

Epidemiology

Beyond the Basics

FOURTH EDITION

Moyses Szklo, MD, DrPH

University Distinguished Service Professor
of Epidemiology and Medicine
Johns Hopkins University

Editor in Chief, *American Journal of Epidemiology*
Baltimore, Maryland

F. Javier Nieto, MD, PhD

Dean and Professor of Epidemiology
College of Public Health and Human Sciences
Oregon State University
Corvallis, Oregon



JONES & BARTLETT
LEARNING



World Headquarters
Jones & Bartlett Learning
5 Wall Street
Burlington, MA 01803
978-443-5000
info@jblearning.com
www.jblearning.com

Jones & Bartlett Learning books and products are available through most bookstores and online booksellers. To contact Jones & Bartlett Learning directly, call 800-832-0034, fax 978-443-8000, or visit our website, www.jblearning.com.

Substantial discounts on bulk quantities of Jones & Bartlett Learning publications are available to corporations, professional associations, and other qualified organizations. For details and specific discount information, contact the special sales department at Jones & Bartlett Learning via the above contact information or send an email to specialsales@jblearning.com.

Copyright © 2019 by Jones & Bartlett Learning, LLC, an Ascend Learning Company

All rights reserved. No part of the material protected by this copyright may be reproduced or utilized in any form, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

The content, statements, views, and opinions herein are the sole expression of the respective authors and not that of Jones & Bartlett Learning, LLC. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not constitute or imply its endorsement or recommendation by Jones & Bartlett Learning, LLC and such reference shall not be used for advertising or product endorsement purposes. All trademarks displayed are the trademarks of the parties noted herein. *Epidemiology: Beyond the Basics, Fourth Edition*, is an independent publication and has not been authorized, sponsored, or otherwise approved by the owners of the trademarks or service marks referenced in this product.

There may be images in this book that feature models; these models do not necessarily endorse, represent, or participate in the activities represented in the images. Any screenshots in this product are for educational and instructive purposes only. Any individuals and scenarios featured in the case studies throughout this product may be real or fictitious, but are used for instructional purposes only.

Production Credits

VP, Product Management: David D. Cella
Director of Product Management: Michael Brown
Product Specialist: Carter McAlister
Production Editor: Vanessa Richards
Senior Marketing Manager: Sophie Fleck Teague
Manufacturing and Inventory Control Supervisor: Amy Bacus
Composition: S4Carlisle Publishing Services

Cover Design: Scott Moden
Rights & Media Specialist: Merideth Tumasz
Media Development Editor: Shannon Sheehan
Cover Image (Title Page, Part Opener, Chapter Opener):
© Enculescu Marian Vladut/Shutterstock
Printing and Binding: Edwards Brothers Malloy
Cover Printing: Edwards Brothers Malloy

Library of Congress Cataloging-in-Publication Data

Names: Szklo, M. (Moyses), author. | Nieto, F. Javier, author.
Title: *Epidemiology : beyond the basics* / Moyses Szklo, F. Javier Nieto.
Description: Fourth edition. | Burlington, Massachusetts : Jones & Bartlett Learning, [2019] | Includes bibliographical references and index.
Identifiers: LCCN 2017061666 | ISBN 9781284116595 (pbk.)
Subjects: | MESH: Epidemiologic Methods | Epidemiologic Factors | Epidemiology
Classification: LCC RA651 | NLM WA 950 | DDC 614.4--dc23
LC record available at <https://lccn.loc.gov/2017061666>

6048

Printed in the United States of America
22 21 20 19 18 10 9 8 7 6 5 4 3 2 1

Contents

Preface vi
Acknowledgments ix
About the Authors x

PART 1 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 15
1.4 Studies Based on Individuals as Observation Units 19
References 41
Exercises 44

PART 2 Measures of Disease Occurrence and Association 49

Chapter 2 Measuring Disease Occurrence 51

2.1 Introduction 51
2.2 Measures of Incidence 52
2.3 Measures of Prevalence 80
2.4 Odds 82
References 82
Exercises 84

Chapter 3 Measuring Associations Between Exposures and Outcomes 87

3.1 Introduction 87
3.2 Measuring Associations in a Cohort Study ... 87
3.3 Cross-Sectional Studies: Point Prevalence Rate Ratio 102
3.4 Measuring Associations in Case-Control Studies 103
3.5 Assessing the Strength of Associations... 115
References 118
Exercises 119

PART 3 Threats to Validity and Issues of Interpretation 125

Chapter 4 Understanding Lack of Validity: Bias 127

4.1 Overview 127
4.2 Selection Bias 129
4.3 Information Bias 135
4.4 Combined Selection/Information Biases... 153
References 168
Exercises 171

Chapter 5 Identifying Noncausal Associations: Confounding ... 175

5.1 Introduction 175
5.2 The Nature of the Association Between the Confounder, the Exposure, and the Outcome 178
5.3 Theoretical and Graphic Aids to Frame Confounding 184
5.4 Assessing the Presence of Confounding ... 186

5.5 Additional Issues Related to Confounding ... 193
 5.6 Conclusion 202
 References 203
 Exercises..... 205

Chapter 6 Defining and Assessing Heterogeneity of Effects: Interaction.....209

6.1 Introduction..... 209
 6.2 Defining and Measuring Effect..... 210
 6.3 Strategies to Evaluate Interaction 211
 6.4 Assessment of Interaction in Case-Control Studies 223
 6.5 More on the Interchangeability of the Definitions of Interaction 231
 6.6 Which Is the Relevant Model? Additive or Multiplicative 232
 6.7 The Nature and Reciprocity of Interaction... 235
 6.8 Interaction, Confounding Effect, and Adjustment..... 239
 6.9 Statistical Modeling and Statistical Tests for Interaction..... 241
 6.10 Interpreting Interaction..... 242
 6.11 Interaction and Search for New Risk Factors in Low-Risk Groups..... 248
 6.12 Interaction and “Representativeness” of Associations 249
 6.13 A Simplified Flow Chart for Evaluation of Interaction..... 251
 References 252
 Exercises..... 254

PART 4 Dealing With Threats to Validity 257

Chapter 7 Stratification and Adjustment: Multivariate Analysis in Epidemiology.....259

7.1 Introduction..... 259
 7.2 Stratification and Adjustment Techniques to Disentangle Confounding..... 260
 7.3 Adjustment Methods Based on Stratification..... 265

7.4 Multiple-Regression Techniques for Adjustment..... 279
 7.5 Alternative Approaches for the Control of Confounding 317
 7.6 Incomplete Adjustment: Residual Confounding 327
 7.7 Overadjustment 329
 7.8 Conclusion 331
 References 335
 Exercises..... 339

Chapter 8 Quality Assurance and Control349

8.1 Introduction..... 349
 8.2 Quality Assurance 351
 8.3 Quality Control 354
 8.4 Indices of Validity and Reliability 364
 8.5 Regression to the Mean..... 400
 8.6 Final Considerations 402
 References 402
 Exercises..... 405

PART 5 Issues of Reporting and Application of Epidemiologic Results 409

Chapter 9 Communicating Results of Epidemiologic Studies411

9.1 Introduction..... 411
 9.2 What to Report..... 411
 9.3 How to Report..... 416
 9.4 Conclusion 431
 References 431
 Exercises..... 433

Chapter 10 Epidemiologic Issues in the Interface With Public Health Policy437

10.1 Introduction..... 437
 10.2 Causality: Application to Public Health and Health Policy 439
 10.3 Decision Tree and Sensitivity Analysis 456

10.4 Meta-analysis..... 461
 10.5 Publication Bias..... 465
 10.6 Translational Epidemiology 469
 10.7 Summary..... 472
 References 473
 Exercises..... 478

Appendix A Standard Errors, Confidence Intervals, and Hypothesis Testing for Selected Measures of Risk and Measures of Association483

Appendix B Test for Trend (Dose Response)509

Appendix C Test of Homogeneity of Stratified Estimates (Test for Interaction).....513

Appendix D Quality Assurance and Quality Control Procedures Manual for Blood Pressure Measurement and Blood/ Urine Collection in the ARIC Study.....517

Appendix E Calculation of the Intraclass Correlation Coefficient.525

Appendix F Answers to Exercises.....529

Index.....565

Preface

This book was conceived as an intermediate epidemiology textbook. Similar to previous editions, the fourth edition discusses key epidemiologic concepts and basic methods in more depth than that found in basic textbooks on epidemiology. For the fourth edition, new examples and exercises have been added to all chapters. In addition, several new topics have been introduced, including the use of negative controls to evaluate for the presence of confounding, a simple way to understand adjustment using multiple regression, use of a single analytic unit, and translational epidemiology. Some concepts that were discussed in previous editions have been expanded, including efficacy and ways to conceptualize a control group.

As an intermediate methods text, this book is expected to have a heterogeneous readership. Epidemiology students may wish to use it as a bridge between basic and more advanced epidemiologic methods. Other readers may desire to advance their knowledge beyond basic epidemiologic principles and methods but are not statistically minded and, therefore, reluctant to tackle the many excellent textbooks that strongly focus on epidemiology's quantitative aspects. The demonstration of several epidemiologic concepts and methods needs to rely on statistical formulations, and this text extensively supports these formulations with real-life examples, thereby making their logic intuitively easier to follow. The practicing epidemiologist may find selected portions of this book useful for an understanding of concepts beyond the basics. Thus, the common denominators for the intended readers are familiarity with the basic strategies of analytic epidemiology and a desire to increase their level of understanding of several notions that are insufficiently covered (and naturally so) in many basic textbooks. The way in which this textbook is organized makes 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 because 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. The strategy of *matching* as an approach by which 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 methods 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. The “natural history” of a study is discussed, which allows distinguishing between these two concepts. In Chapter 4, the most common types of bias are discussed, including selection bias 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, which 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, the role of statistical significance in the evaluation of confounding effects, and the use of negative controls.

Interaction (effect modification) is discussed in Chapter 6. The chapter presents the concept of interaction, emphasizing its pragmatic application as well as the strategies used 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. A new flow chart is presented at the end of the chapter summarizing the main steps in the evaluation of interaction.

The next three chapters are devoted to the approaches used to handle threats to the validity of epidemiologic results. In Chapter 7, strategies for the adjustment of confounding factors are presented, including the more parsimonious approaches (e.g., direct adjustment, Mantel-Haenszel) as well as the more complex approaches (i.e., multiple regression, instrumental variables, Mendelian randomization, and propensity scores). 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 and 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.

Chapter 10 discusses—from the epidemiologist’s viewpoint—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 brief discussions of three topics pertinent to causal inference—sensitivity analysis, meta-analysis, and publication bias—and consideration of the decision tree as a tool to evaluate interventions. A new section reviews the process of translational epidemiology.

As in the previous editions, Appendices A, B, C, and E describe selected statistical procedures (e.g., standard errors and confidence levels, trend test, test of heterogeneity of effects, intraclass 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

viii **Preface**

the Atherosclerosis Risk in 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 advise us of any errors or unclear passages and to suggest improvements. Please email any such suggestions or comments to info@jblearning.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 members, including, among others, George W. Comstock, Helen Abbey, James Tonascia, Leon Gordis, and Mary Meyer. The authors especially acknowledge the late 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 his 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. The original idea for developing a textbook on intermediate epidemiologic methods came from Michel Ibrahim, to whom we are very grateful.

We are indebted to many colleagues, including Leonelo Bautista, Daniel Brotman, Woody Chambless, Steve Cole, Josef Coresh, Rosa Maria Corona, Ana Diez-Roux, Jingzhong Ding, Manning Feinleib, Leon Gordis, Eliseo Guallar, Jay Kaufman, Kristen Malecki, Alfonso Mele, Paolo Pasquini, Paul Peppard, Patrick Remington, Jonathan Samet, Eyal Shahar, Richey Sharrett, and Michael Silverberg. These colleagues reviewed partial sections of this or previous editions or provided guidance in solving conceptual or statistical riddles. We are especially grateful to Blake Buchalter, Gabrielle Rude, Salwa Massad, Margarete (Grete) Wichmann, Hannah Yang, and Jennifer Deal for their careful review of some of the exercises and portions of the text. The authors are also grateful to Lauren Wisk for creating the ancillary instructor materials for this text. Finally, we would like to extend our appreciation to Patty Grubb, Michelle Mahana, and Jennifer Seltzer for their administrative help.

Having enjoyed 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, particularly our wives, Hilda and Marion, we could not have devoted so much time and effort to writing the four editions of this text.

About the Authors

Moyses Szklo, MD, DrPH, is University Distinguished Service Professor of Epidemiology and Medicine (Cardiology) at Johns Hopkins University. His current research focuses on risk factors for subclinical and clinical atherosclerosis. He is also editor in chief of *American Journal of Epidemiology*.

F. Javier Nieto, MD, PhD, is Dean and Professor of Epidemiology at the College of Public Health and Human Sciences at Oregon State University. His current research focuses on epidemiology of cardiovascular and sleep disorders, population-based survey methods, and global health.