

Altman DG (2017). Donald Mainland: anatomist, educator, thinker, medical statistician, trialist, and rheumatologist

© Doug Altman, Centre for Statistics in Medicine, University of Oxford, Botnar Research Centre, Windmill Road, Oxford OX3 7LD, UK. email: doug.altman@csm.ox.ac.uk.

Cite as: Altman DG (2017). Donald Mainland: anatomist, educator, thinker, medical statistician, trialist, and rheumatologist. JLL Bulletin: Commentaries on the history of treatment evaluation

Introduction

“Someone who started teaching as an undergraduate demonstrator in anatomy and continued trying to teach, orally or in writing, for the next 48 years...” (Mainland 1980)

Donald Mainland had a unique career. He graduated in medicine from Edinburgh in 1925 and spent 20 years as Professor of Anatomy in Halifax, Nova Scotia. From the early 1930s he developed a remarkable grasp of statistical principles and methods, and in the second part of his career he was Professor of Medical Statistics at New York University. Here his main focus was obtaining reliable estimates of the effect of health interventions in rheumatology. How did the remarkable transformation from anatomist to statistician come about?

Unusually, Donald Mainland dropped many autobiographical notes into his writings, so we know a lot about his motivations. For example, in 1954 he wrote:

The remarks are those of one who, after graduating in medicine, started research on the embryology of the ferret. He was puzzled by error in cell measurement, by variations in the counts of chromatin particles, and especially by the problems of small samples, because ferrets were expensive. He obtained no answer from biologists, chemists, physicists, or mathematicians, until he was led to the solution by Dr. C. H. Goulden, an agricultural experimenter in Winnipeg, who introduced him to the book by R. A. Fisher (now Sir Ronald Fisher) on Statistical methods for research workers which had appeared three years previously (in 1925).

He then saw that the methods prescribed by Fisher for avoiding bias and allowing for chance, experimental error, biological variation, and sample size, were applicable in all fields of medicine, and that in his own field, anatomy, normal variations in nerve pattern, positions of abdominal viscera, and other organs had often greater clinical importance than the so-called “averages” that he had learned and was teaching. He stole time from the contemplation of the fifteen arteries, which (in those days, at least) arose from the hypogastric artery, and spent this time in applying statistics to his own research and the researches of others, some of it during vacations in Fisher’s laboratory. Then, after World War II he realized that he could not run an anatomy department and keep abreast of statistical developments, and he chose the latter effort (Mainland 1954c).

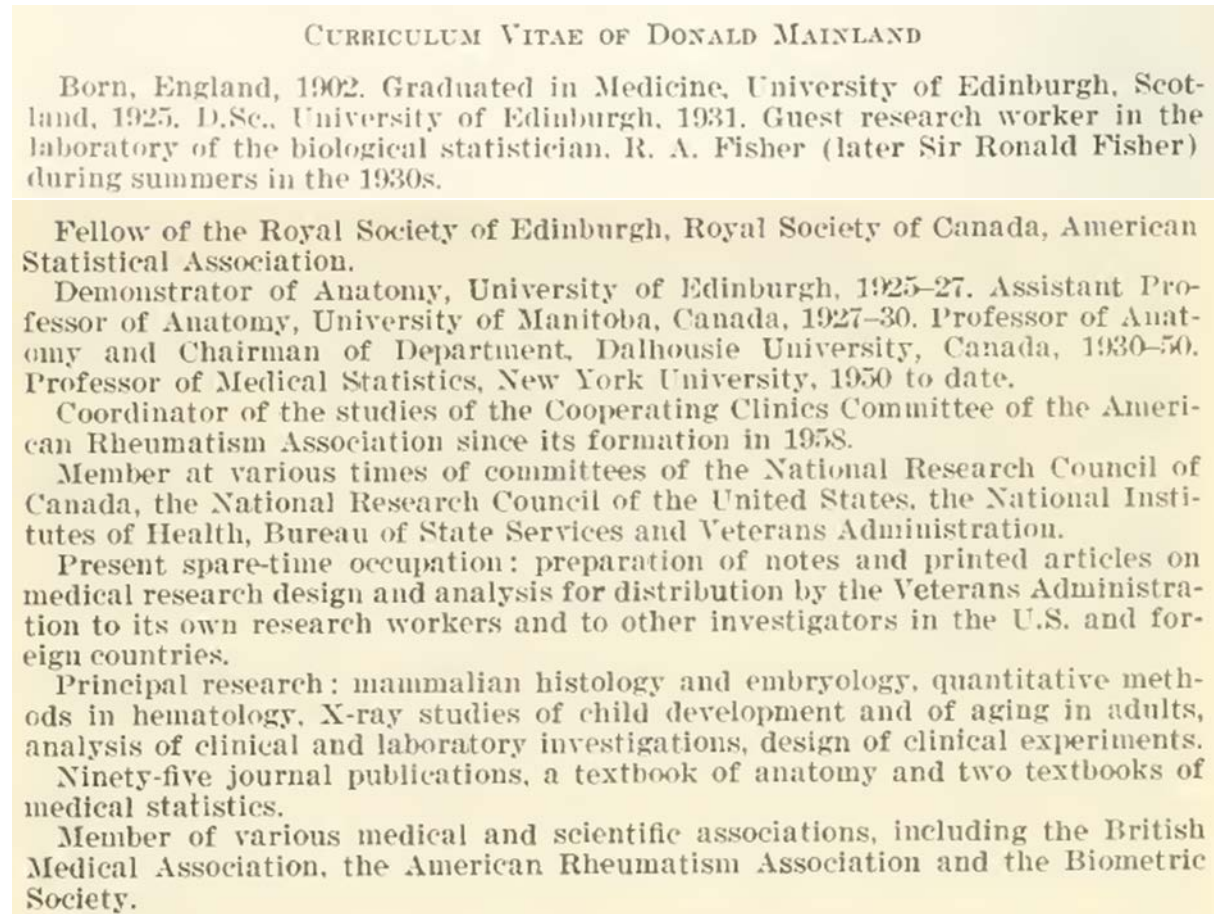
He made rather similar remarks in one of his last publications, in 1980 (Mainland 1980).

Brief Biography

Donald Mainland graduated in medicine at Edinburgh. He taught anatomy in Edinburgh and received a Doctor of Science degree there for his research in embryology and histology. In 1927 he moved to Winnipeg, Manitoba, Canada, and in 1930, at the age of 28, became Professor and Chairman of the Department of Anatomy at Dalhousie University. Even his earliest publications showed an interest in measurement issues, and foreshadowed an increasing interest in statistics. In 1938 he published his

first book on statistics in medicine (Mainland 1938). In 1950 he became Professor of Medical Statistics at New York University and shortly afterwards published his best known book, *Elementary Medical Statistics* (Mainland 1952). Thereafter, Mainland was a prolific and influential writer on statistical topics.

Mainland provided his own CV in the report of some Congressional hearings in 1967, of which more below (U.S. Government 1968).



Assistant Professor in Anatomy, University of Manitoba, Winnipeg

In 1927 Mainland moved to Winnipeg in Canada; his writings do not provide any specific motivation for this move. His three years in Winnipeg yielded 13 published research papers from 1927 to 1931. Five were related to ferrets and one to dogs. One article, in German, was a technical note on staining brains. The jump from this area of work to medical statistics is astonishing. Yet as early as 1931 he began a paper on the sizes of nuclei in ovarian stroma with the following remarks:

In histology two main routes of advance are open. The first is the change and improvement of preparation, fixation, and staining of tissues. The second, less used route is the application of strict quantitative methods to specimens prepared by ordinary technique. This second method has to its disadvantage great laboriousness, the natural antipathy of many toward mathematical methods, and the distrust with which statistical methods especially are viewed. On the other hand, this method is technically simple, consisting chiefly of careful measurement. The statistical treatment of the data may be mastered without great difficulty, and the distrust of statistics diminishes as one becomes familiar with the principles.

Even by 1931, therefore, he had clearly recognised the importance of quantitative methods.

Professor of Anatomy, Dalhousie University, Halifax, Canada

In 1930 Mainland moved to be Professor of Anatomy at Dalhousie University, at the age of 28. Some of his early publications after moving to Halifax showed a concern about measurement issues, which undoubtedly led to his rapidly developing interest in statistics. Mainland had been introduced to the writings of RA Fisher while in Winnipeg, just a few years after Fisher published his famous book *Design of Experiments*. It is clear that Mainland immediately saw the value of this methodology for his own research. Not only did he employ and write about random sampling from an early date, he visited Fisher more than once in London in summers in the 1930s. Those visits presumably followed Fisher's 1934 invitation:

13 April 1934.

Professor D. Mainland,
Department of Anatomy,
Forrest Building,
Dalhousie University,
Halifax, N.S.

Dear Professor Mainland,
It would give me great pleasure, when you are in London, if you would call on me at the Galton Laboratory to discuss what we can do to help you. I think you could be registered if you cared to, as a voluntary worker in my department, which would give you immediate access to our library and equipment.

[Source: Fisher archive, Adelaide]

Fisher is better known for his theoretical developments (in statistics and genetics) than in relation to practical applications of statistics, so Mainland's observations about one of his visits are revealing:

I recall that, during a summer that I spent as a guest in Fisher's laboratory in London, I saw the testing of human olfactory sense, the sampling of human hair, and an investigation of snails. Through personal contact with Fisher I came to know his common-sense attitude in practical research situations, and a desire to find a solution for the experimenter's problem, rather than to use the problem for the indulgence of his mathematical interests (Mainland 1969b).

In 1938 Mainland published his first book on statistics in medicine, *The Treatment of Clinical and Laboratory Data: An Introduction to Statistical Ideas and Methods for Medical and Dental Workers* (Mainland 1938). Fisher was thanked profusely in the preface.

Mainland's 1938 textbook is unusual in its prominent treatment of laboratory data, clearly reflecting Mainland's unusual path into medical statistics. This emphasis is shown in particular by the 75 page chapter on "Data based on measurements" and a further 37 pages on "Errors in measurement". Key ideas include concern about sampling variation, bias, and problems associated with small samples. Another unusual feature was the final chapter "Publication of Data and Results." Here he first discussed the over-summarisation of data in published research papers, and the desirability of others having access to the raw data. He observed that:

"Careful examination of many published articles reveals, however, that, in spite of their containing much numerical information, they do little more than present: (a) an author's opinions, and (b) data that seem to confirm those opinions, that is, do not disagree with them."

Among other topics he addressed in supplementary notes were several issues relating to sample size, including “The size of sample necessary to demonstrate a significant difference” and “The size of sample necessary to estimate a possible difference within desired limits” (pp 305-8). Here he was thinking along lines much as we do now, although the ad hoc trial and error approaches he suggested at that time have been superseded, including in Mainland’s later writings.

Later, Mainland was hugely impressed by the work of Bradford Hill, as exemplified by the 1948 report of the MRC streptomycin trial, which he described as “a beacon” (Chalmers 2005). As he wrote in 1980 (quoted above), the opportunity to work on clinical trials was a major reason for his move from Halifax to New York in 1950. As he reminisced much later:

“... the randomized controlled trial ... is now so familiar as to have lost its glamour, but in the thirties and forties it was tremendously exciting to apply the new methods to research in histology, embryology, blood counts, cell measurement, and my radiological bone and joint studies.” (Mainland 1980)

Mainland was largely self-taught as a statistician, which certainly put him apart from most academic statisticians. Fisher, not noted for his tolerance, clearly saw something in him. He appears to have remained in contact with Fisher, who wrote to Mainland as “My Dear Donald” in 1955. Mainland was conscious of his unorthodox entry into the profession. He began a nice essay in 1950:

“A discussion of statistical methods in clinical research by one who is neither a professional statistician nor a clinician must appear rather incongruous. There may be some advantages in this incongruity, however, for there may be something of interest in the point of view of one who is using statistical methods daily in the laboratory and who is also in daily contact with clinicians. Such a person sees, by contrast, the difficulties of clinical research, and he sees how clinicians who possess keenly critical minds and are interested in scientific medicine are handicapped by lack of acquaintance with statistical ideas.” (Mainland 1950)

Similar comments appear elsewhere in his writings.



Donald Mainland in 1930s. Source: Dalhousie University

Between 1939 and 1946 Mainland published only 2 journal articles. However, it is surely relevant that in 1945 he published his 863-page textbook on anatomy. After the 1938 book he didn't produce another statistical publication until 1948, a 166-page journal article *Statistical methods in medical Research. I. Qualitative Statistics* (Enumeration data), which is, in effect, a mini text book (Mainland 1948). Whether this was intended as the first of a series is unclear, but only one further item

appeared under this heading: part II, on sample size, appeared five years later and was only 11 pages (Mainland and Sutcliffe 1953).

Mainland's 1952 text *Elementary Medical Statistics* (Mainland 1952), although published when he was already in New York, was almost wholly written when he was in Halifax. Indeed, he notes it was based on courses run in Dalhousie. It includes numerous uses of 'random' and 'randomize(d)', but none of 'random allocation', presumably reflecting the fact that while in Halifax he had little or no involvement in clinical trials. There is no doubt that Mainland had strong views on the importance of both random sampling and random allocation.

Professor of Medical Statistics, New York University: Rheumatology trials

The unique switch from Professor of Anatomy to Professor of Medical Statistics came just 2 years after the publication of his huge textbook on anatomy in 1948. It is clear that his earliest work in anatomy sparked his interest in measurements and things statistical, and this book was the watershed between his two very different careers. Just two years later he effectively abandoned anatomy for medical statistics and, in particular, clinical trials. A revealing comment is from 1980:

The possibility of promoting controlled trials in medicine was one of the inducements for me to become a medical statistician in New York in 1950. In those days many clinical and laboratory investigators expected a statistician, after a brief consultation, to do some arithmetic and affix a "certificate of purity" to their reports. Others, who wanted a statistician to co-operate from the planning stage onwards, often failed to realise how few projects a single statistician could thus handle simultaneously... Fortunately, I was able to escape into full-time participation in multicentre therapeutic trials in rheumatology, in which I could visit the clinics and get to know the investigators. Here, as in anatomy teaching, a medical degree (along with my radiological research) gave me insight and entrée, and I even acquired a third label – "rheumatologist" (Mainland 1980).

Clearly he relished the new opportunities. Although he was an author of rather few published reports of clinical trials it seems he was deeply involved in the trials programme. In 1954 he wrote memorably: "To omit randomization because one cannot see clearly how bias could occur is like trusting that glassware in chemistry is clean because it does not look dirty." (Mainland 1954b) Mainland was in the Department of Preventive Medicine at New York University but it seems that most of his clinical projects were in the field of rheumatology.

A clear shift of emphasis can be seen when comparing the two editions of his *Elementary Medical Statistics* (Mainland 1952; Mainland 1963a). The second edition, written after many years in New York, is so different from the first it may fairly be considered a different book. The first 10 chapters in the second edition are written in question and answer format. Unusually, this section includes a chapter on "Lost information" (what we now call 'missing data').



New (2nd) Edition!

Mainland's ELEMENTARY MEDICAL STATISTICS

**A remarkably clear guide to help you
decipher the precise meaning of
statistical reports**

Dr. Mainland shows precisely how to use and interpret statistical methods in medical investigation—methods useful in situations ranging from analysis of results in basic research to evaluation of experimental medical or surgical procedures. Statistical *thinking* is emphasized rather than mere mathematics. You'll find clear de-

lineation of the basis of tests and of the methods of applying tests. Easily understood examples accompany explanations. The first ten chapters are arranged in the form of vital questions to ask yourself regarding your projected, current, or past research and of questions which you can apply in evaluation of others' work. Each such question is fully discussed. In chapters 11 to 16, specific methods of analysis are presented and described. You'll find help on how to understand and use: *random processes—standard deviation—frequency distribution of measurements—causes of bell-shaped distribution—estimation of population percentiles—correlation coefficients*. Dr. Mainland places prime emphasis on rank-order and other nonparametric methods. Electronic computers, attempts at "machine diagnosis," etc., are also evaluated.

CI11-63

W. B. SAUNDERS COMPANY

West Washington Square, Philadelphia 5, Pa.

Please send and bill me: ☐ Check enclosed
(postage prepaid)

☐ Mainland's *Elementary Medical Statistics*...\$9.00

Name

Address

By DONALD MAINLAND, M.B., Ch.B., D.Sc., Professor and Chairman,
Department of Medical Statistics, New York University College of
Medicine. 381 pages, 6½" x 9½", illustrated. \$9.00.
New (2nd) Edition!

Educational writings about medical statistics

Mainland's unique contribution was a long series of writings aimed at explaining and discussing statistical ideas for non-experts. As well as his textbooks he published many educational articles. For example, early in his time in New York he wrote 14 chapters in a volume of *Methods in Medical Research* (with co-authors for a few chapters) (Mainland 1954d), and an article about the problems of being a medical statistician (Mainland 1954c). He wrote about the use of case records, including how to do nonrandomised studies well (Mainland 1955b), and he published 3 articles in the *American Heart Journal* titled "Notes on the planning and evaluation of research, with examples from cardiovascular investigations", a total of 39 journal pages (Mainland 1958a; Mainland 1958b; Mainland 1958c). These were soon followed by two articles on "The use and misuse of statistics in medical publications" (Mainland 1960b) and "The clinical trial – some difficulties and suggestions." (Mainland 1960a)

Most notable, however, were three series of privately circulated "notes". First, between August 1959 and September 1966 he produced a long series of 145 *Notes from a Laboratory of Medical Statistics*. As Mainland later observed (Mainland 1967b): "The 2,500 name mailing list included persons in all areas of medicine and related disciplines, and 25 countries were represented." The series continued under the auspices of the Veterans Administration – 104 items from July 1967 to December 1970 – with a new title, *Notes on Biometry in Medical Research*. Simultaneously, between January 1967 and December 1969, Mainland produced *Statistical Ward Rounds* as a series of 18 long articles in the journal *Clinical Pharmacology and Therapeutics*. These publications together form a substantial body of work – in reprinted book form they comprise 1912 pages (Mainland 1978; Mainland 1979a; Mainland 1979b). At the end of the first of the Ward Rounds he outlined his motivation: "I have to confess that my deepest reason for starting these ward rounds is that I enjoy trying to get my thoughts clearer about statistical ideas and methods, bringing old ideas to new problems, and looking at old problems in new ways."

The *Notes* were often a test bed for new educational material. Later journal articles often drew on the material in the *Notes*. The first 10 items in the series *Notes from a Laboratory of Medical*

Statistics were, however, based on the three articles he had published in the American Heart Journal, and this material then became the basis of the first ten chapters of the second edition of his book *Elementary Medical Statistics* (Mainland 1963a). Most of the Notes addressed issues that arise frequently in everyday applications of statistics, such as sample size, the meaning of significance, multiple testing, causal inference, and identifying outliers.

Not only was Mainland highly prolific, he was influential. As Alvan Feinstein wrote: “With his textbook . . . 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” (Feinstein 1970). Even 25 years after his death and about a half-century after its publication, an editorialist in the *Journal of Rheumatology* recommended Mainland’s *Elementary Medical Statistics* “to all young clinician scientists” (Lassere 2010).

Statistical thinking

A recurring theme of Mainland’s writings was the importance of clear thinking about the problem rather than focusing on the mathematical side of statistics. This idea is the title of a 1982 article, “Medical statistics – thinking vs arithmetic” (Mainland 1982), but thirty years earlier he had addressed the same idea in the opening sentence of the Preface to his text *Elementary Medical Statistics* (Mainland 1952):

Those who have for many years stressed the importance of statistical thinking in medicine cannot be entirely happy to see statistics becoming established as a subject in the undergraduate curriculum, for it thereby becomes liable to the curse that is on all medical “subjects”: the emphasis on memorization, on techniques, and on preparation for board examinations which foster statistical pedagogy.

Similarly,

Research workers’ widespread lack of understanding of the rationale of statistical techniques, and the frequent use of statistical tests as a substitute for thoughtful investigational design, meticulous work, and repetition of experiments, justify the antagonism to statistics exhibited by some experimenters. To one who has had personal experience of the way in which statistical thinking, as distinct from statistical arithmetic, can promote good investigation, this perversion of statistics is lamentable (Mainland 1960b).

Such thinking has its roots in Mainland’s earliest statistical work, as exemplified by this comment from 1936:

What is required in clinical work is not elaborate mathematical tests, but an understanding of the meaning of chance, and adequate precautions that the samples, however small, are unbiased (Mainland 1936).

In 1929, Dunn had reported that over 90% of 200 journal articles required statistical methods but hadn’t used them, and that, in almost 40%, conclusions were drawn which could not have been proved without setting up some adequate statistical control (Altman and Simera 2015). Dunn felt that about half of the papers should never have been published as they stood. Mainland commented on Dunn’s study and noted that his own survey of clinical journals

... revealed the same general verdict, perhaps even a more adverse one, was appropriate in the clinical field ... 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 ... an idea of what results might be expected if the experiment were repeated under the same conditions (Mainland 1938, p 5).

He recognised that many errors are elementary: "...almost any volume of a medical journal contains faults that can be detected by first-year students after only three or four hours' guidance in the scrutiny of reports" (Mainland 1952). In *Note 12 of Notes on Biometry in Medical Research* (Mainland 1978), Mainland discussed in detail the findings of an important review of publications by Schor and Karten (1966).

Poor reporting of research has become a major concern only rather recently (Altman and Simera 2015). So Mainland's 1938 textbook was unusual in concluding with a whole chapter on "*Publication of data and results.*" In the first paragraph, he observed that:

"... incompleteness of evidence is not merely a failure to satisfy a few highly critical readers. It not infrequently makes the data that are presented of little or no value." (Mainland 1938)

Tests of significance

Mainland was well aware of difficulties relating to the wide use and misinterpretation of significance tests and their resulting P values. Of particular concern, then as now, was the wide misinterpretation of a result that was not statistically significant ($P > 0.05$) as evidence of 'no effect'. In 1963 he wrote:

The terminology itself is, I think, largely to blame for the persistent confusion. Our thinking today might have been much clearer if from the beginning the statisticians had used a more specific, more explanatory, and less pretentious term such as "random frequency test" instead of the ambiguous and grandiloquent "significance test." (Mainland 1963a)

He did use the term random frequency test in some of his writings, including in his last publication (Mainland 1985) although not in his contemporaneous textbook (Mainland 1963b). His suggestion had no impact. Mainland recognised the importance of confidence intervals, describing these as early as 1948 (Mainland 1948). Again he observed that their naming was unfortunate:

"... as one of my students remarked, they might be more appropriately called "no confidence limits." Perhaps a still better title would be "minimal estimates of ignorance." This sounds discouraging, but it forces us to recognize that no single study can give us very precise knowledge of the value of a therapy. (Mainland 1963a)

He doesn't seem to have suggested the abolition of p values even though he clearly saw their problems:

"Statistical tests are very dangerous drugs." (Mainland 1969)

"Compulsion to apply a random frequency [significance] test may arise from ignorance or from tribal custom." (Mainland 1978, p77)

This perspective was clearly an important element in his widely stated promotion of statistical thinking rather than statistical arithmetic (Mainland 1982).

Insights

Donald Mainland was not a methodological innovator. Although a method of analysing crossover trials is dubbed the Mainland-Gart test, it is not much used these days. Rather, Mainland addressed common situations in clinical research for which he often demonstrated unusual insights, in some instances many years before the concepts became well-recognised. Some notable examples follow.

Confidence intervals

Confidence intervals were introduced in the 1930s but did not become common in clinical research until the 1980s. Mainland may first have described them in his 1948 mini-text on the analysis of categorical ('enumeration' data) (Mainland 1948). In both editions of his textbook *Elementary*

Medical Statistics (1952 and 1963) confidence intervals were introduced before tests of significance. In the first edition he wrote:

Confidence intervals are required for all estimates, whether of enumeration data or measurements; and they always refer to populations that could be randomly represented by the observed sample (Mainland 1952, p56).

In 1963 he wrote about *The possible demise of "significance tests"*:

In fact there are signs that mechanical "significance" testing, although far from moribund, is not so vigorous as it was a few years ago; and with its death the problem of "non significance" would no longer plague us. It is to be hoped that it would not be replaced by some other misunderstood mechanical trick (Mainland 1963b).

Mainland's optimism was misplaced. Fifty years later the misuse and misinterpretation of significance tests is a bigger concern than ever.

Cluster randomisation

In his 1952 book (Mainland 1952) (p114), Mainland provided:

"... what might be the first attempt in the medical literature to distinguish randomized trials by the unit of allocation. Indeed, in spite of some largely isolated examples, it was not until a brief but seminal article by Cornfield (1978) that these ideas were brought to wide attention to researchers in the health sciences" (Klar and Donner 2004).

Units of analysis

Mainland addressed a related issue, "spurious replication", in his 1952 text (Mainland 1952, p32) and at rather greater length in his second edition (Mainland 1963a). As noted by Bolton:

"Mainland gives other examples, but gets right to the point, defining the analysis of multiple observations from individual units (patients, for example) as mixed sampling. He contends that one has to be very careful about the objective of the experiment. Are we interested in a population of patients or in a population of readings from an individual patient?" (Bolton 1998)

Comparing methods of measurement

Mainland criticized the use of the correlation coefficient for comparing methods of measuring a quantity:

Akin to this misuse of r is the habit of employing it to express the reliability of a technique when an observer has made duplicate measurements or when two observers have measured the same thing. Even when the coefficient is 0.95 or higher, it does not tell us whether, for the purpose in hand, the differences between the duplicate readings are trivial or serious (Mainland 1952, p 334).

A little later, in 1955, he wrote:

From the series of differences (reading by standard method minus reading by easier method) find the mean difference, i.e., the systematic difference between the two methods, and find the standard deviation of the series, i.e., the variable or random difference between the methods (Mainland 1955a).

This is essentially the approach recommended by Bland and Altman 30 years later (Bland and Altman 1986). In fact, Mainland had expressed similar ideas back in 1938 (Mainland and Stewart 1938).

Surrogate measures

Mainland was among the first to draw attention to the issue of ‘surrogate outcomes’, as we now call them (Lassere 2008). In his 1963 textbook he asked “Will the variables that we observe be the variables that we really wish to know about?” (Mainland 1963a). He observed that “if we are substituting something that is easy to observe for something that is difficult to observe, we have no right to do so unless we know the connection between the two things.” Mainland referred to the practice as the ‘substitution game’, attributing this term to Yerushalmy, and discussed the inferential challenges associated with such outcomes.

A few years later he gave similar critical attention to composite indices of disease severity, which were common in studies of rheumatoid arthritis (Mainland 1967a).

Missing data

Mainland’s 1963 textbook includes a whole chapter on “Lost information”, including discussion of both how to minimise losses and how to analyse data when there are missing observations, focusing in the latter on strategies rather than statistical procedures. He observes that the approach should be determined “before our decision can be influenced by the outcome of the trial.” (Mainland 1963a, p 182)

Combining data from multiple studies

In his 1948 mini-text Mainland considered the question of combining data from multiple similar studies (Mainland 1948). He observed that simply collapsing the data as if from a single study “entails a risk of fallacious conclusions from heterogeneous data.” Long before the term meta-analysis was introduced, Mainland used methods based on combining test statistics or P values, rather than effect estimates.

Reporting of research

Mainland was one of the first to recognise the critical importance of how research studies are reported (Altman and Simera 2015). Within the unusual closing chapter on publication in his 1938 textbook, he addressed tabular data, the importance of assessing the reliability of research findings, and the importance of publishing “negative evidence”, and there were two pages addressing summaries and abstracts (Mainland 1938).

Several of Mainland’s Notes addressed deficiencies in publications. *Note 10 of Notes on Biometry in Medical Research* (Mainland 1978) discussed in detail the critical review of a published report of a clinical trial. It includes ten lists of questions (56 items in all) for peer reviewers to consider.

Clinical trials

As noted above, the opportunity to work on clinical trials was a major reason for his radical move from anatomy in Halifax to statistics and clinical trials in New York. In his last publication, he recalled that he had been “enthralled by the earliest controlled therapeutic trials, which applied *statistical thinking* – realistic attention to the effects of variation and the risks of bias” (Mainland 1985). Given his enthusiastic endorsement of Fisher’s methods, and randomisation in particular, this move may perhaps not have been so surprising.

In New York he worked full-time on multicentre therapeutic trials in rheumatology (Mainland 1980). Within a few years he had published an article about clinical trial methods (Mainland 1954a) and more such articles followed (Mainland 1960a; Mainland 1961). Also, many of the statistics notes addressed issues arising in clinical trials. However, over more than 20 years he was co-author of only a handful of publications of the results of trials, the first not until 1962 (Mainland and Sutcliffe 1962) and the last in 1973 (Mainland et al. 1973). This record presumably reflects the collaborative

publication practices of the time as it is clear from his publications and unpublished notes that he spent a lot of time working on the design and analysis of clinical trials.

Congressional hearings (1968)

In 1968, Donald Mainland was an expert witness in Congressional hearings on Competitive Problems in the Pharmaceutical Industry (U.S. Government 1968). Early in his testimony, Mainland offered some general remarks about the value of good controlled trials (experiments):

A word about experiments. We do not use the word "experiment" in our discussion of these trials. We call them trials because experimenting on human beings has a very bad sound. It sounds as if controlled trials are experiments on human beings in the sense of using those people not for their own benefit but for the experimenter's purposes.

Those of us who know what actually goes on in a good drug trial—that is, a trial that is conducted by good investigators who are also good doctors—we see the matter differently. An important point that many people do not realize is that every time a doctor in the ordinary course of his work tries a drug, even a much used drug, on a patient for the first time he is performing an experiment. It may do the patient no good and, with most of the powerful modern drugs, it may well do him harm.

If I were a patient, I would much prefer to be in a good experiment, that is, a good drug trial, even taking a 50–50 chance of a placebo, because I would receive more thorough and systematic examinations than can possibly be given to routine patients. And I would know that, in a double-blind trial, neither I nor the doctor have been fooled by knowing that this was supposed to be a miracle drug. And, I would know that, if things were going wrong, I would be taken out of the trial.

A word about pharmaceutical companies. The executives of the pharmaceutical company may, as individuals, be altruistic. I have met, over the years, a number of them, and I feel they are basically altruistic individuals. I know some of their scientists and often they are of high quality, but we should never forget that the primary aim of a company is commercial. A company can make a quick million or two out of a new drug that is subsequently found to be of no particular value.

Now, some of the more enlightened companies are coming to realize that a succession of those things, a succession of these flash-in-the-pan drugs, does not build a good reputation. We may hope for the spread of such enlightenment, but in the meantime, we should, I think, ask what can be done to counteract the influence of the marketplace in drug evaluation.

It seems to me that the most obviously desirable step would be to take the evaluation of drugs entirely out of producers' hands.

Rodwin noted that the subcommittee heard testimony from several individuals proposing reforms that spanned a continuum from modest changes in current arrangements, to shifting the responsibility for testing drugs' safety and efficacy from drug firms to the federal government" (Rodwin 2012). He summarised Mainland's proposal as follows:

Dr. Donald Mainland, who coordinated research for the American Rheumatism Association's Cooperating Clinics Committee, testified that it was current practice for the drug firm seeking marketing approval to act as an intermediary between the FDA staff and third-party testers. This arrangement allowed drug firms to influence the trials as well as the communication between testers and the FDA. Dr. Mainland suggested that Congress, "take the evaluation of drugs entirely out of the producer's hands," after the completion of toxicological testing on animals, in order to remove the possibility of the producer biasing the process. Dr. Mainland proposed the creation of an independent, not-for-profit drug-testing agency that would provide grants for research in a manner roughly analogous to the NIH. He suggested that a council of experts from universities and research institutions should invite senior investigators to form "working parties" for individual drugs. These "working parties" would then choose teams of suitable investigators to conduct the safety and efficacy studies. The agency should be funded largely by the pharmaceutical industry in a manner that did not allow it to "influence the disposal of the money or interfere in any way with the trials or the results."

It seems likely that his proposal was heavily influenced by bad experiences in the clinical trials he had been involved in, but there are no accounts of such problems among the Notes he published around that time (Mainland 1978).

After retirement

After Mainland retired from his post in New York University in the early 1970s, he and his wife Ruth moved to Kent, Connecticut. Thereafter he published very little and seemingly had little academic activity.

His last publications included a 1982 paper "*Medical statistics – thinking vs arithmetic*", in which he revisited a theme he had long pursued (Mainland 1982), and two personal views in the BMJ in 1984 (Mainland 1984a; Mainland 1984b). His last, posthumous publication was "Some statistical thoughts on neurobehavioral testing", in which he addressed rank methods, multiple regression, and composite indices (Mainland 1985).

Donald Mainland died in July 1985. I have not found any obituary. Harry Marks, medical historian, found no evidence that any of Mainland's papers had been archived (personal communication, 2 February 2009).

Final words

The last of Mainland's *Notes on Biometry in Medical Research* was titled "Envoi, with a vacillating hope" (Mainland 1978). Here he observed that:

The Notes have been a continuation of the attempt, begun in the "Notes from a Laboratory of Medical Statistics", to distinguish between the sense and the nonsense that statistical methods offer to medicine. In preparing the later notes, I have, I think, introduced more unsolved problems, questions without answers, than in earlier ones, in the hope that they will stimulate other research workers to continue the effort."

The closing sentences of this final Note reiterated an important idea from several earlier writings, and appears again in Mainland (1982) – the need for medical statisticians to be embedded in real world research and not (too) mathematical:

What disturbed me, even more, was the suggestion that theory was superior to application, because that is the attitude that will be met by research workers who are trying to think clearly about the subject. They will need the help of professional statisticians well-versed in

theory; and all that I can say is: "Try to get hold of someone who, like some of the professional statisticians mentioned in certain of the Notes, are close to thinking in real world research, and are willing to face frankly the defects of statisticians' efforts. Good luck in your search!" (Mainland 1978) (p676)

Donald Mainland's career was remarkable, with a unique move from anatomy to medical statistics and clinical trials. His background in anatomy was fundamental in shaping his ideas about the place of statistics in clinical research. He exerted a considerable influence on those who came into contact with his writings, either published and privately circulated, or who were lucky enough to have his direct input to their projects.

As reported in the congressional hearings, Mainland had asked Stanley Schor, a medical statistician, to review his published report of a randomised trial (Mainland and Sutcliffe 1962). Schor commented:

"It is, I think, the type of analysis that should be kept as a reference by every clinic investigator. Many times I am asked, are there any studies that have been published that you think are really good in terms of drug trials. Of course, every drug trial has its own bundle of problems, and it is not very useful to set up a step-by-step procedure which is to be used by all drugs in all cases. However, the analysis in this paper brings out so many important points which the usual clinical investigator is either not aware of or simply does not take into consideration that it should be required reading for all the people engaged in clinical trials." (U.S. Government 1968) (p3348)

Donald Mainland could not have been as influential as he was had he not had both the ability to think clearly and also to write clearly. His writing was clear, he eschewed jargon, and he wrote really well for his target readership. His contributions should be remembered. His approach is one to emulate.

Acknowledgements

I thank Ned Glick for helpful discussions and for making available the photographs of Donald Mainland taken by his father, Frederick P Glick. I thank David Colquhoun for helpful comments.



References

- Altman DG, Simera I (2015). A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. JLL Bulletin: Commentaries on the history of treatment evaluation (<http://www.jameslindlibrary.org/articles/a-history-of-the-evolution-of-guidelines-for-reporting-medical-research-the-long-road-to-the-equator-network/>)
- Bland JM, Altman DG (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* i:307-310.
- Bolton S (1998). Independence and statistical inference in clinical trial designs: a tutorial review *J Clin Pharmacol* 38:408-412.
- Chalmers I (2005). Statistical theory was not the reason that randomization was used in the British Medical Research Council's clinical trial of streptomycin for pulmonary tuberculosis. In: Jorland G, Opinel A, Weisz G, eds. *Body Counts: Medical Quantification in Historical and Sociological Perspectives*. McGill-Queen's University Press.
- Colquhoun D (1971). *Lectures on Biostatistics*. Oxford: Clarendon Press.
- Cornfield J (1978). Randomization by group: a formal analysis. *American Journal of Epidemiology* 108:100-102.
- Feinstein AR (1970). Clinical biostatistics. 1. A new name – and some other changes of guard. *Clin Pharmacol Ther* 11:135-148.
- Klar N, Donner A (2004). The impact of E.F. Lindquist's text "Statistical Analysis in Educational Research" on cluster randomisation. JLL Bulletin: Commentaries on the history of treatment evaluation (<http://www.jameslindlibrary.org/articles/the-impact-of-ef-lindquists-1940-text-statistical-analysis-in-educational-research-on-cluster-randomization/>)
- Lassere M (2010). Are reporting standards of diagnostic test evaluation unrealistic? *J Rheumatol* 37:220-222.
- Lassere MN (2008). The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Meth Med Res* 17:303-340.
- Mainland D (1934). Chance and the blood count. *Can Med Assoc J* 30:656-658.
- Mainland D (1936). Problems of chance in clinical work. *Br Med J* 2:221-224.
- Mainland D (1938). *The treatment of clinical and laboratory data: an Introduction to statistical Ideas and methods for medical and dental workers*. Edinburgh: Oliver and Boyd.
- Mainland D (1948). Statistical methods in medical research. I. Qualitative statistics (enumeration data). *Can J Res* 26:1-166.
- Mainland D (1950). Statistics in clinical research: some general principles. *Ann NY Acad Sci* 52:922-930.
- Mainland D (1952). *Elementary medical statistics: the principles of quantitative medicine* Philadelphia: WB Saunders Co

- Mainland D (1954a). The modern method of clinical trial. In Steele JM, ed. *Methods in Medical Research*, Vol 6. Chicago: Year Book Inc, pp152-158.
- Mainland D (1954b). The planning of investigations. *Methods Med Res* 6:138-145.
- Mainland D (1954c). The rise of experimental statistics and the problems of a medical statistician. *Yale J Biol Med* 27:1-10.
- Mainland D (1954d). Statistics in medical research (12 chapters, 3 co-authored by L Herrera) In: Steele JM, ed. *Methods in medical research*. Vol. 6. Chicago: Year Book Publishers Inc, 121-214.
- Mainland D (1955a). An experimental statistician looks at anthropometry. *Ann NY Acad Sci* 63:474-483.
- Mainland D (1955b). Use of case records in the study of therapy and other features in chronic disease. I. Planning the survey. *Ann Rheum Dis* 14:337-352.
- Mainland D (1958a). Notes on the planning and evaluation of research, with examples from cardiovascular investigations. I. *Am Heart J* 55:644-655.
- Mainland D (1958b). Notes on the planning and evaluation of research, with examples from cardiovascular investigations. II. *Am Heart J* 55:824-837.
- Mainland D (1958c). Notes on the planning and evaluation of research, with examples from cardiovascular investigations. III. *Am Heart J* 55:838-850.
- Mainland D (1960a). The clinical trial—some difficulties and suggestions. *J Chronic Dis* 11:484-496.
- Mainland D (1960b). The use and misuse of statistics in medical publications. *Clin Pharmacol Ther* 1:411-422.
- Mainland D (1961). Experiences in the development of multiclinic trials. *J New Drugs* 1:197-205.
- Mainland D (1963a). *Elementary medical statistics: the principles of quantitative medicine* 2nd edn. Philadelphia: WB Saunders Co.
- Mainland D (1963b). The significance of “nonsignificance”. *Clin Pharmacol Ther* 4:580-586.
- Mainland D (1967a). The estimation of inflammatory activity in rheumatoid arthritis. Role of composite indices. *Arthritis Rheum* 10:71-77.
- Mainland D (1967b). Statistical ward rounds—1. *Clin Pharmacol Ther* 8:139-146.
- Mainland D (1969a). Is the difference statistically significant? An innocent request for quack treatment. *Clin Pharmacol Ther* 10:436-438.
- Mainland D (1969b). Statistical ward rounds— 16. *Clin Pharmacol Ther* 10:576-586.
- Mainland D (1978). Notes on biometry in medical research. (Edited by MC Miller III) [104 items in 13 Monographs, 1967 to 1970, published by Veterans Administration, Washington DC]. Ann Arbor, MI: Biometry Imprint Series Press. (Distributed by University Microfilms International).
- Mainland D (1979a). Notes from a laboratory of medical statistics. Volumes I and II. (Edited by MC Miller III) [A series of 145 mimeographed notes distributed by the author (1959 to 1966)]. Ann Arbor, MI: Biometry Imprint Series Press. (Distributed by University Microfilms International).
- Mainland D (1979b). Statistical ward rounds. (Edited by MC Miller III) [18 articles reprinted from *Clin Pharmacol Ther* 1967 to 1969]. Ann Arbor, MI: Biometry Imprint Series Press (Distributed by University Microfilms International).

- Mainland D (1980). Personal view. *BMJ* 280:1269.
- Mainland D (1982). Medical statistics-thinking vs arithmetic. *J Chronic Dis* 35:413-417.
- Mainland D (1984a). Statistical ritual in clinical journals: is there a cure?—I. *BMJ* 288:841-843.
- Mainland D (1984b). Statistical ritual in clinical journals: is there a cure?—II. *BMJ* 288:920-922.
- Mainland D (1985). Some statistical thoughts on neurobehavioral testing. *Obstetric Anesthesia Digest* 5:90-94.
- Mainland D, Herrera L (1956). The risk of biased selection in forward-going surveys with nonprofessional interviewers. *J Chronic Dis* 4:240-244.
- Mainland D, Stewart CB (1938). A comparison of percussion and radiography in locating the heart and superior mediastinal vessels. *Am Heart J* 15:515-527.
- Mainland D, Sutcliffe MI (1953). Statistical methods in medical research. II. Sample sizes required in experiments involving all-or-none responses. *Can J Med Sci* 31:406-416.
- Mainland D, Sutcliffe MI (1962). Hydroxychloroquine sulphate in rheumatoid arthritis, a six-month, double-blind trial. *Bull Rheum Dis* 13:287-290.
- Mainland D, Sutcliffe MI, O'Brien WM (1973). A controlled trial of gold salt therapy in rheumatoid arthritis. *Arthritis Rheum* 16:353-358.
- Rodwin MA (2012). Independent clinical trials to test drugs: the neglected reform. *Saint Louis University Journal of Health Law & Policy* 6:113-166.
- Schor S, Karten I (1966). Statistical evaluation of medical journal manuscripts. *JAMA* 195:1123-1128.
- U.S. Government. (1968) Competitive problems in the drug industry. Hearings before the Subcommittee on Monopoly of the Select Committee on Small Business. United States Senate Ninetieth Congress, first and second sessions. Present status of competition in the pharmaceutical industry. Part 7 Washington DC.

Appendix 1: Some observations made by Donald Mainland

“The way to a more adequate understanding and treatment of medical data would be opened up if all records, articles and even abstracts gave, besides averages, the numbers of observations and the variation, properly expressed . . .” (Mainland 1934).

“What is required in clinical work is not elaborate mathematical tests, but an understanding of the meaning of chance, and adequate precautions that the samples, however small, are unbiased” (Mainland 1936).

“Limitation of journal space and the expense of publishing numerous or elaborate tables usually prohibit full publication of data, except where these are few in number; and yet incompleteness of evidence is not merely a failure to satisfy a few highly critical readers. It not infrequently makes the data that are presented of little or no value” (Mainland 1938, p 283).

“Statistical ideas, to be effective, must enter at the very beginning, i.e., in the planning of an investigation. These ideas, however, lacking in so many published reports, are even less frequent in the unpublished efforts of clinicians to assess the value of their treatments. There must be

something wrong with so-called “scientific” medical education when a young physician says that he has obtained promising results by treating migraine with histamine and yet cannot understand why a professor of pharmacology should ask about controls” (Mainland 1950).

“To omit randomization because one cannot see clearly how bias could occur is like trusting that glassware in chemistry is clean because it does not look dirty” (Mainland 1954b).

“... it is better to refrain from a survey than to produce equivocal results which not only waste effort and money but bring epidemiologic methods into disrepute” (Mainland and Herrera 1956).

“The first step is to realize that the method of clinical trials is still in its infancy, that although the principles are simple, the art is extremely difficult, and that all of us clinicians and statisticians have a very great deal to learn about the art” (Mainland 1960a).

“The only way to learn something about the safety of our numerical findings is by more extensive exploration under other conditions, in other places, and at other times” (Mainland 1960b).

“Research workers’ widespread lack of understanding of the rationale of statistical techniques, and the frequent use of statistical tests as a substitute for thoughtful investigational design, meticulous work, and repetition of experiments, justify the antagonism to statistics exhibited by some experimenters. To one who has had personal experience of the way in which statistical thinking, as distinct from statistical arithmetic, can promote good investigation, this perversion of statistics is lamentable” (Mainland 1960b).

“Statistical tests are very dangerous drugs” (Mainland 1969a).

“Compulsion to apply a random frequency [significance] test may arise from ignorance or from tribal custom” (Mainland 1978, p 77)

“ ‘Double blind’ implies that no one on the physician’s side (himself, assistants, nurses, secretaries, and so forth) and no one on the patient’s side (himself, relatives, friends and so forth) – no one at all who can in any way influence patients’ attitudes, feelings, behaviour, measurements, replies to questions, etc. – shall have any knowledge of the compounds that individual patients are receiving” (Note 98 in *Notes from a laboratory of medical statistics*) (Mainland 1979a).

Appendix 2: Some comments made about Donald Mainland

“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” (Feinstein 1970).

“I have been greatly influenced by the writing of Professor Donald Mainland. *His Elementary medical statistics* (1963), which is much more concerned with statistical thinking than statistical arithmetic, should be read not only by every medical practitioner, but by everyone who has to interpret observations of any sort. If the influence of Professor Mainland’s wisdom were visible in this book, despite my greater concern with methods, I should be very happy” (Colquhoun 1971).

“Dr. Donald Mainland was a man of particular practicality as well as a noted statistician and physician. In his book, *Elementary Medical Statistics*, Dr. Mainland spends some time on this problem and the pitfalls associated with it. He says “... many research workers are rather vague about their sampling units and the need for their independence.” He gives an example of a comparison of two toothpastes (A and B), one given each to two boys. One boy has eight carious teeth and the other none. He concludes that “... this looks like an impressive difference, but we need no profound knowledge of dentistry or of statistics to realize that it provides no adequate evidence that the difference in toothpastes was responsible.” ... Mainland also discusses the oft-

quoted “caged animal” example, using a comparison of diets in pigs in which each treatment group is in a pen. Again, he says, there is no replication. There are only two observations, one on each diet; the pen is the sampling unit, as noted above” (Bolton 1998).

“The history of methods of diagnostic test evaluation not surprisingly includes authorities with an interest in rheumatology. One was Prof. Donald Mainland, a “plain language” medical statistician better known (by some) for the early work he undertook with the Cooperating Clinics Committee of the American Rheumatism Association in developing a composite index of disease activity in rheumatoid arthritis (RA). In *Elementary Medical Statistics* (which I recommend to all young clinician scientists) Prof. Mainland set out questions for use in planning investigations and in evaluating their reports” (Lassere 2010).