
THE ORTHOPAEDIC FORUM



An AAOS-ORS Symposium

Clinical Trials in Orthopaedics Research. Part II. Prioritization for Randomized Controlled Clinical Trials*

By Jeffrey N. Katz, MD, MSc, James G. Wright, MD, MPH, and Elena Losina, PhD

The American Academy of Orthopaedic Surgeons (AAOS) and Orthopaedic Research Society (ORS) Clinical Trials in Orthopaedics Research Symposium had three major themes including barriers to performing clinical trials, methodology of clinical trials, and prioritization of clinical questions in orthopaedics to address in randomized controlled trials. This paper addresses the latter theme. Clinical experts from the major orthopaedic specialties provided presentations on key issues in their respective fields that had been addressed with clinical trials and the clinical questions that were most appropriate and pressing for clinical trials.

Clinicians caring for patients with musculoskeletal disorders must make treatment decisions daily. While the number of clinical trials has increased dramatically in the last two decades, many clinical decisions remain guided by a weak evidence base. The most rigorous evidence for evalu-

ating the efficacy of interventions comes from randomized controlled trials. While it is tempting to envision filling each gap in the clinical evidence base with a randomized controlled trial, trials are resource intensive. The scientific community cannot perform a randomized controlled trial to answer every clinical question. How, then, do clinical scientists prioritize potential trials to select those that yield the most favorable balance between the resources required to conduct the trial and the value of the information a trial would yield?

We suggest approaching these critical questions through three complementary lines of inquiry. The first is evaluation of the information that would be gained if the trial were executed successfully. The second is the feasibility of the trial. The final consideration is the resource cost of performing the trial. We ask: What would be gained from the trial? What would it take

Disclosure: In support of their research for or preparation of this work, one or more of the authors received, in any one year, outside funding or grants in excess of \$10,000 from the R.B. Salter Chair in Surgical Research and from the National Institutes of Health grants AR02123 and AR057827, American Academy of Orthopaedic Surgeons, and Orthopaedic Research Society. Neither they nor a member of their immediate families received payments or other benefits or a commitment or agreement to provide such benefits from a commercial entity.

*This report is based on the Clinical Trials in Orthopaedics Research Symposium sponsored by the American Academy of Orthopaedic Surgeons and the Orthopaedic Research Society, Albuquerque, New Mexico, May 7, 8, and 9, 2009.

for the trial to succeed in answering the study question? What would it cost? The first goal of this paper is to articulate a framework for prioritizing clinical trials using these three considerations. The major features of the framework are outlined in Table I. The second goal of the paper is to evaluate the randomized controlled trial topics proposed by speakers at the AAOS-ORS Clinical Trials in Orthopaedics Research Symposium through the lens of this conceptual framework. These topics are summarized in Table II.

Value of Information Gained

If the trial is executed successfully, will the results be useful to clinicians and patients in their decision to offer a particular treatment? How useful? Does the value of the information gained justify the cost of obtaining it? Several factors influence the value of information to clinicians and patients.

Burden

Is the problem common and disabling? Trials that address rare conditions, or conditions that have trivial effects on patient well-being, will have limited impact at the societal level. We do not suggest that such trials are unimportant or that they should not be done. In fact, we readily acknowledge that trials that address rare problems are critical for the treatment of those conditions to advance. However, these trials will have less public health impact than trials addressing more prevalent and disabling problems, and typically will have a lower priority. It is noteworthy that randomized controlled trials of rare conditions often pose distinctive logistical challenges to obtain an adequate enrollment.

Relevance Over Time

A trial comparing two devices that will both be obsolete before the trial is completed may not have an enduring message. In

contrast, trials comparing two strategies that remain relevant to practice even as the specific technologies change will continue to guide clinical decisions over time. For example, a trial of lumbar fusion versus laminectomy is more likely to influence practice fundamentally over time than a trial of two competing pedicle screw and plate constructs, neither of which is likely to be used in practice for more than a few years. We recognize that there is an important role for trials that compare competing devices for the same indication. Without such trials, choices between competing devices are made on the basis of nonscientific considerations. However, we suggest that, ideally, these trials should be done before the device enters the marketplace or soon thereafter. We also note that such trials should ideally gather appropriate data to support cost-effectiveness analyses, as cost-effectiveness is an important consideration in assessing competing therapies.

Specificity of the Intervention

Are the alternative treatment strategies sufficiently specified to be informative? A trial of a well-defined surgical procedure versus a well-defined exercise protocol (e.g., for rotator cuff tear, meniscal tear, or lumbar disc protrusion) will be more interpretable than trials of “surgery” versus “usual nonoperative care,” for which the actual content of surgery and usual nonoperative care varies widely among trial participants. There is a role for comparisons between new treatments and usual care, as most new interventions must compete with the usual way of treating patients. However, investigators ideally should specify and implement usual care in a standardized manner that permits meaningful inference. In fact, detailed delineation of the intervention is a key element of study design. Consider, for example, a trial of surgery versus usual care in which some of the usual care group receive exercise, others receive corticosteroid injections, some receive nonsteroidal anti-inflammatory drugs (NSAIDs), others receive behavioral cognitive therapy, and many receive various combinations of these approaches. If these diverse treatments are offered haphazardly, without a protocol specifying which treatments should be offered and when, the implications of the trial findings for clinical decision making will be unclear.

Uncertainty

Practice variation provides empirical evidence of uncertainty regarding optimal treatment. If there is no practice variation in the treatment of a particular condition, the clinician community likely will not enroll patients. More importantly, the clinician community might not accept the findings of a trial, even if it shows that the commonly performed treatment is less efficacious than an alternative. On the other hand, if there is considerable variation in treatments used across providers, reflecting uncertainty about appropriate indications and outcomes, a trial is more likely to change practice. The wide variety of approaches to anticoagulation following knee replacement (warfarin, aspirin, heparin, mechanical compression, and combinations of these) offers strong testimony

TABLE I Schema for Prioritization of Randomized Controlled Trials

Value of information gained
Prevalence of the condition and its effect on health status
Uncertainty regarding efficacy and appropriate indications
Variation in practice patterns among clinicians
Preliminary evidence of efficacy
Enduring nature of the intervention
Feasibility
Availability of adequate sample size
Multiple centers
Equipose (community, individual clinician, and patient)
Reliable, valid, responsive measures of outcome
Pilot data to ensure protocol is realistic
Resource cost of trial
Research personnel
Space and equipment
Management of potential conflicts
Assessment of trial costs as weighed against value of information

TABLE II Randomized Controlled Trial Ideas Proposed by Symposium Speakers

Treatments to Be Compared According to Specialty

Hand and upper extremity

- Arthroscopic vs. open rotator cuff surgery for rotator cuff tear
- Proximal interphalangeal joint or wrist arthroplasty vs. nonoperative therapy for advanced arthritis
- Higher vs. lower dose of injectate for steroid injection of carpal tunnel, shoulder, etc.
- Efficacy of cognitive behavioral therapy vs. usual care* for most musculoskeletal conditions

Sports[†]

- Use of single-row vs. double-row anchors in rotator cuff surgery
- Surgery vs. nonoperative care for rotator cuff tendinopathy
- Use of autograft vs. allograft in ACL reconstruction
- Use of single-bundle vs. double-bundle tendon grafts in ACL reconstruction
- ACL reconstruction vs. nonoperative therapy for symptomatic tear of the ACL
- Partial resection of the meniscus vs. specific exercise-based regimen in patients with knee osteoarthritis

Pediatrics

- Sports injury prevention program vs. control program to reduce the prevalence of sports injuries
- Surgical vs. nonoperative management of clavicular fractures in adolescents
- Trial of growth modulation techniques vs. usual care in spinal surgery
- Early vs. delayed reduction of dislocated hips in developmental dysplasia of the hip
- Gait analysis vs. usual process for decision making in cerebral palsy
- Surgical vs. nonoperative intervention for upper extremity impairment in cerebral palsy

Spine[‡]

- Biologic anti-inflammatory medications vs. usual care for sciatica due to disc protrusion
- BMP growth products vs. usual strategy for spine fusion

Lower extremity reconstruction

- Cemented vs. uncemented fixation in total hip replacement for selected patient groups (e.g., older patients)
- Ceramic vs. metal heads in total hip replacement
- Large head diameters (>32 mm) vs. smaller heads in total hip replacement: effect on dislocation and function
- Cemented vs. uncemented fixation in total knee arthroplasty
- Patellar resurfacing vs. nonresurfacing in total knee arthroplasty
- All polyethylene vs. metal-backed tibial components in total knee arthroplasty
- Cross-linked vs. conventional polyethylene in total knee arthroplasty
- Unicompartmental vs. total knee arthroplasty in elderly patients with primarily unicompartmental disease
- Antibiotic-loaded cement vs. usual cement for prophylaxis against infection in total knee arthroplasty

Foot and ankle

- Anticoagulation vs. placebo in foot surgery
- Diabetic ulcer healing: wound adjuvant vs. boot braces
- Total ankle replacement vs. arthrodesis for advanced ankle arthritis
- Toe replacement vs. nonoperative care for advanced toe arthritis

Trauma

- Bone-healing adjuvants vs. usual care for complex fractures
- Ultrasound stimulation vs. usual care for healing of tibial fractures
- Wound closure vs. usual care for open extremity fractures
- Distraction osteogenesis vs. usual care for fracture-healing
- Total hip arthroplasty vs. hemiarthroplasty in displaced hip fracture

Oncology

- Optimal anticoagulation approaches in patients undergoing orthopaedic oncologic procedures

Gene therapy

- Interleukin-1-based intra-articular gene therapy vs. placebo injection for knee osteoarthritis

*As per text, so-called usual care should be specified algorithmically to permit meaningful inference. †ACL = anterior cruciate ligament. ‡BMP = bone morphogenetic protein.

to the uncertainty in the community regarding the optimal anticoagulation approach. We suggest that this level of practice variation and underlying uncertainty makes the clinical com-

munity ready to accept a trial. It would be important in such a trial to understand the minimal clinically important difference in event rates and to power the trial accordingly.

Preliminary Evidence of Efficacy

The case for a randomized controlled trial of an emerging treatment is typically more compelling when there is preliminary evidence of efficacy from smaller controlled studies (strong preliminary evidence), observational studies (less strong), or case series (weak preliminary evidence). If preliminary data do not suggest that the agent (drug, device, or behavioral intervention) is likely to serve as an effective treatment strategy, investment in a randomized controlled trial may not move the field forward.

Role of Negative Trials

Negative trials are extremely useful. If a more expensive or more harmful intervention is used routinely without rigorous evidence of superiority over less expensive or toxic alternatives, a negative trial can change practice. A striking example in the orthopaedic community is arthroscopic debridement and lavage for osteoarthritis. Following the popularization of arthroscopy in the 1980s, arthroscopic lavage and debridement were performed frequently for osteoarthritis, with some support from small nonrandomized studies. However, two high-quality randomized controlled trials performed in the past decade established that arthroscopic lavage and debridement are no more useful than sham surgery or nonoperative therapy for symptoms of osteoarthritis^{1,2}. This conclusion has been integrated into practice guidelines. “Negative” trials must be interpreted carefully. If a trial is powered to detect the minimal clinically important difference in event rates (or in mean outcome scores) and the finding of the trial is that the difference between groups is less than the minimal clinically important difference, then the trial is negative and the null hypothesis can be rejected. However, if the trial does observe a group difference equal to or greater than the minimal clinically important difference, but the difference fails to reach significance, then the trial must be regarded as inconclusive and further research is needed.

Feasibility

The critical issue here is whether the trial will be executed successfully once it is initiated. There are four particularly important aspects of trial feasibility.

Availability of Eligible Subjects

Even if the condition under study is prevalent, the pool of eligible patients may be small. For example, an investigator could readily identify persons with advanced osteoarthritis by enrolling patients in the offices of orthopaedic surgeons who specialize in joint arthroplasty. Persons with early osteoarthritis are more difficult to identify because they are less frequently referred to specialists. Attempts to identify these populations can result in bias if not done carefully. Limitations in the pool of eligible subjects at a single institution may necessitate a multicenter trial, which requires substantially more trial infrastructure, coordination, cost, and more sophisticated statistical considerations than a single-center trial.

Outcome Measures

Every trial requires a reliable, valid, and responsive primary outcome measure. In many conditions, such as cancer and cardiovascular disease, death and salient clinical events—such as myocardial infarction—typically constitute the primary outcomes. However, orthopaedic procedures are often performed to reduce pain or improve function. These domains are more precisely measured with multi-item scales than with single discrete questions³. While the psychometric properties of such outcome measures are beyond the scope of this article, it is important to recognize that an appropriate validated outcome measure must be selected before a trial can proceed. This field is now rather mature, and an appropriate measure often exists for most orthopaedic conditions and interventions. If an appropriate measure does not exist or has not been validated for a particular condition (as may occur in rarer, less studied conditions), a measure must be developed and/or validated before a trial is implemented.

Equipoise

The trial will not enroll adequate numbers of patients successfully if potential subjects and enrolling physicians are not comfortable with both treatment options under study in a randomized controlled trial. The community of clinical scientists involved in the trial defines eligibility criteria in order to include only the patients for whom there is genuine uncertainty about the appropriate management. This consensus on eligibility criteria among clinical scientists is called *clinical equipoise*, or *community equipoise*⁴. However, the participating physician investigator who evaluates an eligible patient may have a strong clinical intuition about which treatment strategy will be most helpful. Despite the fact that the patient is eligible, this clinician may be uncomfortable randomizing. This situation illustrates the tension that may exist between community equipoise and individual equipoise. Community equipoise reflects the judgment of a group of well-informed clinicians, typically based on a critical evaluation of the research literature. Individual equipoise, on the other hand, is the state in which the individual clinician is comfortable with both alternatives. If the clinician wishes to exercise clinical judgment rather than recommend randomization to a patient, despite the lack of evidence supporting a particular treatment for that particular patient, then community and individual equipoise collide^{5,6}. This can dampen enrollment of a randomized controlled trial and also introduce bias and limitations in generalizability, since certain patients who are eligible will not be randomized.

Of course, patients must also experience equipoise—comfort with both options under study—in order to enroll in a randomized controlled trial. If the vast majority of patients have strong preferences for one treatment or the other, the trial will not succeed. In many surgical trials, 20% to 30% of eligible patients enroll, with the remainder either not referred (because of lack of surgeon equipoise) or referred but not enrolled (because of lack of patient equipoise). Thus, individual surgeon equipoise and individual patient equipoise are required for a successful enrollment. Subjects who ultimately enroll in the trial should be compared carefully with those who are eligible

but do not enroll in order to be able to assess generalizability of the enrolled sample to the pool of eligible patients to whom inferences will be made⁵.

Pilot Data

Nothing is more reassuring about the feasibility of a research protocol than direct evidence that all aspects of the protocol can be implemented successfully. Pilot data can also provide key parameters necessary for trial design. For example, sample size calculations require an estimate of the effect of the intervention and the standard deviations of outcome scores associated with the intervention and the control group, along with an estimate of the clinically important difference at the individual and group levels. Pilot studies permit estimation of these parameters in the population of interest.

We note that, even when a trial appears feasible, a skilled team with firm leadership is required to actually perform the trial. That is, a feasible design must be coupled with a competent research team.

Cost

Trials are typically costly. The highly protocolized structure of randomized controlled trials involves multiple personnel. Trials often require direct patient costs such as transportation and incentive stipends. Trials may provide rich mechanistic information if biomarkers such as blood tests and imaging studies are included in the protocol. A multicenter trial requires robust central administration in the principal center, research coordinators in each center, secure data management procedures, a strong analytic group, and consistent regulatory oversight. These various costs should be estimated as accurately as possible in the trial planning stages in order to budget appropriately.

Because trials are costly, they generally require substantial external funding. Many trials of devices, biologics, and pharmaceuticals are sponsored by the company that manufactures the product. Such industry-sponsored trials have potential for bias if the funding organization is permitted to influence the design, execution, and particularly the analysis and interpretation of the trial data. Managing such potential conflicts of interest and maintaining an appropriate distance between the sponsor and the science is critically important. These conflicts may also affect investigator-initiated trials, particularly if they are funded by a commercial party with a vested interest in the trial outcome. The potential for bias in industry-sponsored studies underscores the importance of full disclosure of potential conflicts of interest.

While trials are indeed costly, they also provide our best scientific evidence of efficacy. A well-executed trial of a highly important therapeutic question can change the course of management for the condition internationally. Thus, while the costs are high, so are the potential benefits.

Trial Topics Proposed at the AAOS-ORS Clinical Trials Symposium

Academic orthopaedic surgeons from the major orthopaedic specialty areas made presentations at the symposium that iden-

tified key questions in each specialty area deemed appropriate and timely for randomized controlled trials. Many of these suggested trial ideas are listed in Table II. The list should not be interpreted as the universe of important trial ideas, but rather as a sample that illustrates some potential questions. Some of these questions have been addressed with trials but would benefit from further clarification with additional randomized controlled trials.

While the topics listed in Table II were proposed specifically as potential subjects for clinical trials, we note that the randomized controlled trial is not the only possible design for assessing many of these questions. Particularly in those instances in which preliminary data are lacking or sparse, observational designs can be informative and yield estimates of effect to inform the sample size estimates for the randomized trial. Observational studies are particularly valuable if the goal of the research is not to determine treatment efficacy but rather to evaluate the prognosis of patients who take a particular therapy. As distinct from the randomized controlled trial, however, the observational design opens the risk for confounding by indication.

Viewed from the perspective of prioritization criteria shown in Table I, many of these trial ideas meet some but not all of the high priority criteria. We provide a few examples. A trial of the dose of injectate for injections of the rotator cuff or carpal tunnel would likely be accepted by the clinical community, individual enrolling surgeons, and patients. It would also address an area with considerable practice variation. But the effect on the health status of the population is likely to be modest. A study of unicompartmental versus total knee arthroplasty in select patients would help to resolve an important dilemma but may be limited by a paucity of surgeons willing to randomize in this setting. The trial would require a long length of follow-up in order to detect differences in failure and revision rate. The need to enroll a large cohort and follow them for up to a decade or more creates feasibility challenges.

A comparison of dislocation following implantation of larger versus smaller femoral heads in total hip arthroplasty would address an important, costly problem. Physicians and patients would probably be willing to randomize. The sample size requirements are high because dislocation is an uncommon outcome. The trial would therefore require multiple centers and substantial funding.

Trials of operative versus nonoperative therapy often address highly relevant, important clinical questions. But patients and physicians may have strong beliefs about the appropriateness of surgery, even in the absence of rigorous evidence favoring one strategy over another. Thus, equipoise—for the provider and for the patient—is typically a major challenge for trials of operative versus nonoperative strategies, such as surgical repair versus physical therapy management of non-traumatic rotator cuff tear. As a consequence, these trials will need to approach many eligible patients to fill enrollment goals. Patients with strong a priori preferences are most likely to decline the trial⁷.

In contrast, trials of two surgical techniques are less likely to raise questions of equipoise. Patients may have no a priori opinion about single versus double-bundle anterior cruciate ligament reconstruction or highly cross-linked versus standard polyethylene components for total knee arthroplasty. On the other hand, these trials may compare two alternatives that will both be outmoded by the time of publication, whereas the strategic question—whether to operate or not to operate—will likely remain relevant for years.

In this manner, the framework we have proposed can be used to examine the priority of any randomized controlled trial. Each trial will have one or more features that merit a high priority and others that do not. Ultimately, the stakeholders involved in deciding whether the trial moves forward (e.g., grant reviewers, funding organizations, and professional societies) will need to weigh and balance these priorities to arrive at prioritization decisions. We hope that the perspective offered in the paper helps investigators and other stakeholders to choose among trial ideas both at the idea generation stage and in making the final decisions about which trials to move forward for consideration of funding and ultimate implementation.

NOTE: The authors thank the clinician investigators who made presentations at the symposium. These presentations provided the proposed randomized controlled trial topics listed in Table II. These individuals included, in alphabetical order, Drs. Peter Amadio, Judith Baumhauer, Daniel Berry, David Cohen, Christopher Evans, John Healey, Robert Marx, Unni Narayanan, and Marc Swiontkowski.

Jeffrey N. Katz, MD, MSc
Elena Losina, PhD
Orthopedic and Arthritis Center for Outcome Research,
Department of Orthopedic Surgery
and Division of Rheumatology,
Immunology and Allergy,
Brigham and Women's Hospital,
75 Francis Street,
OBC-4, Boston, MA 02115.
E-mail address for J.N. Katz: jnkatz@partners.org

James G. Wright, MD, MPH
Department of Orthopaedic Surgery,
The Hospital for Sick Children,
555 University Avenue, Toronto,
ON M5G 1X8, Canada

References

1. Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, Feagan BG, Donner A, Griffin SH, D'Ascanio LM, Pope JE, Fowler PJ. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med.* 2008;359:1097-107.
2. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med.* 2002;347:81-8.
3. Daltroy LH. Common problems in using, modifying, and reporting on classic measurement instruments. *Arthritis Care Res.* 1997;10:441-7.
4. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med.* 1987;317:141-5.
5. Katz JN, Wright J, Levy BA, Baron JA, Losina E. Departures from community equipoise may lead to incorrect inference in randomized trials. *J Clinical Epidemiology.* 2010:[Epub ahead of print].
6. Gifford F. So-called "clinical equipoise" and the argument from design. *J Med Philos.* 2007;32:135-50.
7. Creel AH, Losina E, Mandl LA, Marx RJ, Mahomed NN, Martin SD, Martin TL, Millett PJ, Fossel AH, Katz JN. An assessment of willingness to participate in a randomized trial of arthroscopic knee surgery in patients with osteoarthritis. *Contemp Clin Trials.* 2005;26:169-78.