mathematics at both undergraduate and graduate level.

Jordan Stoyanov
Newcastle University
E-mail: jordan.stoyanov@ncl.ac.uk

Applying Quantitative Bias Analysis to Epidemiologic Data
T. L. Lash, M. P. Fox and A. K. Fink, 2009
New York, Springer
xii + 192 pp., £53.99
ISBN 978-0-387-87960-4

Virtually all epidemiological textbooks insist on the importance of quantitative bias analysis for observational studies but none has offered so far a simple and systematic guide on how to do it.

Applying Quantitative Bias Analysis to Epidemiologic Data aims at changing the state of things by

> 'collect[ing] existing methods of quantitative bias analysis, explain[ing] them, illustrat[ing] them with examples and link[ing] them to tools for implementation'

(page 3).

Chapter 1 uses evidence from evolutionary psychology to show that epidemiologists, as all human beings, have intrinsic sources of bias and that 'debiasing', i.e. the sometimes promoted attitude of reducing one's enthusiasm about the study results, simply will not work.

> 'The best option to break the [current] cycle [of reporting only qualitative analyses] is to become more adept at quantifying error, which will mitigate the tendency toward overconfidence'

(page 12). Chapter 2 provides a neat classification of selected techniques for bias analysis (simple bias analysis, multi-dimensional bias analysis, probabilistic bias analysis and multiple-bias analysis) on the basis of the bias parameters, the number of biases that can be modelled, the possibility of calculating a random error and computational intensiveness. Chapter 3 describes internal and external sources for bias analysis. Even when internal data are available, they should not be handled as if they were measured without error and external data may be of value to reinforce or to put into context the study results.

Chapter 4 teaches how to handle selection bias with simple bias analysis. Often, one will have to make an 'educated guess' on the prevalence of the disease and the exposure among non-participants. Although this notion may engender some discomfort, it is better than the alternative of

> 'ignor[ing] the potential for selection bias to affect the study results'

(page 49). Chapter 5 discusses how to handle unmeasured and unknown confounders with simple bias analysis applied to ratio and difference measures with and without effect modification. Chapter 6 is about misclassification bias and provides strong medicine against common errors. For instance, it is a commonly heard mantra that 'non-differential misclassification will bias results toward the null'. However, Dr Lash and colleagues show that this dogma should be replaced by the following evidence (simulation)-based statement:

> 'on average, nondifferential misclassification of a dichotomous variable will bias results toward the null'

(page 94), which clearly does *not* authorize a qualitative approach to the problem. The chapter suggests simple ways of modelling misclassification bias by using sensitivity and specificity or positive and negative predictive values and discusses also the case of dependent misclassification.

Chapter 7 is dedicated to multi-dimensional bias analysis, where

> 'the analytic procedures [of simple bias analysis] are simply repeated with multiple combinations of values assigned to bias parameters'

(page 110). Chapter 8 discusses probabilistic bias analysis, which is a way of overcoming some limitations of simple and multi-dimensional bias analysis. With this approach, a probability distribution is assigned to the bias parameters and a Monte Carlo simulation is performed to obtain a median estimate and 'simulations intervals' corresponding to the 2.5th and 97.5th percentiles. Dr Lash and colleagues discuss the advantages and limitations of five distributions (the uniform, trapezoidal, triangular, Gaussian and beta distributions) and apply two of them (uniform and triangular) to real examples from the epidemiologic literature. An advantage of probabilistic bias analysis is the possibility of calculating a random error in addition to the systematic error. Chapter 9 introduces the more complex topic of multiple bias modelling, i.e. the analysis of all potential sources of bias inside a study. This will require knowledge of the sequence in which threats to validity arise in a given study. The chapter starts with simple methods and then considers multiple-bias analysis and probabilistic bias analysis. Chapter 10 provides very useful guidelines on how to report a quantitative bias analysis. It also makes clear that the inferential approach of probabilistic bias analysis is neither frequentist nor Bayesian. Formal semi-Bayesian methods compare

reasonably well with probabilistic bias analysis but there is no guarantee of total agreement. The authors, who are experienced and supportive users of Bayesian methods, refer readers who are interested in pure Bayesian methods to other texts and I personally agree that

> 'this text is not the forum to continue the [frequentist–Bayesian] debate'

(page 4) as its aim is to introduce the reader to a new way of thinking of bias as a 'quantity' not a 'quality' of an observational study. A great feature of this book is the availability of SAS code to do virtually all calculations and EXCEL spreadsheets to do most of them (`https://sites.google.com/site/biasanalysis/`).

If this very well-written and highly didactic book obtains the popularity that it merits, it might be a strong driving force for the routine implementation of quantitative bias analysis in epidemiological studies. One thing is sure, however: after reading this book, no one could honestly continue to say that he or she does not understand how to perform quantitative bias analysis or has not the right tools to do it.

Giorgio Bedogni
Liver Research Centre
Trieste
E-mail: giorgiobedogni@gmail.com

Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology
A. Lawson, 2009
Boca Raton, Chapman and Hall–CRC Press
xviii + 344 pp., $79.95
ISBN 978-1-584-88840-6

This book provides a technical grounding in spatial models while maintaining a strong grasp on applied epidemiological problems. It is divided into two sections. The first is called 'Background' and is essentially a brisk review of Bayesian theory and computing-intensive methods such as Markov chain Monte Carlo sampling. Important points are made concisely, e.g. the definition of the prior and posterior predictive distributions, illustrated by the negative binomial distribution. A welcome effort is made to clarify concepts which might, in other texts, have been skimmed over in a rush to fit models. For example, 'general clustering', i.e. spatial autocorrelation over the study area, is distinguished clearly from 'cluster detection *per se*', meaning localized patterns, various types of which are considered. Although I believe that cluster detection is overused in the literature, I found this treatment insightful and useful.

The chapter on computational methods describes Metropolis, Metropolis–Hastings and Gibbs algorithms and compares them in practical terms, along with different approaches to assessing convergence. From the start, the concepts are illustrated with disease mapping examples, including R and WinBUGS code. 18 data sets are used in the book. Almost all are from the USA or UK, with most georeferenced as regions rather than points. The majority are said to be available to download from `http://www.sph.sc.edu/alawson/default.htm` although I could not find them there.

The second part of the book, titled 'Themes', occupies about two-thirds of the total and addresses selected topics in spatial analysis. The first of these is 'Disease map reconstruction and relative risk estimation', which introduces some of the most familiar spatial analysis tools such as conditional auto-regressive models, with some useful sections on model specification, e.g. the utility of log-Gaussian as opposed to gamma models, and the role of correlated and uncorrelated random effects for modelling heterogeneity. There is also a very neat section on the use of Moran's I of residuals as a diagnostic for area level data.

A chapter on cluster detection is followed by one on analyses of aggregated data, e.g. those available at an area level (so-called 'ecological' analysis, which in general is not informed by ecology). There is a clear exposition of possible biases in this type of study, and their possible sources, such as measurement error, confounding or the exposure effect varying between areas (interaction).

The following chapter is devoted to problems which may be found in area level ('ecological') studies, namely the 'modifiable areal unit problem ' and the 'misaligned data problem'. One case of the former problem is 'scaling down', in which inferences at the individual level may be sought but are not necessarily reliable. The latter problem—misaligned data—would arise, for example, in analyses linking UK census tracts and postcodes. Although these problems are worthy of attention, I find it odd that no such critique is made of the practice, which is often followed in this book, of treating area level data as if they were located at the regions' centroids.

The following two chapters address topics which are, perhaps, more advanced or specialized, namely multivariate analyses, i.e. of multiple diseases, and spatial survival and longitudinal analyses. The final chapter may be of wider interest, its topic being spatiotemporal modelling, i.e. analyses of data which are not only georeferenced but are also from multiple time points. The models considered include those for log-rates and log-odds, with