PROCEEDINGS OF THE

9th International Biometric Conference

Volume II

BOSTON
AUGUST 22-27, 1976

THE BIOMETRIC SOCIETY

PROCEEDINGS OF THE

9th International Biometric Conference

Invited Papers

Volume II

BOSTON
AUGUST 22-27, 1976

THE BIOMETRIC SOCIETY

ACKNOWLEDGMENTS

The Biometric Society expresses sincere appreciation to the following organizations and individuals for their support toward the costs of the meetings and the publication of the volumes of proceedings:

National Institutes of Health
General Medical Sciences, Bethesda, Maryland
National Cancer Institute, Bethesda, Maryland
National Eye Institute, Bethesda, Maryland
National Heart, Lung, and Blood Institute, Bethesda, Maryland
National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina
National Institute of Neurological Diseases and Stroke, Bethesda, Maryland
National Center for Health Services Research, Rockville, Maryland
The Biometric Society, Raleigh, North Carolina
Abbott Laboratories, North Chicago, Illinois
Arthur D. Little, Inc., Cambridge, Massachusetts
Bristol Laboratories, Division of Bristol-Myers Company, Syracuse, New York
Burroughs Wellcome Company, Research Triangle Park, North Carolina
Carter-Wallace Inc., Cranbury, New Jersey
Ciba-Geigy Corporation, Summit, New Jersey
Energy Research and Development Administration, Oak Ridge, Tennessee
Hoechst-Roussel Pharmaceuticals, Inc., Somerville, New Jersey
Hoffman-La Roche Inc., Nutley, New Jersey
Mr. William R. Holt, Elmsford, New York
ICI United States Inc., Wilmington, Delaware
International Union of Biological Sciences, Paris, France
Lederle Laboratories, Pearl River, New York
Lilly Research Laboratories, Division of Eli Lilly and Company, Indianapolis, Indiana
McNeil Laboratories, Inc., Fort Washington, Pennsylvania
Mead Johnson Research Center, Evansville, Indiana
Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania
Merrell-National Laboratories, Cincinnati, Ohio
Norwich Pharmacal Company, Norwich, New York
Ortho Pharmaceutical Corporation, Raritan, New Jersey
Pfizer Pharmaceuticals, New York, New York
R.J. Reynolds Industries, Inc., Winston-Salem, North Carolina
A.H. Robins Company, Richmond, Virginia
Sandoz, Inc., East Hanover, New Jersey
Smith Kline and French Laboratories, Philadelphia, Pennsylvania
E.R. Squibb and Sons, Inc., Princeton, New Jersey
Syntex (USA) Inc., Palo Alto, California
The Upjohn Company, Kalamazoo, Michigan
USV Pharmaceutical Corporation, Tuckahoe, New York
Vick Chemical Company, Philadelphia, Pennsylvania
Warner-Lambert Research Institute, Morris Plains, New Jersey
Wyeth Laboratories, Philadelphia, Pennsylvania

OFFICERS OF THE BIOMETRIC SOCIETY

President	H.L. LeRoy (ROeS)
Vice-President	C.R. Rao (GInd)
Secretary	J.S. Williams (WNAR)
Treasurer	L.A. Nelson (ENAR)

PROGRAM COMMITTEE, IBC

E.A. Gehan (Chairperson) (ENAR)
L.D. Calvin (WNAR)
L.C.A. Corsten (ANed)
R.N. Curnow (BR)
K. Dietz (Advisor) (ROeS)
J.J. Gart (ENAR)
F.P. Gomes (RBras)
A. Kudô (GJAP)
J. Lellouch (RF)
E. Marubini (RItl)
D. Rasch (RGDR)
E.J. Williams (AR)

PROGRAM CHAIRMAN BIOMETRICS SECTION, ASA

J.S. Williams

PROGRAM COMMITTEE, WNAR

R. Elashoff (WNAR)

LOCAL ORGANIZING COMMITTEE

Y.M. Bishop (Chairperson)
W.G. Cochran
T. Colton
S. Hartz
R. Kent
T. Louis
S. McKinlay
I. Miller
J. Warram
J. Worcester

FINANCE COMMITTEE

W.F. Taylor (Chairperson) (ENAR)
C. Bennett (WNAR)
L.D. Calvin (WNAR)
W. Dixon (WNAR)
S.M. Free, Jr. (ENAR)
E.A. Gehan (ENAR)
D. Gosslee (ENAR)
M. Kastenbaum (ENAR)
L.A. Nelson (ENAR)
M.A. Schneiderman (ENAR)

TABLE OF CONTENTS

THURSDAY, 26 AUGUST, 1976, 8:30 A.M.

STATISTICAL CONTRIBUTIONS TO ENVIRONMENTAL PROBLEMS

(Joint with Biometrics Section, ASA
and Biometric Society (ENAR & WNAR))

ORGANIZER AND CHAIRMAN D.G. Hoel (USA)

SPEAKERS

R. Peto (England)

"The Horse-Racing Effect: Statistical Peculiarities of Rates of Change of Medical Parameters Observed in Prospective Studies with Special Reference to Lung Disease"

W.S. Cleveland and B. Kleiner (USA)

"The Use of Some Graphical Statistical Methods in the Analysis of Photochemical Air Pollution Data" 15

T. Sager (USA)

"Relating Spatial Distributions of Pollutants to Health Effects" 35

J. Neyman (USA)

"Descriptive Statistics vs. Chance Mechanisms and Societal Problems" 59

DISCUSSANTS J. Tukey (USA) and

J. Pratt (USA)

THE HORSE-RACING EFFECT:

Statistical peculiarities of rates of change of medical parameters observed in prospective studies, with special reference to lung disease

Richard Peto

Department of the Regius Professor of Medicine
Radcliffe Infirmary
Oxford University, England

Summary

There are consistent differences between individuals in the average rates of change of certain parameters (e.g. lung function, blood pressure, etc.). These consistent differences have a common effect both on the actual parameter level in middle age and on its rate of change, or slope. Correlations between level and slope due simply to this may be called "horse-racing correlations", and they can cause some curious anomalies in the multiple regression analysis of data from prospective measurement studies. I shall discuss these anomalies and ways of avoiding being misled by them, with particular reference to a recent 8-year prospective study of lung function.

1. MEDICAL BACKGROUND MATERIAL

In Britain, about 10% of male deaths are due, directly or indirectly, to chronic obstructive lung disease (COLD). COLD is unlike most other chronic diseases - except, perhaps, hypertension - in that simple measurement ten, twenty or even thirty years before death (of the amount of air a person can blow out when he tries as hard as possible) can tell us how advanced the degenerative obstructive changes already are. At age 25 or so, most men can blow out a good

four litres of air in the first second of expiration. This excellent Forced Expiratory Volume (FEV) at age 25 won't last, however; even in normal, healthy people the FEV decreases irreversibly with age, until at age 75 the FEV will only be about two litres. In those who are going to die of COLD this lifelong loss of FEV is more rapid: by some time in middle age they will be down to an FEV of perhaps one litre, by which time they start to have to pause for breath in ordinary daily activities like walking around. Progressive loss of FEV will continue for several more years, during which time total disability due to lack of air is likely to develop. Death will probably not occur until years of disability have been suffered. Usually, loss of FEV is irreversible: external factors may influence the rate of loss, but no matter what medical treatment is given substantial increases of FEV will not occur (unless the underlying disease was asthma or something else other than COLD), because reduced FEV is a sign of irreversible destruction of lung tissue. It has been known for at least 20 years that the chief cause of excessively rapid loss of FEV is cigarette smoking, so much so that lifelong non-smokers hardly ever die of COLD. However, it has also been known that although many smokers will die from cigarette-induced COLD (and many will only be saved from death due to COLD by the intervention of premature death from another cause) many smokers will retain perfectly normal lung function, just like that of typical non-smokers, to age 75 and beyond. Moreover,

people in the lower social classes who smoke a given amount are, for reasons which are still unknown, more than five times as likely to die of COLD as are people in the upper social classes who smoke the same amount. Both these facts show that as well as smoking habits there are other critical determinants of the rate of loss of FEV. Obviously, it might be useful if we understood these other processes that determine which smokers will die and which will not.

2. CAUSES OF ABNORMALLY RAPID FEV LOSS

Because low FEV causes unfitness, it leads people to cut down their daily cigarette consumption or to stop smoking altogether. This distorts the relationship between low FEV and cigarette consumption that would be observed in a straightforward survey of middle-aged men; because heavy cigarette consumption both causes, and is prevented by, low FEV, it is difficult to make valid quantitative inferences about the effects of smoking. In some studies this distortion is so strong that the natural relationship, in which the lighter smokers have suffered smaller FEV losses, is actually inverted. Because of this, certain effects can most easily be studied by observing men prospectively, measuring FEV repeatedly over several years. This allows the rate of loss of FEV to be estimated, and this can be used for studies of the causal effects of cigarette smoking and other factors on permanent rates of loss of FEV. Two large prospective studies of FEV have been undertaken in Britain,

and the larger of these has just been published as a monograph (Fletcher et al, 1976). The aim of Fletcher's study was to measure rates of loss of FEV in 1000 ordinary working men (the population being a sample from two engineering works and an office block in London), to record smoking habits in as much detail as possible and to assess by regular questionnaires mucus hypersecretion and various types of infective episode. Hypersecretion and infective episodes may be related to the causal chain leading from cigarettes to permanent loss of FEV, and Fletcher's data are, as expected, sufficiently accurate to test various different theories critically. (Few have survived.) However, the data were collected between 1961 and 1969 and it is now 1976. I was the statistician supposed to collaborate with Professor Fletcher on the analysis and interpretation of his data, and several of those seven years that have elapsed between completion and publication are filled with my mistakes. Some were elementary (like failing to keep a spare copy of the magtape, with disastrous results), but some were not. There are some statistical principles involved in the particular problem of interpreting prospective data on changes in continuous-valued random variables in terms of cause and effect which have no analogue, as far as I know, in other fields, and I want to discuss these particular problems. These problems would also arise in prospective studies of changes in any other continuous biometric variable, such as height, blood pressure, cholesterol, etc. Each of the

large mistakes I made took about six months to put right, since the data always had to be re-analysed; and I hope to present a paper which, had I seen it in 1969, would have saved me one or two years of wasted time with this study. Of the three chief biases I shall discuss, two are already reasonably widely advertised in the statistical literature (although they are not yet widely enough known), while as far as I know almost no explicit attention has yet been drawn to the third bias, which we came to call the "horse-racing effect".

3. THREE PARTICULAR BIASES

(a) Regression to the mean

Suppose we toss two coins, scoring 1 for heads and -1 for tails. The mean score and the difference (first minus second) between the two scores are, in expectation, each zero, and the mean score is uncorrelated with the difference. However, if the first score is positive then the difference can only be zero or negative, while if the first score is negative then the difference must be zero or positive. The first score and the difference are therefore negatively correlated with each other, but it would clearly be foolish to interpret this negative correlation as evidence for any real relationship between the two scores: it is merely a simple example of the classical effect called "regression to the mean". The moral is that it is usually more valid to study the relationship between the rate of change in a set of measurements and the mean of

all of them than the correlation between the initial value and the subsequent rate of change. However, although we avoided these errors, substantial biases which were indirectly due to regression to the mean did invalidate our earlier analyses.

In any prospective study such as ours, some measurements will probably be made only at the initial survey. If the random errors in these ancillary measurements are strongly correlated with the random errors in the "dependent" quantity whose rate of change interests us, then the correlation between this ancillary measurement and the rate of change of the dependent variable will be distorted. The most gross example of such distortion is the well-known correlation between the initial level of the dependent variable and its rate of change, which may easily be so marked as spuriously to suggest the operation of a homeostatic mechanism! Less marked examples of regression to the mean will often arise if, partway through a prospective study, a new ancillary measurement is introduced or an old one is phased out. A correlation between the whole-study mean and rate of change of the dependent variable can also be generated by systematic improvement in measurement techniques at later surveys, due to better apparatus or simply to learning effects.

(b) Incomplete adjustment

This bias affects all statistical investigations which use multiple regression, partial correlation or any of their

analogues: it is not a bias which is restricted to prospective studies, but is a fundamental limitation on scientific inference. If we have three quantities, A, B, and C, then it is commonly argued that if A is still correlated with C given B then the correlation between A and C cannot be merely due to a common correlation with B. This is false, unless B has negligible measurement error. What one should expect is that the F-ratio for the correlation between A and C will be reduced by being multiplied by a factor of approximately S_B^2/σ_B^2 where S_B^2 is the variance of the measurement error of B while σ_B^2 is the total variance of B. The residual correlation of A with C (which may be highly significant) is not, therefore, the critical evidence for an independent association of A with C; what matters is the fraction by which the A-with-C F-ratio is reduced by adjustment for B. If this fraction is large, B may well be the sole cause of the correlation between A and C, even if the residual correlation between A and C is very statistically significant.

(c) The "horse-racing effect"

Discussion of this effect is the chief novelty in my paper. There are consistent differences between individuals in the average rates of loss of FEV which they will suffer during adult life. These consistent differences may be due in part to constitutional factors, to things that happened in childhood or to adult habits, such as the difference between a smoker and a non-smoker. Whatever their causes, they affect both FEV level in middle age and FEV slope (the

rate of loss of FEV). A man whose lifelong rate of FEV loss is more rapid than average will tend to have, in middle age, both a lower than average FEV level and a steeper FEV slope. Conversely, a man whose lifelong rate of FEV loss is shallower than average will tend in middle age to have a high FEV and a shallow slope. If there is substantial heterogeneity in lifelong FEV loss rates, a marked correlation between low FEV and steep FEV slope must exist. Failure to find such a correlation in a large group of men indicates either than the FEV slopes are so inaccurate as to be useless for epidemiological purposes, or that there is no appreciable spread of lifelong loss rates among those men.

At a particular instant in a horse-race, the instantaneous speed of each horse will tend to be correlated with its position, since for obvious reasons the fastest horses will tend to be ahead. By analogy, we have come to refer to the necessary correlation between the value and the rate of change of any quantity as the "horse-racing effect".

Cross-sectional studies, in which each man is seen only once and FEV level is measured, are much easier to do than prospective studies in which each man's FEV is measured several times over a period of 5 or 10 years, and FEV slopes are also calculated. Our prospective study was undertaken to characterise any causal relationships that might exist between FEV and factors such as smoking, infective episodes or chronic phlegm production, each of which might well either cause or be determined by (or both) low FEV. We intended to

do this simply by studying correlations with FEV slope, not realising that the horse-racing correlation between FEV level and FEV slope would mean that this approach would not advance our understanding. What is necessary is to study correlations with FEV slope given FEV level and age, as an approximation to the ideal of studying FEV slope given what kind of a lifelong "FEV loser" each man is by nature.

Example. Suppose that smoking n packs of cigarettes/day just makes each FEV slope n ml/year steeper, and that if smoking ceases the slope reverts to the subject's "natural" slope. (The actual effects of smoking are very much greater than this, but let us first consider the consequences of this simple hypothesis.) Since n ml/year is not very much, if your "natural" FEV slope is shallow you would, under this hypothesis, still have an adequate FEV in old age whether or not you smoke. To get a really low FEV in middle age, you must have a fairly steep "natural" slope. If low FEV caused people to cease smoking, then in middle age we would get not only an inverted relationship between FEV level and current smoking, but also an inverted relationship between FEV slope and smoking. However, among people of the same age who have all lost the same amount of FEV, the continuing smokers will have steeper slopes than those who have recently given up, and so correct inferences might be drawn from the correlation between smoking and FEV slope given FEV level and age.

The above example is not realistic: in reality, the