Marson Smith P, Colquhoun D, Chalmers I (2019). John Henry Gaddum's 1940 guidance on controlled clinical trials

© Iain Chalmers, Centre for Evidence-Based Medicine, Department of Primary Care, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG. Email:ichalmers@jameslindlibrary.org.

Cite as: Marson Smith P, Colquhoun D, Chalmers I (2019). John Henry Gaddum's 1940 guidance on controlled clinical trials JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/john-henry-gaddums-1940-guidance-on-controlled-clinical-trials/)

Introduction

The rationale for and design of controlled clinical trials as we know them today developed during the first half of the 20th century. The use of alternation to generate similar treatment comparison groups was adopted in several countries at the beginning of the century, notably in India and the United States (Chalmers et al. 2011), and the importance of concealing allocation schedules to prevent foreknowledge of allocations and biased inclusion or exclusion became recognised in the 1940s (Chalmers 2010). The need to reduce biased assessment of treatment outcomes by using blinded outcome assessment and placebos increased during the 1930s and 1940s (Kaptchuk 2011; Podolsky 2019), as did the application of statistical methods to assess the results of controlled trials (Matthews 2019).

Although there had been gradual adoption of these principles during the 1920s and 1930s – see, for example, Sinton (1926) in relation to the assessment of antimalarial treatments, and Bullowa (1928) on serum treatment of lobar pneumonia – the only book-length treatment of methodological issues in clinical trials of which we are aware was Martini's Methodenlehre der Therapeutischen Untersuchung [Methodological principles for therapeutic investigations], published in German in 1932 (Martini 1932; Stoll 2004).

Five years later, the British medical statistician Austin Bradford Hill published *Principles of Medical Statistics* (Hill 1937; Farewell and Johnson 2011). Although the concluding chapter of his book contains a 5-page section entitled *The Problems of Clinical Trials*, it was not until the early 1950s that Hill published methodologically-focused articles entitled *The Clinical Trial in the British Medical Bulletin* (Hill 1951) and the *New England Journal of Medicine* (Hill 1952); and it was not until the sixth edition of his book that he introduced a separate chapter entitled *Clinical Trials* (Hill 1955). From the mid-1950s onwards, books and symposia proceedings on clinical trials began to appear (Johnson 2019).

The chapter on 'Therapeutic Trials on Man' in Gaddum's Pharmacology

A contemporary of Hill's who remains insufficiently recognised by historians of clinical trials for his clear thinking about clinical trial design and analysis is John Henry Gaddum, a pharmacologist better known for his contributions to understanding the effects of 5-hydroxytryptamine (5-HT) and lysergic acid diethylamide (LSD) in the control of mood (Feldberg 1967). Gaddum's book *Pharmacology*, published in 1940, contains a remarkable chapter entitled 'Therapeutic Trials on Man'. Given the date of its publication, it deserves wider recognition than it has received. The present article has been written to promote that recognition by reproducing Gaddum's text verbatim.

At the time *Pharmacology* was published, Gaddum was professor of pharmacology in the University of London and based at the Pharmaceutical Society of Great Britain. His book set standards for the discipline, but also for the evidence needed to inform clinical practice. A few editorial notes have been added in square brackets. As he explains in the Preface:

> *This text-book of pharmacology is intended to be used by medical*

RELEVANT TOPICS

Fair tests of treatments
Placebo effects
Principles of Testing
Dramatic effects
Treatment comparisons are essential
Treatment comparisons must be fair
Cluster allocation
Allocation bias
Observer bias
The play of chance
Recording and interpreting numbers
Quantifying uncertainty

TOPICS

Fair tests of treatments
The need to address treatment uncertainties
Principles of Testing
Treatment comparisons are essential
Placebo effects
Dramatic effects
Treatment comparisons must be fair
Biases
Design bias
Allocation bias
N-of-1 crossover
Cluster allocation
Crossover tests
Factorial design
Co-intervention bias
Observer bias
Double dummy
Analysis bias

Biases in judging unanticipated possible effects
Reporting bias
Biases in systematic reviews
Researcher/sponsor bias and fraud
The play of chance
Recording and interpreting numbers
Quantifying uncertainty
Using meta-analysis
Bringing it all together for the benefit of patients and the public
Improving reports of research
Preparing and maintaining systematic reviews
Using the results of research

> *This text-book of pharmacology is intended to be used by medical students at a stage in their education before general principles become obscured by a mass of practical details, but it may also interest others. Facts with immediate practical applications receive especial emphasis, but some other facts are included, since one purpose of the book is to give an account of the experimental methods which have led, and are leading, to the introduction of so many therapeutic measures and to the use of so many potentially dangerous drugs. Medical men [*and women*] are constantly being asked by manufacturers and others to try the effect of new drugs on patients, and it is therefore important that they should know something of the kind of evidence that justifies the trial of new drugs. This book tries to give them this knowledge.*

The chapter entitled 'Therapeutic Trials on Man' is about 2000 words long, and has no subheadings. We have added these in bold in what follows, to indicate the extent to which Gaddum's 5-page coverage of the topic addresses issues seen as relevant today.

The need for controlled experiments of treatments in humans

> *Experiments on man are the only kind of experiment which can give certain evidence of therapeutic action on man. Such experiments are designed to answer the question whether the health of the patients who have taken the remedy is better or worse than it would have been if they had not taken the remedy. This question is not an easy one to answer since it is never possible to know for certain what would have happened without the remedy; patients may recover in spite of drugs or because of them. Any remedy that is used persistently is therefore bound to produce apparent cures fairly often unless it is very toxic.*
>
> *The wise physician can often form a shrewd opinion on the value of remedies, but opinions are not scientific unless the evidence on which they ae based can be written down on paper and survive criticism. The value of objective scientific evidence of this kind lies in the fact that it represents a permanent addition to the common stock of knowledge. Subjective opinions, based on the uncodified experience of the practising physician, are often the best guide to individual treatment, but they form an insecure basis for generalizations. They are influenced to an unknown extent by the subconscious wishes of the doctor, and their authority depends too much on his individual prestige. Objective records of facts have a more permanent value, though they may lead inexperienced persons to false conclusions. Subjective opinions are usually kept in the background of the evidence, but they cannot be entirely eliminated and it will never be possible to make research foolproof.*

The importance of using appropriate measures of treatment outcome

> *Objective scientific evidence in experimental therapeutics depends on something, like the temperature of the patient, that can be measured, something, like death or cure, that can be classified as an all-or-none response and counted, or something, like the contractions of the uterus, that can be recorded. The methods used in the study of remedies of different kinds have been discussed in various chapters of this book, and there is not much to add to these discussions.*
>
> *The proper choice of a method is fundamentally important, since the final argument usually involves the assumption that significant changes in the measured effect represent significant changes in the patient's health. The inexperienced experimenter may prove that some particular treatment increases the weight and wrongly assume that this represents an increase of health, when it is really due to obesity or oedema.*

Experimental design, comparisons, and the need for controls

> *The design of the experiment may make all the difference to the significance of the final result and must be carefully considered before the experiment is started. If the advice of a statistician is likely to be needed in the interpretation of the results, this advice is more likely if taken before the evidence is collected than afterwards; much time has been wasted on badly designed experiments.*
>
> *The evidence is usually based on the comparison of the health of treated patients with that of untreated patients, who act as a control and provide evidence of what the health of the treated patients would have been without the treatment. In some cases it is possible to make each patient*

serve as his own control by making observations before and after treatment. In any case, since the evidence is based on a comparison of the two groups of data, both groups are important, and a given number of observations gives the clearest results when half of them are made on treated patients and half of them are made on control patients.

Distinguishing causation from non-causal association

*The proper choice of controls may convert a vapid theory into a real contribution to knowledge; theories are cheap and ephemeral, but new facts are indestructible. If a patient has been in a steady state, or getting worse, during a preliminary control period and he takes a turn for the better comparatively quickly after the remedy is applied, there is some reason to believe in the remedy. The strength of the evidence depends on the rapidity of the cure compared with the duration of the control period. If a patient who has suffered from myxoedema for years, and has been getting gradually worse, is cured by thyroid in a few weeks, the cure may be considered rapid enough to be convincing, but the revival of a patient who has fainted cannot safely be attributed to any remedy unless it follows within a few seconds of the application of the remedy. Remedies for chronic diseases are more easily studied by this method than remedies for acute diseases, because the control period is usually longer. On the other hand, many chronic diseases have spontaneous remissions during which the patient is temporarily much better. In disseminated [*multiple*] sclerosis, for example, or schizophrenia or lymphadenoma or pernicious anaemia, the disease may practically disappear for months at a time. The study of cures for such diseases requires especial caution.*

The knotty problem of confounding

Evidence regarding the effect of a remedy may be vitiated by the simultaneous application of other remedies any of which may have produced the observed change. Many diseases are cured by rest in bed, and if a patient is put to bed and given medicine at the same time, no one knows whether his cure is due to the bed or to the medicine. For this reason patients are sometimes admitted to hospital and observed during several weeks of control period in bed, before the experimental treatment is applied. The periods before and after the treatment are then compared.

Another method of using control periods is to study the statistics of a disease before and after the introduction of a new remedy. When this is applied to a whole country the results are often difficult to interpret because there are so many factors affecting the result. An increase in the number of registered deaths due to a particular disease, for example, may be due to improved methods of diagnosis, and a decrease may be due to a spontaneous decrease in the virulence of the disease.

The statistics obtained in a single hospital, on the other hand, often provide better evidence. In all methods where control periods are used it is difficult to be quite certain that any observed differences between the patients receiving the experimental treatment and the controls are really due to the remedy being studied. Changes may occur in the cooking arrangements of the hospital or the skill of the nursing staff, or in many other factors which are unknown to the experimenter, but which happen to coincide in time with the start of the experimental period. The reliability of the conclusion that the change is not only post hoc but also propter hoc depends on the skill with which such factors are excluded. For this reason such evidence is never completely objective.

Random allocation to treatment comparison groups

If simultaneous controls are used, it is possible to eliminate the source of error discussed in the last paragraph by selecting the controls at random. If a continuous series of cases obtained from the same source is divided into two groups so that half of them receive the experimental treatment and the others serve as controls, it is sometimes possible to ensure that the only significant difference between the two groups lies in the presence or absence of the experimental treatment. The essential point is that the cases must be selected completely at random.

The probability that the observed difference between the two groups of data would occur by chance can then be calculated by statistical methods, and if this probability is very low the treatment must have had some effect. The use of random controls is almost foolproof.

Concealing allocation schedules to avoid allocation biases

If the cases arriving at a given clinic are assigned alternately to the

If the cases arriving at a given clinic are assigned alternately to the experimental group and the control group, the grouping can be regarded as random, provided that the decision to include each case in the whole series is made by someone who does not know whether the case will be a control or not. Otherwise there is a danger that he may tend to include mild or doubtful cases when he knows that they will be in one or other group, and so tend to produce the result which he subconsciously desires. Methods of randomization depending upon the tossing of pennies and other such mechanical devices are preferable.

Using dummy treatments to control for psychologically-mediated treatment effects

It is always necessary to consider the possible effects of suggestion, which may play a very important part in therapeutics. The patient who has faith in the treatment he receives is more likely to recover than the patient who has no faith, and inert substances may cause dramatic cures if used, in appropriate cases, with suitable suggestion. Cures due to suggestion are, of course, in no way less creditable than cures due to drugs, but unless suggestion is excluded from the experiments on the action of drugs, it is often impossible to know whether the result was due to the suggestion or the drug. Suggestion cannot affect the result if the patient does not know when the treatment starts or whether he belongs to the control group or the experimental group. For this reason it is often best that all the patients should appear to receive the same treatment all the time that they are under observation, but that the pills, mixtures, or injections administered to the control group should not contain the active ingredient whose effect is being studied.

Randomizing groups (clusters)

The actual carrying out of the experiment is often difficult. In studying the effect of milk on schoolchildren it is easy to give milk to all the children in one school and to keep another school as a control, but such controls are not random, since the relative health of the two schools may be affected by many factors besides the milk. On the other hand, if controls are really random, the experiment is likely to be complicated and to lead to jealousies.

The conclusions are, strictly speaking, only valid when the experiment is carried out exactly according to the design, without exceptions for special cases, and this is sometimes difficult to do.

The ethics of acquiescing in therapeutic uncertainties

It is almost impossible to apply the method of random controls to diseases with a high mortality, because the doctor does not feel justified in withholding a remedy which may possibly save lives from any of the patients, even when he does not know definitely whether it is effective or not. On the other hand, it is important to remember that new medicines are often toxic, and may make the disease worse. The maximum number of lives will probably be saved if the true facts are established as rapidly as possible.

Statistical analysis

The methods of determining the significance of the evidence are similar to the methods of calculating the result of a biological assay, but simpler, since quantitative results are seldom sought and definite evidence that the treatment is either a good thing or a bad thing is usually enough. Large numbers of uncritical observations are not necessarily more significant than a small number of carefully controlled experiments. Observations on a dozen treated patients and a dozen random controls may be much more convincing than a series of many thousands of cases with no controls, or even with control periods. If the data are in the form of some quantitative observation, made on each patient, such as the duration of his stay in

hospital, the mean value for the treated and control groups is calculated and the significance between the difference of these two means is estimated by the method given on page 361. If the results have been obtained by counting the number of patients cured, they can be expressed as percentages, and it is then sometimes obvious whether the difference between the two percentages is significant or not. If there is any doubt on this point, the question can be decided in the following way:

> *Let **a** be the number of treated patients cured, and **b** the number of treated patients not cured. Let **c** be the number of control patients cured and **d** the number of control patients not cured.*
> *Calculate χ^2 (chi-square) from the expression*

Calculate χ^2 (chi squared) from the expression

$$\frac{(ad-bc)2(a+b+c+d)}{(a+b)(b+c)(c+d)(d+a)}$$

The probability that random differences, not due to the remedy, would make χ^2 as large as 3.8 is 0.05, or in other words if χ^2 is greater than this, the odds are at least 19 to 1 that the remedy had some effect. If χ^2 = 6.6 the probability is 0.01 and the odds are 99 to 1. This formula is only accurate when the numbers involved are large; there is no simple way of working out the odds accurately when the numbers are small.

The content of almost all of Gaddum's 1940 chapter could have been written yesterday, but there are two respects in which he would probably been more cautious if he had been writing today. The section on random allocation and the last section on statistical analysis betray a common misinterpretation of tests of significance which has existed for over a century. In 1940, RA Fisher dominated statistical thinking, and Gaddum describes Fisher's approach to inference – null hypothesis testing that generates a p value. Gaddum describes the use of statistical methods to calculate "The probability that the observed difference between the two groups of data would occur by chance", but it is now universally recognised by statisticians that the p value is not the "probability that the results occurred by chance", though that remains a very common misapprehension to this day (for example, Gigerenzer et al. 2004; Wasserstein et al. 2016). Unfortunately the calculation of the probability that the results occurred by chance involves making some important assumptions. One suggested solution to this problem is to supplement p values with an estimate of the minimum false positive risk (Colquhoun 2017; 2018). If p equals 0.05 in a single experiment then, under plausible assumptions, the risk that a 'significant' result is a false positive is at least 20–30%. That is one reason why the term 'statistically significant' has given rise to an alarming number of wrong conclusions about the effects of treatments (Matthews 2019).

The other matter on which Gaddum might have used more cautious wording if he had been writing today is his reference to the possible effects of suggestion, which, he asserted, 'may play a very important part in therapeutics'. This view anticipated references in the 1950s to 'the power of the placebo' made by Beecher (1955) and others. Obtaining unbiased estimates of the effects of suggestion (placebo effects) requires comparisons between patients who have been randomly allocated to take placebos and other patients who have been allocated to a 'no-treatment' control group. Without such a comparison, the effect of a placebo intervention cannot be distinguished from the natural course of the disease, and other factors, for example regression to the mean – the tendency for extreme measurements to be closer to the mean (average) when repeated. A systematic review of such comparisons by Hróbjartsson and Gøtzsche (2010) did not suggest that, in general, placebo interventions had important effects, but that they can influence patient-reported outcomes in certain settings and on some outcomes, especially pain and nausea. Even among trials at low risk of bias, however, the estimated effect on pain varied from negligible to important. Debate continues on how to distinguish patient-reported effects of placebo from biased reporting, regression to the mean and other factors.

Who influenced Gaddum's thinking about controlled clinical trials?

Gaddum originally trained as a physiologist, initially at the Wellcome Research Laboratories under JW Trevan (Gaddum wrote his obituary for the Royal Society), then with Henry Dale at the National Institute for Health Research in Hampstead. He accepted chairs in pharmacology from 1934, initially in Cairo, then at University College London. The title page of Pharmacology identifies Gaddum as Professor of Pharmacology in the University of London.

Who influenced Gaddum's thinking about controlled clinical trials? Deeper research than we have done would be needed to understand how his thoughts about controlled clinical trials had been formulated. In the Preface to his book, Gaddum acknowledged comments on the manuscript from Professor GR Cameron and Doctors G. Brownlee, GAH Buttle, KH Coward, R Wien, GH Faulkner, and HR Ing. Bradford Hill's name is notable by its absence, but it seems very likely that he and Gaddum would have been discussing matters of mutual interest in the late 1930s. For example, Hill had written in 1937 that "mistakes, which when pointed out look extremely foolish, are quite frequently made by intelligent persons" (Hill 1937); and three years later, Gaddum suggested that "opinions are not scientific unless the evidence on which they are based can be written down on paper and survive criticism" (Gaddum 1940).

In a paper published two years after Pharmacology (Gaddum 1942), Gaddum pays tribute to Arthur Cushny, who like him, had been a pharmacologist at University

College London before moving to the University of Edinburgh. The notion of controlled comparisons had been illustrated in 1905 by Cushny and Peebles (1905) in an evaluation of the effects of hyoscines on sleep patterns: "as a general rule a tablet was given on each alternate evening, and the duration of sleep and other features were noted and compared to those of the intervening control night on which no hypnotic was given" (Cushny and Peebles 1905). The paper reported their clinical trial in sufficient detail to have been awarded a special place in the history of statistics (Senn 2017) because Student had used the data in his famous paper 'The probable error of a mean' (Student 1908)'.

Nevertheless, Gaddum, using a pharmacological perspective, credited Cushny with having "set new standards of evidence and purged Materia Medica of a large part of the residues inherited from the Middle Ages" (Gaddum 1942). This was a key moment within the field of pharmacology, where the lack of evidence of a treatment being effective caused the treatment to be rejected, despite the fact that it had been used for thousands of years. Gaddum pointed out that although we might have lost some treatments which might have been effective within this purge, we would benefit in future from this new way of evaluating treatments (Gaddum 1942).

Gaddum's thinking is also likely to have been influenced by Alfred Joseph Clark, professor of pharmacology (Materia Medica) in Edinburgh from 1926 to 1942 (Gaddum 1941; Clark 1985). In an article published in 1942, Gaddum drew attention to Clark's "revolutionary ideas of wishing his medical students to understand intelligently the action and fate of a quite small number of basic drugs rather than 'Materia Medica' drudgery of the old school". Gaddum endorsed Clark's proposal that pharmacology "must teach not only the results of pharmacological research, but also the methods by which new discoveries are actually made" (Gaddum 1942).

Later involvement in the adoption of controlled clinical trials

In the light of Gaddum's demonstrably sophisticated understanding of how treatments should be tested, it is unsurprising that Gaddum became a member of the Medical Research Council's committee overseeing its iconic randomised trial of streptomycin for pulmonary tuberculosis (MRC 1948), alongside Hill and Philip D'Arcy Hart (who had been a fellow student at Cambridge). Gaddum later reflected that "Without [the streptomycin trial] the value of these important drugs would still be uncertain, and it is even possible that advertisements for one of them might have persuaded that the other two were comparatively worthless" (Gaddum 1959).

Gaddum worried that published reports of therapeutic research were not affecting how patients were being treated. When reviewing the *British Pharmacopoeia* in 1953, for instance, he mentioned that "some drugs, such as amidopyrine and diamorphine and sulphonal, are now thought to be too dangerous for general use, though some doctors still believe in them." (Gaddum 1953). The following year, Gaddum revisited the topic of controlled clinical trials in his Walter Ernest Dixon Memorial Lecture at the Royal Society of Medicine, addressing in more detail some of the issues he had touched on in 1940 (Gaddum 1954). He re-emphasised the basic principle that "in order to convince the rest of the world it is often necessary to make observations of some kind, not only on the patients who receive the new treatment, but also on a control group who do not" (Gaddum 1954).

Gaddum concluded his lecture by reiterating the same general messages as those he had conveyed in his 1940 chapter on *Therapeutic Trials on Man*:

> *Many factors have contributed to the very rapid advance in therapeutics which has taken place in recent years. Fundamental work in physiology, pharmacology, biochemistry, pathology and bacteriology has increased our knowledge of nature and shown the way to new advances. The pharmaceutical industry has provided us with many new therapeutic tools, but new tools are not much use to those who cannot learn how to use them. Progress would have been less rapid if there had not been parallel advances in the technique of the clinical trial.*
>
> *The examples I have given are from a large number of researches in this field. They were mostly carried out with very simple apparatus, or with no apparatus at all, and illustrate how much can be done with simple equipment, provided that certain general principles are recognised. In all these experiments simultaneous controls are preferable to control periods; errors of allocation must be avoided and randomization achieved with certainty; errors of assessment can best be avoided by the use of the double blind technique, where neither the doctor nor the patients knows which patients receive the dummy treatment; when this is not possible the same object can sometimes be achieved by a compromise, such as that reached with the experiments on tuberculosis where the assessment was*

made by a second doctor who was not responsible for the care of the patients. If these precautions are taken, the subjective opinions of a group of patients can be interpreted with mathematical precision. All these things require careful planning and doctors who are not themselves statisticians should consult a professional statistician before they start their experiment.

Gaddum's chapter on 'Therapeutic Trials in Man' was repeated in the second (1944), third (1948), fourth (1953), and fifth (1959) editions of his book, but it does not appear in the sixth (1968) edition, which appeared under new editors after Gaddum's death.

By the late 1950s, promulgation of these principles in articles and books in English, French and other languages had started to increase (Johnson 2019). John Henry Gaddum was exceptional in having drawn attention to most of these principles as early as 1940.

Acknowledgements
The James Lind Library editors are grateful to Tony Johnson for consulting subsequent editions of Gaddum's Pharmacology and notifying us of their contents.

This James Lind Library article has been republished in the *Journal of the Royal Society of Medicine* 2019;112:394-400. Print PDF

References

Beecher HK (1955). The powerful placebo. JAMA 159:1602-6.

Bullowa JGM (1928). The control. Contribution to a symposium on the use of antipneumococcic refined serum in lobar pneumonia, 15 December 1927. Bulletin of the New York Academy of Sciences 4:339-343.

Chalmers I (2010). Why the 1948 MRC trial of streptomycin used treatment allocation based on random numbers. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/why-the-1948-mrc-trial-of-streptomycin-used-treatment-allocation-based-on-random-numbers/)

Chalmers I, Dukan E, Podolsky SH, Davey Smith G (2011). 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 (https://www.jameslindlibrary.org/articles/the-advent-of-fair-treatment-allocation-schedules-in-clinical-trials-during-the-19th-and-early-20th-centuries/)

Clark DH (1985). Alfred Joseph Clark, 1885-1941. A Memoir, Obituary Notices of Fellows of the Royal Society. Royal Society, pp. 969–984. doi: 10.2307/769190.

Colquhoun D (2017). The reproducibility of research and the misinterpretation of P values. Royal Society Open Science, 4, 171085, DOI: 10.1098/rsos.171085. Available http://rsos.royalsocietypublishing.org/content/4/12/171085

Colquhoun D (2018). The false positive risk: a proposal concerning what to do about p-values (version 2). https://www.youtube.com/watch?v=jZWgijUnIxI

Cushny AR, Peebles AR (1905). The action of optical isomers: II. Hyoscines'. Journal of Physiology 32(5–6), pp. 501–510. doi: 10.1113/jphysiol.1905.sp001097.

Farewell V, Johnson A (2011). The origins of Austin Bradford Hill's classic textbook of medical statistics. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/the-origins-of-austin-bradford-hills-classic-textbook-of-medical-statistics/)

Feldberg W (1967). John Henry Gaddum 1900-1965, Biographical Memoirs of Fellows of the Royal Society 13:56–77.

Gaddum JH (1940). Pharmacology. London: Oxford University Press, p 378-383.
Gaddum JH (1941). Prof. A. J. Clark, F.R.S, Nature 148:189-90.

Gaddum JH (1942). The development of *Materia Medica* in Edinburgh. Edinburgh Medical Journal 49:721–736. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5305916

Gaddum JH (1953). British Pharmacopoeia. Nature 171:990.

Gaddum JH (1954). Clinical pharmacology. Proceedings of the Royal Society of Medicine 47:195–204. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1918604/

Gaddum JH (1959). Quantitative methods in human pharmacology and therapeutics. In: Laurence DR, ed. Proceedings of a symposium held in London on 24th and 25th March 1958. London: Pergamon Press, Series (Biological Council. Co-ordinating Committee for Symposia on Drug Action); v. 3), pp. 3–10.

Gigerenzer G, Krauss S, Vitouch O (2004). The Null Ritual. What you always wanted to know about significance testing but were afraid to ask. In Kaplan D (ed.). The Sage Handbook of Quantitative Methodology for the Social Sciences. Thousand Oaks, CA: Sage Publishing. Available at https://library.mpib-berlin.mpg.de/ft/gg/GG_Null_2004.pdf

Hill AB (1937). Principles of Medical Statistics. London: Lancet. doi: 10.1016/S0140-6736(00)82801-9.

Hill AB (1951). The clinical trial. Brit. Med. Bull. 7:278–282.

Hill AB (1952). The clinical trial. New England Journal of Medicine 247:113-119.

Hill AB (1955). Principles of Medical Statistics, 6th edition. London: Lancet

Hróbjartsson A, Gøtzsche PC (2010). Placebo interventions for all clinical conditions. Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD003974. DOI: 10.1002/14651858.CD003974.pub3.

Johnson A (2019). Textbooks and other publications on controlled clinical trials, 1948 to 1983. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/textbooks-and-other-publications-on-controlled-clinical-trials-1948-to-1983/)

Kaptchuk TJ (2011). A brief history of the evolution of methods to control observer biases in tests of treatments. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/a-brief-history-of-the-evolution-of-methods-to-control-of-observer-biases-in-tests-of-treatments/)

Martini P (1932). Methodenlehre der Therapeutischen Untersuchung [Methodological principles for therapeutic investigations]. Berlin: Springer.

Matthews R (2019, in preparation for James Lind Library).

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 2:769-782.

Podolsky S (2019, in preparation for James Lind Library)

Senn SJ (2017). Cushny and Peebles, optical isomers, and the birth of modern statistics. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/cushny-and-peebles-and-optical-isomers/)

Sinton JA (1926). Studies in malaria, with special reference to treatment. Part I. Introduction and routine methods. Indian Journal of Medical Research 13:565-577.

Stoll S (2004). Paul Martini's methodology of therapeutic investigation. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/paul-martinis-methodology-of-therapeutic-investigation/)

Student (1908). The probable error of a mean. Biometrika 6:1-25.

Wasserstein RL, Lazar NA (2017). The ASA's statement on p-Values: context, process, and purpose. American Statistician 70:129-133. Available at https://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108

Building the Library
Using the Library

COPYRIGHT

Where not otherwise indicated, material in the James Lind Library is licensed under a Creative Commons Attribution 4.0 International License.

FEEDBACK

Contact us

We welcome your suggestions for the design or content of the Library

This website was created by the James Lind Initiative, the Royal College of Physicians of Edinburgh and Minervation Ltd