publisher are not responsible for errors or omissions or for any consequences from application of the information in this book and make no warranty, express or implied, with respect to the content of the publication.

## *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

---

Forewords	xv/xvii
Preface	xviii
Acknowledgements	xx
<b>1 INTRODUCTION AND CAUSAL CONCEPTS</b>	<b>1</b>
1.1 Introduction	2
1.2 A brief history of multiple causation concepts	2
1.3 A brief history of scientific inference	6
1.4 Key components of epidemiologic research	9
1.5 Seeking causes	10
1.6 Models of causation	11
1.7 Counterfactual concepts of causation for a single exposure	18
1.8 Experimental versus observational evidence of causation	22
1.9 Constructing a causal diagram	23
1.10 Causal criteria	25
<b>2 SAMPLING</b>	<b>35</b>
2.1 Introduction	36
2.2 Non-probability sampling	39
2.3 Probability sampling	39
2.4 Simple random sample	40
2.5 Systematic random sample	40
2.6 Stratified random sample	40
2.7 Cluster sampling	41
2.8 Multistage sampling	42
2.9 Targeted (risk-based) sampling	43
2.10 Analysis of survey data	44
2.11 Sample-size determination	48
2.12 Sampling to detect disease	55
<b>3 QUESTIONNAIRE DESIGN</b>	<b>61</b>
3.1 Introduction	62
3.2 Designing the question	64
3.3 Open question	65
3.4 Closed question	65
3.5 Wording the question	69
3.6 Structure of questionnaires	69
3.7 Pre-testing questionnaires	70
3.8 Validation	71
3.9 Response rate	71
3.10 Data-coding and editing	72
<b>4 MEASURES OF DISEASE FREQUENCY</b>	<b>77</b>
4.1 Introduction	78
4.2 Counts, proportions, odds, and rates	78
4.3 Incidence	79
4.4 Calculating risk	80

4.5	Calculating incidence rates	81
4.6	Relationship between risk and rate	83
4.7	Prevalence	84
4.8	Mortality statistics	85
4.9	Other measures of disease frequency	85
4.10	Standard errors and confidence intervals	87
4.11	Standardisation of risks and rates	89
<b>5</b>	<b>SCREENING AND DIAGNOSTIC TESTS</b>	<b>95</b>
5.1	Introduction	96
5.2	Attributes of the test <i>per se</i>	96
5.3	The ability of a test to detect disease or health	104
5.4	Predictive values	107
5.5	Interpreting test results that are measured on a continuous scale	109
5.6	Using multiple tests	115
5.7	Evaluation of diagnostic tests	117
5.8	Evaluation when there is no gold standard	121
5.9	Other considerations in test evaluation	125
5.10	Sample size requirements	127
5.11	Group-level testing	128
5.12	Use of pooled samples	130
<b>6</b>	<b>MEASURES OF ASSOCIATION</b>	<b>139</b>
6.1	Introduction	140
6.2	Measures of association	141
6.3	Measures of effect	144
6.4	Study design and measures of association	147
6.5	Hypothesis testing and confidence intervals	147
6.6	Multivariable estimation of measures of association	152
<b>7</b>	<b>INTRODUCTION TO OBSERVATIONAL STUDIES</b>	<b>155</b>
7.1	Introduction	156
7.2	A unified approach to study design	159
7.3	Descriptive studies	161
7.4	Observational studies	162
7.5	Cross-sectional studies	164
7.6	Estimating incidence from one or more cross-sectional studies	168
7.7	Inferential limitations of cross-sectional studies	169
7.8	Repeated cross-sectional versus cohort studies	170
7.9	Reporting of observational studies	171
<b>8</b>	<b>COHORT STUDIES</b>	<b>179</b>
8.1	Introduction	180
8.2	Selecting the study group	182
8.3	The exposure	186
8.4	Disease as exposure	190
8.5	Ensuring exposed and non-exposed groups are comparable	190
8.6	Follow-up period	191



8.7	Measuring the outcome	191
8.8	Analysis	192
8.9	Reporting of cohort studies	194
<b>9</b>	<b>CASE-CONTROL STUDIES</b>	<b>201</b>
9.1	Introduction	202
9.2	The study base	202
9.3	The case series	205
9.4	Principles of control selection	207
9.5	Selecting controls and data layout in risk-based designs	207
9.6	Sampling controls and data layout in rate-based designs	209
9.7	Other sources of controls	214
9.8	The number of controls per case	215
9.9	The number of control groups	215
9.10	Exposure and covariate assessment	216
9.11	Keeping the cases and controls comparable	216
9.12	Analysis of case-control data	217
9.13	Reporting guidelines for case-control studies	218
<b>10</b>	<b>HYBRID STUDY DESIGNS</b>	<b>223</b>
10.1	Introduction	224
10.2	Case-crossover studies	224
10.3	Case-case studies	228
10.4	Case-case-control studies	229
10.5	Case-series studies	231
10.6	Case-cohort studies	233
10.7	Case-only studies	235
10.8	Two-stage sampling designs	237
<b>11</b>	<b>CONTROLLED STUDIES</b>	<b>243</b>
11.1	Introduction	244
11.2	Background, objectives, and summary trial design	246
11.3	Participants: the study group	247
11.4	Specifying the intervention	250
11.5	Measuring the outcome	251
11.6	Sample size	252
11.7	Allocation of study subjects	254
11.8	Follow-up/compliance	258
11.9	Statistical methods and analysis	259
11.10	Conclusions	262
11.11	Clinical trial designs for prophylaxis of communicable organisms	262
11.12	Reporting of clinical trials	265
<b>12</b>	<b>VALIDITY IN OBSERVATIONAL STUDIES</b>	<b>275</b>
12.1	Introduction	276
12.2	Selection bias	277
12.3	Examples of selection bias	281
12.4	Reducing selection bias	287

12.5	Information bias	288
12.6	Bias from misclassification	290
12.7	Validation studies to correct misclassification	297
12.8	Measurement error	297
12.9	Errors in surrogate measures of exposure	299
12.10	The impact of information bias on sample size	299
<b>13</b>	<b>CONFOUNDING: DETECTION AND CONTROL</b>	<b>307</b>
13.1	Introduction	308
13.2	Control of confounding prior to data analysis	311
13.3	Matching on confounders	311
13.4	Detection of confounding	316
13.5	Analytic control of confounding	322
13.6	Multivariable modelling to control confounding	328
13.7	Other approaches to control confounding and estimate causal effects	328
13.8	Propensity scores for controlling confounding	335
13.9	External adjustment and sensitivity analysis for unmeasured confounders	340
13.10	Understanding causal relationships	342
13.11	Summary of effects of extraneous variables	351
<b>14</b>	<b>LINEAR REGRESSION</b>	<b>359</b>
14.1	Introduction	360
14.2	Regression analysis	360
14.3	Hypothesis testing and effect estimation	362
14.4	Nature of the X-variables	368
14.5	Detecting highly correlated (collinear) variables	374
14.6	Detecting and modelling interaction	376
14.7	Causal interpretation of a multivariable linear model	377
14.8	Evaluating the least squares model	379
14.9	Evaluating the major assumptions	385
14.10	Assessment of individual observations	390
14.11	Time-series data	396
<b>15</b>	<b>MODEL-BUILDING STRATEGIES</b>	<b>401</b>
15.1	Introduction	402
15.2	Steps in building a model	403
15.3	Building a causal model	403
15.4	Reducing the number of predictors	404
15.5	The problem of missing values	408
15.6	Effects of continuous predictors	411
15.7	Identifying interaction terms of interest	418
15.8	Building the model	418
15.9	Evaluate the reliability of the model	423
15.10	Presenting the results	424
<b>16</b>	<b>LOGISTIC REGRESSION</b>	<b>429</b>
16.1	Introduction	430
16.2	The logistic model	430

16.3	Odds and odds ratios	431
16.4	Fitting a logistic regression model	432
16.5	Assumptions in logistic regression	433
16.6	Likelihood ratio statistics	434
16.7	Wald tests	436
16.8	Interpretation of coefficients	436
16.9	Assessing interaction and confounding	439
16.10	Model-building	441
16.11	Generalised linear models	444
16.12	Evaluating logistic regression models	445
16.13	Sample size considerations	455
16.14	Exact logistic regression	456
16.15	Conditional logistic regression for matched studies	456
<b>17</b>	<b>MODELLING ORDINAL AND MULTINOMIAL DATA</b>	<b>461</b>
17.1	Introduction	462
17.2	Overview of models	462
17.3	Multinomial logistic regression	466
17.4	Modelling ordinal data	470
17.5	Proportional odds model (constrained cumulative logit model)	471
17.6	Adjacent-category model	475
17.7	Continuation-ratio model	476
<b>18</b>	<b>MODELLING COUNT AND RATE DATA</b>	<b>479</b>
18.1	Introduction	480
18.2	The Poisson distribution	481
18.3	Poisson regression model	482
18.4	Interpretation of coefficients	483
18.5	Evaluating Poisson regression models	485
18.6	Negative binomial regression	488
18.7	Problems with zero counts	496
<b>19</b>	<b>MODELLING SURVIVAL DATA</b>	<b>501</b>
19.1	Introduction	502
19.2	Non-parametric analyses	507
19.3	Actuarial life tables	507
19.4	Kaplan-Meier estimate of survivor function	510
19.5	Nelson-Aalen estimate of cumulative hazard	512
19.6	Statistical inference in non-parametric analyses	512
19.7	Survivor, failure, and hazard functions	514
19.8	Semi-parametric analyses	519
19.9	Parametric models	536
19.10	Accelerated failure time models	541
19.11	Frailty models and clustering	545
19.12	Multiple outcome event data	551
19.13	Discrete-time survival analysis	552
19.14	Sample sizes for survival analyses	557

<b>20</b>	<b>INTRODUCTION TO CLUSTERED DATA</b>	<b>563</b>
20.1	Introduction	564
20.2	Clustering arising from the data structure	564
20.3	Effects of clustering	570
20.4	Simulation studies on the impact of clustering	574
20.5	Introduction to methods for dealing with clustering	576
<b>21</b>	<b>MIXED MODELS FOR CONTINUOUS DATA</b>	<b>587</b>
21.1	Introduction	588
21.2	Linear mixed model	588
21.3	Random slopes	594
21.4	Contextual effects	598
21.5	Statistical analysis of linear mixed models	601
<b>22</b>	<b>MIXED MODELS FOR DISCRETE DATA</b>	<b>615</b>
22.1	Introduction	616
22.2	Logistic regression with random effects	617
22.3	Poisson regression with random effects	621
22.4	Generalised linear mixed model	623
22.5	Statistical analysis of GLMMs	630
22.6	Summary remarks on analysis of discrete clustered data	639
<b>23</b>	<b>REPEATED MEASURES DATA</b>	<b>645</b>
23.1	Introduction to repeated measures data	646
23.2	Univariate and multivariate approaches to repeated measures data	648
23.3	Linear mixed models with correlation structure	654
23.4	Mixed models for discrete repeated measures data	662
23.5	Generalised estimating equations	665
<b>24</b>	<b>INTRODUCTION TO BAYESIAN ANALYSIS</b>	<b>675</b>
24.1	Introduction	676
24.2	Bayesian analysis	676
24.3	Markov chain Monte Carlo estimation	680
24.4	Statistical analysis based on MCMC estimation	685
24.5	Extensions of Bayesian and MCMC modelling	689
<b>25</b>	<b>ANALYSIS OF SPATIAL DATA: INTRODUCTION AND VISUALISATION</b>	<b>701</b>
25.1	Introduction	702
25.2	Spatial data	702
25.3	Spatial data analysis	705
25.4	Additional topics	711
<b>26</b>	<b>ANALYSIS OF SPATIAL DATA</b>	<b>717</b>
26.1	Introduction	718
26.2	Issues specific to statistical analysis of spatial data	718
26.3	Exploratory spatial analysis	720
26.4	Global spatial clustering	728

26.5	Localised spatial cluster detection	735
26.6	Space-time association	738
26.7	Modelling	742
<b>27</b>	<b>CONCEPTS OF INFECTIOUS DISEASE EPIDEMIOLOGY</b>	<b>753</b>
27.1	Introduction	754
27.2	Infection vs disease	756
27.3	Transmission	758
27.4	Mathematical modelling of infectious disease transmission	760
27.5	Methods of control of infectious disease	763
27.6	Estimating $R_0$ and other parameters	766
27.7	Developing more complex models	771
27.8	Using models	773
27.9	Summary	775
<b>28</b>	<b>SYSTEMATIC REVIEWS AND META-ANALYSIS</b>	<b>779</b>
28.1	Introduction	780
28.2	Narrative reviews	780
28.3	Systematic reviews	781
28.4	Meta-analysis – introduction	785
28.5	Fixed- and random-effects models	786
28.6	Presentation of results	789
28.7	Heterogeneity	791
28.8	Publication bias	798
28.9	Influential studies	801
28.10	Outcome scales and data issues	801
28.11	Meta-analysis of observational studies	804
28.12	Meta-analysis of diagnostic tests	806
28.13	Use of meta-analysis	807
<b>29</b>	<b>ECOLOGICAL AND GROUP-LEVEL STUDIES</b>	<b>813</b>
29.1	Introduction	814
29.2	Rationale for group level studies	815
29.3	Types of ecologic variable	816
29.4	Issues related to modelling approaches in ecologic studies	817
29.5	The linear model in the context of ecologic studies	818
29.6	Issues related to inferences	819
29.7	Sources of ecologic bias	820
29.8	Analysis of ecologic data	825
29.9	Non-ecologic group-level studies	826
<b>30</b>	<b>A STRUCTURED APPROACH TO DATA ANALYSIS</b>	<b>833</b>
30.1	Introduction	834
30.2	Data-collection sheets	834
30.3	Data coding	835
30.4	Data entry	835
30.5	Keeping track of files	836
30.6	Keeping track of variables	836

30.7	Program mode versus interactive processing	837
30.8	Data-editing	838
30.9	Data verification	839
30.10	Data processing—outcome variables	839
30.11	Data processing—predictor variables	840
30.12	Data processing—multilevel data	840
30.13	Unconditional associations	841
30.14	Keeping track of your analyses	841
<b>31</b>	<b>DESCRIPTION OF DATASETS</b>	<b>843</b>

## 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 2<sup>nd</sup> 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).

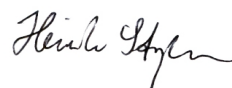
This text focuses on both design and analytic issues. Chapters 1 through 6 focus on basic epidemiologic principles. Chapters 7–11 focus on study design issues for observational studies and controlled trials. There has been much discussion over the past decade about the need for epidemiologists to thoroughly 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](http://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](http://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 2<sup>nd</sup> 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 Epidemiologic 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](http://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.