# The Treatment of Clinical and Laboratory Data

An Introduction to Statistical Ideas
and Methods for Medical
and Dental Workers

By

Donald Mainland, M.B., Ch.B., D.Sc.(Edin.)

Professor of Anatomy, Dalhousie University, Halifax,
Nova Scotia, Canada

With 23 Text Figures



Oliver and Boyd
Edinburgh: Tweeddale Court
London: 33 Paternoster Row, E.C.
1938

PRINTED IN GREAT BRITAIN BY
OLIVER AND BOYD LTD., EDINBURGH

# The Treatment of Clinical and Laboratory Data

# PREFACE

WHILE the ideas that I have tried to set forth in this book were being formed, I received much help from the criticism and questions of colleagues, students and other friends. I wish specially to acknowledge a debt of long standing to Doctor C. H. Goulden of the Dominion Rust Research Laboratory, Winnipeg, for guidance in statistical methods, and for acquainting me with the work of Doctor R. A. Fisher, who has done so much to clarify the logical basis of statistical procedure, as well as to introduce and develop methods that have made that procedure much more useful in clinical and laboratory medicine than it formerly was. In recent years I have had the privilege of Dr Fisher's personal help in several problems, and I wish to thank him for generously giving his time and attention to them. During the preparation of this book I have received much encouragement from him, and where I have secured his help on specific points I have stated it in the text.

If this book should to any extent achieve its purpose, a very large part of its success would be due to the help given by the Banting Research Foundation of Toronto, first to myself and later to some of my research students, in the investigation of quantitative problems in anatomical and associated fields. The trustees of the Foundation have gener-

ously financed these researches, some of which appeared rather remote from medical practice, and it would be fitting if the experience gained in that work could be of some value to clinical workers.

The editor of the *British Medical Journal* kindly permitted me to make use in Chapter II of the material of my article on " Problems of Chance in Clinical Work " (*British Medical Journal*, 1936, ii. 221) and I wish to acknowledge with thanks, also, permission to make quotations granted by the following publishers, editors, authors and others : William Wood & Co., publishers of the *Journal of Pharmacology* (data of Marshall and Barbour, 1933) ; D. Appleton-Century Co. (passages from the White House Conference Report on Child Health—Growth and Development of the Child) ; the Oxford University Press, Professor J. H. Burn and Sir Henry Dale (data and other statements from *Methods of Biological Assay*) ; Professor C. E. Turner, the Massachusetts Institute of Technology and the American Public Health Association (extracts from *Precision and Reliability of Underweight Measurement*) ; the Royal Society (data of Whiteley and Pearson, 1900) ; the American Medical Association (statements of Dr Vanzant and co-workers, 1932) ; the publishers of the *American Journal of Surgery* (data of Dr Der Brucke, 1933) ; the W. B. Saunders Co. (quotations from Dr Raymond Pearl's *Medical Biometry and Statistics*) ; the editor of *Physiological Reviews* (statements from Dr H. L. Dunn's article, 1929) ; the Cambridge University Press, publishers of the *Journal*

*of Hygiene* (data of Dr G. Evans, 1932); the editor of *Nature*, and Dr R. A. Fisher (statements by Dr Fisher, 1929); Messrs Oliver and Boyd (data of Dr W. Q. Wood and Sir David Wilkie, 1933, in the *Edinburgh Medical Journal*, xl. 321); the editor of the *British Medical Journal* (data of Drs Tumarkin, 1936, Black, 1933, Shiskin, 1933, and Millar, 1936); the business manager of the *Journal of Clinical Investigation*, and Dr S. M. Goldhamer (data of Dr Goldhamer, 1933); the editor of the *Journal of the American Dental Association* (data from articles in *Dental Cosmos* by Drs Hellman, 1914, Darlington, 1933, and Arnett and Ennis, 1933); the Deputy Minister, Department of Health and Public Welfare, Winnipeg, Dr A. T. Mathers, Dean of the Faculty of Medicine of the University of Manitoba, and the Great-West Life Assurance Co., Winnipeg (data and comments from the Report on the Poliomyelitis Epidemic of 1928).

Most of the other quotations are much smaller than those just mentioned, but if by oversight I have omitted to obtain permission where it is considered necessary, I extend my apologies.

For the careful typing of my manuscript I am indebted to Miss F. L. Russell, and for the preparation of the diagrams to Messrs G. A. and O. A. Sandoz.

# CONTENTS

*The Examples that are treated in detail are numbered*

# THE TREATMENT OF CLINICAL AND LABORATORY DATA

## CHAPTER I

## INTRODUCTION

When the article that forms the basis of Chapter II of this book appeared in the *British Medical Journal*, I received several letters from clinicians and other medical workers in various parts of the world. These letters confirmed an impression that had been created in discussion with colleagues and students regarding the importance of making adequate allowance for chance in all observations and judgments in medicine and dentistry. This impression was that, in spite of the numerous textbooks on what is commonly called "statistics," something still remains to be done—something that one who is not a professional statistician may do. This book is an attempt to do it—an attempt, primarily, to show why and how allowance should be made for chance, so that a clinician may get the most sound benefit (*a*) from his own observations, and (*b*) from the observations of others. That an anatomist, rather than, perhaps, a physiologist, should make the attempt, appears to be appropriate, because the anatomist, like the clinician, depends for the most part on observation; he cannot isolate the factors that he is studying in the same way or to the same extent as can the

physiologist or biochemist. Like the clinician, the anatomist has frequently to observe the experiments that nature performs.

### Neglect of Chance and of Sampling Errors

Anyone with even a rudimentary idea of the effects of chance will, if he reads medical literature, observe that in a great deal of it there is inadequate recognition and treatment of questions of chance, of the allowance to be made for small samples, and for natural variation from sample to sample of the same material, whether that material is composed of blood specimens, patients with paralysis, or animals subjected to drugs. This statement applies equally to all kinds of unpublished clinical and non-clinical discussions. It is, for example, not uncommon to hear the admission that a certain set of cases is too small to justify conclusions, but it is uncommon to find any attempt to show what conclusions can be reasonably accepted from the given sample. Fairly often the speaker will, in spite of his admission, proceed to generalise and argue from his sample.

To avoid the hasty conclusion that these comments are those of an outsider, a statistical expert unacquainted with clinical problems, it is perhaps desirable to note the stages by which I arrived at this point of view. A medical training, in which no attention was paid to problems of chance and sampling errors, was followed by research in embryology and histology in which these problems at once arose. The adequate teaching of anatomy

entailed wide and continued acquaintance with clinical matters, and it soon became obvious that the same kind of problem was present everywhere, both in medicine and in dentistry.

Certain other things that became obvious may be enumerated thus :

(1) The term " statistics " is unfortunate because it suggests either masses of data, sometimes inaccurate, or highly artificial mathematical tests. The inaccuracy of data, the use of inappropriate tests, and the drawing of unwarranted conclusions have rendered plausible the dictum that " one can prove anything by statistics, except the truth." The very word " statistics " is anathema to some, partly for this reason, partly because of confusion between statistical data and statistical methods, and partly because of ignorance of the basis of the methods and of their legitimate use.

(2) The numerous books on the subject, including those on medical statistics, often give too much and too complicated detail for clinical workers, and commonly fail to show that statistical ideas are of any use to a practitioner.

(3) The most important part, even of elaborate statistical technique, is the critical selection of material. This is illustrated by the title of one of the most recent books on statistics, *The Design of Experiments*, by Dr R. A. Fisher (1937).*

* The numbers inserted after the names of authors indicate the year of publication of a certain article or book. See REFERENCES at the end of this book.

(4) For ordinary clinical work it is the ideas underlying the words "chance" and "sampling variation" that are of most importance, while the necessary technical tests are few and simple. Students have shown that the technique of all the methods detailed in this book, except some of those described in the Supplementary Notes, can be grasped in three or four hours. What takes most time and presents most difficulty is the proper understanding of the problem under investigation, so that the work can be properly planned, and, when the results are being interpreted, appropriate tests can be applied.

(5) This proper understanding of the problem is none other than what is demanded in all good scientific investigation and in all sound observation and judgment by the "practical man." The need would be the same if statistical tests had never been devised.

These five points will be discussed in various parts of this book, but some can be amplified here.

Dunn (1929) reported that out of 200 quantitative medico-physiological papers in current American periodicals over 90 per cent. required statistical methods and did not use them, and that in almost 40 per cent. conclusions were drawn which could not have been proved without setting up statistical control.* My own survey was not numerical and was concerned more with clinical than with laboratory

* Since Dr Dunn wrote this there has in some branches of medical science been a greater realisation of the need for proper methods of analysis, but progress has been much slower than elsewhere, for example in agricultural research.

medicine, but it revealed that the same general verdict, perhaps even a more adverse one, was appropriate in the clinical field, and applicable to articles published in any country. I should, however, prefer to express the verdict in another way, by saying that in large numbers of articles no evidence is given of experimental error, the term "experiment" being used in its widest sense to include observation of material in which nature has performed the experiment. Examination of most of these articles shows that this omission is not simply a failure to state what error should be allowed for, but actual lack of knowledge of the error on the part of the observer. Frequently, indeed, the way in which the observations were planned must have made it impossible for the observer to form a valid estimate of the error. The error of an experiment, in this connection, is simply the means whereby one can form an idea of what results might be expected if the experiment were repeated under the same conditions. Therefore, as Fisher (1937) points out, such an experiment cannot be said, strictly, to be capable of proving anything, and perhaps should not be called an *experiment* at all. In conformity with general usage, the variation between the results of two experiments, conducted under exactly the same conditions, can be attributed to chance. Now, so many everyday judgments, including diagnosis and prognosis, are based on a conception of chance, probabilities, and so on, that the lack of adequate allowance for its effect is curious. Two reasons for it may be mentioned. One is

common to all branches of thought and activity—the fact that the great majority of counts, such as shillings, dollars, books, or patients in the waiting-room, provide definite information, exact in so far as the count has been properly made; and the great majority of everyday measurements, such as heights, distances or weights, are, under ordinary circumstances, exact to the degree required. The other reason is more specially applicable to medicine and dentistry—the fact that quantitative methods have been introduced chiefly from physics and chemistry, and experimental errors in these sciences are usually small relative to the differences that are being measured.

### Mathematical and "Commonsense" Methods

The factors just mentioned are partly responsible for the false antithesis between drawing conclusions by statistical methods and by "commonsense" judgments of the data, the latter method being held to be satisfactory for "practical" purposes. "Commonsense" methods frequently mean no more than acceptance of numerical data at their face value, with perhaps some lip-service to the effects of chance or the size of the sample. Such acceptance of data is dangerous, first, because it does not pay sufficient attention to possible bias, and, secondly, because, even if bias has been eliminated by the design of the experiment, it very frequently fails to make adequate allowance for chance. The tests that are used in making this allowance are naturally mathematical, because counting and measuring are mathematical.

It is a mistake to imagine that they must entail over-refined distinctions or a standard of accuracy greater than the observer wishes or needs. One may use a ruler marked in thirty-seconds of an inch and yet only measure, if one wishes, to the nearest inch. The contrary notions—that mathematical methods (*a*) are unnatural or artificial, and (*b*) are necessarily rigorous—are part of a widespread misunderstanding, which may be due to the mental structure of large numbers of people, but the popularity of Hogben's (1936) *Mathematics for the Million* suggests that the main fault lies in the teaching of mathematics.

Mathematicians themselves are sometimes quoted against those who propose to apply tests to data in medicine. There are several reasons why the evidence of professional mathematicians may have no weight, but it might be thought that experimenters (physicists and chemists) would be in a position to pass a cogent verdict. It should be remembered, however, that medicine and dentistry are essentially parts of biology, a science in which variation between material of all kinds, even obtained under the same conditions, is great and in part uncontrollable. The physicist and chemist use methods in which chance variation plays a small part. The point may be illustrated by the remark made by a professor of chemistry who was invited to sit on a University Board that was examining a student's thesis on the chemical analysis of salmon eggs at various stages of development. One member of the Board suggested that tests might have been applied to determine how far chance could account