

RANDOMIZED CONTROLLED TRIALS IN INTERVENTIONAL SPINE: PERILS AND PITFALLS

David A. Lenrow, MD, JD and Larry H. Chou, MD

It is the responsibility of clinician investigators to advance clinical knowledge and specifically its application to patient care. Randomized controlled trials remain near the top of the hierarchy of evidence based medicine. The acquisition of evidence based medicine by means of randomized controlled trials presents general difficulties and additional pitfalls specific to interventional treatments. The nature of interventional procedures makes the performance of these studies more difficult to plan and execute. To generate clinically useful research results requires an understanding of the

mechanics of performing studies and the reporting of methodologies to ensure appropriate interpretation. Placebo arms and sham interventions present serious ethical issues, which must be analyzed on a case by case basis. The conscientious researcher must always abide by the principles of ethical research and the tenets of human subject protection.

Key words: Randomized controlled trials, CONSORT statement, evidence based medicine, medical ethics

Evidence based medicine continues to evolve as a driving force in clinical practice. This is evident in what third party payers are willing to reimburse and in how and when new technologies and interventions are approved for use in clinical practice. How this hierarchy of evidence is evaluated is somewhat controversial but most clinical scientists continue to believe that randomized controlled trials (RCT) are the gold standard for the evaluation of new treatments (1-4). It has become increasingly clear that the methodology used and reported is critical to how we evaluate specific randomized controlled trials and how or whether we integrate the outcome of specific studies into our clinical practice (5, 6).

RCTs are the gold standard of clinical studies for several reasons. Randomization is essential to remove confounding variables that might otherwise weaken the usefulness of a particular study. Blinding of clinician and subject removes the problem of preconceived notions of subjects or investigators from systematically introducing bias into the outcomes. The goal is to detect the association between

a specific intervention and a predefined outcome.

There are several pitfalls specific to RCTs involving interventions that are technical in nature. The most commonly studied and published of these RCTs are those concerning surgical techniques (1, 7). Interventional spine is analogous to surgical treatment in that it requires a technical procedure that is dependent on the experience and skill of the clinician. Patients often choose a clinician because of the clinician's experience and expertise in a particular procedure, particularly if it is a new procedure. These and other issues make randomization more difficult than in medical interventions such as pharmacologic treatment.

The performance of RCTs in interventional spine care, as in surgical interventions, presents difficult problems. They include ethical issues, clinician and patient preference problems, difficulties with obtaining an adequate sample size and issues of blinding (1). All RCTs need to be evaluated with respect to their methodology, which therefore should be reported so that others can determine the validity of the study and its clinical usefulness. These issues must be balanced while managing the conflict between the treating physician and clinical investigator who are often the same physician. Obtaining scientific evidence remains paramount to the investigator but acting in the best interest of the patient is required of the physician.

HIERARCHY OF EVIDENCE BASED MEDICINE

The value of a particular research study to the clinician is dependent upon the direct application of the results to medical practice (5, 6). In evaluating the medical literature, as well as in selecting submissions for publication, it is essential to understand the hierarchy of evidence (2, 5). It is obvious that the strength of studies varies considerably; some have little if any evidentiary value while others point strongly to the benefit of a particular treatment or device. The clinical inferences that we can draw from a study depend upon the study design and how the methodology of the study was reported. A stellar randomized, controlled, double blind study that was poorly reported should fail to get the recognition that it would otherwise deserve.

The strength of evidence in research methods has been likened to a pyramid. The top of the pyramid is the most clinically relevant and the bottom the least clinically relevant. Systematic Reviews and Meta-Analyses have the most clinical relevance. A systematic review is a comprehensive survey of studies on a particular topic; only studies with the highest level of evidence are included. Meta-Analysis is done by combining the results of several studies providing that the studies results are similar enough, so that the results can be statistically combined. Systematic reviews if done properly have greater

From Department of Rehabilitation Medicine, University of Pennsylvania Medical Center, Philadelphia, PA and Radnor, PA. Address correspondence: David A. Lenrow, MD, JD, University of Pennsylvania Medical Center, 3400 Spruce Street, 5 West Gates Building, Philadelphia, PA 19104. E-mail: david.lenrow@uphs.upenn.edu
Funding: No financial support was obtained for this study.

weight than single studies.

Randomized controlled double blind studies are the next highest on the evidentiary scale. This method is the standard way to analyze the effectiveness of two different therapies or a new treatment. Further down the scale of evidence are Cohort Studies, then Case Control Studies, then Case Series, Case Reports, then Ideas, Opinions, then Animal Research, then In Vitro Research. The importance of these investigations is not questioned as far as the attainment of scientific knowledge is concerned but their direct application to clinical practice is not as strong as RCTs.

GENERAL ISSUES IN RCTs

The methodology of an RCT is (8):

1. To randomize subjects to one of two or more 'Treatment' groups.
2. One of these groups serves as the control to which the other(s) is compared. A 'Treatment' group may be a placebo intervention in some studies.
3. The study is a prospective design and usually is double blind.

The advantages of an RCT design are related to their randomization and blinding (7-9).

1. Randomization removes selection bias since the investigator has no control of which subject enters which treatment arm of the study.
2. Variables that may lead to misinterpretation of results should be equally distributed among the treatment arms of the study. Statistically they should be distributed randomly among the study groups. The ideal study for treatment interventions isolates the treatment. Randomization serves to minimize the influence of other factors.
3. An intention to treat model of analysis should be used to prevent skewing the results in favor of the treatment group. In addition, when evaluating the strength of a study, assume all in the treatment group who drop out did poorly and all in the control group did well. If this changes the results then the study loses power (6).
4. This results in strong evidence that the outcome of the study is secondary to the intervention being studied and not a third or confounding variable that was not equally distributed among the study groups.

PITFALLS IN RCTs

The difficulties in conducting a scientifically valid and clinically useful RCT include ethical concerns, difficulty in randomization, blinding issues, and patient accrual problems. The above categories of difficulties are common to RCTs in general, whether the study concern a drug treatment, surgical or interventional procedure. There are several impediments which are much more prevalent in RCTs involving interventional treatments which will be discussed in a later section.

Ethical issues in research have come into the public view in recent years. Many of the concerns stem from the conflict between treating physician and clinical investigator. Their roles and duties to the patient/subject are at times at odds. To perform a RCT the investigator often believes that one treatment arm of the study is superior to another. This is allowed for an investigator but breaches the ethical duty of the treating physician to act in the best interest of the patient. It has been suggested that to perform a RCT requires a null hypothesis (7). Both treatments should be viewed as equivalent. If the investigator believes one is superior does that make the study unethical? If it does it would impede the performance of many interventional studies. Studies are often pursued when a new intervention is believed to be superior to other available treatments. Subjecting subjects to interventions or treatments that are known to be sub optimal compared to other available treatments is unethical. If we know a treatment is superior or a technique safer we cannot justify using those as treatment arms in an RCT. That would be a violation of research ethics. Merely believing that a particular treatment is superior to another does not make it unethical to use as a treatment arm in an RCT. Belief by an investigator does not rise to the level of evidence required to constitute a violation of research ethics. If a treating physician believes that a particular intervention or treatment is superior their duty is to recommend that particular treatment to the patient. In research if it is known to produce a sub-optimal result it may not be used as a treatment arm in the study.

Clinical equipoise has been suggested as a solution to the requirement of the null hypothesis in an RCT (10). Clinical equipoise requires only that others believe that the other treatment is superior. So even if the investigator believes that one of

the treatments is inferior, as long as others take the contrary view, the study is ethical. That yields essentially the same outcome as saying that belief and knowledge are by nature different. Belief requires only that the individual hold a particular opinion which treatment is superior. Knowledge requires some degree of proof and consensus among investigators and likely clinicians. Of course the issue with evidence based medicine as well as ethics is a matter of degrees. When does a belief become knowledge? If a treatment is widely accepted as inferior or unsafe then it should not be used as a treatment group in an RCT. It then becomes unethical. An obvious example was the experimentation using high dose oxygen on premature infants as a control group and low concentration oxygen as the treatment group (11). This was unethical since it had been previously shown and was widely accepted among clinicians and investigators that high concentration oxygen presented a high risk of adverse affects.

Other ethical issues include removing the ability of the physician to use personal medical judgment to assign treatment leading to the treatment no longer focusing on the good of the individual patient. These are only issues if the researcher is the treating physician. If they are not the treating physician but only the investigator these are not issues. If investigators were required to assign treatment based on the individual subjects best interest (what they believed would be the best for the particular subject) we would not be able to perform Randomized Controlled Trials nor would blinded trials be allowed, unless we believe that in RCTs the null hypothesis is true. In interventional studies we believe that it is usually not the case. Investigators in most studies favor one arm of the study as a better treatment.

The use of placebos as control treatment groups has been common and was believed to facilitate the attainment of new clinical knowledge. With the revision in 2000 of the Declaration of Helsinki by the World Medical Association (12), placebo groups have fallen into disfavor among some countries research regulators. This revision makes it unethical to use placebos in research if there is a known treatment or intervention for that particular disease or illness being studying. If there is no known treatment, placebos are acceptable groups in RCTs. Sub-

jects are to receive at the minimum the best-known treatment in all treatment groups. The issue of placebos remains unsettled in the United States (13). The controversy continues with groups, who have adopted the tenets of the revised Declaration of Helsinki, believing that to protect human subjects all must receive at minimum the best-known treatment. If not the research is unethical.

This leads to the use of active controls in RCTs. This has been termed the active control orthodoxy (14). The comparison is between the accepted treatment and the new treatment. In reality that is what clinicians want to know. Is this new intervention superior to the standard of care? If the answer is yes then the new treatment should be adopted, if not there is no change in clinical practice. The use of placebo controls may allow for the statistically significant results faster and with fewer subjects than if an active control is used. This argument only applies if the difference between outcomes is greater between the placebo and the investigational treatment, when compared to the difference in outcome between the investigational treatment and the active control. In other words the greater the delta in outcome between the control and the experimental treatment the faster the study reaches statistical significance and the less subjects that are exposed to the less effective treatment.

If the accepted current treatment is used as an active control and that treatment has a negative effect on outcome, the delta between that and the experimental treatment (if it has a positive effect) will be greater than between treatment and placebo. In that instance significant results will be attained faster and with a smaller number of subjects. The same logic applies if the experimental treatment has a negative effect the delta between that group and the active control with a positive effect will be greater than if a placebo control were used. So the use of placebos remains an individual decision in RCTs in the U.S. Their use may speed the acquisition of knowledge and prevent the use of ineffective treatments or they may have the opposite effect, depending on the specifics of the study.

The middle ground approach is the reasonable approach until the FDA and OHRP takes a stand on this issue. The recent study on Knee Arthroscopy, in the

New England Journal of Medicine, had a placebo control of sham surgery (15). This was approved and funded by the Veterans Administration Hospital. The individual analysis should be determined by the scientific validity of the study (16, 17). If placebo controls are essential to the study and the potential benefit substantiates the risk, placebo controls are justified. If standard of care controls provide the same scientific validity they should be used. The duty of the investigator is not to provide the best individual care to the subject. If that were required we would not be able to evaluate new and unknown treatments. Research is not governed by the best interest of the subject. The Common Rule requires that the science be valid and the risk be acceptable to society (18). This confusion again stems from the different ethical duties of physicians and investigators and the conflicts when these roles coalesce. Separate roles for treating physician and investigator resolve many of these ethical issues.

Other general issues in RCTs worth mentioning are the high cost and long time span for most studies, inadequate blinding, difficulty recruiting, lack of generalizability of results, failure to use intention to treat analysis and poor reporting. Length of time and cost and difficulty recruiting are essentially self-explanatory. Studies are often not performed because of funding issues, cost to investigator in time and protracted time of study to achieve significance or to assess appropriate outcomes. Inadequate investigator blinding can lead to a large bias in the results changing the outcome of the study (8). Failure to use intention to treat analysis detracts from the value of the study's results. In evaluating the data, if you make the assumption that all who dropped out of the treatment group did poorly and that changes the conclusion, the study lacks power.

Lack of generalizability of results stems from inclusion and exclusion criteria that are too strict. Trials involve the use of a subpopulation of potential patients that may benefit from the particular treatment. The population used in the RCT is chosen to represent, as best as can be projected, the target population of the treatment. The outcome of the RCT can only be applied with any scientific basis to the general population of patients that fulfill the specific inclusion and exclusion

criteria of the study. We must be careful not to narrow our inclusion criteria to the point where we leave only limited value for our results.

Poor reporting leads to improper interpretation and clinical application of studies (19). This may be an ethical breach on the part of both investigator and editor. The publication of studies with inadequate description of methodologies can make the study valueless in clinical application. This would make the scientific value of the study very low and hence the level of acceptable risk would be very low. The analysis of the ethics of performing the study then shifts against its performance. If it is performed and improperly reported the study becomes unethical because the risk to subjects remains but there is no benefit to society.

Uniform reporting of methodology of RCTs has been suggested as the solution to widespread inadequate reporting (19, 20). The CONSORT Statement or the Consolidated Standards of Reporting Trials was developed as one solution. The goal is to have a uniform method of reporting to make evaluation of the methodology and therefore the evidentiary value of RCTs easy to evaluate. These rules were to apply to investigators and be applied by editors of journals. CONSORT was published in 1996 and revised and published in 2001. It consists of a 22-item checklist of requirements for inclusion in reports of RCTs. The second part consists of a flow diagram of subjects through each arm of the study. This allows for easy analysis of where patients dropped out of the study and why. Using this diagram an evaluation of the analysis with specific attention to intention to treat analysis is readily apparent. Much of what is in the CONSORT checklist is already standard in RCT reporting. The information is organized in the checklist and some items are not readily apparent in many studies. Some examples of items often omitted from the methodology section include the specifics of randomization and the specifics of who and how parties were blinded. Application of CONSORT by editors would standardize requirements for publication and ensure that readers have the information necessary to evaluate the evidentiary value of the study and appropriately integrate the outcome into clinical practice (20,21).

SPECIAL ISSUES IN RCTs OF INTERVENTIONAL TREATMENTS

Trials of interventional treatments present additional difficulties, which are specific to the nature of the treatment (1,7). The best-studied interventional treatments are surgical procedures. When reviewing the literature it is apparent that there are far fewer RCTs involving surgical procedures than there are of non-interventional treatments. There are several issues that make trials of interventional treatments more difficult to design and implement. For the most part they stem from issues related to the physician investigator and the patient turned subject. Procedures entail technical expertise and hence a learning curve and variability in proficiency. It is difficult to perform a study using interventional techniques unless the same person performs the treatment or the treatment is somehow standardized. This becomes particularly difficult with new techniques, which are often the focus of investigation. When a new procedure is first used the skill of the physician at performing this intervention is in evolution, they are in the beginning stages of their learning curve. Their proficiency improves with experience and the technique is modified in accordance with their experience. These modifications and improvements make it difficult to accurately compare outcomes of new interventions between investigators and over time with the same investigator (performing the technique). So with technical treatments we would expect that the intervention would be honed over time and practitioners would improve in their proficiency. This would lead to improved outcomes in the study treatment group versus the control group (22). This in contradistinction to drug studies where there is no technique involved. In drug studies where length of use and extent of use increase, side effects become more evident. This leads to better results earlier in medication studies.

Patient preference is a significant impediment to RCTs in interventional procedures. In this day of the Internet and mass media coverage of new medical techniques patients often seek particular treatments. Their choice is most often not scientifically based, but is based on their limited exposure to information and often other people's uneducated preferences. Less invasive procedures with short recovery are more palatable to

many patients and they seek these treatments without regard for the scientific efficacy or safety of the particular treatment. Popularity of a new procedure often is not based on any conclusive scientific evidence. These patients, when seeking medical care, will not be willing to be randomized to the control arm or the current standard of care. They want the new, yet unproven procedure. This makes it very hard to perform RCTs on new popular techniques. An example of this difficulty in the surgical literature is the advent of laparoscopic cholecystectomy. In the early 90's investigators could not adequately recruit subjects to the control arm (open cholecystectomy) because subjects only wanted the new laparoscopic procedure, even though it had not yet been scientifically studied (23). These studies are often forced to utilize historic controls, which have less evidentiary value than RCTs.

Physician preference makes trials involving interventional treatments more difficult to perform. Physicians who perform new techniques often have a preference for that technique. When patients come to them for treatment it is difficult for the physician to randomize these individuals as subjects to the new and control treatments. Physicians appear to become invested in performing specific techniques more than in prescribing particular medications. Hands on treatments are more individualized and a more intimate treatment by an individual physician than are non-interventional treatments.

Patients often are drawn to a particular physician because of their expertise in a particular procedure. When that physician is known for the procedure it is unlikely that they will perform RCTs and be able to recruit subjects to the control arm of the study. Patient preference in choosing a physician by their expertise in a procedure makes recruitment for these RCTs difficult. These conflicts make it difficult to perform RCTs. As the patients perceived difference between treatment alternatives increases it becomes harder to recruit for a trial. These differences need not be real. An example is the current trend towards less invasive techniques of treating spinal disorders. These techniques have quickly gained popularity with the public, faster than outcome studies are performed. This leads to patient withdrawal from the control arm of studies and difficulty with recruitment.

These difficulties are harder to surmount after the public accepts a procedure. It is the investigators and physicians responsibility to evaluate treatments for efficacy and safety prior to offering as routine treatment. Once interventions become the standard of care in treatment, it becomes increasingly more difficult to evaluate them using RCTs.

POTENTIAL BIAS IN RCTs

In utilizing an RCT to base clinical treatment it is important to recognize that there are areas of potential bias. This can occur at the level of the methodology of the specific study or it may occur in the broader context of publication. Lack of blinding leads to investigator bias; a 17% bias has been shown with inadequate investigator blinding (8). Adequate blinding is important for individuals with different roles in the study depending on the particular study (24). Publication bias, language bias and time lag bias relate to bias in the medical literature as a whole (17). Publication and language bias often go together. Many publications are more apt to accept studies for publication that have positive rather than negative outcomes. This can lead to an imbalance in the literature, in favor of positive outcomes for a particular treatment, if negative outcomes are not published. The threshold for evaluation of studies and their acceptance for publication should not depend on whether the outcome was positive or negative. The criteria for relevance and methodology should be the same. Language bias results from the tendency of publications in a particular language to be more likely to accept either positive or negative outcome studies. This leads to skewing of the literature in that particular language with respect to that particular intervention or treatment. An example is that English language journals are more likely to accept positive outcome studies where certain other language journals more favorably view negative outcome studies. Perhaps economics plays a role in how different cultures present scientific studies.

Time lag bias relates to the time between submission of a publication and its eventual journal appearance. Negative studies and positive studies often vary in time to publication. This again leads to a bias in the overall literature with respect to particular interventions. These potential areas of bias make it difficult to accu-

rately evaluate the efficacy of a new treatment. It is to the benefit of patients and clinicians to establish uniformity in how editors and publishers evaluate the scientific value of studies and determine which to publish. Having a broad spectrum of publications does not necessarily correct for the bias in a single publication. Clinicians, investigators and editors need to embrace these issues and forge ahead to facilitate the advancement of evidence based medicine.

CONCLUSIONS

The pursuit of clinically relevant knowledge and new treatments is based on a hierarchy of evidence. Randomized controlled trials remain at or near the top of the evidentiary pyramid for clinical practice. RCTs are important if fields of interventional medicine are to advance. They are more difficult to perform when interventional techniques are being evaluated but these difficulties can be surmounted. This requires that investigators and treating physicians have clear and distinct roles. Subjects and patients must have a separate advocate for treatment and an investigator who performs the study. RCTs must be performed with rigorous standards of design, implementation and reporting. Investigators, editorial boards and clinicians should apply uniform standards of methodology reporting and evaluation. The CONSORT Statement provides a starting point for this systematic process.

The ethics of research are not the same as those in clinical practice. If we erroneously make this assumption it becomes difficult, if not impossible, to perform research on investigational procedures where the outcome is unknown. The nature of clinical research is to obtain new knowledge and evaluate treatments whose efficacy is unknown. That is always in conflict with the duty of a treating physician if there is a known treatment. Their duty is to act in the best interest of the patient, not to act in the interest of science or society at large. The use of placebos, where there is a known treatment, remains an unsettled issue. Perhaps these studies should be evaluated as all others; are they scientifically valid and is the risk acceptable to society. Could they be otherwise performed with less risk to individuals?

Medicine must forge ahead with new interventional procedures with conscience, guidelines and a perpetual focus on improvement. This improvement should encompass techniques of treatment, new methods to acquire better treatments, ways to interpret evidence, reporting methods and protection of individual rights. Our duty above all is to our patients.

Author Affiliation:

David A. Lenrow, MD, JD
Vice Chair of Clinical Services
Assistant Professor, Department of Rehabilitation Medicine, University Of Pennsylvania School of Medicine
3400 Spruce Street., Ground White Building, Philadelphia, PA 19104
Email: david.lenrow@uphs.upenn.edu

Larry H. Chou, MD
Assistant Professor
Director, Physical Medicine and Rehabilitation Services
Penn Medicine at Radnor
Department of Rehabilitation Medicine, University of Pennsylvania School of Medicine, 250 King of Prussia Road, Radnor, PA 19087
Email: larry.chou@uphs.upenn.edu

REFERENCES

1. Lefering R, Neugebauer E. Problems of Randomized Controlled Trials in Surgery. *Nonrandomized Comparative Clinical Studies*. In Abel U, Koch A (eds). *Proceedings of the International Conference on Nonrandomized Comparative Clinical Studies in Heidelberg*, Symposium Publishing, April 10-11. 1997. <http://www.symposium.com/nrccs/lefering.htm>
2. Bigby M. Challenges to the Hierarchy of Evidence: Does the Emperor Have No Clothes? *Arch Dermatol* 2001; 137:345-346.
3. Concato J, Shah N, Horwitz RI. Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Design. *N Engl J Med*. 2000; 1887-1892.
4. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. *N Engl J Med*. 200;342:1878-1886
5. Guyatt GH, Haynes RB, Jaeschke RZ, et al. Users' Guides to the Medical Literature, XXV. Evidence-Based Medicine: Principles for Applying the Users' Guides to Patient Care. *JAMA* 2000; 284:1290-1296.
6. Guyatt GH, Sackett D, Cook DJ. How to Use an Article About Therapy or Prevention. *Users' Guides to Evidence-Based Practice*. In *Centers for Health Evidence, wysiwyg://60/http://www.cche.net/usersguides/therapy.asp*
7. Fung EK, Lore JM. Randomized Controlled Trials for Evaluating Surgical Questions. *Arch Otolaryngol Head Neck Surg*. 2002; 128:631-634.
8. Schulz KF, Chalmers I, Haynes RJ, et al. Empirical Evidence of Bias. Dimensions of Methodological Quality Associated with Estimates of Treatment Effects in Controlled Trials. *JAMA* 1995; 273:408-412.
9. Chalmers I. Unbiased, Relevant and Reliable Assessments in Health Care. *BMJ* 1998; 317:1167-1168.
10. Freedman B. Equipoise and the Ethics of Clinical Research. *N Engl J Med*. 1987; 317:141-145.
11. Burton v. Brooklyn Doctors Hosp., 452 N.Y.S. 2d. 875 (N.Y. App. Div. 1982).
12. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. Edinburgh, Scotland: World Medical Association, October 2000. http://www.wma.net/e/policy/17-c_e.html
13. Vastag B. Helsinki Discord? A Controversial Declaration. *JAMA* 2000; 284:2983-2985.
14. Emanuel EJ, Miller FG. The Ethics of Placebo-Controlled Trials-A Middle Ground. *N Engl J Med* 2001; 345:915-919.
15. Moseley BJ, O'Malley, K, Petersen NJ et al. A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee. *N Engl J Med* 2002; 347:81-88.
16. Horng S, Miller FG. Is Placebo Surgery Unethical? *N Engl J Med* 2002; 347:137-139.
17. Jadad AR, Rennie D. The Randomized Controlled Trial Gets a Middle-aged Checkup. *JAMA* 1998; 279:319-320.
18. 45 C.F.R. 46 (2001)
19. Begg C, Cho M, Eastwood S, et al. Improving the Quality of Reporting of Randomized Controlled Trials. The CONSORT Statement. *JAMA* 1996; 276:637-639.
20. Moher D, Schulz KF, Altman D. The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials. *JAMA* 2001; 285:1987-1991.
21. Rennie D. CONSORT Revised-Improving the Reporting of Randomized Trials. *JAMA* 2001; 285:2006-2007.
22. Bonchek KI. Sounding Board: Are Randomized Trials Appropriate for Evaluating New Operations? *N Engl J Med* 1979; 301:44-45.
23. Barkun JS, Barkun AN, Sampalis JS, et al. Randomized Controlled Trial of Laparoscopic Versus Mini Cholecystectomy. *Lancet* 1992; 340:116-119.
24. Devereaux PJ, Manns BJ, Ghali WA, et al. Physician Interpretations and Textbook Definitions of Blinding Terminology in Randomized Controlled Trials. *JAMA* 2001; 285:2000-2003.

