



## Book Review

### *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*

By Ewout W. Steyerberg

ISBN: 978-0-387-77243-1, Springer, New York, New York (Telephone: 212-460-1500, Fax: 212-460-1575, E-mail: service-ny@springer.com, World Wide Web: <http://www.springer.com>), 2009, 500 pp., \$89.95

On page 2 of this book, Ewout Steyerberg notes that clinical prediction models “combine a number of characteristics (e.g., related to the patient, the disease, or treatment) to predict a diagnostic or prognostic outcome.” This book covers an important topic, because these prediction models are essential for individualizing diagnostic and treatment decision making. The topic is of increased importance as evidence-based medicine is increasingly implemented and as scientific and technologic advances reveal new potential predictors of outcome. This book presents an approach for developing, validating, and updating prediction models. In other words, it provides ways to optimally utilize regression techniques to predict an outcome.

The book is aimed at epidemiologists and applied biostatisticians who want to develop or apply a prediction model. Only a basic knowledge of regression analysis is required, including linear regression, logistic regression, and Cox regression. The book contains few formulas and little theoretical justification, so a strong background in mathematics is not required. Also provided is R code for implementing many of the techniques discussed in the book, which will be very useful for readers to reinforce the ideas being presented and will allow them to modify the code to analyze their own data. In addition, there is a website for the book that contains the data sets referred to in the book, R code used to analyze the data, exercises to reinforce the material, and additional useful material, such as a Microsoft Excel file illustrating a prediction rule for renal artery stenosis.

The author has arranged the book into 4 parts. Part I, “Prediction Models in Medicine,” provides background information on developing and applying prediction models in medicine. Small case studies in public health, clinical practice, and medical research are used for illustration. In part II, “Developing Valid Prediction Models,” issues regarding model development and application are discussed. Of particular note is that use of linear, logistic, and Cox regression models is described for prediction models, as are regression trees and neural networks. The author presents only those features of these methods that are relevant to prediction and refers the reader to other texts for a comprehensive discussion of these methods. Part III, “Generalizability of Prediction Models,” discusses applicability in external settings and advanced issues related to model modification and model extension. Finally, part IV, “Applications,” provides a detailed description of predictive modeling with a binary outcome (30-day mortality after acute myocardial infarction) and survival analysis (prediction of secondary cardiovascular events).

The book includes a significant amount of useful information for anyone interested in developing and validating prediction

models. There is a discussion of issues in designing studies for prediction research, including selecting study subjects and choosing predictors and outcome variables for prediction models. In addition, issues regarding power for studying effects of a specific predictor and for developing a prediction model are discussed. However, formulas for calculating power are not included. The book includes a 7-step checklist to be considered when developing a valid prediction model. This checklist is valuable because it provides structure to model building and validation. The 7 steps are data inspection, coding of predictors, model specification, model estimation, model performance, model validation, and model presentation. Each component of the checklist is discussed.

There are many discussions that will be helpful to the applied biostatistician or epidemiologist who is building a prediction model. In particular, there are enlightening discussions about missing values, selection of main effects and interactions, estimation of model parameters with shrinkage methods, incorporation of external data, internal validation, and evaluation of performance and clinical usefulness. It is refreshing to find a whole chapter devoted to the presentation of a prediction model because a discussion of the advantages and drawbacks to different presentation formats is often overlooked. The book also has a nice discussion about the generalizability of prediction models, including patterns of invalidity that may be encountered in new settings, approaches to modifying and extending a model, and comparisons of centers after case-mix adjustment by a prediction model.

This book is written in a clear and accessible style. Throughout, concepts and techniques are illustrated by examples. It could serve as a textbook for a graduate-level course on predictive modeling in diagnosis and prognosis; prerequisites for such a course would include an understanding of regression techniques, including linear regression, logistic regression, and Cox regression. The book would be a valuable resource for anyone interested in developing or applying a prediction model.

#### ACKNOWLEDGMENTS

Conflict of interest: none declared.

Todd A. Alonzo  
(e-mail: [talonzo@childrensoncologygroup.org](mailto:talonzo@childrensoncologygroup.org))  
*Division of Biostatistics, University of Southern California,  
Los Angeles, CA 90033*

DOI: 10.1093/aje/kwp129; Advance Access publication May 28, 2009