

Editorial

The small clinical trial: is there a better way?

Few readers of medical papers will have failed to notice the enormous rise in the use of statistics in recent years. *Thorax* is by no means alone in referencing statistical guidelines in its instructions to authors.¹ Arguably the biggest influence on improving statistical standards in medicine in the United Kingdom has been the *British Medical Journal*, which has consistently encouraged dissemination of statistical methods with the subject of statistics in medicine frequently among the contents of the journal and with special series of articles, which have formed the basis of two popular books.^{2,3} Despite this heightened statistical awareness in medicine, however, many people still think of statistical methods principally in terms of the analysis of data. Yet ask any statistician which element of his work he considers the most important and the answer will always be the planning of the study. This is reflected in the check lists used by the *British Medical Journal* statistical referees, where around half of the specific checkpoints relate to design.⁴

Even with this recognition of the role of statistics, and the emphasis of statisticians on the importance of a well planned study, many investigations have too few patients. In particular, too many clinical trials are carried out with little apparent consideration of statistical power—that is, the number of patients may be so small that a clinically important difference has little chance of being detected as statistically significant. An accompanying paper in this issue of *Thorax* (p 824) has examined the power of 15 published randomised, controlled, double blind trials in severe acute asthma and has found most to be of inadequate size.

The conventional wisdom is to urge investigators to conduct larger trials with adequate power. For a host of practical reasons, however, it may be difficult or even impossible to do this. What then are we to make of the small trial? Should the editors of journals protect their readers from their malevolent effect by a ruthless editorial policy? When they slip through this net, should we simply consign them to the mental wastepaper bin? This is a tempting philosophy with which to counter the ever expanding publications, but condemning all small trials is too facile. Powell-Tuck

*et al*⁵ provide examples where small trials have been adequate as a guide to clinical practice. Rare conditions will always lead to small clinical trials, and a small randomised controlled trial, presented correctly, will in most circumstances be more informative than its common rival: the uncontrolled trial of a single treatment. Uncontrolled trials, however, remain regrettably popular, particularly in some areas of oncology. Respiratory medicine seems to attract fewer studies of this kind, but if they are encountered there is good evidence that they should be interpreted with extreme caution.⁶

If small trials can be useful in some circumstances, why then do they receive such a bad press, especially from statisticians? This arises largely from misinterpretation of the results on the part of both authors and readers. Statistical significance is often mistakenly taken to be the ultimate arbiter, and a non-significant difference between two treatments is wrongly interpreted as showing that there is *no* difference between the treatments. There is a strong case for suggesting that, where a trial is unavoidably too small to have reasonable power, significance tests should not be reported at all. It is far more appropriate to report, for each outcome measure, confidence intervals for the difference between treatments: that is, we report on the basis of the trial data a range of values in which the population value for such a difference may lie. Indeed for all studies, large and small, the use of confidence intervals could be used far more widely to complement any significance tests that are reported.⁷

Given that the small randomised controlled trial will not actually be harmful if it is reported correctly, how useful is it likely to be? Despite the experience of Powell-Tuck *et al*,⁵ in most cases, when considered in isolation, it will not be very informative. The confidence interval will be wide and the data might conceivably be compatible with an important benefit of either treatment. Its value will be enhanced when it is viewed in relation to the results from similar trials. This is exemplified in this issue's paper by Ward on clinical trials in severe acute asthma (p 824). Combining the confidence intervals from similar randomised controlled trials in a single diagram produces an immediate impact beyond that which can be achieved in a more traditional review paper. The reader is put in a position where he can more easily make an informed assessment of the findings.

The whole approach can be taken one stage further and the results subjected to a formal statistical analysis. It is then often referred to as a meta-analysis. Meta-analyses have been applied in a variety of clinical areas, with notable major examples in breast cancer⁸ and myocardial infarction.⁹ These methods have been recognised as being particularly valuable where there is a simple endpoint such as death and a modest treatment effect is being sought, because the sample size requirements are then more demanding. When the endpoint is continuous, such as the increase in peak expiratory flow rate, proponents of meta-analysis have argued that sufficiently large multicentre studies should be feasible and meta-analyses less of a necessity. Certainly a meta-analysis should not be undertaken lightly. It may seem an easy option but there are myriads of pitfalls. For only if it is comprehensive in its coverage of all trials undertaken can it give valid results. Yet some trials may fail to achieve publication or may appear in very obscure journals, some will be abandoned prematurely because of unpromising early results, and others may still be in progress. Nevertheless, all should be identified and reported if unbiased findings are to be obtained from the meta-analysis.

Should we therefore consider multicentre trials more frequently? Are there better alternatives? If a randomised controlled trial can be carried out with an adequate number of patients in a single centre (and in respiratory medicine this is often the case), there are overwhelming arguments in favour of doing so. Within a single centre there can be tighter control over the conduct of the trial, which may be reflected in a smaller "experimental error." There are limits, however, to the extent to which good experimental design and control of the conduct of the trial can reduce experimental error, and if a convincing single centre trial is not feasible then a multicentre approach should certainly be considered.

Deciding on the desirability of a multicentre trial is, of course, much easier than its implementation. There is the question of which workers to contact, and the difficulty of convincing them that their participation is worthwhile. Different potential participants will want to modify the trial design, probably in totally different ways. Even with goodwill on all sides, multicentre trials need good coordination, which in turn means resources—and this means money. The sheer effort required to mount a multicentre trial and the time that this consumes are powerful disincentives to most researchers. At the selfish personal level, participation in a large multicentre trial might benefit the curriculum vitae less than several small studies. For medicine as a whole, however, there is a considerable advantage in having the collective experience of several units reported together in one major paper. This might be expected to give relatively accurate

answers on the efficacy of particular treatments, and furthermore will be accessible in a single (probably major) journal. In contrast, the piecemeal approach with different centres reporting inadequately sized trials in various journals serves medicine badly.

What then can the profession do to encourage multicentre collaboration when the resources of single centres are insufficient? One approach is through the journals by means of an editorial policy that encourages the reporting of adequately sized trials, even if the results are negative, but of smaller trials only in exceptional circumstances. Another is to have an easily accessible central organisation to consider proposals for multicentre studies. In respiratory medicine the British Thoracic Society Research Committee has such a role. Initial proposals can be quite brief and informally presented, but if they are judged to have merit a study will be taken through the successive stages of detailed planning, implementation, analysis, and reporting. The record of this committee in attracting grant support, encouraging participation, and carrying through studies to a successful completion is impressive. Several further studies are in progress with others at the planning stage, but there is no room for complacency. Despite the work of this and other bodies there are too few multicentre studies and there needs to be greater awareness of the benefits of and the potential for multicentre collaboration—not just in clinical trials but in all areas of medical research.

RJ PRESCOTT

*Medical Statistics Unit
Medical School, Edinburgh EH8 9AG*

References

- 1 Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *Br Med J* 1983;**286**:1489–93.
- 2 Swinscow TDV. *Statistics at square one*. London: British Medical Association, 1976.
- 3 Gore SM, Altman DG. *Statistics in practice*. London: British Medical Association, 1982.
- 4 Gardner MJ, Machin D, Campbell MJ. Use of check lists in assessing the statistical content of medical studies. *Br Med J* 1986;**292**:810–2.
- 5 Powell-Tuck J, MacRae KD, Healy MJR, Lennard-Jones JE, Parkins RA. A defence of the small clinical trial: evaluation of three gastroenterological studies. *Br Med J* 1986;**292**:599–602.
- 6 Pocock SJ. Allocation of patients to treatment in clinical trials. *Biometrics* 1979;**35**:183–97.
- 7 Gardner MJ, Altman DG. Confidence intervals rather than p values: estimation rather than hypothesis testing. *Br Med J* 1986;**292**:746–50.
- 8 Anonymous. Review of mortality results in randomised trials in early breast cancer. *Lancet* 1984;ii:1205.
- 9 Yusuf S, Peto R, Lewis JA, Collins R, Sleight P. Beta blockage during and after myocardial infarction: an overview of the randomized trials. *Progr Cardiovasc Dis* 1985;**27**:335–71.