The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2023, Vol. 116(10) 343–350
DOI: 10.1177/01410768231203922

Trial analysis by treatment allocated or by treatment received? Origins of 'the intention-to-treat principle' to reduce allocation bias: Part 1

I Chalmers[1], R Matthews[2], P Glasziou[3], I Boutron[4] and P Armitage[5]
[1]Centre for Evidence-Based Medicine, University of Oxford, OX2 6HX, UK
[2]Department of Mathematics, Aston University, Birmingham, B4 7ET, UK
[3]Institute for Evidence Based Healthcare, Bond University, Queensland, QLD 4226, Australia
[4]Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Centre for Research in Epidemiology and Statistics (CRESS), F-75004 Paris, France
[5]2 Reading Road, Wallingford OX10 9DP, UK
Corresponding author: I Chalmers. Email: iain@chalmersoxford.net

As ye randomise, so shall ye analyse. (Stephen Senn[1])

Assembling fair treatment comparison groups

Good and bad effects of treatments are sometimes so dramatic that they can be recognised reliably using informal before-and-after treatment comparisons. Relief of pain after spinal injection of a powerful analgesic or allergic reactions to drugs are examples. Much more usually, important wanted and unwanted effects of treatments are not dramatic and cannot be reliably detected using informal comparisons. In this common situation, detecting real treatment effects entails comparing sufficiently numerous people who have received one of two or more alternative treatments, or by comparing contemporaneously those who have received a treatment with those who have not.

If the results of treatment comparisons are to be trustworthy, the comparisons need to be fair. In particular, the people in the different treatment comparison groups need to be alike (differing only by the play of chance), not only in respect of factors known to influence treatment outcomes (chronic illness, for example), but also in other, undocumented ways (genetic predisposition, for example), some of which may not even have been conceptualised.

Ways of creating fair treatment comparison groups began to be applied during the 19th century,[2] and explicit consideration of the methodological principles entailed began to appear during the first half of the 20th century.[3–8] Initially, alternate allocation was used to assign patients to one or other of the different treatments to be compared. By the middle of the 20th century, random allocation to comparison groups had begun to be used in preference to alternate allocation.[2] This was mainly because of the concern that knowledge of upcoming allocations (whether from an allocation schedule based on alternation or on an unconcealed random allocation schedule) could lead those recruiting participants for treatment comparisons to tamper with the schedule, thus undermining its purpose.[9] It may also have reflected an awareness that even strict alternation could not be relied upon to generate unbiased comparison groups (Mainland,[10] p. 268).

In the late 1940s, concealed random allocation was used to generate comparison groups in the British Medical Research Council's (MRC) celebrated multicentre controlled trial of streptomycin in patients with pulmonary tuberculosis.[11] In particular, the details of the allocation schedule were concealed from those entering patients into the trial (Armitage,[12] p. 16). This blinding was to prevent foreknowledge of upcoming allocations, thus reducing any temptation, conscious or unconscious, to undermine strict random allocation and so risk introducing selection bias.[2,13] As one senior contributor to the early history of randomised controlled trials later observed: 'Randomization was introduced to control selection biases, not for any esoteric statistical reason'.[14] Austin Bradford Hill, director of the MRC's Statistical Research Unit, explained that 'our rather elaborate technique of sealed envelopes has been developed merely to ensure that no bias creeps in during this allocation. It has no other magical virtues.' (Hill,[15] p. 170).

The 1948 report of the MRC's streptomycin trial was widely hailed as a landmark in the development

of unbiased treatment comparisons. The trial provided an exemplar for further MRC-controlled trials in an exceptional programme of randomised comparisons assessing a wide variety of interventions during the 1950s and 1960s.[16]

The key figure in these developments was the statistician Austin Bradford Hill. Hill had come to international attention since 1937 through successive editions of his popular textbook *Principles of Medical Statistics*,[5] and through his prominent articles about 'The Clinical Trial' published in the *British Medical Bulletin*[17] and the *New England Journal of Medicine*.[18] A very relevant but less well-known contemporary publication was *Elementary Medical Statistics: The Principles of Quantitative Medicine*.[10] Its author – Donald Mainland – was a British physician who had emigrated initially to Canada, and subsequently to the USA. Doug Altman celebrated this remarkable man in a biographical article commissioned for the James Lind Library.[19]

Maintaining fair treatment comparison groups

Strict random allocation of eligible participants to one of two or more comparison groups provides a foundation for making fair (unbiased) comparisons of the effects of treatments. This random allocation means that any differences in the characteristics of the participants in the comparison groups before treatment can reasonably be assumed to reflect the play of chance. This provides the basis for inferring treatment effects if pre-specified outcomes differ between groups after treatment.

Allocation to treatment comparison groups needs to take place as late as possible before the initiation of treatment. Statistician David Newell once worked with a surgeon who took this advice so literally that he had a silver coin sterilised along with the scalpels, waited until the patient's abdomen had been opened and the diagnosis confirmed before he tossed the coin right there in the operating theatre to decide to which comparison group each patient would be allocated![20]

Random allocation of trial participants to treatment comparison groups abolishes allocation (selection) bias if there is no differential exclusion or loss of participants from the groups (randomised cohorts) as they had been initially constituted. The longer the trial lasts, however, the greater the likelihood that such losses will occur and that they will compromise the unbiased balance achieved using random allocation.

Some of the reasons for biased loss of participants cannot be avoided (death, for example). Other sources of bias could and should be avoided (for example, by strenuous efforts to track down and include missing outcome data, even for participants who have decided to withdraw from further treatment or trial visits).

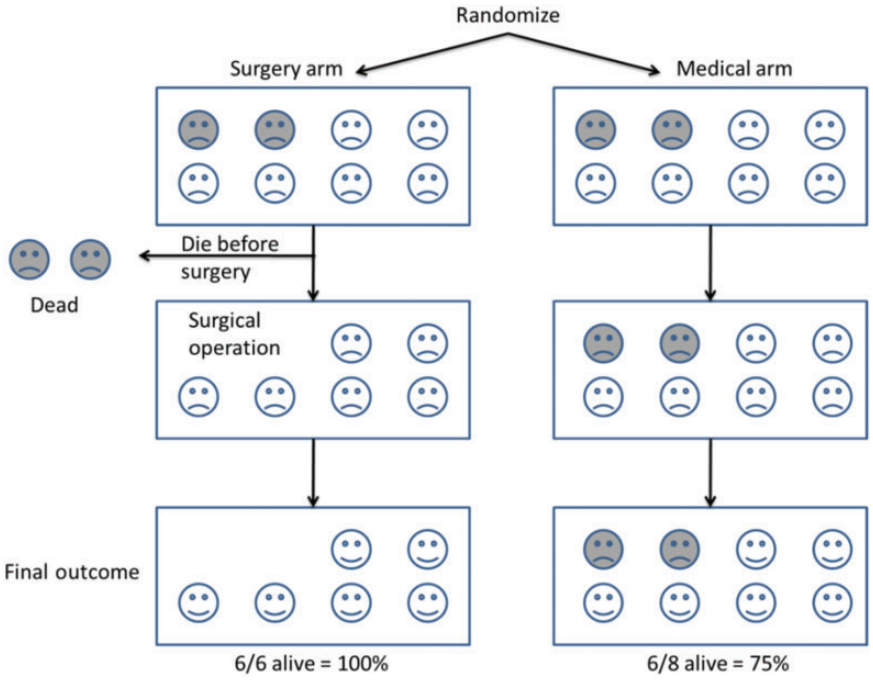
Figure 1 uses a randomised comparison of medical and surgical treatments in which some very ill patients randomised to surgery died before their surgery could be organised. Should these patients be excluded from analysis? As shown in Figure 1, doing that would result in an unfair bias against those allocated medical treatment from which similar very ill patients had not been excluded. To avoid bias, the challenge is to try to preserve the initial comparability of groups created using random allocation – application of a design and analysis strategy that came to be designated 'the intention-to-treat principle'.[21,22]

An early application of the principle was presented by Joseph Bell in his report of a controlled trial of whooping cough vaccine initiated in 1936 in Virginia, USA, with professional support from the Norfolk City Union of King's Daughters Visiting Nurse Association. Bell provides us with an early and remarkable example of the measures he took to establish and maintain fair treatment comparisons.[7] This entailed including all individuals allocated to one or other of the two comparison groups, whether or not they actually received the treatment assigned to them. This was an early example of what came to be recognised by others during the 1950s and which was designated 20 years later by Hill[22] as 'the intention-to-treat principle'.

Furberg[23] and Chalmers[24] have drawn attention to Bell's detailed account of the steps he took to maintain a fair comparison between children allocated to pertussis vaccine and those allocated to a control group. Here is an excerpt from Bell's[7] account:

> The first problem was that of locating for observation a group of children to be vaccinated, identical, in all attributes, which might influence the occurrence and recognition of pertussis, with another group to receive no vaccine. It is impossible to select such identical groups because many of the attributes involved are not known, and many of those that are known cannot be quantitatively assessed; and, furthermore, even if such attributes could be made identical in the two groups at any one moment, they would not remain identical throughout the time necessary for adequate observation. Some attributes without apparent influence on the results may under certain circumstances be of real importance.

Figure 1. If some patients are excluded from the final analysis, such as the two very ill patients who did not survive long enough to have the surgery to which they had been assigned, subsequently observed differences in characteristics may reflect the patients analysed, not differences in the treatments allocated to them (from Evans et al.[21] www.testingtreatments.org).



The only practical approach appeared to rest in the selection of two groups, each of which is a random sample of the combined groups in the exact sense of the term. Thus only can the prediction be made that should the vaccine have no real influence on the occurrence of pertussis, the occurrence in each group will approximate that of the combined group, deviating therefrom strictly within the range of chance sampling variation. On the other hand, if the vaccine confers real protection against the disease, or otherwise really influences its occurrence, the occurrence in each group will differ from that of the combined group outside the range of chance sampling variation. Obviously it is not practically possible to preselect two large strictly random groups of children who are representative of the general population and to ensure that every child in one group receives the vaccine while every child in the other group receives no vaccine during the observation period. Children in the general population have the prerogative to refuse vaccine offered and the liberty to obtain other vaccine when desired. In these premises there is no known way of changing the two groups so that one would include only children actually vaccinated, and the other include only children not vaccinated, without destroying the randomness of the selection and to that extent possibly invalidating the answer to the question asked. After it has been established that the vaccine confers protection, then questions concerning the amount and duration of such protection might in part demand direct comparison of the experience of the children actually vaccinated with those not vaccinated, providing adequate data are at hand to equalize the two groups with respect to attributes which apparently influence the occurrence of the disease.

For this report, the approach to the primary problem involved the preselection of two large strictly random groups of children and the subsequent injection of a large proportion of only one group with the vaccine. All analyses herein presented are a comparison of the experience of such preselected groups regardless of their actual status with respect to receiving the vaccine. The difficulties encountered in this approach are chronologically described in detail so that the reader may evaluate any possible errors involved. (pp. 1536–1537)

The methodological principle – trial planning and analysis by treatment allocated, but not necessarily

received – applied by Joseph Bell may have been applied in other contemporary clinical trials, but it seems that its first mention as a methodological principle in a textbook may have been in 1952, in the first edition of *Elementary Medical Statistics* authored by clinical epidemiologist Donald Mainland. In a section of the book entitled 'On Planning a Simple Experiment', a subsection entitled 'Intercurrent events' addressed the problem resulting from unforeseen events that may occur during any treatment. This may happen after random allocation and be known or suspected to influence the outcome of a clinical trial. Mainland[25] listed five events to illustrate the challenge (p. 109):

1. *The treatment under test may have to be supplemented for the good of the patient. In the investigation of the streptomycin treatment for tuberculosis, some patients had to receive lung collapse treatment by the introduction of air into the pleural or peritoneal cavity.*
2. *Complete change of treatment may be necessary in one or more patients.*
3. *A patient may incur an accident or disease, which may be (a) obviously associated with the original disease or with the treatment under test, (b) obviously not associated with either, or (c) have a possible association with them.*
4. *Treatment may be temporarily suspended owing to the patient's business or domestic affairs.*
5. *A patient may be lost, e.g., by removal from the city, by his failure to persevere with the treatment, or by death.*

> *In deciding what should be done with data from any such patients the criterion must always be whether their inclusion or omission would introduce bias. Unless the appropriate decision is obvious, the best plan is to analyse all the data together, then to analyse the special cases and the main series separately.* [our emphasis] An effort should be made to see if useful information can be obtained from the whole series up to the time that the special event occurred. (In the more complex methods of analysis of mensuration data there are ways of estimating the most likely values of missing items and then allowing for the defectiveness of such estimates). (Mainland,[25] p. 109)

This topic is one of the many matters on which Donald Mainland expanded in the second edition of his textbook *Elementary Medical Statistics*.[26] This included a chapter on 'Lost information', focusing on strategies rather than statistical procedures, and discusses how to minimise losses and how to analyse data when there are missing observations.

Publication of the first edition of Mainland's important book[19,25] coincided with the publication of two prominent articles by Hill entitled 'The Clinical Trial', one published in the *British Medical Bulletin*,[17] the other in the *New England Journal of Medicine*.[18] These two papers attracted a great deal of attention. Hill notes in the former that 'The aim is to allocate the patients to the 'treatment' and 'control' groups in such a way that the two groups are *initially* [our emphasis added] equivalent in all respects relevant to the enquiry'. However, neither of Hill's two articles explains how to address possibly biased departures from the comparison groups that had been generated using random allocation.

Hill's articles and Mainland's book ushered in what Shapiro and Shapiro[27] have dubbed 'an avalanche' of American and British books and symposia devoted to clinical trials'. Initially, this consisted of books and meetings focusing on statistical methods – for example, the fifth (1952) and sixth (1955) editions of Hill's *Principles of Medical Statistics*, and books authored by Leon Bernstein and Miles Weatherall[28] and Gustav Herdan.[29]

Other substantive documents focused almost entirely on preclinical research. Commenting on the report of a conference entitled *Experimental Methods for the Evaluation of Drugs in Various Disease States*,[30] Harry Gold[31] pointed out that among 21 papers covering nearly 300 pages of text, there was only one report of an experiment describing an attempted controlled evaluation of drugs in human disease. Gold did not specify the report to which he was referring, but having looked through all the conference papers, it seems very likely to have been the report by Stamler et al.[32] on the effects of oestrogen in atherosclerotic disease. It was exceptional among the research work presented at the conference in having been based on coordinated participation among seven US Veterans' Hospitals, and cooperative working with British colleagues in Edinburgh Royal Infirmary.

In the early 1950s, the wider-ranging contents of Mainland's book contrast with the focus in other books on statistical methodology rather than control of biases. However, in the late 1950s, things began to change. In 1958, a 1-day *Symposium on Clinical Trials* was held at the Royal Society of Medicine in London. It was attended by 38 people, all of them British, nine of whom presented papers. The symposium was chaired by Sir Charles Dodds and supported by the drug company Pfizer. The presentations covered the Aims and ethics of clinical trials; Controls; Criteria for measurement in acute diseases; Criteria for measurement in chronic diseases; Statistical aspects of clinical trials; Clinical

management; Presentation of results; and Financing of clinical trials. Importantly, the MRC statistician Ian Sutherland[33] summarised the challenge of maintaining fair treatment comparison groups generated by random allocation:

> In any trial, but particularly in the case of long-term treatment for a protracted disease such as tuberculosis, changes of the prescribed treatment may occur. Such changes may indicate a genuine failure of one of the treatments, perhaps due to a lack of clinical effect, or to excessive toxicity, which makes it essential for the clinician to depart from the protocol in the interests of the patient, but they may also reflect a lack of faith in one of the treatments, which may not really be justified. Substantial losses from the latter cause may disturb the similarity of the residual series of patients, and consequently bias the assessment, or even make it impossible to draw valid conclusions. The same applies to losses from observation, whether these are complete, the patient refusing to co-operate further, or partial, when necessary, observations on the progress of the patients have been missed. Both sources of bias are less potent if treatment has been blind. But the risk emphasises the general principle that, once allocated to treatment, every patient must be accounted for in the results, and changes of treatment or losses from observation kept to the unavoidable minimum. (p. 53)

These points were made the following year in one of 16 chapters in *Clinical Evaluation of New Drugs*, for which all the contributors were American.[34] Two of the contributors – statisticians Louis Lasagna and Paul Meier[35] – co-authored a chapter (pp. 37–60) addressing the difficulties resulting from loss of patients from initially randomised cohorts. Thus:

> Perhaps the most common difficulty is the loss of some patients from observation. Such losses may occur for many reasons – uncooperativeness, toxicity of the drug, death, etc. This situation is generally covered in articles by a remark such as the following: 'Seventeen patients failed to complete the course of treatment – fifteen on Regimen A and two on Regimen B. The analysis is restricted to the 113 patients who completed the treatment schedule'. The analysis then presented usually takes no account of the lost patients. Now it may be that this type of analysis is the most reasonable under the circumstances, but such a study is by no means equivalent to a study that began with 113 patients and had no losses. The fact of losses introduces a new source of bias, possibly great enough to vitiate the results completely. Worse still, the experimenter may sometimes be unable to tell if his results remain valid or not. For example, if significantly more subjects are lost from the group treated with Drug A than from the group treated with Drug B, one may suspect that Drug A is in some way more objectionable than Drug B and, since more of the likely-to-be-affected group has been selected out (lost) from the group on Drug A, the remaining parts of the two groups are not strictly comparable, and no amount of statistical manipulation will make them so . . . *There is no wholly satisfactory method for dealing with losses other than to avoid them*. [our emphasis] Thus, although complete freedom from losses is often an impossible goal, it is worth great effort and expense to keep the number of losses at the absolute minimum.' (p. 56)

At about the same time, a committee of the British Pharmacological Society convened a 2-day meeting in London to address aspects of *Quantitative Methods in Human Pharmacology and Therapeutics*.[36] Of the 64 people who attended, 13 (3 of whom were invited presenters) came from outside Britain (from USA, Sweden, Denmark, Netherlands and Germany). The meeting was supported by the Wellcome Trust, the Wellcome Foundation and the Ciba Foundation. Most of the contributions to the meeting were only indirectly relevant to the design and analysis of clinical trials, but one element of a presentation made by Richard Doll addressing *Practical Problems of Drug Trials in Clinical Practice* is relevant to the focus of the current article, that is, non-random losses from randomised cohorts[37]:

> Another point which may give rise to difficulty is what to do with the patient with whom, for one reason or another, it is impossible to complete the proposed treatment. It may be, for example, that the patient will die after the decision has been made to take him into the trial, but before the treatment has been properly begun. In this case, it is tempting to exclude him from the trial on the grounds that the trial drug has not had an opportunity to exert its effect. This is, however, quite unjustifiable as there would be no similar inclination to remove a similar patient from the control series. In general once a patient has been included in the trial, his fate must not be omitted from the results, unless it has been decided before the trial begins to exclude all patients dying within (say) the first 12 hr after the decision has been made to bring them into the trial – irrespective of which treatment they are in. A somewhat similar situation occurs if treatment has to be stopped because of side-effects. It might be argued that patients with coronary thrombosis who, in a trial of anticoagulants, could not proceed with the

> treatment should be excluded from the trial group. It is, however, not impossible that the factors which made them unduly susceptible to the drug would also affect the prognosis and the only proper thing to do is to stop the test drug, but to continue observations on them and retain them in the trial group. If this is not done and the number of patients withdrawn from treatment is appreciable, it becomes impossible to attach any meaning to the result of the trial. If they are retained in the trial group, the result of the trial group at least answers the question 'Is any benefit obtained if I try to treat all my patients with anticoagulants?' If the patient is lost sight of in a long continued trial, the position is more difficult. The question then arises whether there is any reason to suppose that the loss of the patient is related in any way with the result of the treatment. It is seldom possible to answer this question firmly in the negative and the best thing to do is to make various postulates about the reasons why the patients have stopped attending and to see whether any of them would necessitate altering the conclusions to be drawn from the trial. (Doll,[16] pp. 218–219)

In the same year, a 328-page book entitled *Medical Surveys and Clinical Trials* was published[38] and its 18 chapters cover a lot of ground. However, we have identified only one short passage relevant to our documentation of the development of thinking about application of the bias-reducing 'intent-to-treat principle'. The passage that follows was contributed by John Knowelden[39] (p. 129) and emphasises the importance of making exhaustive efforts to achieve as complete follow-up as possible to reduce bias to the greatest extent possible:

> Sometimes the period of observation after establishing the protected and unprotected groups is relatively short, as in the trial which showed a reduction in paralytic poliomyelitis two to eight weeks after giving gamma globulin, or it may extend for years, as in B.C.G. and pertussis-vaccine trials. Whatever the duration, it is vital that the same degree of observation is paid to both groups, and this is best achieved by regular visits or follow-up examinations. This was particularly important in the Medical Research Council's B.C.G. vaccine trial in school leavers where the control group among the negative reactors received no specific treatment. If subsequently a higher proportion of these had been lost sight of than of the vaccinated, it might have been difficult to say whether the defaulters were on average similar to the remaining controls who were observed; defaulting might have been the result of becoming infected or dying from tuberculosis, or on the other hand, because being perfectly fit, there seemed no point in returning for further examination. One of the strongest features of the B.C.G. vaccine trial was that by a combination of postal enquiries, visits by health visitors and re-examinations at mobile radiographic units 94 per cent of the 56,000 participants had been brought into contact with the teams within the first 18 months, and this proportion was virtually the same in vaccinated and unvaccinated children. (Knowelden,[39] p. 129)

Practical application of the key methodological principles emerging in meetings and in publications during the 1950s was manifested in the UK in 25 large controlled trials being reported between 1944 and 1960.[16] Statistician Sheila Bird, commenting on what had been achieved, has written 'The exposition of the British concept of the controlled clinical trial is astonishing for just how much had been got right within barely two decades'.[16]

Dedication

We are indebted to the late and much missed Tony Johnson (1943–2022) and his colleague Vern Farewell for creating an invaluable annotated bibliography of early textbooks and other publications on controlled clinical trials.

Author's note

This paper is the first part of a two-part series.

Declarations

Competing Interests: None declared.

Funding: None declared.

Ethics approval: Not applicable. Research for this paper did not entail any direct investigation of humans or animals.

Guarantor: IC.

Contributorship: All named authors were involved in conceptualising, drafting, and finalising the article for publication.

Acknowledgements: We thank Sheila Bird, Diana Elbourne, Klim McPherson, David Nunan, Andy Oxman, Stephen Senn, John Simes and Steven Woloshin for providing relevant information and commenting on earlier drafts of this article.

Provenance: Invited article from the James Lind Library.

References

1. Senn S. Added values. Controversies concerning randomization and additivity in clinical trials. *Stat Med* 2004; 23; 3729–3753.
2. Chalmers I, Dukan E, Podolsky SH and Davey Smith G. *The advent of fair treatment allocation schedules in clinical trials during the 19th and early 20th centuries*.

JLL Bulletin: Commentaries on the history of treatment evaluation. 2011. See www.jameslindlibrary.org/articles/the-advent-of-fair-treatment-allocation-schedules-in-clinical-trials-during-the-19th-and-early-20th-centuries/ (last checked 4 October 2023).
3. Sinton JA. Studies in malaria, with special reference to treatment. Part II. The effects of treatment on the prevention of relapse in infections with *Plasmodium falciparum*. *Indian J Med Res* 1926; 13: 579–601.
4. Martini P. *Methodenlehre der Therapeutischen Untersuchung [Methodological Principles for Therapeutic Investigations]*. Berlin: Springer, 1932.
5. Hill AB. *Principles of Medical Statistics*. London: Lancet, 1937.
6. Gaddum JH. Therapeutic trials on man. In: JH Gaddum, ed. *Pharmacology*. London: Oxford University Press, 1940: 379–383.
7. Bell JA. Pertussis prophylaxis with two doses of alum-precipitated vaccine. *Public Health Rep* 1941; 56: 1535–1546.
8. Medical Research Council. Clinical trial of patulin in the common cold. *Lancet* 1944; 2:373–375.
9. Chalmers I. *Why the 1948 MRC trial of streptomycin used treatment allocation based on random numbers*. JLL Bulletin: Commentaries on the history of treatment evaluation. 2010. See www.jameslindlibrary.org/articles/why-the-1948-mrc-trial-of-streptomycin-used-treatment-allocation-based-on-random-numbers/ (last checked 4 October 2023).
10. Mainland D. Intercurrent events. In: *Elementary Medical Statistics: The Principles of Quantitative Medicine*. London: WB Saunders Company, 1952: 109.
11. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *Br Med J* 1948; 2: 769–782.
12. Armitage P. The construction of comparable groups. In: AB Hill, ed. *Controlled Clinical Trials*. Oxford: Blackwell, 1960: 14–18.
13. Chalmers I. Statistical theory was not the reason that randomisation 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*. Montreal: McGill-Queens University Press, 2005: 309–334.
14. Doll R. The role of data monitoring committees. In: Duley L and B Farrell, eds. *Clinical Trials*. London: BMJ Books, 2002: 97–104.
15. Hill AB. *Controlled Clinical Trials*. Oxford: Blackwell Scientific Publications, 1960.
16. Chalmers I. *UK Medical Research Council and multicentre clinical trials: from a damning report to international recognition*. JLL Bulletin: Commentaries on the history of treatment evaluation. 2013. See www.jameslindlibrary.org/articles/uk-medical-research-council-and-multicentre-clinical-trials-from-a-damning-report-to-international-recognition/ (last checked 4 October 2023).
17. Hill AB. The clinical trial. *Br Med Bull* 1951; 7: 278–282.
18. Hill AB. The clinical trial. *N Engl J Med* 1952; 247: 113–119.
19. Altman DG. *Donald Mainland: anatomist, educator, thinker, medical statistician, trialist, rheumatologist*. JLL Bulletin: Commentaries on the history of treatment evaluation. 2017. See www.jameslindlibrary.org/articles/donald-mainland-anatomist-educator-thinker-medical-statistician-trialist-rheumatologist/ (last checked 4 October 2023).
20. Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol* 1992; 21: 837–841.
21. Evans I, Thornton H, Chalmers I and Glasziou P. *Testing Treatments*. 2nd edition. London: Pinter and Martin, 2011: 77.
22. Hill AB. *Principles of Medical Statistics*. 7th edition. Oxford: Oxford University Press, 1961: 258–259.
23. Furberg CD. *How should one analyse and interpret clinical trials in which patients don't take the treatments assigned to them*? JLL Bulletin: Commentaries on the history of treatment evaluation. 2009. See www.jameslindlibrary.org/articles/how-should-one-analyse-and-interpret-clinical-trials-in-which-patients-dont-take-the-treatments-assigned-to-them/ (last checked 4 October 2023).
24. Chalmers I. *Joseph Asbury Bell and the birth of randomized trials*. JLL Bulletin: Commentaries on the history of treatment evaluation. 2010. See https://www.jameslindlibrary.org/articles/joseph-asbury-bell-and-the-birth-of-randomized-trials/ (last checked 4 October 2023).
25. Mainland D. *Elementary Medical Statistics: the Principles of Quantitative Medicine*. India: Pranava Books, 1952.
26. Mainland D. Lost information. In: *Elementary Medical Statistics: The Principles of Quantitative Medicine*. London: WB Saunders Company, 1963: 178–186.
27. Shapiro AK and Shapiro E. *The Powerful Placebo: From Ancient Priest to Modern Physician*. Baltimore: Johns Hopkins University Press, 1997.
28. Bernstein L and Weatherall M. *Statistics for Medical and other Biological Students*. London: E & S Livingstone, 1952.
29. Herdan G. *Statistics of Therapeutic Trials*. London: Elsevier, 1955.
30. Whitelock OV, Furness FN and Craver BN, eds. Experimental methods for the evaluation of drugs in various disease states. *Ann N Y Acad Sci* 1956; 64: 463–731.
31. Gold H. Experiences in human pharmacology. In: DR Laurence, ed. *Quantitative Methods in Human Pharmacology and Therapeutics*. London: Pergamon Press, 1959: 40–54.
32. Stamler J, Pick R and Katz LN. Experiences in assessing estrogen antiatherogenesis in the chick, the rabbit, and man. *Ann NY Acad Sci* 1956; 64: 596–619.
33. Sutherland I. Statistical aspects of clinical trials. In: Dodds C, ed. *Report of a Symposium on Clinical Trials held at The Royal Society of Medicine*, London, 25 April, 1958:53; 50–55.

34. Waife SO and Shapiro AP. *The Clinical Evaluation of New Drugs*. New York: Hoeber-Harper, 1959.
35. Lasagna L and Meier P. Experimental design and statistical problems. In: Waife SO, AP Shapiro, eds. *The Clinical Evaluation of New Drugs*. New York: Hoeber, 1959: 37–60.
36. Laurence DR, ed. *Quantitative Methods in Human Pharmacology and Therapeutics*. London: Pergamon Press, 1959.
37. Doll R. Practical problems of drug trials in clinical practice. In: DR Laurence, ed. *Quantitative Methods in Human Pharmacology and Therapeutics*. London: Pergamon Press, 1959: 213–221.
38. Witts LJ, ed. *Medical Surveys and Clinical Trials*. London: Oxford University Press, 1959.
39. Knowelden J. Prophylactic trials. In: Witts LJ, ed. *Medical Surveys and Clinical Trials*. London: Oxford University Press, 1959: 118–133.