

Controlled trials testing two or more treatments simultaneously

Richard Doll

J R Soc Med 2005;98:479–480

INTRODUCTION

In 1950, when I started clinical research in gastroenterology, the treatment of gastric ulcers was far from satisfactory. The role of *Helicobacter pylori* had not been discovered and the symptoms, although they could be relieved, kept on recurring irrespective of treatment and often eventually became so severe that the ulcer had to be resected with much of the acid secreting part of the stomach.

Orthodox treatment consisted of five elements, which were combined with varying emphasis, depending on the views of the individual physicians. All of them prescribed alkalis for the relief of pain and all of them recommended bed rest if symptoms persisted. Nearly all advised a bland diet, varying from 2-hourly milk feeds to a convalescent diet that excluded fried foods, pastry, various meats, and raw vegetables. Nearly all sought to treat the postulated underlying emotional factors by discussion, reassurance, and a sedative. Lastly, in an attempt to reduce acid secretion and inhibit gastric tone, many also prescribed atropine or one of its synthetic analogues.

To this schedule there was often added some new treatment that became popular for a while before being replaced by another: I had no difficulty in drawing up a list of remedies beginning with each letter of the alphabet. If, therefore, any substantial proportion of even the most promising remedies were to be properly evaluated it would take a very long time, so there would be considerable advantage in testing two or more in the same group of patients.

THE FACTORIAL DESIGN

As it happened a technique for doing this had been devised at least as early as the late 1920s, when Wyckoff and his colleagues¹ tested the value of digitalis in pneumonia by grafting it on to a trial of antipneumococcus serum that was then being carried out in three New York hospitals. Alternate patients were treated with or without serum and within each of these two groups alternate patients were also

given digitalis. The trial was less than ideal as different doses of digitalis were given at the different hospitals, serum was omitted from some patients in the second year of the trial, and a substantial number of patients scheduled to receive digitalis did not receive it, which was probably just as well as those who did get digitalis had the higher fatality rate.

A much more satisfactory trial was carried out 15 years later, when Wilson *et al.*² sought to test simultaneously the separate effects of supplements of cysteine and reduced dietary fat on the course of infective hepatitis, albeit treating only 103 patients. As had become standard scientific practice, alternate patients were consequently prescribed different treatments, with or without a supplement of 5 g cysteine a day, but alternate patients in each group (with and without cysteine supplements) were additionally prescribed either a low fat or a high fat diet, the patients on the two different fat diets being nursed in separate wards. When, therefore, the patients given supplementary cysteine were compared with those not given it, each group had had comparable diets, in that half had had a high fat diet and half a low fat diet. The same comparability held with regard to supplementary cysteine when the patients on the two fat diets were compared. The results suggested some possible benefits from cysteine, in that jaundice, liver enlargement, and biliuria did not last so long, but no difference was observed in the course of the disease between those given high and low fat diets.

EXTENSION TO TEST THREE TREATMENTS WITH RANDOMIZATION

With such trials as precedents, my colleagues and I decided to adapt the method to test three therapies at the same time, giving successive patients one of eight possible combinations (*a*, *b* and *c*; *a* and *b*; *a* and *c*; *b* and *c*; *a* alone; *b* alone; *c* alone; or none of them). By then, however, Bradford Hill had introduced the principle of randomization in place of a fixed schedule of alternation³ and the particular therapies for each patient were decided by opening numbered envelopes which contained the appropriate instruction, successive groups of eight including all the possible combinations. This, it has to be admitted, sacrificed the principal advantage of randomization, namely, the avoidance of any possibility of bias in deciding whether the

Clinical Trial Service Unit & Epidemiological Studies Unit, Harkness Building, Radcliffe Infirmary, Oxford OX2 6HE, UK

The author died 14 July 2005

next patient presenting in the clinic was suitable for inclusion, as towards the end of each group of eight patients it was known what the treatments were likely to be. To diminish this risk, strict criteria were laid down about the characteristics of the patients to be included in, or excluded from, the trial.⁴

The first trial tested the effect of bed rest in hospital against ambulant treatment, of phenobarbitone to relieve anxiety, and of vitamin C (which had recently been popularized as a therapy). It found that of the three treatments only bed rest hastened healing.⁵ Subsequently, 15 other treatments were tested using the same technique. Most trials included only 64 patients and no useful result was likely to have been obtained if the effect of the treatment had been judged simply by, for example, the proportion of ulcers healed. The radiologist collaborating in the trial was, however, at pains to obtain a picture showing the maximum size of the ulcer profile and this enabled the patient's response to be assessed quantitatively, by measuring the change in the area of the ulcer silhouette over a standard period of 4 weeks.

A similar method for testing three therapies at once was adopted independently by Thomas Chalmers and his colleagues in a series of trials of therapy for infectious hepatitis in the US Army.⁶ In their trial, three dietary regimens were tested: a high (4000) calorie diet against a standard (3000) calorie diet: a high (19%) protein diet against a standard (11%) protein diet: and supplements of choline and multivitamins against no supplement. Of the three comparisons a statistically significant difference was found only with the different protein diets, the high protein diet being associated with a shorter duration of illness.

EXTENSION TO LARGE TRIALS

The desirability of factorial designs has become of increasing importance because of the cost of trials, as well as the time involved in conducting them, both of which inhibit repetition. They are particularly needed to provide clear information about the benefit of new treatments that have only moderate effects and need to be assessed by the frequency of relatively uncommon outcomes (such as fatality may be). These needs have been met by the development since the 1980s of really large controlled trials after the successful conduct of a trial of the treatment of

myocardial infarction in over 16 000 patients.⁷ Subsequent trials of this size have often had a factorial design testing two therapies⁸ or three.⁹ The clarity of the results so obtained has, in some instances, quickly changed standard medical practice, as with the demonstration of benefit from both aspirin and streptokinase in the treatment of myocardial infarction.¹⁰

CONCLUSION

The use of a factorial design in controlled trials has a history of only seven decades. Within this period it has become established as a valuable technique that has enabled conclusions to be drawn about the benefit, or lack of benefit, of controversial treatments much more quickly and more cheaply than would otherwise have been the case.

Acknowledgment This paper was previously published by The James Lind Library [www.jameslindlibrary.org]. Accessed Friday 19 August 2005.

REFERENCES

- 1 Wyckoff J, Dubois EF, Woodruff IO. The therapeutic value of digitalis in pneumonia. *JAMA* 1930;**95**:1243–9
- 2 Wilson C, Pollock MR, Harris AD. Diet in the treatment of infectious hepatitis. *Lancet* 1946;**i**:881–3
- 3 Medical Research Council Streptomycin In Tuberculosis Trials Committee. Streptomycin treatment for pulmonary tuberculosis. *BMJ* 1948;**2**:769–2
- 4 Doll R. Medical treatment of gastric ulcers. *Scott Med J* 1964;**9**:183–96
- 5 Doll R, Pygott F. Factors influencing the rate of healing of gastric ulcers: admission to hospital, phenobarbitone, and ascorbic acid. *Lancet* 1952;**i**:171–5
- 6 Chalmers TC, Eckhart RD, Reynolds WE, *et al.* The treatment of acute infectious hepatitis. Controlled studies of the effects of diet, rest, and physical reconditioning on the acute course of the disease and on the incidence of relapses and residual abnormalities. *J Clin Invest* 1955;**34**: 1163–234
- 7 ISIS-1 Collaborative Group. Randomised trial of intravenous atenolol among 16,027 cases of suspected acute myocardial infarction: ISIS-1. *Lancet* 1986;**ii**:57–66
- 8 ISIS-2 Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;**ii**:349–60
- 9 ISIS-4 Collaborative Group. A randomised factorial trial assessing early captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995; **345**:669–85