

CLINICAL BIOSTATISTICS

Alvan R. Feinstein

CLINICAL BIOSTATISTICS

Alvan R. Feinstein, M.D.

Professor of Medicine and Epidemiology,
Yale University School of Medicine,
New Haven, Conn.

The C. V. Mosby Company

SAINT LOUIS 1977

Copyright © 1977 by The C. V. Mosby Company

All rights reserved. No part of this book may be reproduced in any manner without written permission of the publisher.

Printed in the United States of America

Distributed in Great Britain by Henry Kimpton, London

The C. V. Mosby Company

11830 Westline Industrial Drive, St. Louis, Missouri 63141

Library of Congress Cataloging in Publication Data

Feinstein, Alvan R

Clinical biostatistics.

Originally appeared as essays in the journal *Clinical pharmacology and therapeutics*.

Includes bibliographies.

1. Medical statistics—Addresses, essays, lectures.
2. Medical research—Statistical methods—Addresses, essays, lectures. I. Title. [DNLM: 1. Biometry. 2. Epidemiologic methods. HA29 F299c]
RA409.F36 610.1'5195 77-3703
ISBN 0-8016-1563-1

*For my mother, Bella, and my wife, Linda,
who have given me
roots, wings, and love*

PREFACE

When I began writing the series of essays called “Clinical biostatistics” in 1970, I thought I would run out of material in about a year. I knew I had some unorthodox things to say and some unconventional viewpoints to develop, but I believed the development would require only six or seven essays. As I became more deeply immersed in biostatistical ideas, however, I constantly found more to do. At every level of contemplation—ranging from the massive logistics of large-scale clinical trials, to the elaborate complexities of “retrospective case-control” studies, to such apparently simple questions as how (and why) to calculate a standard deviation—the world of biostatistics seemed beset with scientific problems that had received unsatisfactory solutions because the *-statistics* had been given more attention than the *bio-*. As I began grappling with those challenges, the essays continued to proliferate, until almost 40 of them have now appeared.

While the essays were making their bimonthly and later trimonthly appearances in *Clinical Pharmacology and Therapeutics*, readers of the journal were highly complimentary; and many urged me to publish the series as a book. At first I resisted this suggestion, mainly because I have never liked this type of autoanthology. The author of a book, I thought, should prepare a suitably renovated text, not just an unaltered collection of previously published papers. My resistance to the suggestion eventually collapsed, however, under two sets of pressures: time and audience. As I kept wanting to prepare that new text while never finding the necessary time to do so, I realized that the only hope for achieving a book in the imminent future was to preserve the original essays. Furthermore, many readers kept assuring me that the original essays should remain intact—that their “spirit” might be lost in a revision and that the book would be more enjoyable to read if it preserved the informality of the original prose. Adding to these incentives for an “anthology” format were the fiscal concerns of The C. V. Mosby Company, publishers of the journal where the “Clinical biostatistics” essays have appeared. The cost of publishing the book could be substantially reduced if the essays were maintained in their original form, with each text and bibliography unchanged.

Accordingly, this book contains a collection of original essays from the “Clinical biostatistics” series. They have been rearranged as chapters, into a logical pattern that differs from the chronologic sequence in which they first

appeared. A few of the original titles and many of the identifying numbers of the essays have been changed to conform with the current sequence of chapters. Otherwise, the texts and lists of references for each essay remain intact. To keep the book from being too large, I have omitted a few of the essays that contained quantitative surveys of the medical and statistical literature, digressions into the ethics of research and the teaching of statistics, or specific critiques of individual clinical investigations. The remaining 29 essays have been divided into an introductory chapter that is followed by five major sections, each preceded by a brief commentary.

The essays are intended for people who have already developed an interest in biostatistical issues. The interest may have arisen spontaneously; it may have been necessitated by the demands or comments of a manuscript reviewer; or it may have been provoked by the efforts needed to understand the many mathematical machinations that are used in published reports of current research. I assume that the reader is aware of rudimentary statistical tactics, but is otherwise not particularly adept in mathematics and is possibly frightened by it. With that assumption, the goal is to enlighten and perhaps to entertain with the style of an essay, not to educate with the formality of a textbook.

Conventional textbooks and courses in biostatistics are usually devoted to the theoretical processes that produce such mathematical calculations as P values, confidence intervals, correlation coefficients, and regression equations. Amid the mathematical emphasis, almost no attention has been given to the basic scientific procedures used for planning research, obtaining data, and analyzing results. The aim of these essays is to provide supplemental reading for the many important topics that are omitted from conventional textbooks, and also some remedial reading for topics that usually receive inadequate consideration.

Because of the way the book has been assembled, it contains three features for which I apologize. The first is that the text regularly contains references to previous or forthcoming essays in the originally published series. Although useful liaisons for essays that were dispersed in time, many of the references will now appear in the wrong places in the rearranged series. Since the current text is identical to what originally appeared in the journal publications, these references could not be changed. The second flaw is that the original bibliographic citations have also, of necessity, been preserved at the end of each essay. This process, while making the citations easy to find, produces frequent redundancy in some of the listings. The third infelicitous feature is that certain ideas are mentioned repeatedly in different locations of the text. The repetition seemed desirable in a succession of individual essays spread over a 6-year period, but may be less appealing if the essays are read contiguously. I hope that readers will find these sporadic repetitions instructive rather than irritating.

In discussing the various challenges and imperfections of biostatistics, I have tried to keep the prose lively and have occasionally made it deliberately provocative. Most readers have said they enjoy this approach, but it has sometimes led to the accusation that I am antistatistical. This accusation has probably been received by anyone who has ever been discontent with the defects of any status quo. Like the established tenets of clinical medicine and epidemiology, the

established creeds of statistics contain many infirmities. In pointing out the infirmities, I have always tried to offer constructive suggestions for improvement; and I would hardly want to spend so much effort working in the domain of clinical biostatistics if I did not respect both the *clinical bio-* and the *-statistics* portions.

To do the kind of thinking and writing that have produced these essays, I have had many sources of support for which I want to express thanks. The Veterans Administration, at its hospital in West Haven, provided research aid for many years while I was Chief of the Eastern Research Support Center and, later, of the Cooperative Studies Program Support Center. For my activities at the Yale University School of Medicine, the National Center for Health Services Research and Development supplied grants for several projects from which many of these essays emerged as by-products. During a highly productive period from 1971-1973, as a visiting professor, I received professional hospitality and illuminating stimulation from the Department of Clinical Epidemiology and Biostatistics at McMaster University Medical Center in Hamilton, Ontario, Canada. For the past few years, the essays have been composed in my work as Director of the Yale Clinical Scholar Program, which is sponsored by the Robert Wood Johnson Foundation.

In addition to this institutional aid, I have been greatly helped by human talents and contributions. Before submitting the prepared essays for publication, I have relied on thoughtful appraisal and stringent evaluation from critics who are clinicians, epidemiologists, statisticians, or computer experts. In acknowledging my gratitude for their valuable help, I also herewith absolve them of any responsibility for the contents. They are Linda Marean Feinstein, Michael Gent, Charles A. Goldsmith, Moreson H. Kaplan, Donald Mainland, Walter A. Ramshaw, David L. Sackett, Helen L. Smits, Walter O. Spitzer, and Carolyn K. Wells. I am also especially grateful to Dr. Walter Modell, Editor of *Clinical Pharmacology and Therapeutics*, for his constant encouragement and for the editorial freedom he has provided. For excellent performance in the tasks of typing the difficult combinations of prose and mathematical symbols, I thank Elizabeth Tartagni, Carrol Ludington, and Pamela Rowe.

Finally, I want to thank my wife, Linda, and our children, Miriam and Daniel, who have gently tolerated the many hours in which I was absent or secluded while working on these essays, and who have filled the nonwriting hours with warmth, affection, and joy.

Alvan R. Feinstein
New Haven, 1977

CLINICAL BIOSTATISTICS

CONTENTS

- 1 Introduction and rationale, *1*

SECTION ONE

THE ARCHITECTURE OF COHORT RESEARCH

- 2 Statistics versus science in the design of experiments, *17*
- 3 Components of the research objective, *28*
- 4 Intake, maintenance, and identification, *38*
- 5 Subsequent implementation of the objective, *54*
- 6 Sources of 'transition bias,' *71*
- 7 Sources of 'chronology bias,' *89*
- 8 Credulous idolatry and randomized allocation, *105*
- 9 Consequences of 'compliance bias,' *122*

SECTION TWO

OTHER ARCHITECTURAL PROBLEMS

- 10 Statistical malpractice—and the responsibility of a consultant, *137*
- 11 Random sampling and medical reality, *154*
- 12 The rancid sample, the tilted target, and the medical poll-bearer, *169*
- 13 Ambiguity and abuse in the twelve different concepts of 'control,' *186*
- 14 The epidemiologic trohoc, the ablative risk ratio, and 'retrospective' research, *197*
- 15 On the sensitivity, specificity, and discrimination of diagnostic tests, *214*

SECTION THREE

PROBLEMS IN MEASUREMENT

- 16 On exorcizing the ghost of Gauss and the curse of Kelvin, *229*
- 17 The derangements of the 'range of normal,' *243*

- 18 How do we measure 'safety' and 'efficacy'? 256
- 19 The difficulties of pharmaceutical surveillance, 271

SECTION FOUR

MATHEMATICAL MYSTIQUES AND STATISTICAL STRATEGIES

- 20 Permutation tests and 'statistical significance,' 287
- 21 The direction of relationships, hypotheses, and probabilities, 305
- 22 Sample size and the other side of 'statistical significance,' 320
- 23 Problems in the summary and display of statistical data, 335

SECTION FIVE

THE ANALYSIS OF MULTIPLE VARIABLES

- 24 On homogeneity, taxonomy, and nosography, 353
- 25 A primer of multivariate analysis, 369
- 26 The purposes of prognostic stratification, 385
- 27 The process of prognostic stratification, 398
- 28 Evaluation of a prognostic stratification, 414
- 29 Additional tactics in prognostic stratification, 430

INDEXES

Index of Authors, 447

Index of Methodologic Topics, 453

Index of Clinical and Other Practical Examples, 465

CHAPTER 1

Introduction and rationale

The "Clinical biostatistics" series began when I was invited to succeed Dr. Donald Mainland, who had retired from writing a bimonthly "column" on statistics for Clinical Pharmacology and Therapeutics. The first essay in the series contains my tribute to Dr. Mainland's many previous contributions to biostatistics and also describes the background philosophy with which the new series would be approached. The text was as follows.

—ARF

Donald Mainland can be succeeded but not replaced. His training, timing, and temperament have made him a unique figure in the domain of medical statistics, and a tough act to follow.

In training, he was graduated in medicine with honors in 1925 at Edinburgh, where he was later awarded the Doctor of Science degree for his research in embryology and histology. After finishing medical school, he taught anatomy at Edinburgh for several years and then went to Canada. He worked at Manitoba from 1927 to 1930, when he left to become Professor and Chairman of the Department of Anatomy at Dalhousie University. His first publication in 1927—dealing with an uncommon abnormality in a muscle¹³—was a harbinger of his subsequent concern with frequency distributions in biology. Within the next two years, he was evaluating the accuracy of techniques for estimating irregular anatomic areas.¹⁴ Later on, during various embryologic in-

vestigations, he contemplated methods for assessing the size and volume of cellular structures.¹⁵⁻¹⁹ Over the next few years, he applied his quantitative interests to measuring the forces of muscles²⁰⁻²¹; and then, in 1934, with a paper on "Chance and the blood count,"²² the quantitative anatomist started his metamorphosis into medical statistician. By 1936, after some additional research on blood cells and blood counts, he had begun to write on "Problems of chance in clinical work."²³ In 1938, he produced his first book on quantitative medicine,²⁴ and twelve years later, after continuous productivity in both biologic research and medical statistics, he accepted New York University's invitation to become Professor of Medical Statistics. From that position, with persistent intellectual growth and enormous practical experience, he has continued to provide enlightenment to his colleagues, students, consultees, and readers.

In timing, Dr. Mainland became interested in biologic statistics during an era when the analytic techniques were in primitive stages of conception and dissemination. He knew many of the early

This chapter originally appeared as "Clinical biostatistics—I. A new name and some other changes of the guard." In Clin. Pharmacol. Ther. 11:135, 1970.

heroes in the contemporary statistical pantheon, and he became a pioneer physician in developing the modern relationship between statistics and medicine. After the first edition of his classic book, *Elementary Medical Statistics*²⁵ in 1952, he continued to produce a powerful array of creative, didactic, expository, and polemic publications on the use of statistics in medicine. With his textbook, now in its second edition,²⁶ and his many other writings, he has probably contributed as much as any single person to the statistical sensibility of clinical investigators in North America today.

In temperament, he has managed to preserve the extraordinary virtue of common sense, despite his constant exposure to the abstract concepts, arcane models, and intellectual folderol that lurk in the statistician's world. Part of this virtue is attributable to Mainland's firm rooting in the realities of medical biology. He has not merely preached about biostatistical research; he has practiced it. During the development of his statistical interests, he maintained his activities in biologic investigation—contributing, among other items, a textbook on anatomy²⁷—and he currently continues an active role in several large-scale clinical research projects, chiefly in rheumatoid arthritis.

But the greater part of Mainland's virtue is probably attributable to the man himself. Now near the age of retirement, he remains young in mind, in spirit, and in outlook. What other "older man," venerated and respected as he nears completion of his major work, is ready to recognize that "repetition of this theme during two or three decades, by others as well as myself, has had very little effect"³¹; to confess that he is "technically unsophisticated"³²; to solicit disagreement and rebuttals to all of his comments; and to be constantly receptive to new approaches for old problems. How many established "authorities" are brave enough to appraise their previous work with comments like these: "Grading all four items together

today, I would award a C, or perhaps a C+, but nothing higher,"²⁸ or "I sometimes wonder how many more instances of stupidity I might dig up from the days when I was hypnotized by statistical techniques applied to pooled data."³⁰

My own first encounter with Mainland came in about 1960, when I discovered his publications entitled "Notes from a Laboratory of Medical Statistics"—a group of documents still cherished by the recipients who were lucky enough to learn about the "Notes," and to satisfy Mainland's hardy standards for the mailing list. ("There is a limit of 3,000 to the number of 'Notes' that can be issued. . . We are sorry that we can no longer replace 'Notes' that have disappeared after they have been received . . . Agencies that require formal invoicing are also too much trouble to deal with.") These "Notes," which Mainland issued periodically whenever he found time to do so, were the ancestors of the more recent "Statistical ward rounds" in this JOURNAL, and the "Notes on Biometry in Medical Research," which have appeared under the sponsorship of the Veterans Administration.

I still remember the enchantment of discovering those early "Notes." For several years previously, my own clinical research had brought me increasingly in contact with statistical procedures, and my manuscripts were being frequently sent to statisticians for review. Since many of the reviewers' comments were either clinically absurd or statistically incomprehensible, I had begun, in self-defense, to read textbooks on statistics. Like Mainland's, my education in statistics is largely self-acquired; but unlike most physicians, I was not intimidated by the arithmetic, since I had done graduate work in pure mathematics before entering medical school. What I found in the textbooks was sometimes enlightening, but more often appalling.

From my previous activities in pure mathematics and in biologic science, I had become accustomed to a rigorous type

of either logical or empirical documentation for any assertion. In pure mathematics, such an assertion was called a *theorem*, and the rigorous documentation was a sequence of logically cohesive statements called a *proof*. In biologic science, the assertion was called a *hypothesis*, and the rigorous documentation was a collection of empirical data called *observed evidence*. But most of the statistical textbooks seemed to contain neither a logical nor an empirical documentation for the assertions. The texts were often like cookbooks, containing a series of instructive recipes on how to tabulate data and perform certain "tests of significance." These instructions were seldom accompanied by a proof of their validity, by any references to where a proof might be found, or by any empirical data to demonstrate that the procedures would remain valid when their prerequisite conditions were violated. From time to time, I would explore the literature of mathematical statistics, looking for either the rational logic or the scientific evidence to support what appeared in the "cookbooks," but I was seldom successful. Even with a mathematical background, I could not understand many of the esoteric formulations; and my biologic background made me wary of the unrealistic assumptions that underlay many of the mathematical arguments.

I did not know at the time that some of these mathematical defects were so commonplace as to arouse public lament by distinguished statisticians. Said Harold Hotelling¹¹ in 1960:

The custom of omitting proofs, which would not be tolerated in pure mathematics beyond a very limited extent, is common in the teaching of statistics, and is excused on the grounds that the students do not know enough mathematics to understand the proofs. Perhaps in some cases a better reason is that the teachers, and the authors of the textbooks, do not understand the proofs. In some instances no proofs exist, and in some instances no genuine proofs can exist, because the methods taught are demonstrably wrong.

Aware of some of the many intellectual problems that pervade work in clinical medicine, I had expected to find that the cerebral grass would be greener in the statistician's yard. To my dismay, I found many weeds being cultivated and labeled as flowers. Apart from my dissatisfactions with the absence of proofs for didactic assertions, I was disturbed by the lack of real attention to the consequences of the biologic component of "biostatistics." Here were men of high professional and intellectual competence. How could they so blithely ignore the effects of their erroneous assumptions that most clinical data came from "random samples" with "normal distributions" and "continuous variables"? How could they discuss the design of clinical experiments by extrapolating from a brewery vat or an agricultural field to a human population? How could they give so much emphasis to procedures for purely statistical analysis, while showing so little rigorous concern for such basic issues in scientific logic as specifying the fundamental question, determining whether the research would answer that question, choosing an appropriate control group, checking the reliability of the data, establishing reproducible criteria for subjective evaluations, and ascertaining whether the investigated population was both homogeneous enough for everyone to be "lumped" together and selected in a manner that justified the idea of "randomness"?

Wandering among statistical doctrines that often seemed neither mathematically validated, biologically cogent, nor intellectually challenged, I came upon Mainland's "Notes." The man seemed to know that biostatistics ought to pertain to biology, and he seemed to know about biology. He sounded like someone who had learned about research not by agglomerating theories of probability, or massaging data whose origins he had never observed, but by actually feeling a tissue, handling an animal, calibrating an instrument, looking through a microscope, or talking to a

patient. An effective stimulant for the intellectual torpor of the textbooks, Mainland's "Notes" made biostatistics vivid, vital, and exciting. He brought into open view many of the critical issues that lay hidden beneath glib traditional preconceptions; he helped demonstrate that many statistical models were inappropriate and misleading for biology; and he provided a medium in which biologic scientists—insecure and anxious in their heretical suspicions about conventional statistical dogmas—could take comfort from seeing that other scientists and statisticians shared the same heresies. Here, at last, was a statistician who could talk sensibly about clinical research. (He could also sometimes talk too long, but verbosity is an accepted occupational hazard of biostatisticians. Mainland was readily forgiven for occasional ventures into prolix prose, and his successor in these columns hopes that future readers will be equally tolerant.)

Mainland was the first medical statistician I had encountered who acted as though "bio" were an integral part of "biostatistics," instead of a prefix attached casually to "statistics" for the sake of an occasional teaching exercise or a book intended for graduate students in biologic domains. Since that time, I have met a few other statisticians who have truly become biostatisticians, but Mainland remains a pioneer both in migrating in the unusual profession direction from biology to statistics, and in exemplifying the modern fusion of biology with statistics. Clinical investigators today owe him an inestimable debt of gratitude for the contributions he has made to our domain by preserving thoughtful realism in our statistical outlook. Many people believe that his book, *Elementary Medical Statistics*, could benefit from tighter organization and greater succinctness, but it is still the only such publication that gives at least as much attention to the medical issues of medical statistics as to the statistical maneuvers. As clinical investigators be-

come increasingly involved in biostatistics, and as we begin to appreciate its scope, accept its challenges, educate our statistical co-workers in its problems, and contribute creative solutions to those problems, Donald Mainland will remain one of our honored "founding fathers." He is a physician who helped establish the basic concept on which we must now build—the concept that biostatistics can best be developed neither from abstract theory in statistics nor from imprecise anecdotal evidence in biology, but from a coordinated integration of perceptive observation and thinking in both.

In succeeding Dr. Mainland as master of ceremonies for these columns, I hope to preserve his basic outlook and philosophic standards, although I shall undoubtedly introduce some deviations of my own, because our specialized interests and training have been so different. His prestatistical domain was anatomy; mine has been clinical medicine. His basic work for almost two decades has been centered in a department of medical statistics; mine has been (and remains) centered in a department of internal medicine. He became intimately familiar with many aspects of the basic mathematical precepts of statistical tactics; my acquaintance with some of these precepts is tenuous, and I shall regularly ask my statistical colleagues for help when the discussions get into issues with which I am relatively unfamiliar. Since Dr. Mainland's packs of cards and barrels of discs have not been transmitted as a legacy of this job, I shall probably use a computer for many of the exercises in random number selection that he would have consigned to his trusty manual companions. I shall probably also call upon the computer for certain new activities that it now makes possible in modern biostatistics.

One of the main challenges will be to keep the column as least as interesting, informative, and provocative as Mainland made it. Connoisseurs of the Mainland style will recall that he often goads his

readers deliberately, hoping to elicit further discussion. (Example: "I hope that . . . I have trodden on some official corns hard enough to initiate a foot-to-brain-to-hand reflex that will produce a defense of their . . . methodology."²⁹) My own tactics in provocation may be somewhat different and occasionally inadvertent, but I hope to preserve the principle that all of us need vigilant prodding to avoid or destroy complacency. The pace of science and technology has become too rapid for anyone to maintain prolonged intellectual comfort about any established axioms, concepts, or other beliefs that have not been regularly subjected to intensive scrutiny and skeptical reappraisal.

To help augment the role of this column as a medium of vigorous communicative exchange and intellectual growth, I plan to invite various guests, either as proponents of their views or in rebuttal of mine, to become the "columnist" from time to time. The columns will be titled in a numbered sequence for my own papers, but a different designation will be used to accommodate other authors. I also hope that readers will frequently write to express their agreement or dissent about anything that appears here, and I would plan to have the letters (with the author's name omitted, if so desired) become a source of lively discourse in future columns. The medical statistical problems with which we struggle are too numerous and too important to be resolved without an abundance of argument. I hope that the arguments from all the people who contribute to these proceedings will be responsible, thoughtful, clearly written, and prepared in an atmosphere of light rather than heat—but arguments nonetheless.

Readers are invited not only to express opinions about what has appeared, but also to make suggestions about topics for future discussion. My ideas about choice of topics will come from several sources: (1) personal adventures during my own research activities; (2) review and occasional revival of ideas expressed pre-

viously by Dr. Mainland and by other people who have written about clinical biostatistics; (3) new stimulation from projects encountered in work at the Veterans Administration Research Support Centers, which are currently asked to help prevent or remedy biostatistical maladies in hundreds of biomedical research tasks each year; and (4) comments from readers. Those first three sources of input have already provided the topics planned for discussion in the next few columns, but the slots are open thereafter, and suggestions will be happily received.

One immediately obvious change is in the title of this column. Many leaders are "done in" by their successors, and Dr. Mainland should not have to worry about being blamed for my mistakes, misconceptions, or mischief. To give him that freedom of responsibility, and also to allow him to use "Statistical ward rounds" as a title for a possible book, the name of these essays has been changed to "Clinical biostatistics." I know that many readers will prefer the older title; the new one seems more formal, somber, and sesquipedalian, but it was the least of the available evils in nomenclature, and I hope that these new "rounds" will retain the appealing informality and free-wheeling intellectual fun of their ancestors.

. . .

There are more profound reasons, however, for a change that brings *clinical* into juxtaposition with *biostatistics*. In quantitative nomenclature, the *statistics* part of *biostatistics* occupies ten letters and the *bio* only three; the addition of eight *clinical* letters to the total phrase may help restore nominal as well as conceptual balance. More importantly, however, the domain of biostatistics is currently beset with many intellectual maladies that I believe can be remedied only if clinical biologists begin to make active contributions to the domain. These maladies, which arise not in the contents of statistical thinking but in the way statistical concepts are applied to

other disciplines, have recently become subjects of public comment by leading statisticians:

John W. Tukey³⁸—

A teacher of biochemistry does not find it intolerable to say, "I don't know." Nor does a physicist. . . . Why should not . . . statisticians do the same? . . . Far better an approximate answer to the *right* question, which is often vague, than an exact answer to the *wrong* question, which can always be made more precise.

J. G. Skellam³⁷—

. . . Valuable information which affects common-sense judgments tends to be ignored when formal statistical tools are employed along conventional lines. . . . Surprisingly little attention is normally given to what is often a much more serious source of error and deception, the defects of the model itself. . . . There is an important difference of emphasis between the application of mathematics to biology, and the mathematization of biology, and it is the latter which needs the most encouragement, for it is here that the real difficulties lie. . . . I am somewhat disturbed by the thought that the exalted status of mathematics . . . might possibly . . . exercise . . . (an) unintentional brand of tyranny over other ways of thinking.

William Feller⁵—

I do not criticize statistical theory as such or the proper uses of statistics. . . . The trouble is that these methods are often used thoughtlessly and routinely by researchers for purposes for which they were not intended. . . . In biologic experimental work, for instance, a . . . common abuse is to use a statistical test to try to "prove" a hypothesis. . . . The scandal is that the "significant" results are published as though they had meaning. . . . All too frequently statisticians impose all kinds of nonsensical conditions on the poor biologist or psychologist—conditions which, although they produce unequivocal statistical results, actually hinder him in his research.

Complaints about the status quo have also come from computer experts trying to implement some of the existing mathematical approaches and models. Said R. W. Hamming¹⁰:

I have been repeatedly shocked to find out how often I thought I knew what I was talking about; but that in the acid test of describing

explicitly to a machine what was going on I was revealed to have been both ignorant and extremely superficial. It is this many-times-repeated experience that has led me to assert that mathematics has often chosen to ignore the careful examination and exposition of the methods it uses.

Other discontents have recently been expressed by statisticians about the methods used to prepare consultants for the role they have been playing for more than 40 years, ever since R. A. Fisher's⁶ epochal book made biologists begin the frequent search for statistical advice. Among the comments have been the following:

J. G. Skellam³⁷—

I attribute this (undesirable) attitude largely to the way that statistics is usually taught—as a mathematical discipline of great intrinsic interest imparted to talented students who unfortunately have rarely had proper training in natural science or first hand experience of scientific research.

M. Zelen³⁹—

The statistical design of experiments . . . as taught in most schools, seems so far removed from reality, that a heavy dose may be too toxic with regard to future applications. . . . It is absolutely vital that the future biometrician spend part of his training as a biometrician-in-residence at an institution or laboratory where active scientific work is being conducted.

C. P. Cox¹—

If the discipline of statistics is to retain its identity . . . the inter-connections of statistics and research in "user" disciplines must be continually developed. . . . Besides training in statistics, an aspirant statistical consultant should receive complementary and systematized, as distinct from casually acquired, training in the disciplines in which he is expected to consult. . . . All the scientists concerned may be advantageously encouraged to scrutinize and clarify their ideas on scientific method and to challenge purely statistical inferences whenever these are unconvincing.

The direction of communication

What is surprising about all these critical comments is that they have come from statisticians rather than from clinicians or

other recipients of the statistical consultations. In an era in which patients have been increasingly vocal in complaints about the services received from their clinical consultants in problems of medicine, clinicians have been notably silent in commenting on the "quality of care" in the interchange that occurs when the clinician* is a "patient" coming to a "doctor" who is a statistical consultant in problems of research.

Clinicians have had many of these consultative encounters. Courses on statistics are now offered in the curricula of most medical schools. Editors of many medical, psychological, and other journals will routinely request that suitable manuscripts be passed by a statistical censor. Proposals for large-scale clinical trials are not only often designed by statisticians, but also must be approved by statisticians before the project is funded.

During these many interplays of statistics and medicine, however, the path of consultative enlightenment has remained unidirectional. Thus, although the manuscripts and contents of medical journals are regularly subjected to critical statistical appraisal, almost no evidence of exposure to clinical reviewers can be found in the many medically oriented papers that regularly appear in such periodicals as *Biometrics*, *Biometrika*, and the *Journal of the American Statistical Association*. Many critiques have been published on the unsuitable or incorrect statistical

methods used for papers that have appeared in the medical literature,^{2-4, 8, 33-35} but I am not aware of any comparable critiques of the inappropriate or sometimes bizarre medical assumptions contained in papers that appear in the statistical literature. Innumerable books have been written on the general topic of elementary statistics for clinicians, but no one has written a clinical primer for statisticians.

It is time that clinicians began to widen this narrow path of communication and to inform our colleagues of the statistically pertinent knowledge that was learned during those many years in hospital wards and clinical practice. Although the statistician has spent many years in graduate school getting his Ph.D. degree, and can tell us a great deal about what he discovered during that time and afterward, the clinician has spent many years getting not only his M.D. degree, but also such additional postdoctoral "degrees" as F.A.C.P. or F.A.C.S. If clinicians need to know about the mystic statistic, statisticians might benefit from discovering the clinical pinnacle. We all have much to teach each other.

The composition of "statistics"

One of the main first steps in this process of mutual education is to recognize that "statistics" is a composite domain, containing at least two distinctly different intellectual activities: (1) the acquisition, logical organization, and numerical presentation of data, and (2) the analysis of the data to arrive at decisions about degrees of variation, interrelation, and difference. The first type of activity is often called *descriptive statistics*; it produces the collections of data that appear in baseball batting averages, in financial charts, in the birth rates and death rates of "vital statistics," and in the many graphs, tables, and other numerical expressions of biomedical projects ranging from molecular explorations to therapeutic surveys. The second type of activity,

*To avoid ambiguity, let me define a clinician as a member of one of the healing professions—such as medicine, osteopathy, and clinical psychology—who takes direct responsibility for the care of living patients, or who has spent substantial amounts of postgraduate time (more than an internship) in developing his skillful knowledge of such activities. The clinician may be in private practice, academic research, or administrative work, but his distinguishing characteristic is a background of observational and therapeutic experience in dealing with sick people. Although an M.D. degree is sometimes regarded as the hallmark of a clinician, many M.D.'s—such as anatomists, biochemists, "clinical" pathologists, epidemiologists, microbiologists, pathologists, pharmacologists, and physiologists—may have neither the training nor the functional responsibilities of clinicians. This definition is intended only to clarify what I am talking about, and has no pejorative connotations in any direction.