SEQUENTIAL PROCEDURES IN PERSPECTIVE

There are not just 'sequential trials' and other trials. There is a sequential aspect within every trial; sometimes it happens to be the most important part of the design.

Fourteen years ago Armitage (1960) presented a strong case for the use of sequential statistical methods in clinical studies. In this issue of the journal Huskisson & Grayson describe a study in which an analysis of this type was used to compare the night-time relief provided by indomethacin or amylobarbitone in patients with rheumatoid arthritis. Despite their beguiling simplicity, and the small number of patients required to obtain a definitive result in many instances, these tests have not gained wide popularity. In part, this reflects a lack of understanding of the principles on which they are based. Two points are of crucial importance: firstly, sequential methods cannot be substituted indiscriminately for traditional fixed sample-size methods since only a single predetermined response to the contrasted treatments is analysed, and statistical evaluation of other factors is usually precluded. Secondly, the test of significance employed is based not on the usual hypothesis that contrasted treatments are equivalent until shown otherwise, but on the assumption that a predetermined difference in their performance does exist, i.e., the test is of a 'structured hypothesis', not of the 'null hypothesis'.

In many situations only the demonstration of a material difference between competing treatments would have important clinical implications. Thus the design selected by Huskisson & Grayson tests whether at least 85% of the target population is likely to favour one of the selected treatments. By aiming to detect a treatment difference only if it is of this order, a statistically valid result, whether positive or negative, can be obtained from a small number of patients. It is, however, important to appreciate that a negative result in this context does not exclude the existence of a lesser difference between the treatments.

But how much confidence can be placed in a generalization derived from a consideration of only eight patients? In general, the smaller the sample, the greater the danger that it is not representative of the target population. But all clinical trials are vulnerable to the operation of bias in addition to random variation: very frequently, owing to shortage of acceptable patients, later entrants are less well suited to the trial, and,

particularly in prolonged studies, the attitudes and decisions of the assessors are liable systematically to change (Dunne, 1970a). In consequence, since treatment differences are likely to be most faithfully reflected by the responses of early entrants to a trial, the use of sequential methods holds some advantage. It follows that, whether or not formal sequential analyses are employed, sequential display of trial data is a prudent precaution.

If this is done, time dependent changes in the pattern of results will emerge from time to time which cannot reasonably be ascribed to chance and this suggests that factors other than treatments are exerting an important influence on the results. On other occasions a careful search fails to disclose any systematic cause for these changes. For example, in the M.R.C. multicentre trial of corticosteroid drugs in the nephrotic syndrome (Black, Rose & Brewer, 1970) a 'run against entropy' occurred early in the trial with four deaths in succession within one treatment group. A sequential boundary would have been reached by one more case. Then, the pathway turned back within the boundaries and never crossed them. Similarly, despite a random treatment allocation which could not be faulted, a significant over-allocation to one group appeared at one stage in a major rheumatological trial (Beveridge (1971), personal communication). The effect later vanished; it may have been caused by some transient imbalance between entries and withdrawals affecting one treatment. In another trial conducted by ourselves (Dunne, 1970b) the sequential pathway crossed a boundary. Then the physicians learned by accident that one kind of tablet had been used much more than the other, a piece of misleading information as it happened. The trial proceeded, but the sequential results turned back and re-entered the boundaries.

Time dependent bias can never be eliminated from clinical trials; its impact can be reduced by using suitable block designs because these have the effect of averaging time-dependent effects throughout the block. But it is also necessary to plot sequential charts to demonstrate the existence of time-dependent changes and investigate their causes. This can only increase the information derived from the trial, and reveal major systematic bias as it occurs. This is a good reason for having a trial controller, someone independent of the physicians who knows the code and can follow the

trial's progress. Sequential charts are useful to him, particularly when deciding to stop a trial prematurely because of adverse drug effects. However, trial controllers may find themselves facing some very difficult decisions, as in the examples just cited. The criteria for deciding the significance of associated events listed by Hill (1965) may then prove useful.

It is not necessary to plot full sequential diagrams in many cases. Control charts (Moroney, 1951), or better cumulative sum charts are often equally effective (Chaput de Saintonge & Vere, 1974). Several charts can be run together for different variables; often sequential trials are made using only one criterion of preference to gain the advantage of economy but this need not be the rule. When several charts are run synchronously it should be remembered that at conventional significance levels 1 in 20 of them will, on average, be misleading. The significance of multiple significances is well discussed by Sakoda, Cohen & Beall (1954).

In summary, there is a sequential element in all trials in which results are collected over time. The conventional advantage of economy, for which sequential methods were developed, need not be regarded as the only criterion for their use; they provide a means to look at the progress of a block design trial too. Formal sequential designs do not test a null hypothesis, but a pre-specified preference ratio. The number of patients can, in general, be kept small only when that ratio is large. But when it is large a non-significant result may conceal smaller but important effects which could readily have been shown using more subjects.

J.F. DUNNE
D.W. VERE

*Department of Pharmacology and Therapeutics,
The London Hospital Medical College,
London E1 2AD*

References

- ARMITAGE, P. (1960). *Sequential Medical Trials*. London: Blackwell.
- BLACK, D.A.K., ROSE, G. & BREWER, D.B. (1970). Controlled trial of prednisone in adult patients with nephrotic syndrome. *Br. med. J.*, 3, 421-426.
- CHAPUT DE SAINTONGE, D.M. & VERE, D.W. (1974). Why don't doctors use cusums? *Lancet*, i, 120-121.
- DUNNE, J.F. (1970a). Time dependent variation in clinical trials. Ph.D. Thesis, London University.
- DUNNE, J.F. (1970b). How to improve at clinical trials. In: *Principles and Practice of Clinical Trials*, eds. Harris, E.L. & Fitzgerald, J.D., p. 235-241. London: Livingstone.
- HILL, A. BRADFORD. (1965). The environment and disease: association or causation? *Proc. Roy. Soc. Med.*, 58, 295.
- MORONEY, M.J. (1951). *Facts from figures*. Middlesex: Penguin.
- SAKODA, J.M., COHEN, B.H. & BEALL, G. (1954). Tests of significance. *Psychol. Bull.*, 51, 172-175.