

independent inquiry at Glasgow, however, the "serum" cases were treated in the Royal Infirmary, and a series of patients of the same social stratum", admitted during the same period to the Belvedere Isolation Hospital under the care of one physician, served as the control group. It is clear that there may be serious fallacies in any system which contrasts a group of serum-treated patients with a control group drawn from a different stratum of the population, or with a control group in a previous year, when the severity of the prevailing pneumonia might have been different."

Later paragraphs in this section of the paper are also important. The second paragraph:

- (i) addresses eligibility criteria;
- (ii) describes measures to exclude, before allocation, patients deemed unlikely to benefit from serum;
- (iii) describes the exclusion from the analysis of patients who died less than 24 hours after allocation;
- (iv) expresses concern that the sample may have been too small "for statistical purposes".

The second paragraph reads as follows:

"Certain principles of selection were laid down so as to make the data derived from the centres homogeneous, and to exclude from the comparison patients in whom the serum could not be expected to have any effect. For the latter reason all patients admitted later than the fifth day of illness were excluded from the inquiry. Also all patients dying within twenty-four hours of admission to hospital were taken out of the series, though the evident severity of their illness would not have prevented their inclusion at first, either in the control or in the serum group. No case of pneumonia complicated by other obvious disease, such as gross nephritis, advanced heart disease, diabetes, etc., was accepted for either group. All forms diagnosed as bronchopneumonia were also excluded. That these limitations were desirable was agreed upon by all the workers at a preliminary conference on the subject. It will be appreciated, however, that, with such restrictions, it was difficult in three years to obtain fully adequate data for statistical purposes."

The final paragraph in this section of the paper is also important in the methodological history of clinical trials. It begins by noting the higher case-fatality rate in older patients, and (i) describes steps taken to reduce the likelihood of chance imbalances in the numbers of older patients between the serum and control groups, and (ii) ends by noting the importance of large numbers for reducing chance imbalances in important prognostic factors. The paragraph reads as follows:

"Sex was disregarded, but the question of age was too important to be neglected. Table II from the present series illustrates afresh the well-known fact that the fatality of lobar pneumonia tends to be much greater over the age of 40 than in younger persons. The fortuitous inclusion of a few more elderly patients in one group than the other might influence unfairly the final figures for comparison. It was therefore decided to omit from the series all patients under the age of 20 and over the age of 60, and to classify the remainder into broad age groups. It will be noted that this plan still left altogether unregulated the chance scatter of distribution of patients with severe or mild pneumonia into either the serum or the control groups, and also of those for treatment early or relatively late in the progress of the disease. It was thought better not to attempt a deliberate sorting of cases in respect of mildness or severity, but to trust that the distortion of chance scatter would become almost negligible in a fairly large number of cases. Reference to a possible influence of the "severity factor" on the results is, however, made later in the report."

It is not clear what is implied in the above paragraph by the reference to classifying eligible patients "into broad age groups", but the next section of the report - entitled 'Statistics of Results' - suggests that it did not involve alternate allocation within strata defined by age. This section also draws attention to the fact that only in Aberdeen and London was there strict adherence to an allocation schedule based on alternation:

"Subject to the criteria mentioned above, patients at London and Aberdeen were placed in the groups for serum treatment, or for control, alternately in the order of their admission to hospital without selection as to age or severity. At Edinburgh the same general rules and criteria were observed, and there was no selection of cases for serum treatment. But in some wards of the General Infirmary serum was not used throughout the whole period of the inquiry, and consequently the patients from these wards overload the number of controls. In the other wards the alternate case

plan was maintained to the end. At Glasgow the alternate case plan was not used, but patients in one hospital were treated with serum and those in another hospital served as controls. Hence it is only at Aberdeen and London that the serum treated cases equal the control cases in number."

The analysis of the results involved comparing observed and expected numbers of deaths - the latter being defined as "those which would have been recorded if the serum treated groups had died at the same percentage rates as the corresponding controls". Observed and expected numbers of deaths were compared in strata defined by centre, age of patients, and type of pneumonia. Inspection of these many subgroup analyses suggested variations in effects. The report notes, however: "This raised the question of the chance scatter of patients with poorer prognosis from any cause into either the serum or the control group preponderantly", and goes on to conclude:

"The variation in results at the different centres cannot be explained, but they show the difficulties in the way of accurately evaluating a treatment of this nature on the basis of small numbers of cases."

In brief, this report reveals a clear appreciation of (i) possible sources of allocation bias, and ways to reduce it; (ii) concern about the danger of being misled by the play of chance.

Who was responsible for drafting these methodologically sophisticated passages in the paper? Papers available at the Public Record Office provide some clues (FD1/2372. Serum treatment of pneumonia). The four centres (Aberdeen, Edinburgh, Glasgow and London) submitted data to staff at the Medical Research Council, who forwarded them to Professor TR Elliott, director of the Medical Unit at University College Hospital in London, who chaired a conference to discuss the results at the different centres on 10 November 1933. Professor Elliott appears to have been asked to preparing a report for publication, and this would have had to take into account the variety of interpretations of the data expressed by the investigators. For example, after the conference Dr John Cowan of Glasgow (where there had been 17 deaths compared with 23 expected) expressed a clear view in a letter written to the Secretary of the Medical Research Council:

"On the facts available in U.S.A. and at home serum seems to me to be proved to be beneficial in I and probably proven in II. It should be available in consequence in ALL hospitals. Why have so many folk - in London and here too - fought shy of it? Why are not Barts etc all using it? The days of controls are no longer possible: it is not fair to them." (John Cowan to FHK Green, 17 November 1933).

By contrast, a letter from Stanley Davidson, Professor of Medicine at Aberdeen, the centre which had the most striking result, to Dr A Landsborough Thompson at the MRC revealed the more cautious interpretation subsequently shown in the published report:

"Nine cases of Type I pneumonia in the control series died, and only one in the treated. We are unable to explain these excellent results except: (1) on the grounds of the beneficial effects of serum, and (2) that these results are exaggerated by chance, owing to the small number of cases involved." (LSP Davidson to A Landsborough Thompson, 24 November 1933)

FHK Green responded:

"I have told Professor Elliott that the evidence for the 'miracle of Aberdeen' appears to be unassailable and he has replied that, this being so, it must clearly be a case of 'go to Peebles for pleasure: go to Aberdeen if you get pneumonia'" (FHK Green to LSP Davidson, 28 November 1933)

There is no correspondence in the file to indicate how the findings were interpreted by the clinicians responsible for the study's London arm, where there were 11 deaths from Type I pneumonia in the serum group compared with only 6 expected.

The 'contrasting' results in the two centres (Aberdeen and London) which had apparently used alternate allocation throughout the recruitment phase must have been one of the challenges faced by those who drafted the report for publication. It seems very probable that Austin Bradford Hill had a hand in this. Two months before the trial was published in the BMJ, Bradford Hill wrote a detailed critique of it in an internal and unpublished report for the Medical Research Council (Bradford Hill 1933).

It is very frustrating that this report, which likely occupies a key place in the history of controlled trials, appears now to have been mislaid. The historian Joan Austoker was able to inspect it at MRC Headquarters during the 1980s (Austoker and Bryder 1989), and reported that Bradford Hill had questioned the methods used in allocating cases into serum and control groups and stressed that greater effort should be taken "that the division of cases really did ensure a random selection" (Bradford Hill 1933, quoted in Austoker and Bryder, 1989).

The limitations of this study - with its relatively small numbers and mixture of ways of generating control groups - are likely to have been very important in leading Bradford Hill to go on to design large trials using concealed allocation schedules. In an article published in 1988, Jan Vandenbroucke (1988) noted that the 1934 report contains "a beautiful discussion of selection and comparability of treatment groups and that this came before the publication of Fisher's *Design of Experiments*." I sent Bradford Hill a copy of Vandenbroucke's article, and he responded in a letter to me as follows (Hill 1988):

"Thank you for sending me the Dutch article on the history of the R.C.T. I am interested in his comment on the M.R.C. Therapeutic Committee's report on the serum treatment of lobar pneumonia which contains "a beautiful discussion of selection and comparability of treatment groups & that this came before the publication of Fisher's *Design of Experiments*." I feel certain that I wrote that para and I had learned from Pearson & Greenwood & Yule (vide the references No 21&22). I had applied that teaching to the M.R.C's trial of a vaccine against whooping cough and was itching to apply it in the clinical field. Streptomycin provided the opportunity. Of course later I may have been influenced by Fisher but not very much - in fact in his famous 'tea and milk' experiment I think he was wrong."

In discussing the planning and interpretation of experiments in the first (1937) edition of his book 'Principles of Medical Statistics', Bradford Hill states that the allocation of alternate cases to the treated and control groups "is often satisfactory" because "*in the long run* (emphasis in the original) we can fairly rely upon this *random allotment* (my emphasis) of the patients to equalise in the two groups the distribution of other characteristics that may be important." (Bradford Hill 1937, p 5) He goes on to outline how allocation can be done within strata defined by characteristics (such as age) known to have an influence on the results of treatment. In later editions of the book, Bradford Hill used data from the MRC trial of serum treatment for pneumonia to illustrate how this might be important. In the opening chapter of the 1946 edition of the book, for example, he tabulates data (which cannot be derived from the published report) revealing differences in the age distributions of patients in the serum and control groups in the two centres in the study (Aberdeen and London) where alternation was said to have been used throughout the period of recruitment (Bradford Hill 1946, pp 6-7). Bradford Hill's text implies that these differences reflect chance; but he may have suspected that they reflected biases introduced by failure to adhere strictly to the alternate allocation scheme. Indeed, in spite of his insistence that alternate allocation must be strictly applied, from the first edition of his book onwards, Bradford Hill does not comment on the fact that the totals of 159 and 163 patients in the serum and control groups are clearly incompatible with strict alternate allocation (Bradford Hill 1946, p 7).

Bradford Hill's missing 1933 critique of the MRC study would almost certainly shed more light on this speculation. Although some subsequent single centre trials supported by the MRC continued to use alternation during the 1930s (for example, Snodgrass and Anderson 1937; Anderson 1939), the steps subsequently taken to conceal allocation schedules from those recruiting participants in the MRC multicentre trials of whooping cough vaccine (Medical Research Council 1951) and streptomycin for pulmonary tuberculosis (Medical Research Council 1948) conducted after World War II seem likely to reflect Bradford Hill's concerns about uncontrolled allocation biases in the MRC study of serum treatment for pneumonia. Thus, this carefully reported but imperfectly conducted study early in the life of the MRC's Therapeutic Trials Committee seems likely to have played a key role in one of the most important methodological advances in the history of clinical trials (D'Arcy Hart 1999; Chalmers 1997; 2000; 2001).

Just as departures from strict alternation can introduce bias, so also can departures from schedules based on a list of random numbers or coin tosses. The methodological advance manifested in the MRC trials of whooping cough vaccine (Medical Research Council 1951) and streptomycin (Medical Research Council 1948) was that both trials were designed with a view to preventing those involved in allocating people to the comparison groups knowing or predicting correctly which allocation was next in line. It is because this concealment of allocations was more likely to be achieved using random allotment than with alternation that randomisation was adopted (D'Arcy Hart 1999; Chalmers 2001). As Richard Doll has observed, randomization was introduced "to control allocation biases, not for any esoteric statistical reason." (Doll 2000)

It seems likely that Bradford Hill drew some key methodological lessons from the Medical Research Council trial of serum treatment of lobar pneumonia, which thus played a key role in the evolution of the rigorous methods that he applied in described in the MRC trials of whooping cough vaccine and streptomycin conducted in the 1940s.

References

Anderson T (1939). Sulphanilamide in the treatment of measles. *BMJ* 1:716-718.

Austoker J, Bryder L (1989). The National Institute for medical research and related activities of the MRC. In: Austoker J, Bryder L, eds. *Historical perspectives on the role of the MRC*. Oxford: Oxford University Press, pp 35-57.

Bradford Hill A (1933). Medical Research Council 1487, VI: A. Serum treatment of pneumonia. 22 December 1933. Cited in: Austoker J, Bryder L. *The National Institute for medical research and related activities of the MRC*. In: Austoker J, Bryder L, eds. *Historical perspectives on the role of the MRC*. Oxford: Oxford University Press, 1989: 35-57.

Bradford Hill A (1937). Principles of medical statistics. London: Lancet, p 5.

Bradford Hill A (1946). Principles of medical statistics. London: Lancet, p 7.

Bradford Hill A (1988). Letter to Iain Chalmers, 7 August.

Bradford Hill A (1990). Memories of the British Streptomycin Trial in Tuberculosis. *Cont Clin Trials* 11:77- 79.

Chalmers I (1997). Assembling comparison groups to assess the effects of health care. *J Roy Soc Med* 90:379-386.

Chalmers I (1999). Why transition from alternation to randomisation in clinical trials was made. *BMJ* 319:1372.

Chalmers I (2001). Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. *International Journal of Epidemiology* 30:1170-78.

D'Arcy Hart P (1999). A change in scientific approach: from alternation to randomised allocation in clinical trials in the 1940s. *BMJ* 1999; 319: 572-573.

Doll R (2000). The role of data monitoring committees. In: Duley L, Farrell B, eds. *Clinical trials*. London: BMJ Books, p 97.

Green FHK (1933). Letter to Dr Richard Armstrong, 28 November 1933. Public Record Office, FD1/2372.

Medical Research Council (1934). The serum treatment of lobar pneumonia. *BMJ* 1:241-45.

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: *BMJ* 2:769-82.

Medical Research Council (1951). The prevention of whooping-cough by vaccination. *BMJ* 1:1463-71.

Snodgrass WR, Anderson T (1937). Prontosil in the treatment of erysipelas: a controlled series of 312 cases. *BMJ* 2:101-104.

Vandenbroucke JP (1987). A short note on the history of the randomized controlled trial. *J Chron Dis* 40:985-987.

[Key Passages](#)

[Whole Article](#)

[Commentary](#)

[Acknowledgements](#)

[Context](#)

[Back](#)

[Home](#)

[Search](#)

