Education and debate

Need for expertise based randomised controlled trials

P J Devereaux, Mohit Bhandari, Mike Clarke, Victor M Montori, Deborah J Cook, Salim Yusuf, David L Sackett, Claudio S Cinà, S D Walter, Brian Haynes, Holger J Schünemann, Geoffrey R Norman, Gordon H Guyatt

Surgical procedures are less likely to be rigorously evidence based than drug treatments because of difficulties with randomisation. Expertise based trials could be the way forward

Although conventional randomised controlled trials are widely recognised as the most reliable method to evaluate pharmacological interventions,[1 2] scepticism about their role in non-pharmacological interventions (such as surgery) remains.[3–6] Conventional randomised controlled trials typically randomise participants to one of two intervenions (A or B) and individual clinicians give intervention A to some participants and B to others. An alternative trial design, the expertise based randomised controlled trial, randomises participants to clinicians with expertise in intervention A or clinicians with expertise in intervention B, and the clinicians perform only the procedure they are expert in. We present evidence to support our argument that increased use of the expertise based design will enhance the validity, applicability, feasibility, and ethical integrity of randomised controlled trials in surgery, as well as in other areas. We focus on established surgical interventions rather than new surgical procedures in which clinicians have not established expertise.

Use of expertise based trials

Investigators have used the expertise based design when conventional randomised controlled trials were impossible because different specialty groups provided the interventions under evaluation—for example, percutaneous transluminal coronary angioplasty versus coronary artery bypass graft surgery.[7–9] In 1980, Van der Linden suggested randomising participants to clinicians committed to performing different interventions in an area in which a conventional randomised controlled trial was possible.[10] Since that time, however, the expertise based design has been little used, even in areas where it has high potential (such as, surgery, physiotherapy, and chiropractic).

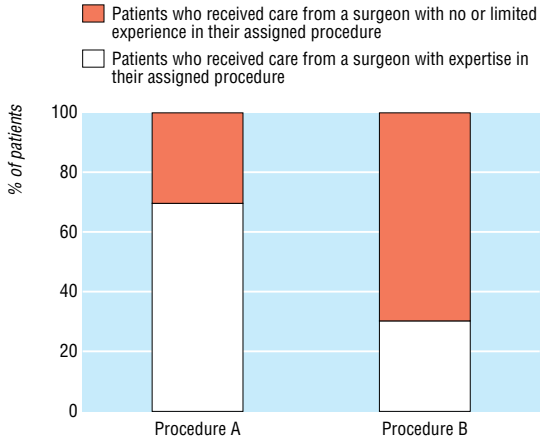
Problems with validity of conventional randomised controlled trials

Differential expertise between procedures

Because it takes training and experience to develop expertise in surgical interventions, individual surgeons tend to solely or primarily use a single surgical approach to treat a specific problem.[10 11] The restricted expertise that results can compromise the validity of conventional randomised controlled trials. For example, in a conventional randomised controlled trial, if surgeons with expertise in intervention A treat 70% of the patients in both groups A and B, and surgeons with expertise in intervention B treat 30% of those in both groups A and B, the trial results will be biased towards intervention A (fig 1). We will refer to this type of bias as differential expertise bias. The more disproportionate the number of cases being performed by surgeons with expertise in procedure A compared with surgeons with expertise in procedure B, the greater the impact of differential expertise bias on the trial results.

We estimated the potential for differential expertise bias through a survey of 139 surgeons in a large (>1000 patients) conventional randomised controlled trial comparing two surgical procedures for treating a tibial shaft fracture (reaming versus no reaming before insertion of an intramedullary nail).[12] Seventy four surgeons completed the survey. Significantly more surgeons had no or limited experience with the non-reamed procedure (which is more technically challenging) than the reamed procedure in the year before they joined the randomised controlled trial (table 1). The median number of cases surgeons performed in the year before randomised controlled trial participation was 12 reamed procedures and 2 non-reamed procedures (median difference 7 procedures, 95% confidence interval 5 to 11).

This example shows the potential for differential expertise bias. Three key considerations suggest that this problem is likely to be common in surgical trials. Firstly, trialists rarely, if ever, institute measures to ensure that the number of participating surgeons with expertise in each procedure is equal. Secondly, although some conventional randomised controlled trials try to reduce bias by requiring participating surgeons to perform a minimum number of both the experimental and control proce-



Differential expertise bias in a conventional randomised controlled trial in which 70% of surgeons are expert in procedure A and 30% in procedure B

Table 1 Experience of 74 surgeons with reamed and non-reamed procedure in the year before participating in the randomised controlled trial

	No (%) of surgeons	
No of cases	Reamed procedure	Non-reamed procedure
0	7 (9)	26 (35)
1-4	8 (11)	22 (30)
5-9	18 (24)	11 (15)
10-19	15 (20)	4 (5)
20-40	17 (23)	7 (9)
> 40	9 (12)	4 (5)

dures before participating in the trial, this measure is unlikely to eliminate bias because outcomes often improve with extensive experience with a procedure. Thirdly, even if these two problems are overcome, one of the procedures (let us say procedure A) may be more technically challenging. If this is the case, after doing the required numbers of unfamiliar procedures, surgeons who have to acquire expertise in procedure A will remain more technically challenged than those who have to acquire skills in procedure B. In this situation, the trial will be biased towards B, the less technically challenging procedure.

The second of these three issues deserves further discussion. All surgical interventions have a learning curve. Depending on how technically challenging the procedure is the number of cases required to achieve basic and advanced competence will vary. For many procedures, surgeons are unlikely to reach the plateau of the learning curve until they undertake many more procedures than are required to participate in a conventional surgical randomised controlled trial. For example, an observational study of 15 427 segmental colon resections showed that surgeons with five or more years' experience after surgical residency had lower risk adjusted death rates than surgeons during their first four years after residency training.[13] Similarly, an observational study of 5678 isolated coronary artery surgeries showed a progressive decrease in observed and risk adjusted mortality during the first four years of independent practice.[14] Therefore, differential expertise bias probably persists even when surgeons perform a minimum number of cases with the procedure in which they lack expertise before participating in a conventional randomised controlled trial. Exceptions may exist for procedures so technically simple that experienced surgeons acquire expertise after a small number of cases.

Potential problems related to unblinded surgeons

Surgeons participating in conventional surgical randomised controlled trials usually have opinions about the relative effectiveness of the procedures under investigation. Surgeons solely or primarily using procedure A probably do so because they believe it gives better outcomes. As a result, they probably expect and hope that the randomised controlled trial testing the outcomes of procedure A versus procedure B will affirm their belief.

Thus, surgeons, who are necessarily unblinded to the procedure they perform, may subconsciously systematically bias trial findings in a conventional randomised controlled trial. This bias may manifest itself through several mechanisms, including surgeons being more meticulous when performing one procedure than the other or differentially prescribing effective cointerventions.[2] Although it is preferable for independent blinded individuals to collect data and assess outcomes, in some trials it is done by the surgeons. When outcome evaluation is open to judgment and surgeons are involved in the process, they may differentially record data, repeat measurements, or interpret outcomes depending on whether a patient received procedure A or procedure B.[15]

We asked surgeons participating in the randomised controlled trial of different strategies for nailing tibial fractures whether they thought a reamed procedure or non-reamed procedure was superior before participating in the randomised controlled trial and at the time of the survey (that is, when about 900 patients had been randomised).[12] Surgeons rated their confidence about the superiority of the procedure they selected on a seven point scale, with 1 representing no confidence, 4 representing moderate confidence, and 7 representing extreme confidence. Before participating in the randomised controlled trial, 87% (95% confidence interval 77% to 94%) of respondents believed that a reamed procedure was superior and 86% of respondents indicated their confidence about the superiority of a reamed procedure was in the moderate to extreme range. After 900 patients were randomised, responses remained similar.

The results of this survey reflect the possible magnitude of treatment preference among surgeons participating in a randomised controlled trial comparing surgical procedures. This may lead to bias for reasons outlined above. As is the case with balancing expertise, trialists are unlikely to be able to ensure the

Table 2 Conditions for pragmatic and explanatory randomised controlled trials using conventional and expertise based methods

	Pragmatic trial	Explanatory trial
Conventional	All surgeons in routine clinical practice setting	Surgeons with advanced expertise in ideal clinical settings
Expertise based	All surgeons with expertise in procedure A or procedure B in routine clinical practice settings	Surgeons with advanced expertise in procedure A or B in ideal clinical settings

absence of a dominant treatment preference among participating surgeons.

Procedural crossovers

Our ability to determine if patients have a better outcome when they receive one of two procedures will be enhanced if patients actually receive the procedures to which they were randomised. If this is not the case because of procedural crossovers, the trial's ability to determine the true effect will be compromised.

We evaluated the number of crossovers in the reamed and non-reamed groups in the trial we surveyed. Of the 510 patients allocated to a reamed intervention, five received a non-reamed procedure, whereas of the 498 patients allocated to a non-reamed intervention, 40 received a reamed procedure ($P<0.0001$). These findings show the large potential for differential crossovers in a conventional randomised controlled trial. Procedural crossovers initiated by surgeons are more common when surgeons have limited experience with a procedure than when they have more extensive experience.[16–18] Except for the unlikely event that exactly the same number of participating surgeons have expertise in the experimental and control procedures (and both groups are allocated to perform an equal number of procedures A and B), there is a potential for differential crossover in the two arms.

Validity of surgical expertise based randomised controlled trials

In the surgical expertise based randomised controlled trial, patients are randomised to different surgeons with expertise in the relevant intervention. The first advantage of the expertise based randomised controlled trial is that surgeons will perform only the procedure in which they have expertise, avoiding the problem of differential expertise.

As in the conventional randomised controlled trial, surgeons in the expertise based randomised controlled trial will be unblinded. However, in the expertise based randomised controlled trial surgeons are likely to be subconsciously biased toward the procedure in which they have expertise. Consequently, the likelihood of differential procedural performance, cointerventions, data collection, and outcome assessment decreases. A third advantage of the expertise based randomised controlled trial is that procedural crossovers are less likely to occur because surgeons are doing the procedures with which they are most comfortable.[16–18]

Applicability of expertise based randomised controlled trials

Enrolling a large number of centres and surgeons will enhance the applicability of both expertise based and conventional randomised controlled trials. Surgical randomised controlled trials, which often include only a few surgeons at a single centre, frequently neglect this issue.[19 20] It is, however, encouraging that many of the expertise based randomised controlled trials performed when this was the only viable design have recruited many centres and included large numbers of surgeons.[7–9] If this capacity for multicentre trials with large surgeon samples can be extended to expertise based randomised controlled trials when both designs are possible, this will enhance their applicability.

If an expertise based randomised controlled trial shows that one procedure is superior to another, it does not follow that all surgeons with expertise in the less effective procedure and little or no experience in the more effective procedure can expect their patients to have better outcomes if they immediately start performing the superior procedure. Rather, if these surgeons acquire the same skill set and expertise as the surgeons who participated in the randomised controlled trial, they can expect their patients to have improved outcomes when they switch procedures.

The applicability of the results of a surgical randomised controlled trial further relates to whether a trial is an explanatory trial that uses only surgeons with advanced expertise in ideal clinical settings or if it is a pragmatic trial that uses surgeons with at least basic competence in routine clinical practice settings. Both conventional and expertise based randomised controlled trials can be explanatory or pragmatic trials (table 2).

Table 3 Surgical expertise based randomised controlled trials

Trial	Year of publication	No of patients	Surgical discipline	Surgical problem	Interventional groups evaluated	No of hospitals
Finkemeier[24]	2000	94	Orthopaedics	Tibial shaft fracture	1. Nail insertion with reaming (3 surgeons) 2. Nail insertion without reaming (3 surgeons)	1
Machler[25]	1999	120	Cardiac	Need for aortic valve replacement	1. Minimally invasive aortic valve surgery (2 surgeons) 2. Conventional aortic valve surgery (2 surgeons)	1
Wyrsch[26]	1996	39	Orthopaedics	Tibial plafond fractures	1. Open reduction and internal fixation of the tibia and fibula (2 surgeons) 2. External fixation with or without limited internal fixation (4 surgeons)	3 (both interventions at all sites)
Wihlborg[27]	1990	200	Orthopaedics	Femoral neck fractures	1. Rydell four-flanged nail (4 surgeons) 2. Gouffon pins (3 surgeons)	1
Phillips[28]	1985	138	Orthopaedics	Severe ankle fractures	1. Open reduction and internal fixation ASIF technique (1 surgeon) 2. Closed cast treatment or open reduction and internal fixation of medial malleolus (1 surgeon)	1

ASIF=Association for the Study of Internal Fixation.

Feasibility

Randomised controlled trials done in the past few decades show that conventional surgical randomised controlled trials are feasible. However, surgical research is still dominated by observational studies,[19 21] and general surgery interventions are half as likely to be based on the results of randomised controlled trials as medical interventions.[22 23] To investigate the feasibility of the expertise based randomised controlled trial in areas in which both designs are an option we searched three electronic bibliographic databases: MEDLINE (1966 to September Week 2 2003), EBM Reviews—Cochrane Central Register of Controlled Trials (issue 2, 2003), and EMBASE (1980 to 2003 Week 38). Complete listings of search strategies are available from the authors.

This search identified 162 citations, and two researchers independently reviewed each citation to determine if the report was a surgical expertise based randomised controlled trial. We contacted one of the trial authors in cases of disagreements. We identified five surgical expertise based randomised controlled trials that were conducted in an area in which both designs were a potential option.[24–27] Table 3 summaries the characteristics of these expertise based randomised controlled trials, showing that such trials are feasible in both emergency and elective surgery.

Surgical expertise based randomised controlled trials may be more feasible than conventional randomised controlled trials. Surgeons may be more willing to participate in an expertise based randomised controlled trial because they have to perform only the procedure for which they have developed expertise. Furthermore, surgeons do not have to do a minimum number of operations with the unfamiliar intervention before participating in the trial. This is likely to appeal to both surgeons and investigators and could prevent delays in starting trials.

The feasibility of expertise based randomised controlled trials will be enhanced when an intermediary physician randomises patients before they are seen by the participating surgeon. For example, in trials of patients admitted to accident and emergency, a surgical resident or an emergency physician could randomise the patient. Expertise based randomised controlled trials may also work well in group practices. Some cardiac surgeons, for instance, have group practices and accept patients for surgery without assignment to a specific surgeon. Patients at such practices could be randomised into an expertise based randomised controlled trial when they are accepted for surgery. Greater practical experience with the expertise based design is needed to evaluate the effectiveness of these approaches to recruitment.

Enrolling patients into an expertise based randomised controlled trial may be more challenging after a patient has seen a specific surgeon because randomising such a patient to another surgeon may be awkward. Even under these circumstances, an expertise based randomised controlled trial may prove feasible. For instance, in the largest randomised trial of treatment for subarachnoid haemorrhage, in which 2143 patients were randomised to neurosurgical clipping versus endovascular coiling,[29] patients were seen first by neurosurgeons, who then randomised them to neurosurgery or endovascular coiling by an interventional radiologist.

A potential challenge to undertaking an expertise based randomised controlled trial is that it is highly desirable to have surgeons from both treatment groups at all participating hospitals. This avoids the possible negative impact on recruitment resulting from patients having to travel to other hospitals and the potential confounding related to variations in non-surgical care in the different hospitals. If, in an expertise based randomised controlled trial, it is not possible to have surgeons from both treatment groups at all participating hospitals, the effect of centre is likely to vary depending on the nature of the procedure. The centre may not have an important effect for outpatient procedures, but the effect may be large for major inpatient procedures. For trials of major procedures, trialists will have to weigh the potential influence of differential expertise bias, surgeon unblinding, and differential crossover versus centre effect when deciding on the optimal trial design. The need to have two groups of surgeons on call presents a further challenge to undertaking expertise based randomised controlled trials of 24 hour acute surgical care.

A surgical expertise based randomised controlled trial must ensure satisfactory competence among the surgeons doing each procedure. Strategies to achieve this goal will include selecting qualified surgeons who have attained a specified level of post training experience, who fulfil requirements established by professional guidelines, or who have documented their expertise is at the plateau of the learning curve.

Ethics

Although the medical community accepts conventional surgical randomised controlled trials as ethical, some surgeons may have ethical problems with enrolling patients in a trial when they know they may have to do a procedure with which they feel inexperienced.[10 11] This problem does not arise in expertise based randomised controlled trials because surgeons perform only the procedures in which they have established expertise.

The consent process for expertise based randomised controlled trials can inform patients that, regardless of the procedure to which they are allocated, a surgeon with specific expertise will do the assigned intervention. Although rarely acknowledged, this is not the case for most conventional surgical randomised controlled trials. Obtaining consent for the reamed versus non-reamed trial in which we conducted our survey might have been problematic had patients been informed that they might be randomised to a procedure in which their surgeon was both inexperienced and sceptical of its effectiveness.

Applicability to non-surgical areas

The issues of validity, applicability, feasibility, and the ethics of expertise based randomised controlled trials relative to conventional randomised controlled trials are also relevant to the evaluation of interventions in many other fields including rehabilitation, behaviour modification, physiotherapy, chiropractic, radiation oncology, occupational therapy, and education. The issues are relevant to any area in which the skill set that a clinician requires to perform the experimental and control interventions varies importantly. The issues become increasingly germane when clinicians administering the interventions cannot be blinded.

We thank the study to prospectively evaluate reamed intramedullary nails in tibial shaft fractures (SPRINT) investigators for completing our survey and providing data related to their trial. We also thank D Altman, I Boutron, M Gent, C Meinert, K Schulz, and W Van der Linden for helpful comments and Neera Bhatnagar for the electronic database search.

Contributors and sources: PJD originated the idea for this paper and brought together all the authors to formulate and debate the points included in the text. PJD, MB, SDW, and GHG conducted the survey presented in this paper, PJD undertook the data analysis, and all authors contributed to the data interpretation. PJD wrote the first draft of the paper and all authors provided critical revisions to the manuscript. All authors have expertise in randomised controlled trial methodology and most have

Summary points

Questions remain about the use of randomised controlled trials to evaluate non-pharmacological interventions such as surgery

An alternative is to use expertise based randomised controlled trials, in which participants are randomised to clinicians with expertise in intervention A or intervention B

Interventions are performed only by clinicians with expertise in the procedure, which reduces both bias and ethical concerns

Expertise based randomised controlled trials may have greater applicability and feasibility than conventional trials

led large international randomised controlled trials, including surgical randomised controlled trials. PJD will act as the guarantor.

Funding: PJD is supported by a Canadian Institutes of Health Research senior research fellowship award, MB holds a Canada research chair in surgical outcomes, DJC holds a Canada research chair, SY holds an endowed chair of the Heart and Stroke Foundation of Ontario and is a senior scientist of the Canadian Institues of Health Research, GRN holds a Canadian research chair from the Social Sciences and Humanities Research Council. VMM is a Mayo Foundation scholar.

Competing interests: None declared.

1 Chalmers I. Unbiased, relevant, and reliable assessments in health care: important progress during the past century, but plenty of scope for doing better. *BMJ* 1998;317:1167-8.
2 Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
3 Russell R. Surgical research. *Lancet* 1996;347:1480.
4 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-8.
5 Wehbe MA. The prospective, randomized, double-blind clinical trial in orthopaedic surgery. *J Bone Joint Surg Am* 1998;80:1395.
6 Black N. Evidence-based surgery: A passing fad? *World J Surg* 1999;23:789-93.
7 Coronary angioplasty versus coronary artery bypass surgery: the Randomized Intervention Treatment of Angina (RITA) trial. *Lancet* 1993;341:573-80.
8 CABRI Trial Participants. First-year results of CABRI (coronary angioplasty versus bypass revascularisation investigation). *Lancet* 1995;346:1179-84.
9 Bypass Angioplasty Revascularization Investigation (BARI) Investigators. Comparison of coronary bypass surgery with angioplasty in patients with multivessel disease. *N Engl J Med* 1996;335:217-25.
10 Van der Linden W. Pitfalls in randomized surgical trials. *Surgery* 1980;87:258-62.
11 Rudicel S, Esdaile J. The randomized clinical trial in orthopaedics: obligation or option? *J Bone Joint Surg Am* 1985;67:1284-93.
12 Devereaux PJ, Bhandari M, Walter S, Sprague S, Guyatt G. Participating surgeons' experience with and beliefs in the procedures evaluated in a randomized controlled trial. *Clin Trials* 2004;1:225.
13 Prystowsky JB, Bordage G, Feinglass JM. Patient outcomes for segmental colon resection according to surgeon's training, certification, and experience. *Surgery* 2002;132:663-70 (discussion pp670-2).
14 Bridgewater B, Grayson AD, Au J, Hassan R, Dihmis WC, Munsch C, et al. Improving mortality of coronary surgery over first four years of independent practice: retrospective examination of prospectively collected data from 15 surgeons. *BMJ* 2004;329:421.
15 Devereaux PJ, Bhandari M, Montori VM, Manns BJ, Ghali WA, Guyatt GH. Double blind, you are the weakest link—good-bye! *ACP J Club* 2002;136:A11.
16 DeTurris SV, Cacchione RN, Mungara A, Pecoraro A, Ferzli GS. Laparoscopic herniorrhaphy: beyond the learning curve. *J Am Coll Surg* 2002;194:65-73.
17 Menon VS, Manson JM, Baxter JN. Laparoscopic fundoplication: learning curve and patient satisfaction. *Ann R Coll Surg Engl* 2003;85:10-3.
18 Lobato AC, Rodriguez-Lopez J, Diethrich EB. Learning curve for endovascular abdominal aortic aneurysm repair: evaluation of a 277-patient single-center experience. *J Endovasc Ther* 2002;9:262-8.
19 Bhandari M, Richards RR, Sprague S, Schemitsch EH. The quality of reporting of randomized trials in the Journal of Bone and Joint Surgery from 1988 through 2000. *J Bone Joint Surg Am* 2002;84A:388-96.
20 Solomon MJ, McLeod RS. Clinical studies in surgical journals—have we improved? *Dis Colon Rectum* 1993;36:43-8.
21 Pollock AV. Surgical evaluation at the crossroads. *Br J Surg* 1993;80:964-6.
22 Howes N, Chagla L, Thorpe M, McCulloch P. Surgical practice is evidence based. *Br J Surg* 1997;84:1220-3.
23 Ellis J, Mulligan I, Rowe J, Sackett DL. Inpatient general medicine is evidence based. A-Team, Nuffield Department of Clinical Medicine. *Lancet* 1995;346:407-10.
24 Finkemeier CG, Schmidt AH, Kyle RF, Templeman DC, Varecka TF. A prospective, randomized study of intramedullary nails inserted with and without reaming for the treatment of open and closed fractures of the tibial shaft. *J Orthop Trauma* 2000;14:187-93.
25 Machler HE, Bergmann P, Anelli-Monti M, Dacar D, Rehak P, Knez I, et al. Minimally invasive versus conventional aortic valve operations: a prospective study in 120 patients. *Ann Thorac Surg* 1999;67:1001-5.
26 Wyrsch B, McFerran MA, McAndrew M, Limbird TJ, Harper MC, Johnson KD, et al. Operative treatment of fractures of the tibial plafond. A randomized, prospective study. *J Bone Joint Surg Am* 1996;78:1646-57.
27 Wihlborg O. Fixation of femoral neck fractures. A four-flanged nail versus threaded pins in 200 cases. *Acta Orthop Scand* 1990;61:415-8.
28 Phillips WA, Schwartz HS, Keller CS, Woodward HR, Rudd WS, Spiegel PG, et al. A prospective, randomized study of the management of severe ankle fractures. *J Bone Joint Surg Am* 1985;67:67-78.
29 Molyneux A, Kerr R, Stratton I, Sandercock P, Clarke M, Shrimpton J, et al. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial. *Lancet* 2002;360:1267-74.

(Accepted 24 October 2004)

bmj.com 2005;330:88

Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main Street, West Hamilton ON, Canada L8N 3Z5
P J Devereaux *assistant professor*
Deborah J Cook *professor*
S D Walter *professor*
Brian Haynes *professor*
Geoffrey R Norman *professor*
Gordon H Guyatt *professor*

Department of Surgery, McMaster University
Mohit Bhandari *orthopaedic surgeon*
Claudio S Cinà *vascular surgeon*

UK Cochrane Centre, Oxford
Mike Clarke *director*

Department of Medicine, Mayo Clinic College of Medicine, Rochester, Minnesota, United States
Victor M Montori *assistant professor*

Population Health Research Institute, McMaster University
Salim Yusuf *professor*

Trout Research and Education Centre of Irish Lake, Markdale, Ontario, Canada
David L Sackett *director*

Departments of Medicine and Social and Preventive Medicine, University at Buffalo, Buffalo, New York, United States
Holger J Schünemann *associate professor*

Correspondence to: P J Devereaux, philipj@mcmaster.ca