Chapman & Hall/CRC Biostatistics Series

Essentials of a Successful Biostatistical Collaboration

Arul Earnest

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC Biostatistics Series

Essentials of a Successful Biostatistical Collaboration

Arul Earnest
Monash University
Australia

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper
Version Date: 20160808

International Standard Book Number-13: 978-1-4822-2698-0 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: Earnest, Arul, author.
Title: Essentials of a successful biostatistical collaboration / Arul Earnest.
Description: Boca Raton : Taylor & Francis, 2016. | Series: Chapman & Hall/CRC biostatistics series | Includes bibliographical references and index.
Identifiers: LCCN 2016014111 | ISBN 9781482226980 (alk. paper)
Subjects: LCSH: Clinical medicine--Research--Statistical methods. | Sampling (Statistics)
Classification: LCC R853.S7 E27 2016 | DDC 616.0072/7--dc23
LC record available at https://lccn.loc.gov/2016014111

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Printed and bound in the United States of America by Publishers Graphics, LLC on sustainably sourced paper.

Essentials of a Successful Biostatistical Collaboration

Chapman & Hall/CRC Biostatistics Series

Published Titles

Adaptive Design Methods in Clinical Trials, Second Edition
Shein-Chung Chow and Mark Chang

Adaptive Designs for Sequential Treatment Allocation
Alessandro Baldi Antognini and Alessandra Giovagnoli

Adaptive Design Theory and Implementation Using SAS and R, Second Edition
Mark Chang

Advanced Bayesian Methods for Medical Test Accuracy
Lyle D. Broemeling

Applied Biclustering Methods for Big and High-Dimensional Data Using R
Adetayo Kasim, Ziv Shkedy, Sebastian Kaiser, Sepp Hochreiter, and Willem Talloen

Applied Meta-Analysis with R
Ding-Geng (Din) Chen and Karl E. Peace

Basic Statistics and Pharmaceutical Statistical Applications, Second Edition
James E. De Muth

Bayesian Adaptive Methods for Clinical Trials
Scott M. Berry, Bradley P. Carlin, J. Jack Lee, and Peter Muller

Bayesian Analysis Made Simple: An Excel GUI for WinBUGS
Phil Woodward

Bayesian Designs for Phase I–II Clinical Trials
Ying Yuan, Hoang Q. Nguyen, and Peter F. Thall

Bayesian Methods for Measures of Agreement
Lyle D. Broemeling

Bayesian Methods for Repeated Measures
Lyle D. Broemeling

Bayesian Methods in Epidemiology
Lyle D. Broemeling

Bayesian Methods in Health Economics
Gianluca Baio

Bayesian Missing Data Problems: EM, Data Augmentation and Noniterative Computation
Ming T. Tan, Guo-Liang Tian, and Kai Wang Ng

Bayesian Modeling in Bioinformatics
Dipak K. Dey, Samiran Ghosh, and Bani K. Mallick

Benefit-Risk Assessment in Pharmaceutical Research and Development
Andreas Sashegyi, James Felli, and Rebecca Noel

Published Titles

Benefit-Risk Assessment Methods in Medical Product Development: Bridging Qualitative and Quantitative Assessments
Qi Jiang and Weili He

Biosimilars: Design and Analysis of Follow-on Biologics
Shein-Chung Chow

Biostatistics: A Computing Approach
Stewart J. Anderson

Cancer Clinical Trials: Current and Controversial Issues in Design and Analysis
Stephen L. George, Xiaofei Wang, and Herbert Pang

Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation
Mikel Aickin

Clinical and Statistical Considerations in Personalized Medicine
Claudio Carini, Sandeep Menon, and Mark Chang

Clinical Trial Data Analysis using R
Ding-Geng (Din) Chen and Karl E. Peace

Clinical Trial Methodology
Karl E. Peace and Ding-Geng (Din) Chen

Computational Methods in Biomedical Research
Ravindra Khattree and Dayanand N. Naik

Computational Pharmacokinetics
Anders Källén

Confidence Intervals for Proportions and Related Measures of Effect Size
Robert G. Newcombe

Controversial Statistical Issues in Clinical Trials
Shein-Chung Chow

Data Analysis with Competing Risks and Intermediate States
Ronald B. Geskus

Data and Safety Monitoring Committees in Clinical Trials
Jay Herson

Design and Analysis of Animal Studies in Pharmaceutical Development
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bioavailability and Bioequivalence Studies, Third Edition
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bridging Studies
Jen-pei Liu, Shein-Chung Chow, and Chin-Fu Hsiao

Design & Analysis of Clinical Trials for Economic Evaluation & Reimbursement: An Applied Approach Using SAS & STATA
Iftekhar Khan

Design and Analysis of Clinical Trials for Predictive Medicine
Shigeyuki Matsui, Marc Buyse, and Richard Simon

Design and Analysis of Clinical Trials with Time-to-Event Endpoints
Karl E. Peace

Design and Analysis of Non-Inferiority Trials
Mark D. Rothmann, Brian L. Wiens, and Ivan S. F. Chan

Difference Equations with Public Health Applications
Lemuel A. Moyé and Asha Seth Kapadia

DNA Methylation Microarrays: Experimental Design and Statistical Analysis
Sun-Chong Wang and Arturas Petronis

DNA Microarrays and Related Genomics Techniques: Design, Analysis, and Interpretation of Experiments
David B. Allison, Grier P. Page, T. Mark Beasley, and Jode W. Edwards

Dose Finding by the Continual Reassessment Method
Ying Kuen Cheung

Dynamical Biostatistical Models
Daniel Commenges and Hélène Jacqmin-Gadda

Elementary Bayesian Biostatistics
Lemuel A. Moyé

Empirical Likelihood Method in Survival Analysis
Mai Zhou

Essentials of a Successful Biostatistical Collaboration
Arul Earnest

Published Titles

Exposure–Response Modeling: Methods and Practical Implementation
Jixian Wang

Frailty Models in Survival Analysis
Andreas Wienke

Fundamental Concepts for New Clinical Trialists
Scott Evans and Naitee Ting

Generalized Linear Models: A Bayesian Perspective
Dipak K. Dey, Sujit K. Ghosh, and Bani K. Mallick

Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries
Daryl S. Paulson

Inference Principles for Biostatisticians
Ian C. Marschner

Interval-Censored Time-to-Event Data: Methods and Applications
Ding-Geng (Din) Chen, Jianguo Sun, and Karl E. Peace

Introductory Adaptive Trial Designs: A Practical Guide with R
Mark Chang

Joint Models for Longitudinal and Time-to-Event Data: With Applications in R
Dimitris Rizopoulos

Measures of Interobserver Agreement and Reliability, Second Edition
Mohamed M. Shoukri

Medical Biostatistics, Third Edition
A. Indrayan

Meta-Analysis in Medicine and Health Policy
Dalene Stangl and Donald A. Berry

Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools
Marc Lavielle

Modeling to Inform Infectious Disease Control
Niels G. Becker

Modern Adaptive Randomized Clinical Trials: Statistical and Practical Aspects
Oleksandr Sverdlov

Monte Carlo Simulation for the Pharmaceutical Industry: Concepts, Algorithms, and Case Studies
Mark Chang

Multiregional Clinical Trials for Simultaneous Global New Drug Development
Joshua Chen and Hui Quan

Multiple Testing Problems in Pharmaceutical Statistics
Alex Dmitrienko, Ajit C. Tamhane, and Frank Bretz

Noninferiority Testing in Clinical Trials: Issues and Challenges
Tie-Hua Ng

Optimal Design for Nonlinear Response Models
Valerii V. Fedorov and Sergei L. Leonov

Patient-Reported Outcomes: Measurement, Implementation and Interpretation
Joseph C. Cappelleri, Kelly H. Zou, Andrew G. Bushmakin, Jose Ma. J. Alvir, Demissie Alemayehu, and Tara Symonds

Quantitative Evaluation of Safety in Drug Development: Design, Analysis and Reporting
Qi Jiang and H. Amy Xia

Quantitative Methods for Traditional Chinese Medicine Development
Shein-Chung Chow

Randomized Clinical Trials of Nonpharmacological Treatments
Isabelle Boutron, Philippe Ravaud, and David Moher

Randomized Phase II Cancer Clinical Trials
Sin-Ho Jung

Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research
Chul Ahn, Moonseong Heo, and Song Zhang

Published Titles

Sample Size Calculations in Clinical Research, Second Edition
Shein-Chung Chow, Jun Shao, and Hansheng Wang

Statistical Analysis of Human Growth and Development
Yin Bun Cheung

Statistical Design and Analysis of Clinical Trials: Principles and Methods
Weichung Joe Shih and Joseph Aisner

Statistical Design and Analysis of Stability Studies
Shein-Chung Chow

Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis
Kelly H. Zou, Aiyi Liu, Andriy Bandos, Lucila Ohno-Machado, and Howard Rockette

Statistical Methods for Clinical Trials
Mark X. Norleans

Statistical Methods for Drug Safety
Robert D. Gibbons and Anup K. Amatya

Statistical Methods for Healthcare Performance Monitoring
Alex Bottle and Paul Aylin

Statistical Methods for Immunogenicity Assessment
Harry Yang, Jianchun Zhang, Binbing Yu, and Wei Zhao

Statistical Methods in Drug Combination Studies
Wei Zhao and Harry Yang

Statistical Testing Strategies in the Health Sciences
Albert Vexler, Alan D. Hutson, and Xiwei Chen

Statistics in Drug Research: Methodologies and Recent Developments
Shein-Chung Chow and Jun Shao

Statistics in the Pharmaceutical Industry, Third Edition
Ralph Buncher and Jia-Yeong Tsay

Survival Analysis in Medicine and Genetics
Jialiang Li and Shuangge Ma

Theory of Drug Development
Eric B. Holmgren

Translational Medicine: Strategies and Statistical Methods
Dennis Cosmatos and Shein-Chung Chow

I would like to show appreciation to my wife, Josephine, and daughters, Megan and Emma, for their support. I also dedicate the book to my mum, Mariammah, and my late father, Stephen Monickam, for their unfailing belief in the value of education.

Preface

The biostatistics profession is relatively new in the field of medicine. Most biostatisticians learn their tools of the trade through on-job training, and they either pick up the skills on the job through experience gained from a large number of collaborative projects or learn from an experienced mentor if available. Current resources on 'consulting' for biostatisticians is targeted at statistical consulting, which is different from a 'collaboration', and they include resources that provide examples from a diverse range of applications with few medical examples, as they are not specifically targeted at biostatisticians (Daniel, 1969; Demming, 1965; Hahn and Hoerl, 1998), or contain a collection of papers on statistical consulting in diverse areas such as finance, science, medicine and marketing with each study exemplifying a sample of issues, including crossover designs and ethical issues with clinical trials (Hand and Everitt, 1987) and case studies, including battery failure analysis, job promotion discrimination, lost mail analysis and food experiments (J. and Mc Dougall, 2010). A comprehensive resource covering all the aspects of collaboration with clinicians, especially writing for a publication or a grant proposal, oral and written communication and presentation skills, and managing projects and collaborators is currently lacking. Those that endeavour to cover communication aspects, for instance, unfortunately do not cover the technical and methodological aspects, which are equally important in collaboration (Derr, 2000) or focus on writing skills for reports, not publications (J. and Mc Dougall, 2010), the latter being more critical in a collaboration. This book will provide valuable insights for a practising biostatistician on 'hard skills' like methodological concepts revolving around study designs, form and database design, statistical analysis plan (SAP) and sample size calculations and non-methodological 'soft skills', such as how best to communicate with clinicians, as well as the best practices to adopt in terms of project, time and data management, writing skills and managing relationships with collaborators.

A unique aspect of this book is the dissemination of results from a survey among practising biostatisticians from a diverse background of education and working institutions, to gauge their responses on important issues such as problems ever faced in collaborating with clinicians, training on collaboration/consultation skills, frequency of performing selected tasks as a biostatistician, issues to address and important skills to gain in order to enhance greater collaborations and so on. The results of this survey are presented in Chapter 11, and they also provide context upon which many of the other chapters are developed, including communication skills and project management expertise. In the same chapter, we also present 'views

from the other side', that is, qualitative interviews with experienced clinician researchers who have worked with a number of biostatisticians in the past and have published extensively in a range of clinical fields, including infectious disease, rheumatology and psychiatry. Thoughts on how biostatisticians have added value to their research projects, challenges faced in collaborating with biostatisticians and specific knowledge or skill sets that they think biostatisticians should acquire to raise the standard of collaborative input are presented.

Chapter 8 addresses one major current gap in the training needs for academic biostatisticians, namely best practices in project management. Topics such as the creation of a project file to keep track of and manage important collaborative projects, guidelines on the naming and organisation of files and folders, database security and confidentiality, ensuring consistency and reproducibility in the statistical analysis, tips for working on multiple collaborations successfully as well as choosing an ideal mentor and working on a successful relationship with the mentor are discussed.

Some biostatisticians work in isolation in their institutions with few or even no colleagues to check with at their workplace, and this book intends to serve as a guide that they can use in their everyday work, and hopefully shorten their learning curve. The book will also be an ideal resource for postgraduate biostatistics and epidemiology students, to prepare them for a career as a collaborating biostatistician in an academic institution/hospital. The ultimate objective is to raise the quality and quantity of collaborative research that biostatisticians are involved in. It has been generally recognised that globally there is a shortage of academic biostatisticians, especially in hospitals, medical schools and research institutions, and this has been attributed to the increased recognition of the contribution statistics has made to all fields of health research, insufficient training opportunities in applied and medical statistics and inadequate supply of capable interest to make a career in this field (Pocock, 1995). In addition to the existing biostatisticians based around the various hospitals and research institutions around the world, the number of new biostatisticians entering the field is thus likely to increase, and this book will serve as a useful resource for them.

In terms of academic level, it is anticipated that this book will appeal to readers with at least a bachelor's degree in a quantitative field, including statistics. Those with a postgraduate degree in biostatistics and epidemiology will also benefit. This book can be used as a teaching module for postgraduate courses such as MSc and PhD in biostatistics, epidemiology and public health. This book is primarily targeted for teaching purposes, and intended for biostatisticians, epidemiologists and any quantitative scientists working in academic institutions/hospitals and who are involved in collaborative research with clinicians or other researchers.

This book should be useful for clinicians who are involved in some form of clinical and health services research and public health work, and who would like to collaborate with a biostatistician. It should serve its purpose

of informing how an effective working relationship can be fostered between the two groups. It is also possible that the book can be used as a didactic teaching tool for clinical research programmes designed for clinicians who are interested in a career in research.

This book is organised along the following framework: the first five chapters provide a summary of the essential methodological and analytical skills that a practising biostatistician needs to have in order to perform his/her roles and functions as a statistical collaborator effectively, particularly within an academic or hospital setting. This is naturally organised with the design of the study presented first (e.g. observational and randomised controlled trial, data collection form and database design, sample size and power calculations) followed by the chapter on SAP. The SAP chapter is exemplified with clinical examples using the Stata statistical software, with an aim towards introducing to readers some of the common statistical techniques encountered in the clinical research setting. This will provide guidance to the biostatistician on the appropriate statistical tests to employ depending on the aims and hypotheses. Although Stata has been used to exemplify the statistical technique and interpretation of the output with the clinical examples provided, the key learning points are applicable to other software as well. This chapter contrasts with current resources that aim to present the use of Stata to analyse selected projects and presents a brief description of the data, model and exercises provided at the end of each chapter (Rabe-Hesketh and Everitt, 2000) and another which is targeted at data analysts with guidelines including those on organising data and performing and reporting analysis using programmes written in Stata (Long, 2009).

Chapters 6 through 9 encompass the non-methodological 'soft skills' required in collaborations, namely effective communication, effective writing, how to successfully manage collaboration, good practices in project management and so on. Chapter 10 highlights the 'how not to', that is, the pitfalls in the design, analysis and presentation of data. This chapter endeavours to provide biostatisticians with knowledge on the critical snares in the various stages of the research collaboration cycle. Once again, the focus here is on health care-related examples and projects relevant for the practising biostatistician, compared to other resources aimed towards the discussion of more general statistical errors in hypothesis testing, modelling and interpretation of reports (Good and Hardin, 2001). The final chapter provides results from a survey among practising biostatisticians as well as interviews with experienced clinicians from a range of medical specialties, gauging their views on the collaboration process. The main objective is to present views from biostatisticians with a wide range of experience and expertise on the collaboration process. Another common thread underlying in all the chapters is the provision of 'questions to ask clinicians', which are useful checklists for biostatisticians to bring to the table when meeting their collaborators. These documents should prove useful when discussing with the collaborator about the study design, sample size calculation or when

deliberating on the SAP. In addition to the many clinical examples used in the book to exemplify statistical and methodological ideas and concepts, the chapters also end off with a real-life collaborative project, where possible. Valuable lessons are then drawn from the use of actual collaborative projects.

The key features of this book are as follows:

1. Provides an overview of common functions expected from a collaborating biostatistician
2. Identifies tools and resources that a biostatistician needs to work effectively
3. Uses selected published articles from the author's work with clinicians' to highlight key statistical techniques and learning points
4. Asserts best practices in terms of project management and use of gold standard templates in the design and reporting of studies
5. Presents technical statistical concepts in simpler terms that can be conveyed to clinicians
6. Employs actual interviews with biostatisticians and clinicians from a broad spectrum of medical specialities to convey key learning points regarding collaboration with biostatisticians

In this book, statisticians and biostatisticians have been used interchangeably, and they represent the same potential audience who may benefit from this book. In addition, the book does not intend to cover the wide range of statistical applications in the health care field, as there are already books available that cover specific topics in more detail (Altman, 1991; Armitage et al., 2002; Pocock, 2000). The statistical techniques described in the book are based on the more common clinical examples a biostatistician usually encounters in the hospital setting. Therefore, it is not possible to cover all the statistical techniques, e.g. population health methods, such as statistical process control charts, and survey data analysis, the design and analysis issues related to *in vitro* and *in vivo* experiments, the design and data analysis issues related to omics data and issues related to developing predictive models in this book.

This book was motivated by the author's personal journey as a practising collaborative biostatistician for more than 15 years. Apart from obtaining a degree and two postgraduate degrees from three different countries including the United Kingdom, Australia and Singapore, the author also has experience working in various academic and hospital settings in Australia and Singapore, from which real-life examples described in this book are drawn from. This includes selected examples from more than 125 publications in peer-reviewed medical journals, 2 book chapters, 13 successful grant applications, more than 130 conference proceedings, as well as more than 50 invited talks and presentations that he has been involved in over the past 15 years. The author has also taught in formal courses on

epidemiology, health statistics, multi-variate analysis, grantmanship and statistical software classes. In addition to mentoring statisticians, clinicians in residency programmes and clinician-researchers, he has also taught successive batches of medical students and master's in public health students both in Australia and in Singapore. The author has reviewed numerous grant proposals and protocols while serving as a regular member of local review panel for grants submitted to the National Medical Research Council in Singapore, as well as the Institutional Review Board of the National Healthcare Group.

The author was previously an academic editor of the editorial board of *PLoS ONE*, a non-profit, open-access, online publication that reports on original research from all disciplines within science and medicine, and regularly reviews articles for a number of journals including *Journal of American Medical Association* (JAMA); *BMC Health Services Research, Journal of Urban Health, Annals of the Academy of Medicine, Singapore, Australian Journal of Rural Health, International Journal of Infectious Diseases, Latin American Applied Research – An International Journal*, and *Medical Journal of Australia*. The reviews have provided useful information that has been presented in this book. He has also been invited to sit in the judging panel in numerous competitions, including the Singapore General Hospital Annual Scientific Meeting, Singapore Health and Biomedical Congress and Pitch for Fund, Medicine Academic Clinical Program at Singapore General Hospital, where he has had the opportunity to critically evaluate various clinical research projects.

From April 2011 to August 2014, the author was the director of the Centre for Quantitative Medicine at the Duke-NUS Graduate Medical School in Singapore. During this time, he was responsible for the organisation and allocation of faculty time to various activities, such as teaching of medical students, clinician research programmes such as the Master of Clinical Investigation, the Khoo Scholars and the Research Development Seminars. During the same period, he also worked with senior clinicians, hospital administrators and chief executive officers from various hospitals and research institutions, to further extend the reach of biostatistical collaborations of the Centre to other health care institutions such as the National Neuroscience Institute and the SingHealth Polyclinics, Singapore. He was also involved in the organisation and successful execution of two workshops designed to help equip biostatisticians with the skills necessary to collaborate successfully with clinicians. Input for this book was also obtained from fellow biostatisticians and clinician collaborators through a structured self-administered survey and qualitative interviews. A number of collaborators (both biostatisticians and clinicians) from the United States, Australia and Singapore have also looked at the draft chapters and provided critical input, and they have been recognised in the 'Acknowledgements' section.

References

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall/CRC Press.

Armitage, P., Berry, G., & Matthews, J. (2002). *Statistical Methods in Medical Research* (4th ed.). Oxford: Blackwell Science.

Cabrera, J., & Mc Dougall, A. (2010). *Statistical Consulting*. New York: Springer-Verlag.

Daniel, C. (1969). Some general remarks on consulting in statistics. *Technometrics, 11*, 241–245.

Demming, W. E. (1965). Principles of professional statistical practice. *Ann Math Stat, 36*, 1883–1900.

Derr, J. (2000). *A Guide to Effective Communication* (1st ed.). Pacific Grove, CA: Duxbury Press.

Hahn, G., & Hoerl, R. (1998). Key challenges for statisticians in business and industry. *Technometrics, 40*, 195–200.

Hand, D. J., & Everitt, B. S. (1987). *The Statistical Consultant in Action*. Cambridge: Cambridge University Press.

Long, J. S. (2009). *The Workflow of Data Analysis Using Stata*. College Station, TX: Stata Press.

Pocock, S. J. (1995). Life as an academic medical statistician and how to survive it. *Statistics in Medicine, 14*, 209–222.

Pocock, S. J. (2000). *Clinical Trials – A Practical Approach*. Chichester: John Wiley & Sons.

Rabe-Hesketh, S., & Everitt, B. S. (2000). *A Handbook of Statistical Analyses using Stata* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC Press.