

## Controlled trials: the 1948 watershed

Richard Doll

### Clinical trials before 1946

Important scientific advances seldom occur out of the blue but can be seen in retrospect to have been the culmination of processes that have built up over the years. This was certainly true of the introduction of the new method of conducting clinical trials, first reported in 1948, that has played such a major part in the progress of clinical medicine in the last half century. When I qualified in medicine in 1937, new treatments were almost always introduced on the grounds that in the hands of professor A or in the hands of a consultant at one of the leading teaching hospitals, the results in a small series of patients (seldom more than 50) had been superior to those recorded by professor B (or some other consultant) or by the same investigator previously. Under these conditions variability of outcome, chance, and the unconscious (leave alone the conscious) in the selection of patients brought about apparently important differences in the results obtained; consequently, there were many competing new treatments. The treatment of peptic ulcer was, perhaps, more susceptible to claims of benefit than most other chronic diseases; so that in 1948, when I began to investigate it, I was soon able to prepare a list of treatments beginning with each letter of the alphabet. Standard treatments, for their part, tended to be passed from one textbook to another without ever being adequately evaluated.

A few clinicians had realised that this was unsatisfactory and had called for the use of concurrent controls, usually suggesting that alternate patients should be treated by one or other of the two methods being compared. Fibiger is recorded as having used this method in a trial of sensitised serum for the treatment of diphtheria in 1898,<sup>1</sup> but use of the method spread slowly. By the 1930s it was used regularly by the Medical Research Council, presumably on the advice of its statistical committee which was under the chairmanship of Major Greenwood. It was used to evaluate the serum treatment of lobar pneumonia in 1934,<sup>2</sup> and by D'Arcy Hart in a trial of patulin for treatment of the common cold in 1944.<sup>3</sup> This was expanded by Wilson et al in 1946 to a factorial design to enable two comparisons to be made in the same group of patients. This necessitated four treatment groups in order to test both the value of a low fat diet and dietary supplementation with di-cysteine in the treatment of acute hepatitis.<sup>4</sup> Patulin, which had been discovered by a chemist who had failed to isolate penicillin, had no apparent effect on the common cold, and a low fat diet

### Summary points

Claims of therapeutic benefit are often misleading unless there are concurrent control groups

When treatments are allocated to alternate patients there is a risk of bias in the selection of patients

Randomly allocating individuals after entry into a trial eliminates bias and provides a proper estimate of random error

Randomisation can be done within strata of the likely response to treatment, if clinicians can define the strata sufficiently clearly

Ethical standards should be the same for therapeutic trials and routine care

Large numbers of participants may be needed in a trial if moderate yet important effects are to be detected

Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE  
Richard Doll,  
*honorary member*

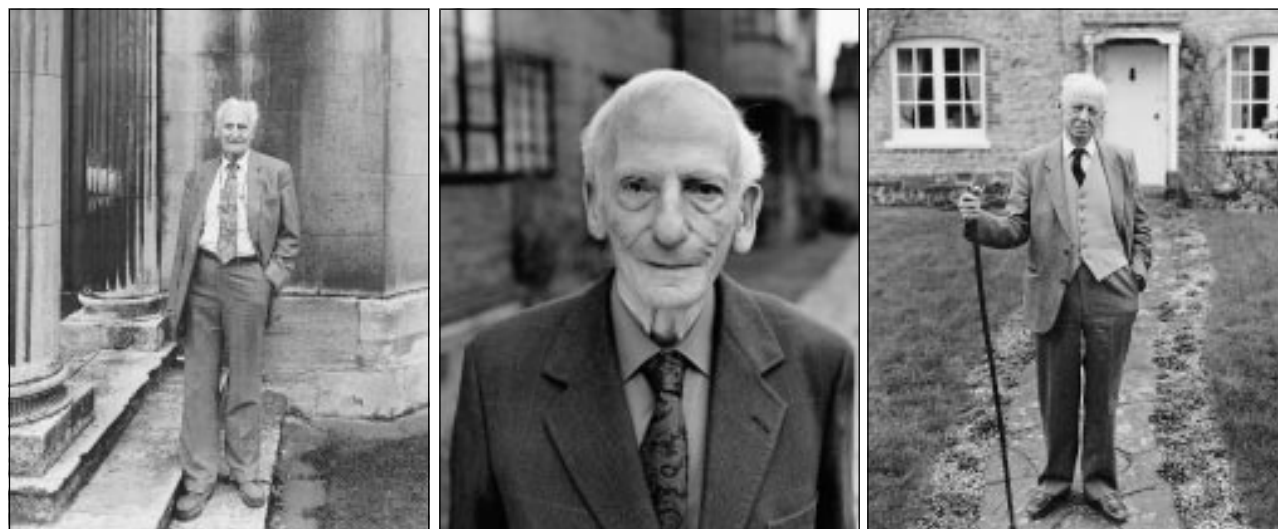
*BMJ* 1998;317:1217-20

had no apparent effect on acute hepatitis, but di-cysteine seemed to reduce the duration of acute hepatitis by a few days.

### The introduction of randomisation

The technique of alternate allocation had one major disadvantage: the investigator knew which treatment the next patient was going to receive and could be—and indeed often was—biased by knowing what the next treatment would be when deciding whether or not a patient was suitable for inclusion in the trial. Even blinding the investigator to the nature of the given treatment, which was often possible, by presenting the treatments in similar forms labelled A and B did not get over the difficulty completely; the investigator could quickly get the impression that one treatment was superior to the other and subsequently be biased in deciding on the next patient's eligibility.

One way of avoiding such biases would be to divide the patients into two similar groups before it was known which group would get which treatment and then, at the last minute, allocate one whole group to one treatment and the other to another by tossing a coin. This method was proposed by van Helmont, a



Richard Doll, Philip D'Arcy Hart, and Francis Avery Jones. Portraits by Nick Sinclair (published in *With Head and Heart and Hand: Photographs of 20th Century British Doctors*, BMJ Books, 1997: see [www.bmj.com/collections/century/data/exhibit.htm](http://www.bmj.com/collections/century/data/exhibit.htm))

medicinal chemist, in 1662 when he challenged the academics of the day to compare their treatments based on theory with his based on experience. "Let us take out of the hospitals, out of the Camps, or from elsewhere, 200, or 500 poor People, that have Fevers, Pleurisies, etc. Let us divide them into half, let us cast lots, that one half of them may fall to my share, and the other to yours . . . We shall see how many funerals both of us shall have. But let the reward of the contention or wager, be 300 florens, deposited on both sides."<sup>5</sup> Sadly, the challenge was not accepted. The technique was, however, actually put into practice by Amberson et al 260 years later to assess the value of sodium gold thio-sulphate in the treatment of pulmonary tuberculosis.<sup>6</sup> Amberson et al divided 24 patients into two groups of 12, the members of each group being "individually matched" in pairs. They then tossed a coin to decide which treatment each group should get.

This technique suffers, as Armitage has pointed out,<sup>5</sup> not only from the virtual impossibility of truly being able to match cases, but also because it provides no means of measuring the relevant random error. Both difficulties are overcome within quantifiable limits by the randomisation of individuals. This technique had been used in agricultural experiments described by Fisher in 1926 (when plots of land were individually randomised)<sup>7</sup>; Bradford Hill had recognised the desirability of using it in clinical medicine when he published a series of articles on the principles of medical statistics in 1937.<sup>8</sup> He did not recommend the randomisation of individuals then, preferring that the two treatments for comparison be allocated to alternate patients because, as he wrote in 1990, by referring to the randomisation of treatments he might have scared doctors off any use of concurrent controls.<sup>9</sup> In 1946, when he judged the time was right, he recommended the randomisation of individual patients and this rapidly gained acceptance among medical scientists.<sup>10 11</sup> He advocated it not so much because it provided a proper estimate of random error, which was the principal reason it was advocated by Fisher, but on the practical grounds that it eliminated bias in selection.

The trial in which treatments first began to be allocated randomly to individuals was one designed to test the efficacy of immunisation against whooping cough,<sup>10</sup> not the trial of streptomycin for treating pulmonary tuberculosis.<sup>11</sup> The latter trial, organised by D'Arcy Hart and Daniels, started in September 1946, a few months after the whooping cough trial; but it was reported in 1948, three years earlier than the results of the whooping cough trial. Consequently, although it was certainly the first to be reported, it undeservedly earned the reputation of being the first truly randomised trial. In both cases efforts were made to blind the assessor to the participant's treatment and, when practicable as in the first trial, to blind the participants. In both cases ethical considerations played a major part.

### Ethical considerations

At that time there were not any ethical committees to consult nor were there any ethical criteria laid down by the Medical Research Council or any other responsible body. Medical ethics were primarily defined by the Hippocratic oath, which all newly qualified doctors were required to swear. Bradford Hill was not a doctor, but he had such a deep understanding of the nature of medical practice that the lecture that he gave at the Royal College of Physicians on medical ethics and controlled trials was listened to with respect and was widely acclaimed.<sup>11</sup>

In the trial of streptomycin, the first issue that had to be faced was whether it was ethical to withhold from any patient a drug that had been effective in animal experiments and had had encouraging clinical results in the few published reports. There was, however, only a small amount of the drug in Britain and it was not possible to buy more from abroad. It was agreed to use the limited supplies to treat patients with two conditions that had previously been invariably fatal: military tuberculosis and tuberculous meningitis. The amount of streptomycin left over was insufficient to treat more than a tiny proportion of the people desperately ill with other types of tuberculosis. The Medical Research Council's Streptomycin in Tubercu-

loysis Trials Committee agreed that “it would have been unethical not to have seized the opportunity to design a strictly controlled trial which could speedily and effectively reveal the value of the treatment.”<sup>12</sup> The question of whether it was ethically justifiable to withhold the drug from any patient was, therefore, answered with an unhesitating “Yes.”

Two other questions that the committee considered were whether the doctors involved could modify the treatment schedule and whether the control patients should be given apparently similar placebos. It was agreed that the doctor must always “do for his patient whatever he really believes to be essential for that patient to return him to health.”<sup>12</sup> This meant that if any patient seemed likely to benefit from an induced pneumothorax—the only specific treatment available for pulmonary tuberculosis before the introduction of streptomycin—the treatment must be given irrespective of whether it upset the balance of the streptomycin and the control groups, as in fact it proved to do. The use of a placebo was ruled out in the interest of the patients because it would have required an intramuscular injection four times a day for four months. The response to treatment could be assessed objectively without it: psychological factors would have little impact on such a serious disease, and there was “no need in the search for precision to throw common sense out of the window.”<sup>12</sup>

The committee did not discuss the need to obtain informed consent; the overriding issue was the welfare of the patient. Bradford Hill inveighed strongly against the compulsion to obtain formal consent if this required giving a frightening account of the risks associated with the patient’s condition. “Does the doctor invariably seek the patient’s consent before using a new drug alleged to be efficacious and safe? If the answer is No, then what process, one may ask, makes it needful for him to do so if he chooses to test the drug in such a way that he can compare its effects with those of the previous orthodox treatment?”<sup>12</sup>

The question he asked might be answered differently today, but the principle with which he was concerned—that there should not be one standard for the ethics of therapeutic trials and another for routine medical care—is still valid and ought to be a major determining factor in our approach to patients.<sup>13</sup> To seek informed consent when it went against the patient’s interest was, in Bradford Hill’s opinion, unethical and should be dispensed with subject, as would now generally be agreed, to approval by an appropriate independent committee.

The situation in the whooping cough prevention trial was different. Parents of children aged 6-18 months were asked to volunteer to have their children entered into the trial; they were given a pamphlet describing the study, which included the information that half the inoculations would not be against whooping cough but would be “anti-catarrhal.” No child was entered until a consent form had been received; this condition was emphasised in the report by the acknowledgment of the “many parents who, in the full knowledge that their children would not necessarily receive pertussis vaccination, consented to take part in the investigation.”<sup>9</sup>

## Clinical opposition

The spread of randomisation, until it became an essential element of trials submitted to licensing authorities for the approval of new drugs, was initially slow and not without opposition. This was most commonly expressed along the lines of Lewis’s criticism of what he called “the statistical method of testing treatment.”<sup>14</sup> Lewis was the doctor in charge of the department of clinical research at University College Hospital, London, and the doyen of clinical research in the United Kingdom in the 1930s. He died in 1945 before randomised trials were introduced but I can imagine what his reaction would have been. Lewis thought that when testing treatments for acute diseases two groups of patients that were as similar as possible should be treated in exactly the same way and concurrently, except that one group should receive the remedy and the other should not. However, he added that “it is to be recognised that the statistical method of testing treatment is never more than a temporary expedient, and that but little progress can come of it directly: for in investigating cases collectively, it does not discriminate between cases that benefit and those that do not, and so fails to determine criteria by which we may know beforehand in any given case that treatment will be successful.”<sup>14</sup>

Lewis’s objection was repeated many times in the first few years after randomisation was introduced. Bradford Hill would reply: “Tell me the criteria to distinguish patients who will respond from those who don’t and we will build this into the trial” (A Bradford Hill, personal communication). There was never any serious response to Bradford Hill’s challenge, and it came to be accepted that randomisation was appropriate within strata defined by the clinician—if the clinician could define them sufficiently clearly for practical use.

## Modern developments

Early randomised trials can properly be criticised on the grounds that they were often too small to have any chance of detecting moderate effects. Small trials can be successful when the effect is large but this seldom occurs. They can also be successful when the effect is moderate and the outcome is measured quantitatively, as in the series of trials of treatment for gastric ulcer conducted by Avery Jones and me in the 1950s and ’60s, which recorded the percentage reduction in the size of the ulcer over a fixed time.<sup>15</sup> However, when the outcome is qualitative rather than quantitative, moderate but important effects will often be missed unless the number of patients treated is large. For example, many early trials of the treatment of myocardial infarction, stroke, and cancer, which were not large by modern standards, consequently led to the misleading conclusion that there was no benefit. Bradford Hill and his students, such as myself, were primarily concerned with getting the principle adopted, and to have pressed for trials on thousands of patients would have been self defeating. There were too few physicians, leave alone surgeons, who were willing to expose their theories to cold scientific investigation. Multicentre trials were organised from the beginning (the streptomycin trial involved six centres), several centres were successfully

involved in trials like those of adrenocorticotropic hormone and cortisone for the treatment of ulcerative colitis, and more centres became involved in the Medical Research Council's trials of treatment for different forms of leukaemia.<sup>16</sup> It was many years before randomisation was accepted as such a normal procedure. Only then did it become possible to organise the groundbreaking international study of infarct survival (ISIS) trials for the treatment of myocardial infarction, which involved hundreds of centres and randomly allocated tens of thousands of patients, and thereby showed the value of moderate improvements in the treatment of common diseases.<sup>17</sup>

Without Bradford Hill, randomisation would have come about sooner or later, perhaps introduced by Rutstein in the United States. Rutstein collaborated with Bradford Hill in the design of an Anglo-American trial of adrenocorticotropic hormone, cortisone, and aspirin in the treatment of acute rheumatic fever.<sup>18</sup> Randomisation would have been adopted much more slowly, however, without Bradford Hill's understanding of medical susceptibility and medical ethics and without his concern for simplicity of design and clarity of presentation. Modern authors please note.

Competing interests: None declared.

- 1 Fibiger J. Om serumbehandling af Difteri. *Hospitalstidende* 1898;6:309-25.
- 2 Medical Research Council Therapeutic Trials Committee. The serum treatment of lobar pneumonia. *BMJ* 1934;i:241-5.

- 3 Medical Research Council Patulin Trials Committee. Clinical trial of patulin in the common cold. *Lancet* 1944;ii:373-4.
- 4 Wilson C, Pollock MR, Harris AD. Diet in the treatment of infectious hepatitis. *Lancet* 1946;i:881-3.
- 5 Armitage P. The role of randomisation in clinical trials. *Stat Med* 1982;ii:345-52.
- 6 Amberson JB Jr, McMahon BL, Pinner M. A clinical trial of sanocrysin in pulmonary tuberculosis. *Amer Rev Tuberc* 1931;24:401-35.
- 7 Fisher RA. The arrangement of field experiments. *Journal of the Ministry of Agriculture* 1926;33:503-13.
- 8 Hill AB. *Principles of medical statistics*. London: Lancet, 1937.
- 9 Hill AB. Memories of the British streptomycin trial in tuberculosis. *Control Clin Trials* 1990;11:77-9.
- 10 Medical Research Council Whooping-Cough Immunization Committee. The prevention of whooping-cough by vaccination. *BMJ* 1951;ii:1463-71.
- 11 Medical Research Council Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment for pulmonary tuberculosis. *BMJ* 1948;ii:769-82.
- 12 Hill AB. Medical ethics and controlled trials. *BMJ* 1963;i:1043.
- 13 Collins R, Doll R, Peto R. Ethics of clinical trials. In: CJ Williams, ed. *Introducing new treatments for cancer: practical, ethical, and legal problems*. Chichester: John Wiley, 1992.
- 14 Lewis T. *Clinical science illustrated by personal experiences*. London: Shaw and Sons, 1934:178-83.
- 15 Doll R. Medical treatment of gastric ulcer. *Scott Med J* 1964;9:183-96.
- 16 Medical Research Council Working Party on Leukaemia in Childhood. Improvement in treatment for children with acute lymphoblastic leukaemia. *Lancet* 1986;i:408-11.
- 17 Second International Study of Infarct Survival Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;ii:349-60.
- 18 Rheumatic Fever Working Party of the Medical Research Council of Great Britain, Subcommittee of Principal Investigators of the American Council on Rheumatic Fever and Congenital Heart Disease, American Heart Association. Treatment of acute rheumatic fever in children: a cooperative clinical trial of ACTH, cortisone, and aspirin. *BMJ* 1955;i:555-74.

(Accepted 6 October 1998)

## Use of randomisation in the Medical Research Council's clinical trial of streptomycin in pulmonary tuberculosis in the 1940s

Alan Yoshioka

Centre for History of Science, Technology and Medicine, Imperial College, London SW7 2AZ  
Alan Yoshioka, historian

Correspondence to: Mr Yoshioka, PO Box 166, Picou, NS B0K 1H0 alan\_yoshioka@hotmail.com

*BMJ* 1998;317:1220-3



The MRC trial report is reproduced on our website, together with its two accompanying editorials

A slight mystery surrounds the clinical trial that this special issue of the *BMJ* commemorates.<sup>1</sup> It is now widely recalled that pulmonary tuberculosis patients in the Medical Research Council (MRC) trial were allocated to a streptomycin treatment group and a control group by a process using random sampling numbers and sealed envelopes (box). The editorial that introduced the MRC's report to *BMJ* readers on 30 October 1948 called attention to this new scheme, distinguishing it from the older practice of taking alternate cases in order of admission to hospital as the method of creating a control group.<sup>2</sup> The statistician involved, Professor (later Sir) Austin Bradford Hill, had been promoting the use of random allotment since before the second world war.<sup>3</sup> Remarkably, however, the word "random" appeared nowhere in the MRC's files on streptomycin for 1946. During that year, the now famous scheme was explicitly mentioned in a single letter. So why did the MRC use randomisation in this clinical trial?

Any answer is tentative as contemporary evidence about the reasoning of the members of the MRC's committee is thin. We can make some inferences from the arguments they used after the fact to justify the random allocation scheme to the medical profession. Their design, and how it was presented, should be understood

### Summary points

Randomised allocation of patients is rarely mentioned in the Medical Research Council's documents on streptomycin clinical trials

The meaning of the term "randomisation" has shifted over time; justifications for using randomisation have also changed

Streptomycin was isolated in 1943 at Rutgers University in New Jersey; the MRC began planning clinical trials in 1946; the first patients with tuberculosis entered the trials in 1947

The British government initially purchased 50 kg of American streptomycin for the MRC; most of the supply went to the clinical trial in pulmonary tuberculosis

Public demand for streptomycin was far in excess of supplies in Britain. Randomisation relieved the MRC's clinicians of responsibility for deciding who would be treated