

Large Multicenter Trials: What Do They Achieve and What Should Be Done in Perfusion?

Paul Myles, FANZCA

Department of Anaesthesia and Perioperative Medicine, Alfred Hospital and Monash University, Melbourne, Australia

Presented at the Perfusion Downunder Meeting, Hayman Island, Queensland, Australia, August 9–11, 2007.

Abstract: There have been a vast number of publications in the perfusion and cardiac surgical literature suggesting possible therapeutic benefits from many perfusion interventions. Most of the reports are case series and other observational studies; few are randomized trials, and most of these are small, focusing on surrogate endpoints. We know there are many factors that can affect outcome after cardiac surgery, and some of these can bias results of clinical studies. Evidence-based medicine has highlighted the importance of avoiding bias with good study design, critical appraisal, and careful application into clinical practice. Associations shown in observational studies do not provide reliable evidence of effect (causation). Random allocation to treat-

ment groups accounts for many sources of bias, but small randomized trials can still be unreliable because they may identify a spurious positive finding by chance (type I error), as well as providing imprecise estimates of effect, as shown by wide confidence intervals. Obtaining data on actual outcomes with enough study power requires a large number of patients. Meta-analysis of small randomized trials can increase power, but this introduces other sources of bias. Large randomized pragmatic trials, using straightforward interventions reflecting routine clinical practice, can optimize the ability to generalize and therefore are clinically relevant and reliable. They thus provide the best evidence of effectiveness. *JECT. 2007;39:274–277*

BEWARE OBSERVATIONAL STUDIES

A confident demeanor is a valued element of being a good clinician, but in reality, we practice in an environment of great uncertainty. Individual patient and clinician characteristics, and a whole host of perioperative factors, can affect outcome after cardiac surgery. Similarly, interpretation and clinical application of published research is hindered by uncertainty because of numerous alternative explanations of observed findings: is a positive result “true” or is it a spurious finding? Evidence-based medicine has highlighted the importance of avoiding bias with good study design, critical appraisal, and careful application into clinical practice (1). Many sources of bias have been known for a long time (2), but others have become more widely recognized because of new study design techniques (3,4).

Large case series, registries, and other prospective observational studies can provide a lot of useful data concerning patient and surgical characteristics and their relationship with patient outcome. Being large, such studies can provide sufficient power to identify possible associations between many specific factors and complications, and so it is a useful first step when studying potential

improvements in care. However, strong associations have many possible explanations (2,4).

An illustrative example in perfusion is the relationship between red cell transfusion and poor outcome after coronary artery bypass graft surgery (CABG) (4). A large observational study in 11,963 patients undergoing CABG, in which one half had received a transfusion, found that red cell transfusion was associated with an increased risk of numerous serious postoperative complications. This included renal failure [odds ratio (OR), 2.1; 95% confidence interval (CI): 1.9–2.3; $p < .001$], serious infection (OR, 1.8; 95% CI: 1.7–1.8; $p < .001$), cardiac complications (OR, 1.6; 95% CI: 1.5–1.6; $p < .001$), neurologic events (OR, 1.4; 95% CI: 1.3–1.4; $p < .001$), and death (OR, 1.8; 95% CI: 1.7–1.9; $p < .001$). These associations persisted after accounting for patient and perioperative factors. Also, there was a clear dose response shown with each unit of red cells transfused being associated with an incrementally increased risk for complications (5). However, does this mean that red cell transfusion increase death and disability after CABG? Should we restrict red cell transfusion in this setting? Such studies have led to calls for restrictive transfusion practices and in particular use of a transfusion protocol (5), but there is no compelling evidence that this will reduce complications after CABG. There are many possible explanations for the observed association between

Author for correspondence: P. Myles, P.Myles@alfred.org.au

red cell transfusion and poor outcome (4). Some can be tested and/or controlled for with multivariate analyses, but experience has taught us that these cannot be depended on (6,7). Associations shown in observational studies do not provide reliable evidence of effect (causation). They are hypothesis-generating, and stronger levels of evidence should be sought (7).

STUDY DESIGN

Reliability, or precision, is important to clinicians because we want to know whether likely effects of any new treatment are clinically useful in any particular circumstance (7). If uncertainty exists, a change in practice is unlikely until further studies are done. These issues are highlighted when considering the potential benefits and risks of off-pump coronary artery surgery (OPCAB). Early reports and case series published in the early 1990s suggested OPCAB reduced complications traditionally associated with cardiopulmonary bypass (CPB). These positive reports fostered great interest in off-pump techniques, but many cardiac surgeons were cautious because of concerns about inadequate revascularization. At present, the key questions seem to be (i) does OPCAB improve postoperative outcomes, and if so, (ii) is there sufficient evidence to support more widespread use; but if not, (iii) what type of studies are required to provide compelling evidence? A similar series of questions surrounded warm heart surgery in the 1980s, with the eventual widespread use of 32–34°C for most routine cardiac surgery (8), and laser myocardial revascularization with the eventual decline in interest for the technique in most settings (9).

WHAT OUTCOMES SHOULD WE BE MEASURING?

Research in CPB and other perfusion techniques frequently uses surrogate endpoints—biochemical markers such as blood gases, electrolytes, and hematologic results, urine flow, myocardial ischemia, embolic load, and cerebral blood flow. Some of these have no meaningful relationship to “actual” outcomes of concern to patients (10).

Clinical researchers typically use surrogate endpoints because obtaining data on actual outcomes with enough study power requires a larger number of patients. It is common for there to be an underlying assumption that the surrogate endpoint relates directly to the actual outcome. For example, troponin release is a marker of myocardial damage and therefore is a surrogate marker of myocardial infarction (MI); similarly, for delta creatinine and renal failure, cerebral oximetry and stroke, embolic load and stroke, and so on. In some circumstances, this can be accepted, but in others, the relationship is more tenuous. Experience in other clinical settings tells us that studies

showing a positive effect on a surrogate endpoint can be quite misleading when a definitive outcome study is done (11). Thus, there should be greater efforts at following up initial positive studies based on surrogate endpoints with true outcome studies.

Because serious adverse outcomes after surgery are rare, outcome studies need to be large. For example, the incidence of stroke, renal failure, or death after CABG is mostly <4%. Study power is determined by the number of trial events, and therefore, power can be increased by focusing on high-risk patients and/or by using a combined endpoint (7,12). Both these approaches have been used in the Aspirin and Tranexamic Acid for Coronary Artery Surgery (ATACAS) trial (13).

WHY DO LARGE RANDOMIZED TRIALS?

It is far simpler, and far more common, for large clinical studies to be observational (non-randomized), and therefore, they may be biased (2,14). We know outcome after cardiac surgery is dependent on many factors, and therefore a new treatment being studied may have a spurious association with a good outcome unrelated to any true effect. Random allocation to treatment groups accounts for many such sources of bias (7).

The major source of bias overcome by randomization is treatment bias, whereby allocation to treatment groups is not decided by the clinician or patient. However, small randomized trials can still be unreliable because they may identify a spurious positive finding by chance (type I error), as well as providing imprecise estimates of effect, as shown by wide confidence intervals (7,8). For example, a small randomized study in 80 patients undergoing CABG surgery compared on-pump and off-pump CABG (15). Myocardial injury was assessed using serial troponin release, and this was significantly lower in the OPCAB group for up to 24 hours postoperatively. In addition, inotropic requirements were less in the OPCAB group, but this did not reach statistical significance. These findings offered promise but do not convincingly show improved outcome (less myocardial injury or deaths) with OPCAB.

Another small trial addressed the effect of OPCAB on long-term graft patency in 197 patients (16). Graft patency was similar for OPCAB and conventional CABG using CPB at 30 days ($p = .19$) and at 1 year (absolute difference, -2.2% ; 95% CI: -6.1% to 1.7% ; $p = .27$). Rates of death, stroke, myocardial injury, and reintervention were also comparable. The authors concluded that OPCAB provided comparable rates of complete revascularization and was cost effective (16). Is this enough to change practice? Given that the author had extensive experience in OPCAB surgery, does this allow an ability to generalize to other settings? Do the 95% CIs suggest a clinically important increase or decrease in graft patency, myocardial in-

jury, or death? For example, the risk of death was slightly higher in the OPCAB group, but this was not statistically significant (OR, 1.6; 95% CI: 0.3–9.8). Would a 1.6-fold (or 9.8-fold) increase risk of death temper enthusiasm in OPCAB? If the study was multicenter (numerous surgeons and settings), and the outcomes being studied were myocardial injury, stroke, and death, a change in practice would be justified. To be fair to the authors of this study, it was not their stated intention to address these latter issues. However, that is what is required to change practice (17).

Small trials are still prone to imbalances in prognostic factors that can have a potent effect on outcome: a special type of bias known as confounding (7,18). The larger the sample size in a randomized trial, the less likely it is that confounding can occur. A large randomized trial will equalize both known and unknown confounders between groups (6,7). Large trials are usually multicentered, and sometimes multinational, to maximize recruitment and enable early conclusion (6,7). This provides a broad range of settings and offers an opportunity to identify other patient, clinician, and institutional factors that may influence outcome. Large trials with straightforward requirements reflecting standard practice are sometimes called effectiveness, pragmatic, or practical trials (6,17). They thus optimize the ability to generalize their findings and therefore are clinically relevant. This is often not the case, with interested researchers studying select groups of patients in specialized settings. This is particularly relevant in perfusion studies, including use of off-pump surgical techniques, because of specific expertise at some centers—positive results may not be reproduced in other settings.

Another area of interest to perfusionists is neurologic injury associated with CPB, and in particular, stroke. A small trial reported a reduction in cognitive deficits with OPCAB at 3 months but no difference at 6 months after surgery (19). There were no significant differences in stroke or death rates, but the study was not designed or adequately powered to reliably address these issues. Several other trials have been done on this topic, and therefore, a pooled analysis can be done. This increases sample size and therefore increases study power. Such meta-analyses can provide least-biased estimates, but there are some weaknesses with this approach, particularly when meta-analysis is limited to small trials (20,21).

Sedrakyan et al. (22) identified 41 trials of OPCAB that had enrolled 3996 patients and reported a 50% reduction in the relative risk of stroke (95% CI: 7%–73%), 30% reduction in atrial fibrillation (95% CI: 16%–43%), and 48% reduction in wound infection (95% CI: 26%–63%). These studies represent a diverse range of clinical settings. Here, for the first time, we have strong evidence that OPCAB can significantly reduce serious complications associated with CPB. The outcomes are serious and

have a real impact on patient's lives. This evidence might affect a surgeon's practice, but the authors caution over-interpretation because of limitations of meta-analyses, variations in surgical expertise, increased need for conversion to on-pump surgery, and a lack of long-term outcome data (graft occlusion and reoperation). The current state of evidence clearly supports a definitive large randomized trial comparing OPCAB and on-pump CABG with short-term and long-term follow-up.

What should be done to improve the evidence base of perfusionist practices? Collecting accurate perfusion and perioperative data is a good first step. Ideally such data collection should be coordinated and extensive (multi-center), using agreed data definitions. Considering possible improvements in care (from such data or from any positive publications) should lead to testing new interventions with randomized trials. Irrespective of the size of the clinical trial, meaningful outcome data should be routinely collected and reported (to enable meta-analyses), and definitive large trials should eventually be done. Large trials rightly deserve the mantle of "gold standard" in providing evidence of effectiveness, because they provide reliable and relevant information to guide clinical practice (6,17,23,24).

REFERENCES

1. Evidence-based medicine working group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992; 268:2420–5.
2. Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32:51–63.
3. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc*. 1984;79:516–24.
4. Koch CG, Li L, Duncan AI, et al. Morbidity and mortality risk associated with red blood cell and blood-component transfusion in isolated coronary artery bypass grafting. *Crit Care Med*. 2006;34: 1608–16.
5. Ferraris VA, Ferraris SP, Saha SP, et al. Perioperative blood transfusion and blood conservation in cardiac surgery: the Society of Thoracic Surgeons and The Society of Cardiovascular Anesthesiologists clinical practice guideline. *Ann Thorac Surg*. 2007;83:S27–86.
6. Myles PS. Why we need large trials in anaesthesia and analgesia. In: Tramer MR, ed. *An Evidence-Based Resource in Anaesthesia and Analgesia*, 2nd ed. London: BMJ Publishing Group; 2003: 12–21.
7. Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: Clinical trials. *Lancet*. 2001;357:373–80.
8. Salerno TA. Warm heart surgery: Reflections on the history of its development. *J Card Surg*. 2007;22:257–9.
9. Bridges CR, Horvath KA, Nugent WC, et al. The Society of Thoracic Surgeons practice guideline series: Transmyocardial laser revascularization. *Ann Thorac Surg*. 2004;77:1494–502.
10. Fisher DM. Surrogate outcomes: Meaningful not! *Anesthesiology*. 1999;90:355–6.
11. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med*. 1991;324:781–8.
12. Ioannidis JPA, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol*. 1997;50:1089–98.

13. Myles PS, Smith J, Knight J, Cooper DJ, McNeil J, Silbert B. Aspirin and tranexamic acid for coronary artery surgery: a randomized controlled trial. Available at <https://www.atcas.org.au>. Accessed June 8, 2007.
14. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. *Am J Med.* 1982;72:233–40.
15. Ascione R, Lloyd CT, Gomes WJ, Caputo M, Bryan AJ, Angelini GD. Beating versus arrested heart revascularization: Evaluation of myocardial function in a prospective randomized study. *Eur J Cardiothorac Surg.* 1999;15:685–90.
16. Puskas JD, Williams WH, Mahoney EM, et al. Off-pump vs conventional coronary artery bypass grafting: early and 1-year graft patency, cost, and quality-of-life outcomes: A randomized trial. *JAMA.* 2004; 291:1841–9.
17. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA.* 2003;290:1624–32.
18. Rothman KJ. Epidemiologic methods in clinical trials. *Cancer.* 1977; 39:S1771–5.
19. Van Dijk D, Jansen EWL, Hijman R, et al. Cognitive outcome after off-pump and on-pump coronary artery bypass graft surgery: A randomized trial. *JAMA.* 2002;287:1405–12.
20. Cappelleri JC, Ioannidis JP, Schmid CH, et al. Large trials vs meta-analysis of smaller trials: How do their results compare? *JAMA.* 1996;276:1332–8.
21. LeLoerier J, Gregoire G, Benhaddad A, Lapierre J, Derderian FL. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med.* 1997;337:536–42.
22. Sedrakyan A, Wu AW, Parashar A, Bass E, Treasure T. Off-pump surgery is associated with reduced occurrence of stroke and other morbidity as compared with traditional coronary artery bypass grafting: A meta-analysis of systematically reviewed trials. *Stroke.* 2006; 37:2759–69.
23. Myles PS. Why we need large randomised studies in anaesthesia. *Br J Anaesth.* 1999;83:833–4.
24. Peto R, Baigent C. Trials: the next 50 years. Large scale randomised evidence of moderate benefits. *BMJ.* 1998;317:1170–1.