Book Review

Epidemiology and Biostatistics
Edited by Tongzhang Zheng, Paolo Boffetta, and Peter Boyle


The International Prevention Research Institute in Lyon, France, recently published the first edition of an introductory textbook, Epidemiology and Biostatistics. The Institute created the book by collating teaching materials used in its summer course, which is offered in conjunction with the University of Dundee (Dundee, United Kingdom). The summer course aims to build capacity in public health and epidemiology worldwide, particularly in lower-resource settings. Consistent with that emphasis, this text provides a comprehensive and practical overview of design and analysis topics fundamental to the conduct and interpretation of public health research.

The text begins with a chapter that reviews the history of epidemiologic involvement in public health, dating to 350 BC. While I have never seen this distant history so thoroughly and entertainingly integrated, I would have appreciated an extension into the 20th and 21st centuries, with perhaps mention of such landmarks as the Framingham Heart Study, the Nurses’ Health Study/Physicians’ Health Study, the American Cancer Society’s Cancer Prevention Studies, and the various registry-based studies carried out in Europe, Canada, and elsewhere. These more current studies are important because they illustrate the dramatic influence that epidemiology can have on public health.

The main strength of the text is the opening section, which focuses on age, cohort, and period analyses. This topic often receives too little attention in methodological texts, but here each piece receives an individual chapter, and then a fourth chapter ties all the pieces together. This treatment of age-period-cohort effects is the most clear and expansive explanation that I have read, and in my view it is the best rationale for adding this particular volume to one’s collection. By comparison, the third edition of Modern Epidemiology (1) devotes 1 page to the topic, and Koepsell and Weiss’s Epidemiologic Methods (2) has 4. In this section, I also first encountered a potential limitation for some readers that runs throughout most of the book: The examples come predominantly from cancer epidemiology. While cancer prevention and control is a worldwide problem, I suspect that epidemiologists practicing in lower-resource settings would appreciate use of a wider portfolio of health topics as examples.

Other chapters I considered to be especially well-written were the chapters titled “Overview of Study Designs,” “Cohort Study,” “Cross-Sectional Study,” “Case-Crossover Study,” “Meta and Pooled Analysis,” and “Screening.” Chapters that would benefit from expansion in the next edition include the chapters on statistical analysis, where there is too little emphasis on single population estimates of disease frequency and association and very little attention paid to analyzing risk and rate data as opposed to odds ratios.

Following the section on age-period-cohort effects, the text parallels the outlines of many other methodological texts, with sections on study design, biases, and the role of chance. After that comes the section on data analysis, including univariate analysis, stratified analysis, fundamentals of regression modeling, and meta-analysis. Chapters tend to be concise, clearly written, and well-illustrated with examples, and they make good use of figures, tables, and equations to reinforce the main points. A main strength of most chapters is the straightforward and practical advice for good practice. For example, a section of the chapter titled “Exposure” provides excellent descriptions and examples of exposure dose, exposure duration, induction period, cumulative exposure, time since first exposure, and time since cessation of exposure.

The main limitation of most chapters is the lack of a theoretical foundation for many concepts. The very section that provides excellent descriptions of exposure characteristics omits the importance of defining the absence of exposure (a reference group). As a second example, the chapter titled “Case-Control Study” describes the case-cohort design and explains that controls represent the distribution of exposure in the source population. It also says that odds ratios from case-control studies estimate the risk ratio only if the disease is rare, when in fact the case-cohort odds ratio estimates the risk ratio regardless of whether the disease is rare. Had the chapter included a theoretical foundation for control selection, this methodological disconnect might have been avoided. I frequently encountered similar examples as I read the text, leading me to conclude that the text is better suited to guiding practice grounded in example rather than in theory.

Anyone who has taught introductory epidemiology has faced the challenge of organizing the flow of information across a course. How does one explain measures of disease frequency without first explaining populations? But it is very hard to explain the relevance of a population’s person-time without first describing rates, which are a measure of disease frequency. In this text, readers will likely find it frustrating to be introduced to concepts without a clear
definition at first introduction. This frustration could be somewhat alleviated in a second edition by using major section headings to subset the chapters, by using a better organized outline for the sequence of chapters, and by including cross-references between chapters that present related material. These needed editorial additions are perhaps missing from this text because chapters were authored by many different contributors.

The preface advises readers that the text is aimed at students between an introductory level and an intermediate level. This target manifests as somewhat uneven coverage of some topics. An entire chapter is devoted to proportional mortality studies, a design seldom used these days. On the other hand, the chapter on the case-control design provides little coverage of the case-cohort design and no coverage of incidence-density sampling or risk-set sampling, all of which are more commonly used today. Two chapters on molecular epidemiology address an important and growing topic in epidemiology, but other currently important topics like social epidemiology are not addressed.

As I read this text, I noticed multiple occurrences of statements that I recall reading in my introductory and intermediate texts, only to later learn that they were not made with sufficient rigor. Examples include: “[the] ‘P value … corresponds to the probability that an observed association is due to chance alone” (p. 234); “the confidence interval … provide[s] a range of plausible values that could be found if a similar study [were] conducted again” (p. 236); “non-differential misclassification (of dichotomous variables) predictably biases associations towards the null” (p. 201); and “we … produce attributable fractions for one risk factor as the percentage of disease that could be avoided in the absence of exposure” (p. 186). Statements such as these are easy to find in many texts and lecture notes, particularly those designed for introductory and intermediate courses. Consequently, I do not wish to offer them as specific criticisms of this text. I wonder, however, whether as a college of research practitioners and educators we should be teaching incomplete or inaccurate concepts to students who may never learn them correctly, or may unlearn them and relearn them, or will supplement their course work to learn more advanced and complete concepts. Undoubtedly there are topics that are too complex to fully digest when first encountered, and some have argued that examples like those above are among these topics (3). I disagree (4), and urge that these concepts and others like them more often receive a full treatment even in introductory and intermediate teaching materials.

The inaugural edition of Epidemiology and Biostatistics provides a comprehensive review of study design, analysis, and interpretation. Particular strengths include the sections on age-period-cohort effects, data collection, and study design. The text is generally well-written and I expect will be well-understood by readers for whom English is a second language. The authors have successfully endeavored throughout to illustrate concepts with figures, tables, and examples. A second edition would benefit from expansion of the examples to topics outside of cancer epidemiology, somewhat better organization of the bookwide outline, cross-referencing between related sections, and advancement of the theoretical basis and completeness of some concepts. The target audience of public health practitioners in lower-resource regions would, I think, appreciate the addition of chapters on topics such as environmental and occupational epidemiology, infectious disease epidemiology, social determinants of health, and reproductive epidemiology. Nonetheless, most such practitioners will find the practical emphasis of the text and the illustration of most topics with good examples to be of real utility.

ACKNOWLEDGMENTS

Conflict of interest: none declared.

REFERENCES


Timothy L. Lash1,2 (e-mail: tlash@wakehealth.edu)

1 Department of Epidemiology and Prevention, School of Medicine, Wake Forest University, Winston-Salem, NC 27106
2 Department of Clinical Epidemiology, Aarhus University Hospital, DK-8200 Aarhus N, Denmark

DOI: 10.1093/aje/kws339; Advance Access publication: September 18, 2012