Springer, New York, 2009. xii + 194 pp. $79.95/€64.15. ISBN 9780387879604.

Bias analysis is an emerging area of innovation in epidemiology methods. It is concerned with the formal assessment of uncertainty from bias, which the authors define as uncertainty that does diminish with increasing sample size. Common examples include the effects of missing confounders, measurement error, and the choice of model specification. Bias analysis is conceptually similar to sensitivity analysis, where the investigator studies the robustness of study findings to small departures in modeling assumptions; for example, by adding or dropping covariates in a regression model.

Formal bias analysis techniques often get short shrift in epidemiological texts. Nonetheless observational research is increasingly concerned with scientific problems that are plagued with biases, such as in meta analysis and studies using large administrative databases and registries. This shifting research focus, coupled with innovation in computation, has set the stage for surging interest in bias analysis techniques.

Applying Quantitative Bias Analysis to Epidemiologic Data is the first text of its kind to give a comprehensive overview of the field. As the title suggests, it focuses on epidemiological applications, although many of the methods carry over to other disciplines such as the social sciences and economics. The text is aimed at a readership that has a firm grounding epidemiologic methods and intermediate data analysis typical of a graduate program in epidemiology. It could serve as a reference manual or a companion text for an advanced methods course.

The first three chapters provide an overview of bias analysis. Chapter 1 outlines the objectives of the text and gives historical perspective on the treatment of uncertainty in epidemiologic research. It also reviews the limitations of different heuristics and qualitative approaches to uncertainty assessment, thereby motivating quantitative methods. Chapter 2 is a guide to choosing and implementing bias analysis tools. It gives an overview of the steps in planning and conducting a bias analysis including whether or not it is worthwhile. As background they discuss design and analysis strategies for reducing random and systematic errors. Chapter 3 discusses data sources for bias analysis with special emphasis on validity data. The authors distinguish between internal validity data, where information is obtained from a subsample of the study, versus external validity data, which uses alternative study populations or guess work.

Chapters 4 to 6 concern the key types of bias in epidemiology, namely selection bias, confounding, and measurement error. Each chapter is dedicated to one of these biases and presents models and parameterizations with a focus on 2×2 tables for exposure and disease relationships. A strength of the presentation is that the model setups are fairly exhaustive and cover a variety of settings. For example, Chapter 6 reviews models for misclassification of the outcome, exposure, or covariates. The authors also describe how practitioners can elicit inputs in bias analysis. Conceptualizing and quantifying bias can be fairly abstract and the authors give a clear exposition of the process. A great strength of the text is the detailed examples and calculations. The authors supply software in the form of Excel spreadsheets and `SAS` programs that can be downloaded from the text website. I myself tried using the software and found it straightforward to manipulate and well laid out.

Chapters 7 and 8 motivate multidimensional bias analysis and probabilistic bias analysis. In practice there may be uncertainty in the user inputs for bias adjustment. We can explore sensitivity to a range of possible values. A different approach is to supply probability distributions that model beliefs about the inputs in the spirit of a Bayesian analysis. The resulting analysis provides a synthesis of bias uncertainty with the usual uncertainty from random error.

Chapter 9 discusses multiple bias modeling, which is perhaps the most realistic setting for observational data. It concerns the case where there are multiple sources of uncertainty that can impact results. The authors describe sequential bias analysis techniques where the investigator can "unpack" the analysis results by removing biases one at a time. Uncertainty in the bias parameters plays a crucial role in this case, and the authors illustrate how probabilistic approaches can reduce the burden of presentation of results. I particularly enjoyed reading the discussion of the ordering of bias corrections. Chapter 10 concludes with a discussion of presentation and inference.

This book fills an important gap among epidemiology texts. It provides a unified reference for the myriad of bias analysis methods that appear in the literature. It is broad and thorough in scope, and yet easily accessible and well organized. I would recommend it to health researchers who seek to keep abreast of the latest research in epidemiology methods.

Lawrence McCandless
Faculty of Health Sciences
Simon Fraser University
Burnaby, British Columbia, Canada

LEMEY, P., SALEMI, M., and VANDAMME, A.-M. (eds). **The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing**, 2nd edition. Cambridge University Press, Cambridge, UK, 2009. xxvi + 723 pp. $125.00/£70.00. ISBN 9780521877107.

Phylogenetics is the reconstruction and analysis of phylogenetic trees and networks from genetic data. It is a rapidly growing field that has become central to systematic biology, viral evolution, conservation biology, and the study of evolutionary processes.

This impressive handbook (more than 700 pages) provides unparalleled coverage of the practice of turning genetic data into sensible statements about evolutionary history and processes. With its 21 chapters organized into seven sections, and 37 pages of references, it is a sort of "Hitchhiker's Guide" to phylogenetics, as suggested by the Douglas Adams quote (in Edward Holmes' Foreword section) that finishes: "DON'T PANIC."

A notable aspect of the book is the clean division of each chapter into two parts: "Theory" followed by "Practice." The "Theory" describes the underlying mathematical, statistical,

or computational ideas, while "Practice" illustrates the use of the techniques, typically with real datasets, and one or more software packages. Each chapter is written by different authors; in many cases, they are precisely the developers of the software that is demonstrated in the practice sections.

The book is attractively presented, from the Dali-themed design of a melting clock on the cover (a "relaxed molecular clock?") to the extensive 18-page glossary of technical terms near the back. The figures include helpful screen shots (some in color) showing the output of various phylogenetic packages.

Over the last decade, a number of books have told different aspects of this story—most notably, the comprehensive *Inferring Phylogenies* (Felsenstein, 2004) and the popular hands-on manual *Phylogenetic Trees Made Easy* (Hall, 2007), now in its third edition. But the style and content of this book is quite different to these, or to other more technical books such as *Computational Molecular Evolution* (Yang, 2006).

The first edition of the current book (2003, with a different subtitle, and one fewer editor) pioneered the theory-application format, but lacked coverage of some relevant or recent developments. The second edition provides a significant boost and update, with six new chapters with a stronger focus on statistical aspects, the description of new software, and more attention to areas like recombination analysis, Bayesian phylogenetics, and genealogy-based population genetics.

The book starts with a concise overview by one editor (Vandamme) of molecular data and evolution, the nature of phylogenetic trees, and the sorts of problems that make inferring phylogenetic trees fascinating yet complex. It sets the stage for the seven sections that follow.

The second section has two chapters that deal with sequence database searching and multiple sequence alignment—here the reader will learn the basic theory and practice of BLAST and FASTA along with an array of methods for aligning sequences, with particular emphasis on CLUSTAL.

Section 3 is arguably the core of the book, with six chapters on phylogenetic inference. These cover nucleotide substitution models, phylogenetics using distance methods, maximum likelihood, Bayesian methods, parsimony, and protein sequences. Standard phylogenetic software—`PAUP*`, `RAxML`, `MrBayes`, and `PHYLIP`—along with more specialized packages are illustrated on various datasets.

Methods for testing models and trees follow in Section 4, with chapters on model selection, molecular clock tests, and posterior statistical tests on trees, such as bootstrap resampling. The other substantial section is a collection of three chapters on aspects of population genetics, which gives an overview of Coalescent and Bayesian methods for sampling trees (illustrated using BEAST), and the estimation of population genetic parameters (illustrated by the LAMARK software package).

The remaining handful of chapters throughout the latter half of the book deal with a variety of relevant topics such as molecular adaption, recombination, site saturation (using DAMBE), and split networks (using SPLITSTREE).

What is missing? The book covers a great deal but is not all encompassing, nor is it intended to be. There is relatively little coverage of some recent approaches in phylogenetics for which some software exists—these include phylogenomics (where the data consists of gene content, gene order, rare genomic events such as SINEs, and alignment-free phylogenetic methods), metagenomics, gene-tree species tree resolution (lineage sorting), explicit methods for phylogenetic network reconstruction (e.g., for lateral gene transfer), and applications of phylogenetics in biodiversity conservation. These new areas—and some more classic approaches (such as phylogenetic comparative methods)—might be considered, perhaps, in a third edition.

There is also relatively little historical overview, but this is not the focus of this book. Rather it is intended as a hands-on resource for the student or professional scientist who wishes to master basic phylogenetic concepts and get to grips with the software—in this regard it succeeds admirably.

References

Felsenstein, J. (2004). *Inferring Phylogenies*. Sunderland, Massachusetts: Sinauer Associates.

Hall, B. G. (2007). *Phylogenetic Trees Made Easy*. Sunderland, Massachusetts: Sinauer Associates.

Yang, Z. (2006). *Computational Molecular Evolution*. New York: Oxford University Press.

Mike Steel
Biomathematics Research Centre
University of Canterbury
Christchurch, New Zealand