
Preface

Biostatistical applications in molecular biology have increased tremendously in recent years. For example, a search of the Current Index to Statistics indicates that there were 62 articles published during 1995–1999 that had “marker” in the title of the article or as a keyword. In contrast, there were 29 such articles in 1990–1994, 17 in 1980–1989, and only 5 in 1970–1979. As the number of publications has increased, so has the sophistication of the statistical methods that have been applied in this area of research.

In *Biostatistical Methods*, we have attempted to provide a representative sample of applications of biostatistics to commonly occurring problems in molecular biology, broadly defined. It has been our intent to provide sufficient background information and detail that readers might carry out similar analyses themselves, given sufficient experience in both biostatistics and the basic sciences. Not every chapter could be written at an introductory level, since, by their nature, many statistical methods presented in this book are at a more advanced level and require knowledge and experience beyond an introductory course in statistics. Similarly, the proper application of many of these statistical methods to problems in molecular biology also requires that the statistical analyst have extensive knowledge about the particular area of scientific inquiry. Nevertheless, we feel that these chapters at least provide a good starting point, both for statisticians who want to begin work on problems in molecular biology, and for molecular biologists who want to increase their working knowledge of biostatistics as it relates to their field.

The chapters in this volume cover a wide variety of topics, both in terms of biostatistics and in terms of molecular biology. The first two chapters are very general in nature: In Chapter 1, Emmanuel Lazaridis and Gregory Bloom provide an historical overview of developments in molecular biology, computational biology, and statistical genetics, and describe how biostatistics has contributed to developments in these areas. In Chapter 2, Gregory Bloom and his colleagues describe a new paradigm linking image quantitation and data analysis that should provide valuable insight to anyone working in image-based biological experimentation.

The remaining chapters in *Biostatistical Methods* are arranged in approximately the order in which the corresponding topic or methods of analysis would

be utilized in developing a new marker for exposure to a risk factor or for a disease outcome. The development of such a marker would most likely begin with an examination of the genetic basis for one or more phenotypes. Chapters 3 and 4 deal with two of the most fundamental aspects of research in this area: microarray analysis, which deals with gene expression, and proteomics, which deals with the identification and quantitation of gene products, namely, proteins. Research in either or both of these areas could produce a biomarker candidate that would then be scrutinized for its clinical utility.

Chapters 5 and 6 deal with issues that arise very early in studies attempting to link the results of experimentation in molecular biology with exposure or disease in human populations. In Chapter 5, I discuss many of the issues associated with determining whether a new biomarker will be suitable for studying a particular E-D association. Jane Goldsmith, in Chapter 6, discusses the importance of designing studies with sufficient numbers of subjects in order to attain adequate levels of statistical power.

Chapters 7 and 8 are concerned with genetic effects as they relate to human populations. In Chapter 7, Peter Jones and his colleagues describe statistical models that have proven useful in studying the associations between disease and the inheritance of particular genetic variants. In Chapter 8, Stan Young and his colleagues describe sophisticated statistical methods that can be used to control the overall false-positive rate of the perhaps thousands of statistical tests that might be performed when attempting to link the presence or absence of particular alleles to the occurrence of disease.

Jim Dignam and his colleagues, in Chapter 9, describe the statistical issues that one should consider when evaluating the clinical utility of molecular characteristics of tumors, as they relate to cancer prognosis and treatment efficacy. Finally, in Chapter 10, Greg Rempala and I describe methods that might be used to validate statistical methods that have been developed for analyzing the E-D association in specific situations, such as when the exposure has been characterized poorly.

I would like to express my sincere appreciation to the reviewers of the various chapters in this volume: Rich Evans of Iowa State University, Mario Cleves of the University of Arkansas for Medical Sciences, Stephen George of the Duke University Medical Center, Ralph O'Brien of the Cleveland Clinic Foundation, and Martin Weinrich of the University of Louisville School of Medicine. I am also indebted to John Walker, Series Editor for *Methods in Molecular Biology*, and to Thomas Lanigan, President, Craig Adams, Developmental Editor, Diana Mezzina, Production Editor, and Mary Jo Casey, Manager, Composition Services, Humana Press.

Stephen W. Looney



<http://www.springer.com/978-0-89603-951-3>

Biostatistical Methods

Looney, S.W. (Ed.)

2002, XII, 216 p., Hardcover

ISBN: 978-0-89603-951-3

A product of Humana Press