

Editorial

Randomised trials of new procedures: problems and pitfalls

Forty years after the introduction of the randomised trial to assess the effectiveness of streptomycin for the treatment of tuberculosis, new drugs cannot gain clinical acceptance or regulatory approval without being tested in randomised trials. This experience has resulted in the general impression that new procedures and techniques must be validated similarly. In regard to coronary bypass surgery Hiatt suggested that "well designed trials should precede widespread dissemination, as is done to a considerable extent for drugs."¹

That recommendation reveals misconceptions about the differences between trials of new drugs and trials of new procedures. Indeed, with interventional procedures becoming ever more important in the treatment of cardiovascular disorders, discussions that once seemed pertinent only to surgical therapy are now pertinent to interventional cardiovascular treatment as well.

Drugs versus procedures

Table 1 lists several obvious differences between drugs and procedures. Drugs have an unchanging composition, but as drug use increases, additional side effects and complications become apparent. When used in a trial, a drug's effectiveness is unrelated to the physician's skill; not only are results generally consistent among collaborating institutions, but they are also applicable to non-participating institutions. Finally, a placebo is usually available, and crossover between treatment groups is exceptional.

In contrast, new procedures are introduced while they are imperfect. The indications are uncertain and the risks are high. As the procedure becomes more widespread, refinements occur and the risks decline, often dramatically. In early trials, results vary considerably in the hands of different operators. Placebos are obviously not available for any invasive procedure, and crossover from medical to procedural treatment groups is commonplace.

Bias in randomised studies

Although randomisation should eliminate bias, it often cannot do so, particularly with studies of new procedures. The first type of bias can occur even before the referral of patients to the study. When two different drug protocols or two surgical protocols are being compared with each other, referring physicians may not strongly prefer one to the other and are willing to randomise patients. However, when new procedures are being compared to established medical therapy, there are more likely to be opinionated adherents and detractors who divert high risk patients (about whom the physicians feel strongly) away from randomisation. It is commonly observed that patients in ran-

domised trials fare better with either treatment than do historical controls, and this observation is usually attributed to the fact that trial patients receive more attentive, more standardised, and more effective therapy. The better results with either treatment may result instead from referral bias that causes the trial group to be a low risk subset. Such bias makes it more difficult to prove the efficacy of a new procedure. During the early 1970s, the randomised trial of coronary bypass for stable angina at the University of Oregon experienced a striking diversion of referrals to private surgical teams who readily accommodated the referring physicians' growing preference for surgical therapy. The trial terminated with smaller patient subgroups than originally planned, and with statistically inconclusive results.²

Second, even participating investigators have human instincts and biases that can influence the selection of patients for randomisation. In the US Veterans' Administration study of coronary bypass surgery, patients were excluded because of subjective criteria such as unstable angina or poor left ventricular function. This led to a tendency at many participating hospitals to include only the most stable and willing patients.³

A third source of bias is non-adherence to the assigned therapy, or crossing over. In studies of a procedure, crossover can occur only in one direction. If crossover is frequent, the medical group is eventually composed only of patients in whom medical therapy is successful, thus confounding statistical analyses.

Timing and significance of procedural trials

It has been suggested that randomised studies should be performed from the very first clinical trial of costly procedures of all kinds,^{1,4} and that scientific journals should reject reports of non-randomised trials; however, the rapid evolution of procedural techniques usually invalidates early trials. The US Veterans' Administration study of coronary bypass surgery began in the early 1970s, but it was intensely criticised for poor operative results compared with outcomes at the time results were finally reported.³ Indeed, initial entries (1970-72) were discarded because results were poor. Though it made many newspaper headlines the VA study had little impact on the practice of coronary surgery, which progressed rapidly while the study was underway. In a more recent example, randomised trials of intracoronary stents had no sooner been reported^{5,6} than the results were superseded by dramatic changes in anticoagulant therapy. Furthermore, as progress in procedural techniques comes from many investigators, early trials that restrict procedures to a few participating institutions would stifle technical progress.

On the other hand, if a randomised trial is delayed until a procedure is mature, and good results are the norm, it will be difficult or impossible to randomise a broad spectrum of patients. In the Coronary Artery Surgery Study (CASS) begun in the United States after coronary bypass was well established, some institutions did not randomise any patients and participated only in a registry.⁷

The incentive to participate in early randomised trials comes from grant funds that are often vital to the participating institutions. Even when a proposed trial is

Table 1 Obvious differences between randomised trials for drugs and procedures

Drug	Procedure
Unchanging compound	Evolves continuously
Complications increase with use	Complications decrease with use
Results unrelated to physician skill	Results vary with operator
Placebo usually available	No placebo
Crossover rare	Crossover common

premature, a decision not to participate or to abandon an ongoing trial that is becoming obsolete and pointless, is a form of bureaucratic and academic Hari-Kari. Large collaborative studies generate their own central bureaucracy and peripheral constituency. When the US Veterans' Administration proposed a randomised study of the Björk-Shiley mechanical heart valve versus the Hancock bioprosthetic valve in 1976, the study was doomed to premature obsolescence by the predictably rapid development of prosthetic heart valves. This irrelevant study has continued stubbornly onward at great public expense, even though the Björk-Shiley valve has not been sold in the US for many years, and the standard Hancock model has been supplanted by the hydraulically superior modified orifice valve in most centres.

Summary

Randomised trials that compare new procedures with established ones must avoid prerandomisation bias and must allocate patients to treatment groups based on objective or quantitative criteria, not on subjective clinical judgment. Risk, length of follow up, and sample size must be used to calculate the statistical power of the study, so that a significant difference between treatments does not remain undetected (a type II error). There should already be sufficient experience with the new procedure so that complication rates have stabilised, and participating operators are equally comfortable with all procedures being studied.

Even with the above stipulations, randomised trials that compare medical with procedural treatment pose additional problems⁸ (many of which have been omitted from

this necessarily brief discussion); few such studies have had a major impact on clinical practice. The most useful randomised studies of procedures are those that compare one procedure with another, or those that assess a specific refinement in an established procedure, such as the use of different anticoagulation regimens for coronary stents. Fortunately, clinically useful information has always been available from non-randomised studies. The recent trend towards meta-analysis of large clinical series can substitute for those randomised studies that are unlikely to be helpful.

LAWRENCE I BONCHEK

*Surgical Director,
Mid-Atlantic Heart Institute at Lancaster General Hospital,
555 North Duke Street, Lancaster, PA 17603, USA*

- 1 Hiatt HH. Lessons of the coronary-bypass debate. *N Engl J Med* 1977;297:1462-4.
- 2 Kloster FE, Kremkau EL, Ritzman LW, Rahimtoohal SH, Rosch J, Kanarek PH, et al. Coronary bypass for stable angina: a prospective randomized study. *N Engl J Med* 1979;300:149-57.
- 3 Lawrie GM, Morris GC Jr, Howell JF, Chapman DW. A debate on coronary bypass. *N Engl J Med* 1977;297:1464-70.
- 4 Spodick DH, Aronow W, Barber B, Blackburn H, Boyd D, Conti R, et al. Standards for surgical trials [editorial]. *Ann Thorac Surg* 1979;27:284.
- 5 Serruys PW, Dejaegere P, Kiemeneij F, Macaya C, Rutsh W, Heyndrickx G, et al. A comparison of balloon-expandable-stent implantation with balloon angioplasty in patients with coronary artery disease. *N Engl J Med* 1994;331:489-95.
- 6 Fischman DL, Leon MB, Baim DS, Schatz RA, Savage MP, Penn I, et al. A randomized comparison of coronary-stent placement and balloon angioplasty in the treatment of coronary artery disease. *N Engl J Med* 1994;331:496-501.
- 7 CASS Principal Investigators and their Associates. Myocardial infarction and mortality in the coronary artery surgery study (CASS) randomized trial. *N Engl J Med* 1984;310:750-8.
- 8 Bonchek LI. The role of the randomized clinical trial in the evaluation of new operations. *Surg Clin North Am* 1982;62:761-9.