

Special problems associated with surgical trials in inflammatory bowel disease

ROBIN S MCLEOD MD FRCSC FACS

RS MCLEOD. Special problems associated with surgical trials in inflammatory bowel disease. *Can J Gastroenterol* 1995;9(7):393-396. It is well accepted that the randomized controlled trial (RCT) is the best design for determining treatment effectiveness. Despite this, the surgical literature, particularly in the area of inflammatory bowel disease, is replete with case series while there is a paucity of RCTs assessing surgical interventions. There certainly is no lack of questions to be answered and there is a lack of consensus on many issues. Why then are there so few RCTs? Unfortunately, it appears that surgeons in general are more accepting of evidence from case series. However, it also appears that there are certain issues that are unique to, or at least occur more frequently in, clinical trials assessing surgical interventions. The issues of special concern in surgical trials can broadly be categorized as methodological and feasibility issues. Methodological issues include difficulties in standardization of the procedure, deciding on the timing of the trial and prevention of bias when blinding may be difficult or impossible. All these are challenges, but do not preclude the performance of RCTs. On the other hand, feasibility issues may pose more of a challenge. Patients may refuse entry, particularly if they view the two treatments as 'unequal' (eg, medical versus surgical therapy) and irreversible (as they usually are in surgery). Patient accrual may be difficult given the variable patterns of disease and indications for surgery. Finally, unlike medical therapies which are controlled by regulating agencies, surgical procedures can simply be performed and there may be little incentive, and even some disincentives, for the surgeon to participate in a trial. Despite these difficulties, better evaluation of surgical therapies is required. Where possible, RCTs should be performed, but in situations where this is not possible other trial designs must be employed that may lack some of the rigour of an RCT but that are more rigorous than the uncontrolled case series. (*Pour le résumé, voir page 394*)

Key Words: Clinical trials, Inflammatory bowel disease, Surgery

Departments of Surgery and Preventive Medicine and Biostatistics, University of Toronto, Toronto, Ontario

Correspondence: Dr RS McLeod, Suite 451, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5. Telephone 416-586-8347, fax 416-586-8644, e-mail mcleod@mshri.on.ca

This paper was presented at the Trends in Inflammatory Bowel Disease Therapy meeting, April 6 to 9, 1994, held in Victoria, British Columbia. This paper has also been published in Sutherland LR, et al, eds. *Inflammatory Bowel Disease: Basic Research, Clinical Implications and Trends in Therapy*. Boston, Dordrecht and London: Kluwer Academic Publishers, 1994

DECISIONS ABOUT THE EFFECTIVENESS of a treatment may be based on careful observations. This approach may be appropriate if the condition is universally fatal or has a high mortality rate, and the treatment effect is large. In these situations, the value of the treatment is self-evident and it is unlikely that other factors are responsible for differences in outcome. However, in modern surgical practice these situations are uncommon. Technological developments or surgical interventions lead to small improvements in survival, symptoms or quality of life, and it is necessary to control for extraneous factors to be certain that the observed difference is indeed due to the treatment. It is accepted that the randomized controlled trial (RCT) is the best trial design to determine treatment effectiveness. Several attributes of the RCT minimize the risk of random error and systematic (bias) error, and thus minimize the risk of making an incorrect conclusion about treatment effectiveness. First, subjects are randomly allocated to two groups: a treatment group in which the new treatment is tested and a control group in which standard therapy or placebo is administered. Second, the interventions and follow-up are standardized and performed prospectively. Thus, both groups should be similar in all respects except for the single factor being stud-

Problèmes spéciaux associés aux essais chirurgicaux dans la maladie inflammatoire de l'intestin

RÉSUMÉ : De nos jours, l'essai contrôlé randomisé (ECR) est généralement reconnu comme le meilleur modèle pour déterminer l'efficacité d'un traitement. Malgré cela, la littérature portant sur la chirurgie, particulièrement dans le domaine de la maladie inflammatoire de l'intestin, abonde de séries de cas, mais fait preuve d'une pauvreté manifeste en fait d'ECR pour évaluer les interventions chirurgicales. De nombreuses questions restent en suspens. Pourquoi y a-t-il si peu d'ECR ? Malheureusement, il semble que les chirurgiens acceptent davantage les preuves tirées des séries de cas. Toutefois certains aspects restent uniques, ou du moins plus fréquents, lors des essais cliniques qui évaluent des interventions chirurgicales. Les principales préoccupations des essais chirurgicaux peuvent, de façon sommaire, être catégorisées comme des problèmes de méthodologie et de faisabilité. Pour ce qui est de la méthodologie, la standardisation de l'intervention pose des problèmes, ainsi que le choix du moment où elle a lieu, sans compter qu'il peut être difficile, voire impossible de prévenir les biais lors d'interventions à l'insu. Voilà des obstacles importants, qui n'empêchent toutefois pas la réalisation d'ECR. Par ailleurs, les questions de faisabilité peuvent être plus difficiles à régler. Les patients peuvent refuser de participer, particulièrement s'ils perçoivent les deux traitements comme non équivalents (par ex., médicamenteux versus chirurgicaux) et irréversibles (comme c'est habituellement le cas lors de chirurgies). Les progrès des patients peuvent être difficiles à mesurer, compte tenu des manifestations variables de la maladie et des indications de la chirurgie. Finalement, contrairement aux traitements médicamenteux qui sont contrôlés par des agences de réglementation, les interventions chirurgicales sont simplement effectuées et le chirurgien ne verra pas d'avantages à y participer, au contraire. Malgré ces difficultés, il faut se donner les moyens de mieux évaluer les traitements chirurgicaux. Autant que possible, il faut choisir des ECR, mais dans les situations où cela n'est pas possible, d'autres modèles d'essais doivent être employés, qui n'auront peut-être pas la rigueur de l'ECR, mais qui seront plus rigoureux que les séries de cas non contrôlées.

ied. Not only does this guard against differences in factors known to be important, but it also ensures that there are not differences due to other factors that have not yet been identified.

Although the RCT has been accepted by physicians, there is a relative paucity of RCTs published in the surgical literature. In a review of three surgical journals, *British Journal of Surgery*, *Diseases of the Colon and Rectum* and *Surgery*, only 6% of all clinical articles published in 1980 were found to be RCTs (1). Only 16% of articles reported comparative studies. Unfortunately, the proportion of RCTs published in these journals in 1990 was unchanged.

The lack of RCTs does not appear to be due to the lack of controversy regarding surgical procedures. In the surgical management of inflammatory bowel disease (IBD), there is controversy regarding the necessity of a mucosectomy and a defunctioning ileostomy

when performing a pelvic pouch procedure. The role of stricturoplasty in the management of Crohn's disease is unresolved. The relative merits of surgical and medical treatments have been debated. For instance, do patients with limited terminal ileal Crohn's disease do better with early surgery or with medical therapy?

In this report, problems precluding or hindering the performance of RCTs testing surgical therapies are categorized and discussed under the general headings of methodological issues and feasibility issues.

METHODOLOGICAL ISSUES

The methodological issues that must be addressed by an investigator studying a surgical procedure are similar to those encountered by the investigator testing a medical therapy (Table 1). However, because of the nature of surgery, these issues may present more

TABLE 1
Methodological issues of concern in surgical trials

1. Placebo effect of surgery
2. Blinding
3. Standardization of the procedure
4. Timing of the trial
5. Obsolete procedure

of a problem to the surgical investigator. While these issues make the performance of a trial more difficult, none of them precludes the performance of an RCT.

Placebo effect of surgery: Any intervention may have a placebo effect and thus placebo medication is administered to the control group in medical trials. While this is possible in a trial assessing medical therapy, it is usually not feasible or ethical to perform a sham operation when testing a surgical procedure. It appears, also, that the placebo effect of surgery may be greater than with medication. For example, Dimond et al (2) reported a series of 18 patients in which 13 had ligation of their internal mammary artery for coronary artery disease and five had a sham operation. All of the latter group reported subjective improvement in their symptoms.

The placebo effect may be difficult to overcome, especially in situations where patient assessment is the important outcome. An example of such a situation is a trial comparing surgery with medical management in Crohn's disease where quality of life is the primary outcome. In this situation the patient's assessment of quality of life may be altered by the fact that the patient had a surgical procedure. In situations where outcome is measured with 'hard' outcomes (eg, development of a cancer, death) or where outcome is assessed by a blinded observer, the placebo effect is less of a concern.

Blinding: The issue of blinding is of importance because of the placebo effect of surgery. Blinding of the patient and investigators may be difficult if not impossible if a surgical therapy is being compared with medical therapy. It is less difficult if two surgical therapies are

being compared, although even in this situation the scars or the side effects of the two procedures may differ, so the patient is aware of the procedure performed.

Blinding is critical in procedure assessment. If blinding does not occur, there may be a bias in favour of one of the treatment arms. In some situations, even though the patient and investigator are unblinded it may be possible to have the outcome assessed by a blinded assessor. In circumstances where the primary outcome is a change in symptoms or quality of life as assessed by the patient, it may be possible to measure a 'hard' outcome in addition, and if it correlates with the patient's assessment then there is less concern about the possibility of bias due to a placebo effect. An example might be to correlate the endoscopic appearance with the patient's symptoms in patients with recurrent Crohn's disease.

Standardization of the procedure: Standardization of the procedure is difficult because surgeons may vary in their experience with and their ability to perform a surgical procedure; there may be individual differences in performing the procedure; and there may be technical modifications as the procedure evolves. There are strategies, however, to ensure that critical aspects of the procedure are standardized. These include ensuring that all surgeons agree on the performance of these aspects of the procedure, providing teaching sessions and obtaining documentation that the procedure has been performed satisfactorily (for example, resection margins may be assessed to ensure that they are free from macroscopic evidence of Crohn's disease). Finally, one can limit the number of surgeons participating in the trial. That decision may vary depending on the complexity of the surgical procedure and the concern about the generalizability of the results, because limiting the number of surgeons may decrease the generalizability.

Timing of the trial: Chalmers (3) has argued that the first patient in whom a procedure is performed should be randomized. Most surgeons would argue, however, that there is a learning curve

in any procedure and that modifications to the technique are made frequently at its inception. By including these early patients, one would bias the results against the new procedure. On the other hand, it may be difficult to initiate a trial when the procedure is widely accepted by both the patient and surgical communities.

Obsolete procedure: There are concerns that, given the time necessary to accrue patients and complete a trial, the results of the trial may not be relevant because modifications may have been made to the procedure or other procedures may have been developed in the interim. This argument, however, is equally pertinent to medical trials where there may be more information available on the optimal dose or new drugs may have become available since the start of the trial.

FEASIBILITY ISSUES

While methodological issues may challenge the investigator, feasibility issues may preclude the performance of an RCT. These issues include uncommon condition, patient preferences and surgeon preferences.

Uncommon condition: The prevalence of IBD, including both ulcerative colitis and Crohn's disease, in North America is approximately 50 cases per 100,000 population (4). Although a large proportion of these patients may require surgery at some time, there may be relatively few patients who fit the entry criteria for a particular trial given the various sites, severity and manifestations of the disease. For instance, trials assessing surgery for perianal disease, segmental colitis and gastroduodenal Crohn's disease would be difficult to perform because of the relative rareness of these sites of disease. Similarly, the indications for stricturoplasty are relatively uncommon.

Patient preferences: Patient preferences may preclude the performance of an RCT. In a medical trial, patients who are randomized to one treatment arm have the possibility of being offered the effective treatment at the conclusion of the trial. Thus, they may benefit directly from the results of the trial. For example, following a trial assessing

maintenance therapy for either Crohn's disease or ulcerative colitis, patients may be eligible to receive the more efficacious treatment. Surgical procedures, however, are usually permanent, and so there is little chance of the patient receiving the more effective treatment at the conclusion of the trial and thus benefitting from the results of the treatment. As a result, patients may be reluctant to enter a trial particularly if the treatments are viewed as being unequal. For example, if a trial were performed to compare recurrence rates following total proctocolectomy or colectomy and ileorectal anastomosis in Crohn's disease, patients randomized to the total proctocolectomy group would be left with a permanent ileostomy regardless of the trial's results. In such situations, if there is genuine equipoise about the effectiveness of the two treatments, it is likely that patients will have a preference for one treatment and refuse randomization. However, there are trials that have compared surgery with a medical therapy (EC-IC Bypass procedure) (5) and surgical procedures of differing magnitude (lumpectomy versus mastectomy for breast cancer) (6). Thus, more research is required to determine whether patient preferences preclude the performance of RCTs and what factors affect their decision-making. Physicians' opinions on the reasons for patients not participating in trials may not necessarily reflect patients' opinions.

Surgeon preferences: Perhaps one of the major reasons for the lack of surgical trials is the greater acceptance by surgeons of results from case reports and series. There is little incentive for surgeons to do RCTs and perhaps some disincentives. Unlike medications, surgical procedures are not regulated by a regulatory agency such as the Health Protection Branch (HPB) in Canada or the Federal Drug Administration (FDA) in the United States. Surgeons can perform a new operation with little constraint even from their hospital or local ethics committee. On the other hand, if a trial is initiated, then the surgeon must seek approval from the ethics committee and consent from the pa-

tient. A good example of the lack of regulation has been the explosion of laparoscopic techniques in general surgery without any substantive trials having been performed. Economic disincentives may also make participation in trials unattractive. In the case of laparoscopic cholecystectomy, surgeons who failed to perform this procedure suddenly began noticing the effects on their practices. Many surgeons felt it necessary to begin performing this procedure even though the results and complications were not fully evaluated. Another disincentive is the lack of funding for surgical trials. Because of HPB and FDA regulations that require investigation of drugs before their release, industry is interested in providing funding to perform medical trials. These sources are generally not available for funding surgical trials and funding must come from other sources.

Finally, the lack of professional acclaim for clinical as opposed to laboratory research for academic surgeons may be another factor.

CONCLUSIONS

Should more RCTs be performed to test surgical therapies in IBD? The answer is unequivocally yes. However, these trials must address important, clinically relevant questions, have appropriate outcome measures and be methodologically sound. In most instances this will mean multicentre trials. There must be commitment from surgeons and adequate funding must be available. Finally, there are some situations where RCTs cannot be performed, and in these situations trial designs that are less rigorous than the RCT but more rigorous than the case series may need to be adopted.

REFERENCES

1. Solomon MJ, McLeod RS. Clinical studies in surgical journals – Have we improved? *Dis Colon Rectum* 1993;36:43-8.
2. Dimond EG, Kittle CF, Crockett JE. Evaluation of internal mammary artery ligation and sham procedure in angina pectoris. *Circulation* 1958;18:712-3.
3. Chalmers TC. Randomization of the first patient. *Med Clin North Am* 1975;59:1035.
4. Donaldson RM. Crohn's disease. In: Sleisenger MH, Fordtran JS, eds. *Gastrointestinal Disease. Pathophysiology, Diagnosis, Management*. Philadelphia: WB Saunders Co, 1983:1088.
5. The EC/IC Bypass Study Group. Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke: Results of an international randomized trial. *N Engl J Med* 1985;313:1191-200.
6. Fisher B, Redmond C, Fisher ER, et al. Eight year results of an RCT comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. *N Engl J Med* 1989;320:822.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

