after counterfactuals were introduced into the contemporary statistics literature. Pearl's own framework is heavily dependent on and closely related to the prior work. Pearl nonetheless engages closely with other causal frameworks and the additional comments that are included in the second edition help to clarify what is at stake (see especially the engagement with Heckman on pages 374–380). More discussion along these lines will surely be desirable. Further consideration of the trade-off between the strength of assumptions and the scope of the framework would also have been desirable: Pearl's structural account is remarkably versatile but it makes strong assumptions; the advantage of the potential outcomes framework and alternative graphical and non-graphical causal models is that fewer assumptions are required (see Robins and Richardson (2010)). Pearl would perhaps rightly point out that the assumptions that are made in these alternative frameworks are less transparent. This trade-off between transparency, versatility and strength of assumptions arguably merits more attention than is given in the book. Pearl states in the preface to the second edition that he hopes his 11th chapter will

> 'clear the province of causal thinking from the last traces of controversy'.

This he has not accomplished but the new material in the second edition of Pearl's text clearly has moved the discussion forward. It is moreover perhaps testament to the success of the first edition that the additions to the second are mostly in the form of commentary rather than structural change.

References

van der Laan, M. J. and Robins, J. M. (2003) *Unified Methods for Censored Longitudinal Data and Causality*. New York: Springer.
Morgan, S. L. and Winship, C. (2007) *Counterfactuals and Causal Inference*. Cambridge: Cambridge University Press.
Robins, J. M. and Richardson, T. S. (2010) Alternative graphical causal models and the identification of direct effects. In *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures* (ed. P. Shrout). Oxford: Oxford University Press.
Rosenbaum, P. R. (2002) *Observational Studies*. New York: Springer.
Rubin, D. B. (2006) *Matched Sampling for Causal Effects*. Cambridge: Cambridge University Press.

Tyler VanderWeele
Harvard University
Boston
E-mail: tvanderw@hsph.harvard.edu

Design of Observational Studies

P. R. Rosenbaum, 2010
New York, Springer
xviii + 386 pp., £72.00
ISBN 978-1-441-91212-1

William G. Cochran defined an observational study as an empirical investigation of treatment effects when random assignment to treatment or control is not possible and suggested that it should be designed to resemble a randomized experiment. *Design of Observational Studies* is a brilliantly written and thought-provoking book teaching how to do just that. It is structured in four parts: 'Beginnings', 'Matching', 'Design sensitivity' and 'Planning analysis'.

Part I is a general introduction to causal inference in observational studies. Observational studies are contrasted with randomized experiments to make clear how they can be designed 'as if' they were experimental studies. Two conceptual models of observational studies are discussed: a naive model assuming that individuals who look comparable are comparable and a more realistic model addressing the fact that individuals who look comparable for the observed covariates may differ in other ways. The task that is associated with the naive model is mostly mechanical and consists of matching treated and control subjects for observed covariates. 'The second task is not a mechanical but rather scientific task; this task is, therefore, more challenging and hence more interesting' (page 75). Sensitivity analysis is the central instrument to perform the second task, where sensitivity is to be interpreted 'as the magnitude of the departure from the naïve model that would need to be present to materially alter the study's conclusions' (page 76).

Part II is dedicated to matching and opens with an example of how a well-designed matched observational study may be presented in a scientific journal. The propensity score, i.e. the conditional probability of exposure to treatment given the observed covariates, is then introduced and it is shown how it can be improved by using distance matrices, callipers with penalty functions, optimal matching (pairing of each treated subject with a different control), matching with multiple controls and optimal full matching (pairing of each treated subject with a different control and vice versa). If matching is not satisfactory, almost exact matching, exact matching or the use of small penalties may be employed. Other kinds of matching that are discussed are fine balance (forcing a nominal variable to be balanced), matching without groups (a 'non-bipartite' matching where subjects are divided into pairs to minimize the total distance within pairs)

and risk set matching (exemplified by the study of a treatment given at various times). Part II is closed by a useful chapter teaching how to use the open source software R to build a distance matrix and to add a propensity score calliper to the matrix, and finding an optimal match.

Part III discusses sensitivity analysis and its central role in observational studies. This part teaches how to use sensitivity analysis to compare alternative designs for an observational study ('power analysis'), discusses the role of heterogeneity in observational studies (making the important observation that reducing heterogeneity confers benefits that cannot be obtained by increasing the sample size), suggests how to evaluate uncommon but dramatic responses to treatment (by contrasting different approaches) and shows how expected patterns of responses (e.g. a dose–effect relationship) may affect sensitivity to unmeasured biases.

Part IV is the shortest part of the book but conveys a very important message by showing how to improve design sensitivity and how to apply Ronald A. Fisher's suggestion 'to make theories elaborate' to disambiguate the association between treatment and outcome.

In my opinion, *Design of Observational Studies* is a brilliantly written and thought-provoking book that should be read by every researcher who is involved in the design and analysis of observational studies. This book is ideally suited to bridge the gap between standard epidemiological textbooks and more advanced books such as *Observational Studies* (Rosenbaum, 2002). Taking the liberty to answer the question which closes the preface of *Design of Observational Studies*, 'To end well, how should we begin?', I wish to reply 'By reading this book!'.

Reference

Rosenbaum, P. R. (2002) *Observational Studies*. New York: Springer.

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

ggplot2: Elegant Graphics for Data Analysis
H. Wickham, 2009
London, Springer
210 pp., £31.37
ISBN 978-0-387-98140-6

Hadley Wickham has recently made available the `ggplot2` R package, and this publication aims to explain the principles underlying this graphical language. The term language is here apposite, as Wickham has indeed striven to create a flexible graphical toolbox composed of major building blocks assembled by using a specific grammar. The package proposes two main functions: `qplot()`, which stands for quick plot and `ggplot()`, which is a more extensive graphical facility. The book parallels this dichotomy by providing the reader with an introductory chapter, demonstrating the basic usage of `qplot()`. For the unaccustomed reader, this first chapter constitutes a showcase of the possibilities that are offered by this R package.

The rest of the book tackles the different aspects of the `ggplot2` grammar. This includes scales and legends, positioning, statistical summaries and faceting. Of particular interest is Chapter 3, which describes the grammar of `ggplot2`, and the logic behind it. Overall, the book is well written and contains numerous mouth watering figures, which should motivate most readers to get started with the package.

Reading this publication, however, is not sufficient to become completely proficient in using this R package, and several posts to the `ggplot2` mailing list may be necessary to solve specific graphical problems. The success of this package is directly demonstrated by the fact that more than 1000 R users have subscribed to the `ggplot2` mailing list. As a result, I have found that most of the queries that I had were answered within an hour from being posted.

One criticism, however, is that it seems to be quite difficult to find an exhaustive list of the R functions and options that are available to the user. This is a lacuna, which is present in the book, in the R help pages and on the Web site of the package. A repository of all `ggplot2` functions with specific descriptions of the general options that are available, or some more comprehensive help pages, would be very much welcomed.

It is naturally difficult to disentangle the `ggplot2` package from the publication bearing its name. Overall, I wholly recommend the use of the package. Whether the book itself is a necessary purchase will then depend on the learning style of the package user. This publication may be well suited for those who prefer a step-by-step approach to learning R, as opposed to searching discussion groups, to find reams of code that can be tailored to a particular graphical problem.

Cedric Ginestet
King's College
London
E-mail: cedric.ginestet@kcl.ac.uk