

STATISTICS IN MEDICAL RESEARCH

Methods and Issues,
with Applications in Cancer Research

Edited by

K VALERIE MIKE

KENNETH E. STANLEY

STATISTICS IN MEDICAL RESEARCH

Methods and Issues,
with Applications in Cancer Research

Edited by

VALERIE MIKÉ

Memorial Sloan-Kettering Cancer Center
and Cornell University

KENNETH E. STANLEY

Harvard University School of Public Health
and Sidney Farber Cancer Institute



JOHN WILEY & SONS

New York Chichester Brisbane Toronto Singapore

Copyright © 1982 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

Main entry under title: Statistics in medical research.

(Wiley series in probability and mathematical statistics.

Applied probability and statistics. ISSN 0271-6356)

Includes index.

1. Medical research—Statistical methods—Congresses.
2. Cancer—Research—Statistical methods—Congresses.
3. Biometry—Congresses. I. Miké, Valerie, 1934- . II. Stanley, K.E. (Kenneth E.) III. Series. [DNL.M: 1. Biometry—Methods—Congresses. 2. Neoplasms—Congresses. 3. Research—Congresses. QZ 206 C748s 1981]

R853.S7S78 1982 610'.7'2 82-10871

ISBN 0-471-86911-2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

There is a growing need in medical research for the contributions of professionals trained in biostatistics and epidemiology, and there is a nationwide shortage of adequate manpower. This was also a conclusion of recent studies carried out by the National Research Council of the National Academy of Sciences. Biostatistics and epidemiology were specifically identified as areas in which educational programs should be supported.

The subject was discussed at length in the course of a workshop on epidemiology and biostatistics organized by the National Cancer Institute in 1979. Since there are not enough professionals in these fields to meet the needs of medical schools and cancer centers around the country, the recommendation was made that special seminars and teaching materials be developed to enhance the effectiveness of individuals now filling many of the positions. The statisticians attending this meeting included Theodore Colton, Valerie Miké, and Marvin Zelen, who was then Chairman of the Biometrics Section of the American Statistical Association. In response to the general recommendation of the workshop, Marvin Zelen subsequently asked Valerie Miké to chair a new committee of the Biometrics Sec-

tion that was being established to plan a summer conference on statistics in cancer research. Theodore Colton and Kenneth Stanley agreed to serve as committee members, and were later joined by David Braun, Richard Gelber, and Martin Lesser. A grant proposal for support of this activity was submitted to the NCI in 1980 and it was awarded in early 1981.

The conference was held at Memorial Sloan-Kettering Cancer Center during the week of June 22-26, 1981. Lectures and panel discussions covered a broad spectrum of topics, from study design and statistical methodology to medical ethics and aspects of teaching and consulting in the medical community. In addition to statisticians from major institutions around the country, the faculty included an attorney and practicing physicians representing the fields of medicine, surgery, pathology, psychiatry, and medical ethics.

This book is based on the lectures and discussions presented at the conference. Many participants expressed the view that the material covered would be of interest to a wider audience and should therefore be published as a volume.

The book offers a comprehensive overview of the field of biostatistics. Although the emphasis of the meeting was on cancer research, nearly everything discussed is applicable to other areas of medical investigation. Since there are currently many opportunities in biostatistics, this volume can provide useful information for classically trained statisticians interested in entering the field, and it can help those new to the field to become more effective collaborators. It can serve as stimulus for graduate students in statistics, to nurture their interest and to prepare them for careers in biostatistics. It can also be read with benefit by clinical investigators seeking a better understanding of statistical concepts and related multidisciplinary aspects of medical research.

The chapters in the book have been grouped according to content. The introductory paper on the role

of statistics in medical research is followed by a section on epidemiology. An overview is given of both descriptive and analytic epidemiology in the context of cancer research, and there is a thought-provoking article concerned with analyzing recent trends in cancer mortality and incidence.

The next section focuses on important issues pertaining to clinical studies. Chapter 4 presents a historical perspective and reports on the results of a survey in which senior investigators and statisticians were asked to assess the impact of statistics on advances made in the treatment of cancer. An interesting outcome was the high response rate and the nearly complete consensus of the respondents. Chapters 5 and 6 are edited transcripts of panel discussions exploring ethical, legal, and psychological issues as well as statistical problems related to the planning of clinical studies.

A section on practical considerations provides "how-to" information for the newcomer to clinical research. Chapter 7 defines basic terms and outlines procedures for the design of clinical trials; useful tables for sample size calculations are also included. Specific guidelines on data management and quality control are offered in Chapter 8. Chapter 9 provides the beginner with detailed instructions on acquiring efficient and cost-effective computing capability. It contains an annotated listing of major data base management systems and statistical software and graphics packages, complete with addresses and telephone numbers of vendors.

The papers dealing with statistical methodology are grouped in Chapters 10-15. These are meant to provide an introduction to the relevant material, with an overview of current developments. The chapters present basic concepts of the analysis of survival data, starting with a discussion of survival curves and progressing to group comparisons and model building. Two chapters are devoted to the analysis of

categorical data, using the approach of log-linear and logistic models. The section concludes with an assessment of methods for monitoring and stopping clinical trials. This is the only part of the book that contains material requiring more than an elementary background in statistics. But the definitions, examples, and extensive bibliographies given in each chapter offer worthwhile information for any serious reader.

The final section contains two panel discussions on communication, and these should be of equal interest to statisticians and biomedical investigators. Chapter 16 explores problems of interaction with the medical community in the statistician's triple role as consultant, collaborator, and teacher. Chapter 17 is concerned with the interpretation and presentation of statistical results.

The four multidisciplinary panel discussions help to convey the full flavor of medical research, including the complex and essentially open-ended nature of questions pertaining to research involving human subjects. Treatment of controversial matters is reported as it took place in the various sessions. References include guides to the vast literature of the new field of bioethics. The texts of the Nuremberg Code and the Declaration of Helsinki are reproduced in full at the end of Chapter 5.

For simplicity of style the pronoun "he" has been used throughout.

We would like to thank all those who contributed to the preparation of this volume. In particular, help with editing the panel discussions and reading the papers was provided by Jenny Baglivo, David Byar, Theodore Colton, Mitchell Gail, Edmund Gehan, Nancy Geller, Susan Groshen, Martin Lesser, and Howard Thaler. Blanche Sherman gave essential administrative support. The most extensive contribution was made by Larry Fishman, who first served as conference coordinator and subsequently as editorial assistant. He

carried much of the overall responsibility for communicating with the authors and assembling the material; he then also typed the camera-ready version of the manuscript.

Finally, we would like to express our deep appreciation to Dr. Marthana Hjortland, Program Director for Biometry, Special Programs Branch, Division of Cancer Cause and Prevention, National Cancer Institute, who played a significant role in providing the advice and encouragement needed to ensure the successful completion of this entire project. The work was supported by NCI Grant CA-29801.

Valerie Miké
Kenneth E. Stanley

New York, New York
Boston, Massachusetts
April 1982

Contributors

GEORGE J. ANNAS, J.D., M.P.H.
Department of Social Medical Sciences
and Community Medicine
Boston University School of Public Health
Boston, Massachusetts

JOHN C. BAILLAR III, M.D., Ph.D.
U.S. Environmental Protection Agency
Washington, D.C. and
Department of Biostatistics
Harvard University School of Public Health
Boston, Massachusetts

DAVID W. BRAUN, JR., Ph.D.
Biostatistics Laboratory
Memorial Sloan-Kettering Cancer Center
and Cornell University
New York, New York

BYRON W. BROWN, JR., Ph.D.
Division of Biostatistics
Stanford University School of Medicine
Stanford, California

DAVID P. BYAR, M.D.

Biometry Branch

National Cancer Institute

Bethesda, Maryland

ERIC J. CASSELL, M.D.

Department of Public Health

Cornell University Medical College

New York, New York

THEODORE COLTON, Sc.D.

Epidemiology and Biostatistics Section

Boston University School of Public Health

and Hubert H. Humphrey Cancer Center

Boston, Massachusetts

JEROME J. DECOSSE, M.D., Ph.D.

Department of Surgery

Memorial Sloan-Kettering Cancer Center

New York, New York

MITCHELL H. GAIL, M.D., Ph.D.

Biometry Branch

National Cancer Institute

Bethesda, Maryland

EDMUND A. GEHAN, Ph.D.

Department of Biomathematics

The University of Texas System Cancer Center

Houston, Texas

RICHARD D. GELBER, Ph.D.

Department of Biostatistics

Harvard University School of Public Health

and Sidney Farber Cancer Institute

Boston, Massachusetts

E. ROBERT GREENBERG, M.D.

Department of Community and Family Medicine

Dartmouth Medical School

Hanover, New Hampshire

LAWRENCE E. HINKLE, JR., M.D.
Department of Medicine
Cornell University Medical College
New York, New York

JAMES F. HOLLAND, M.D.
Department of Neoplastic Diseases
Mount Sinai School of Medicine
New York, New York

JIMMIE C. B. HOLLAND, M.D.
Psychiatry Service
Department of Neurology
Memorial Sloan-Kettering Cancer Center
New York, New York

STEPHEN W. LAGAKOS, Ph.D.
Department of Biostatistics
Harvard University School of Public Health
and Sidney Farber Cancer Institute
Boston, Massachusetts

MARTIN L. LESSER, Ph.D.
Biostatistics Laboratory
Memorial Sloan-Kettering Cancer Center
and Cornell University
New York, New York

ROBERT J. LEVINE, M.D.
Department of Internal Medicine
Yale University
The School of Medicine
New Haven, Connecticut

THOMAS A. LOUIS, Ph.D.
Department of Biostatistics
Harvard University School of Public Health
Boston, Massachusetts

xviii Contributors

VALERIE MIKÉ, Ph.D.
Biostatistics Laboratory
Memorial Sloan-Kettering Cancer Center
and Cornell University
New York, New York

FREDERICK MOSTELLER, Ph.D.
Department of Health Policy and Management
Harvard University School of Public Health
Boston, Massachusetts

JUDITH R. O'FALLON, Ph.D.
Cancer Center Statistics
Mayo Comprehensive Cancer Center
Rochester, Minnesota

CARL M. PINSKY, M.D.
Department of Medicine
Memorial Sloan-Kettering Cancer Center
New York, New York

PAUL P. ROSEN, M.D.
Department of Pathology
Memorial Sloan-Kettering Cancer Center
New York, New York

MARVIN A. SCHNEIDERMAN, Ph.D.
Clement Associates
Arlington, Virginia

DAVID A. SCHOENFELD, Ph.D.
Department of Biostatistics
Harvard University School of Public Health
and Sidney Farber Cancer Institute
Boston, Massachusetts

KENNETH E. STANLEY, Ph.D.
Department of Biostatistics
Harvard University School of Public Health
and Sidney Farber Cancer Institute
Boston, Massachusetts

Contents

PART I INTRODUCTION

1. The Role of Statistics in Medical Research

Frederick Mosteller

3

PART II EPIDEMIOLOGY

2. Cancer Epidemiology

Theodore Colton and E. Robert Greenberg

23

3. Trends in Cancer Mortality and Incidence in the United States: Is the Future Clear or Clouded?

Marvin A. Schneiderman

71

PART III ISSUES IN CLINICAL STUDIES

4. Clinical Studies in Cancer: A Historical Perspective

Valerie Miké

111

5. Clinical Trials: Exploring Ethical, Legal, and Psychological Issues -- Panel Discussion	
<i>Valerie Miké, Chairman</i>	156
6. Issues in the Design of Clinical Trials -- Panel Discussion	
<i>John C. Bailar III, Chairman</i>	189
PART IV PRACTICAL CONSIDERATIONS	
7. Design and Implementation of Clinical Trials	
<i>Martin L. Lesser</i>	225
8. Data Management and Quality Control	
<i>Judith R. O'Fallon</i>	254
9. Statistical Software: Data Base Management, Statistical Packages, and Graphics	
<i>David W. Braun, Jr.</i>	280
PART V STATISTICAL METHODOLOGY	
10. Estimation in Survival Analysis: Parametric Models, Product-Limit and Life-Table Methods	
<i>Byron W. Brown, Jr.</i>	317
11. Inference in Survival Analysis: Nonparametric Tests to Compare Survival Distributions	
<i>Stephen W. Lagakos</i>	340

12. Analysis of Survival Data: Cox and Weibull Models with Covariates <i>David P. Byar</i>	365
13. Analysis of Categorical Data: Exact Tests and Log-Linear Models <i>Thomas A. Louis</i>	402
14. Analysis of Categorical Data: Logistic Models <i>David A. Schoenfeld</i>	432
15. Monitoring and Stopping Clinical Trials <i>Mitchell H. Gail</i>	455
PART VI COMMUNICATION	
16. Interacting with the Medical Community: Consulting, Collaboration, Teaching -- Panel Discussion <i>Theodore Colton, Chairman</i>	487
17. Interpretation and Presentation of Statistical Results -- Panel Discussion <i>Kenneth E. Stanley, Chairman</i>	513
INDEX	543

Part I

Introduction

