

To Hilda and Marion





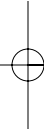
Table of Contents

Preface	ix
Acknowledgments	xiii
PART I—INTRODUCTION	1
Chapter 1—Basic Study Designs in Analytical Epidemiology	3
1.1 Introduction: Descriptive and Analytical Epidemiology	3
1.2 Analysis of Age, Birth Cohort, and Period Effects	4
1.3 Ecologic Studies	14
1.4 Studies Based on Individuals as Observation Units	19
PART II—MEASURES OF DISEASE OCCURRENCE AND ASSOCIATION	45
Chapter 2—Measuring Disease Occurrence	47
2.1 Introduction:	47
2.2 Measures of Incidence	49
2.3 Measures of Prevalence	71
2.4 Odds	73
Chapter 3—Measuring Associations Between Exposures and Outcomes	77
3.1 Introduction	77
3.2 Measuring Associations in a Cohort Study	77
3.3 Cross-Sectional Studies: Point Prevalence Rate Ratio	89
3.4 Measuring Associations in Case-Control Studies	89
3.5 Assessing the Strength of Associations	101

PART III—THREATS TO VALIDITY AND ISSUES OF INTERPRETATION	107
Chapter 4—Understanding Lack of Validity: Bias	109
4.1 Overview	109
4.2 Selection Bias	112
4.3 Information Bias	116
4.4 Combined Selection/Information Biases	133
Chapter 5—Identifying Noncausal Associations: Confounding	151
5.1 Introduction	151
5.2 The Nature of the Association Between the Confounder, the Exposure, and the Outcome	154
5.3 Theoretical and Graphical Aids to Frame Confounding	161
5.4 Assessing the Presence of Confounding	163
5.5 Additional Issues Related to Confounding	169
5.6 Conclusion	180
Chapter 6—Defining and Assessing Heterogeneity of Effects: Interaction	183
6.1 Introduction	183
6.2 Defining and Measuring Effect	184
6.3 Strategies To Evaluate Interaction	184
6.4 Assessment of Interaction in Case-Control Studies	194
6.5 More on the Interchangeability of the Definitions of Interaction	202
6.6 Which Is the Relevant Model: Additive Versus Multiplicative Interaction?	204
6.7 The Nature and Reciprocity of Interaction	206
6.8 Interaction, Confounding Effect, and Adjustment	210
6.9 Statistical Modeling and Statistical Tests for Interaction	212
6.10 Interpreting Interaction	212
6.11 Interaction and Search for New Risk Factors in Low-Risk Groups	217
6.12 Interaction and “Representativeness” of Associations	218
PART IV—DEALING WITH THREATS TO VALIDITY	223
Chapter 7—Stratification and Adjustment: Multivariate Analysis in Epidemiology	225
7.1 Introduction	225
7.2 Stratification and Adjustment Techniques To Disentangle Confounding	226

7.3	Adjustment Methods Based on Stratification	230
7.4	Multiple Regression Techniques for Adjustment	244
7.5	Incomplete Adjustment: Residual Confounding	280
7.6	Overadjustment	282
7.7	Conclusion	284
Chapter 8—Quality Assurance and Control		295
8.1	Introduction	295
8.2	Quality Assurance	297
8.3	Quality Control	300
8.4	Indices of Validity and Reliability	311
8.5	Regression to the Mean	341
PART V—ISSUES OF REPORTING AND APPLICATION OF EPIDEMIOLOGIC RESULTS		349
Chapter 9—Communicating Results of Epidemiologic Studies		351
9.1	Introduction	351
9.2	What To Report	351
9.3	How To Report	356
9.4	Conclusion	369
Chapter 10—Epidemiologic Issues in the Interface with Public Health Policy		375
10.1	Introduction	375
10.2	Causality: Application to Public Health and Health Policy	376
10.3	Sensitivity Analysis	393
10.4	Meta-Analysis	395
10.5	Publication Bias	398
10.6	Summary	402
Appendix A—Standard Errors, Confidence Intervals, and Hypothesis Testing for Selected Measures of Risk and Measures of Association		411
Appendix B—Test for Trend (Dose Response)		437
Appendix C—Test of Homogeneity of Stratified Estimates (Test for Interaction)		441
Appendix D—Quality Assurance and Quality Control Procedures Manual for Blood Pressure Measurement and Blood/Urine Collection in the ARIC Study		443
Appendix E—Calculation of the Intraclass Correlation Coefficient		453
Appendix F—Exercises—Answers		457

Index ???
About the Authors ???



Preface to the 2nd Edition

This book was conceived as an intermediate epidemiology textbook. As in its 1st edition, the 2nd edition explores and discusses key epidemiologic concepts and basic methods in more depth than that found in basic textbooks on epidemiology.

As an intermediate methods text, this book is expected to have a heterogeneous readership. Some potential readers are epidemiology students who may use it as a bridge between basic and more advanced epidemiologic methods. Other readers may include those who would like to advance their knowledge beyond the basic epidemiologic principles and methods but who are not statistically minded and are thus reluctant to tackle the many excellent textbooks that strongly focus on epidemiology's quantitative aspects. Although the demonstration of many epidemiologic concepts and methods needs to rely on statistical formulations, these formulations are extensively supported in this book by real-life examples throughout, hopefully making their logic intuitively easier to follow. The practicing epidemiologist, too, may find selected portions of this book useful for an understanding of concepts beyond basics. Thus, the common denominators for the intended readers are familiarity with the basic strategies of analytic epidemiology and a desire to increase the level of understanding of several notions that are more or less (and naturally so) insufficiently covered in many basic textbooks. The way in which this textbook is organized should make this readily apparent.

In Chapter 1, the basic observational epidemiologic research strategies are reviewed, including those based on studies of both groups and individuals. Although descriptive epidemiology is not the focus of this book, birth cohort analysis is discussed in some depth in this chapter, as this approach is rarely covered in detail in basic textbooks. Another topic in the interface between descriptive and analytical epidemiology, namely ecological studies, is also discussed, with a view toward extending its discussion beyond the possibility of inferential (ecological) bias. Next, the chapter reviews observational studies based on individuals as units of observation—that is, cohort and case-control studies. Different types of case-control design are reviewed (case-based, case-control studies within a defined cohort). The strategy of matching as an approach to achieve comparability prior to data collection is also briefly discussed.

Chapters 2 and 3 cover issues of measurement of outcome frequency and measures of association. In Chapter 2, absolute measures of outcome frequency and their calculation

methods are reviewed, including the person–time approach for the calculation of incidence density and both the classic life table and the Kaplan-Meier method for the calculation of cumulative incidence. Chapter 3 deals with measures of association, including those based on relative (e.g., relative risk, odds ratio) and absolute (attributable risk) differences. The connections between measures of association obtained in cohort and case-control studies are emphasized. In particular, a description is given of the different measures of association (i.e., odds ratio, relative risk, rate ratio) that can be obtained in case-control studies as a function of the control selection strategies that were introduced in Chapter 1.

Chapters 4 and 5 are devoted to threats to the validity of epidemiologic studies, namely bias and confounding. In Chapter 4, the most common types of bias are discussed, including both selection and information bias. In the discussion of information bias, simple examples are given to improve the understanding of the phenomenon of misclassification resulting from less than perfect sensitivity and specificity of the approaches used for ascertaining exposure, outcome, and/or confounding variables. This chapter also provides a discussion of cross-sectional biases and biases associated with evaluation of screening procedures; for the latter, a simple approach to estimate lead time bias is given that may be useful for those involved in evaluative studies of this sort. In Chapter 5, the concept of confounding is introduced, and approaches to evaluate confounding are reviewed. Special issues related to confounding are discussed, including the distinction between confounders and intermediate variables, residual confounding, and the role of statistical significance in the evaluation of confounding effects.

Interaction (effect modification) is discussed in Chapter 6. The chapter discusses the concept of interaction, emphasizing its pragmatic application as well as the strategies to evaluate the presence of additive and multiplicative interactions. Practical issues discussed in this chapter include whether to adjust when interaction is suspected and the importance of the additive model in public health.

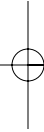
The next three chapters are devoted to approaches to handle threats to the validity of epidemiologic results. In Chapter 7, strategies for the adjustment of confounding factors are presented, including both the more parsimonious approaches (e.g., direct adjustment, Mantel-Haenszel) and the more complex (e.g., multiple regression). Emphasis is placed on the selection of the method that is most appropriate for the study design used (e.g., Cox proportional hazards for the analysis of survival data or Poisson regression for the analysis of rates per person-time). Chapter 8 reviews the basic quality control strategies for the prevention and control of measurement error and bias. Both qualitative and quantitative approaches used in quality control are discussed. The most often used analytic strategies for estimating validity and reliability of data obtained in epidemiologic studies are reviewed (e.g., unweighted and weighted kappa, correlation coefficients) in this chapter. In Chapter 9, the key issue of communication of results of epidemiologic studies is discussed. Examples of common mistakes made when reporting epidemiologic data are given as a way to stress the importance of clarity in such reports.

For the 2nd edition, we have added Chapter 10, which discusses epidemiologic issues relevant to the interface between epidemiology, health policy and public health, such as Rothman's causality model, proximal and distal causes, and Hill's guidelines. This chapter also includes a brief discussion of three topics pertinent to causal inference, namely, sensitivity analysis, meta-analysis and publication bias. Another new feature of the 2nd edition is the addition of exercises at the end of each chapter.

As in the first edition, Appendixes A, B, C, and E describe selected statistical procedures (e.g., standard errors and confidence levels, trend test, test of heterogeneity of effects, intra-class correlation) to help the reader more thoroughly evaluate the measures of risk and association discussed in the text and to expose him or her to procedures that, although relatively simple, are not available in many statistical packages used by epidemiology students and practitioners. Appendix D includes two sections on quality assurance and control procedures taken from the corresponding manual of the Atherosclerosis Risk and Communities (ARIC) Study as examples of real-life applications of some of the procedures discussed in Chapter 8. Finally, Appendix F provides the answers to the exercises.

We encourage readers to let us know of any errors or unclear passages, and to suggest improvements. Please e-mail any such suggestions or comments to: info@jbpub.com/.

All significant contributions will be acknowledged in the next edition.





Acknowledgments

This book is an outgrowth of an intermediate epidemiology course taught by the authors at the Johns Hopkins Bloomberg School of Public Health. Over the years this course has benefited from significant intellectual input of many faculty, including, among others, George W. Comstock, Helen Abbey, James Tonascia, Leon Gordis and Mary Meyer. The authors especially acknowledge George W. Comstock, a mentor to both of us, who was involved with the course for several decades. His in-depth knowledge of epidemiologic methods and wisdom over the years have been instrumental to our professional growth. Dr. Comstock also kindly provided many of the materials and examples used in chapter 9 of this book.

Our present and past chairmen, Jon Samet and Leon Gordis, have supported this endeavor consistently, for which we are deeply appreciative. We are indebted to many colleagues, including Woody Chambless, Joseph Coresh, Ana Diez-Roux, Jingzhong Ding, Rosa Maria Corona, Manning Feinleib, Eliseo Guallar, Jay Kaufman, Alfonso Mele, Paolo Pasquini, Michael Silverberg, and Eyal Shahar who reviewed partial sections of the text or provided guidance in solving conceptual or statistical riddles. For this second edition, our colleagues Daniel Brotman, Steve Cole, Leon Gordis, Paul Peppard, Patrick Remington, and Richey Sharrett provided insightful suggestions. We are especially grateful to our students Gabrielle Detjen, Salwa Massad, and Hannah Young for their careful review of the exercises. Lastly, we would like to extend our appreciation to Patty Grubb, Rodney Palmer, and Jennifer Seltzer for the technical and administrative help.

To have had the privilege of teaching intermediate epidemiology for so many years made us realize how much we have learned from our students, to whom we are deeply grateful. Finally, without the support and extraordinary patience of all members of our families, we could not have devoted so much time and effort into writing both editions of this book.

