

## The MeTeOR Trial (Meniscal Tear in Osteoarthritis Research): Rationale and design features

Jeffrey N. Katz <sup>a,\*</sup>, Christine E. Chaisson <sup>b</sup>, Brian Cole <sup>c</sup>, Ali Guermazi <sup>d</sup>, David J. Hunter <sup>e</sup>, Morgan Jones <sup>f</sup>, Bruce A. Levy <sup>g</sup>, Lisa A. Mandl <sup>h</sup>, Scott Martin <sup>a</sup>, Robert G. Marx <sup>h</sup>, Clare Safran-Norton <sup>a</sup>, Frank W. Roemer <sup>d,i</sup>, Debra Skonieczki <sup>a</sup>, Daniel H. Solomon <sup>a</sup>, Kurt P. Spindler <sup>j</sup>, John Wright <sup>a</sup>, Rick W. Wright <sup>k</sup>, Elena Losina <sup>a</sup>

<sup>a</sup> Brigham and Women's Hospital, Boston, MA, USA

<sup>b</sup> Boston University School of Public Health, Boston, MA, USA

<sup>c</sup> Rush University Medical Center, Chicago, IL, USA

<sup>d</sup> Boston University Medical Center, Boston, MA, USA

<sup>e</sup> University of Sydney, Sydney, Australia

<sup>f</sup> Cleveland Clinic Sports Health, Cleveland, OH, USA

<sup>g</sup> Mayo Clinic, Rochester, MN, USA

<sup>h</sup> Hospital for Special Surgery, New York, NY, USA

<sup>i</sup> Klinikum Augsburg, Augsburg, Germany

<sup>j</sup> Vanderbilt University Medical Center, Nashville, TN, USA

<sup>k</sup> Washington University School of Medicine, St. Louis, MO, USA

### ARTICLE INFO

#### Article history:

Received 14 February 2012

Revised 27 August 2012

Accepted 28 August 2012

Available online 5 September 2012

#### Keywords:

Randomized controlled trial

Methodology

Osteoarthritis

Efficacy

Meniscal tear

Arthroscopic partial meniscectomy

### ABSTRACT

This paper presents the rationale and design features of the MeTeOR Trial (Meniscal Tear in Osteoarthritis Research; ClinicalTrials.gov NCT00597012). MeTeOR is an NIH-funded seven-center prospective randomized controlled trial (RCT) designed to establish the efficacy of arthroscopic partial meniscectomy combined with a standardized physical therapy program as compared with a standardized physical therapy program alone in patients with a symptomatic meniscal tear in the setting of mild to moderate knee osteoarthritic change (OA). The design and execution of a trial that compares surgery with a nonoperative treatment strategy presents distinctive challenges. The goal of this paper is to provide the clinical rationale for MeTeOR and to highlight salient design features, with particular attention to those that present clinical and methodologic challenges.

© 2012 Elsevier Inc. All rights reserved.

## 1. Introduction

Symptomatic, radiographic osteoarthritis affects over 27 million people in the United States, including over 9 million with knee osteoarthritis [1]. Meniscal tears are also highly prevalent,

as evidenced by community-based magnetic resonance imaging (MRI) studies which show that 35% of individuals greater than 50 years of age have evidence of a meniscal tear on MRI, even though two thirds of these tears are asymptomatic [2]. Meniscal damage is especially prevalent among persons with osteoarthritis. In fact, 75% of persons with radiographically documented symptomatic knee OA have evidence of a meniscal tear on MRI [3]. Prior studies have shown that patients with a meniscal tear and concomitant symptomatic OA have similar levels of pain as patients with OA and no meniscal tear [3]. Thus, although

\* Corresponding author at: Orthopedic and Arthritis Center for Outcomes Research, Brigham and Women's Hospital, 75 Francis Street, OBC-4016, Boston, MA 02115, USA. Tel.: +1 617 732 5338; fax: +1 617 525 7900.

E-mail address: [jnkatz@partners.org](mailto:jnkatz@partners.org) (J.N. Katz).

structural evidence of a meniscal tear is common, especially in patients with osteoarthritis, these tears are often asymptomatic.

The high prevalence of meniscal tear in patients with knee OA and the observation that these lesions are often asymptomatic create challenges in clinical decision-making. It is difficult to determine if the tear is an independent source of symptoms. This is an important differentiation since clinicians who have the clinical impression that the tear is symptomatic may recommend arthroscopic partial meniscectomy (APM).

The history of arthroscopic approaches in patients with knee osteoarthritis is noteworthy. Over the last decade, two pivotal trials have been published on the efficacy of arthroscopic surgical approaches in patients with advanced OA in whom there is no clinical suspicion of a symptomatic meniscal tear. Moseley and colleagues [4] compared arthroscopic debridement and arthroscopic lavage to a sham surgical procedure. Kirkley and colleagues [5] compared debridement to a nonoperative regimen. In each of these trials, both the surgical intervention and the control intervention resulted in substantial improvement in pain and functional status over 24 months of follow-up, with no meaningful differences between arms. Thus, consistent evidence from two high-quality randomized controlled trials suggests that arthroscopic surgery is not effective for the treatment of severe knee osteoarthritis.

However, the most frequent indication for arthroscopic surgery in patients with knee OA is a symptomatic meniscal tear. These tears are frequently addressed surgically with arthroscopic partial meniscectomy (APM), in which the surgeon trims the torn meniscus back to a stable rim. This procedure is performed on over 300,000 persons with concomitant OA and meniscal pathology annually in the United States [6]. At over \$5,000 per case, these procedures account for about \$1.5 billion in direct medical costs in the US annually. Despite the frequency of APM, the efficacy of this procedure has not been evaluated rigorously among patients with concomitant mild or moderate OA. There is only one published randomized controlled trial of the efficacy of APM, involving 90 patients followed for 5 years at one Swedish center. This study did not identify a difference in pain relief or functional status between APM plus an exercise regimen and the exercise regimen alone [7,8]. Given the frequency and cost of APM and the lack of evidence supporting utilization of this procedure, this question requires further rigorous clinical evaluation in a multicenter randomized trial. In particular, research is needed to determine whether APM is efficacious in the short term (e.g. 6 months) and over a longer period of time.

This rationale gives rise to the specific aims of the MeTeOR trial: (1) to evaluate the efficacy after 6 months of follow-up of arthroscopic partial meniscectomy, as compared with a standardized nonoperative regimen in the management of symptomatic meniscal tear occurring in patients with concomitant osteoarthritic change; (2) to follow the randomized groups through 24 months of follow-up to evaluate longer-term efficacy of APM; and (3) to assess the cost-effectiveness of APM in this clinical setting.

## 2. Design justification and the importance of equipoise

The RCT is the most rigorous test of efficacy and is often used to address significant public health problems. RCTs are often more resource-intensive than cohort studies, and they also

require that both patients and physicians must be comfortable with randomization. This critical feature of impartiality between two (or more) alternative therapies is known as equipoise [9,10].

To assess physician equipoise, the trial design was presented to potentially interested surgeons. Those who were not comfortable with randomization chose not to participate in the trial. At the time MeTeOR began, the literature did not identify distinct subgroups of patients for whom which one treatment or the other was superior. Discussions were held with surgeon investigators to identify the clinical circumstances in which they would or would not feel comfortable with both surgical and nonoperative options. These entry and exclusion criteria were further supported by our survey of orthopedic sports surgeons, among whom the most important factors prompting the decision to perform APM are the presence of a normal radiograph, failure of nonoperative therapy and physical examination findings suggestive of symptomatic meniscal tear, particularly the McMurray test [11].

We assessed patient equipoise by field-testing pilot enrollment well before MeTeOR was launched. In this pilot study, we identified patients who met the eligibility criteria for MeTeOR and summarized the study design and protocol to them. We explained that we were planning a study in which patients would be assigned randomly to receive either APM or nonoperative therapy. We then asked subjects if they would agree to participate in such a study. Of 88 patients in three centers included in this pilot study, 22% reported they would definitely agree to the trial and another 24% reported they would probably agree [12]. On the basis of these data, we anticipated that 20–25% of eligible patients would participate and planned sample size projections accordingly. We plan to compare subjects who refused participation with those who accepted to assess for potential selection bias. These projections of 20–25% acceptance of randomization are consistent with the rate of enrollment into other surgical RCTs such as the randomized components of SPORT (Spine Outcomes Research Trial), a set of trials contrasting surgical and nonoperative management of symptomatic disc herniation, spinal stenosis and spondylolisthesis [13–15].

## 3. Setting and sample

We designed an efficacy study in which the interventions were administered under ideal circumstances. Accordingly, the MeTeOR trial was performed in academic referral centers and with participating surgeons considered to be leaders in their fields. The trial was originally set in five centers: Brigham and Women's Hospital, Boston; Hospital for Special Surgery, New York; Cleveland Clinic, Cleveland; Vanderbilt University, Nashville; and Mayo Clinic, Rochester. In order to accelerate the pace of enrollment, two additional centers were added in the second year of enrollment: Rush University Medical Center, Chicago, and Washington University, St. Louis.

The entry and exclusion criteria are provided in Table 1. Our goal was to enroll patients with a meniscal tear in the setting of mild to moderate OA. The definition of osteoarthritis posed a design challenge: we did not wish to restrict the sample to patients with radiographically documented osteoarthritis (requiring a definite osteophyte, joint space narrowing, or both) because radiographic findings reflect intraarticular tissue

**Table 1**  
Entry and exclusion criteria for the MeTeOR Trial.

| Entry criteria  | Rationale  |
|---|--|
| Age 45 years or greater   | Higher likelihood of coexisting OA   |
| Symptoms for at least 4 weeks, managed with one or more of: medications, activity limitations or PT   | Unreasonable to operate prior to trying conservative therapy               |
| Symptoms consistent with torn meniscus (at least one of: clicking, catching, popping, giving way, pain with pivot or torque, pain that is episodic, pain that is acute and localized to one joint line) | Surgery generally not done in absence of symptoms suggesting meniscal tear |
| Availability of knee radiograph and MRI   | Required to ascertain eligibility  |
| Evidence on knee MRI of osteophytes or full-thickness cartilage defect; or plain radiographic evidence of osteophytes or joint space narrowing  | Documents signs of early OA  |
| Evidence on knee MRI of a meniscal tear that extends to the surface of the meniscus   | Documents meniscal tear  |
| Willingness to undergo randomization and ability to understand and sign an informed consent document  |  |
| Exclusion criteria  | Rationale  |
| A chronically locked knee (e.g. patient cannot reduce locking; a clear-cut indication for APM)  | Locked knee is an indication for surgery                                   |
| Kellgren–Lawrence grade 4 (far advanced OA)   | TKR typically more appropriate than APM in this setting                    |
| Inflammatory arthritis or clinically symptomatic chondrocalcinosis  | Alternative source of pain and swelling                                    |
| Injection with viscosupplementation in past 4 weeks in index knee   | Could obscure treatment effects  |
| Contraindication to surgery or physical therapy   |  |
| Bilateral symptomatic meniscal tears  | Difficult to follow response to index knee                                 |
| Prior surgery on same knee  | Complex pathoanatomy   |

damage only in advanced stages of joint damage [16]. There is general recognition that osteoarthritis begins well before signs of radiographic osteoarthritis become apparent. We considered cartilage defects on MRI to be a marker of early disease and therefore appropriate for trial eligibility [17].

The definition of a symptomatic meniscal tear also posed methodological challenges [18]. There is general agreement among clinicians that the symptoms of isolated knee OA differ from those arising specifically from a meniscal tear, which frequently have a more mechanical quality [18]. Presumably, these more mechanical symptoms reflect interference with smooth joint motion as a result of the torn meniscal fragment. However, there is no standard, valid definition of meniscal symptoms and there is a paucity of literature on the sensitivity and specificity of the symptoms typically regarded as consistent with a torn meniscus such as locking, clicking, popping, pain with pivot, and pain localized to one spot. Acknowledging this limitation in our understanding of meniscal symptoms, we adopted a broad definition: patients had to have one or more of the checklist of symptoms listed in the third “Entry Criteria” bullet point in Table 1. In exploratory analyses of the trial data we will be able to determine whether one or more of these symptoms is associated with response to surgery.

#### 4. Enrollment and randomization

At this writing enrollment and randomization are complete with follow-up ongoing. Research coordinators in all centers screened the schedules of participating surgeons to identify patients who met screening criteria (see Fig. 1). These include age 45 or greater, knee pain and no prior surgery on the index knee, as well as a documented meniscal tear on MRI. The research coordinator flagged these patients’ charts when they were seen in clinic. The surgeon completed an eligibility checklist that included the eligibility and exclusion criteria shown in Table 1. If patients were eligible, the surgeon informed them of the trial and asked if they were potentially interested in participating. If so, the surgeon referred them to the research coordinator, who provided a detailed description of the study, using a script to ensure that these conversations were standardized across all centers. Patients who wished to participate completed an informed consent form and baseline questionnaire. The research coordinator received training from a study physical therapist in order to perform a standardized physical examination, which includes measures of knee range of motion, quadriceps and hamstring strength, leg length, varus and valgus knee deformity, pes planus deformity and a timed up and go test.

Once the baseline assessment was completed, patients were randomized. Randomization was performed in real time on MeTeOR’s secure website. Subjects were randomized in blocks of varying size within each site, stratified by the extent of osteoarthritis on the baseline plain radiograph (Kellgren–Lawrence grade 0–2 (no joint space narrowing) versus Kellgren–Lawrence grade 3 (<50% joint space narrowing)).

#### 5. Interventions

##### 5.1. Surgical arm

In the planning phase of the trial, surgical investigators discussed the surgical protocol in two face-to-face meetings and arrived at a standardized approach. The surgeon performed an arthroscopic partial meniscectomy, trimming the damaged meniscus back to a stable rim. Surgeons trimmed loose fragments of cartilage and bone but they did not attempt to stimulate a healing response by penetrating the subchondral bone (microfracture technique). Intraarticular corticosteroid injections were not permitted at the time of surgery. Preoperative antibiotics were used routinely. Patients were scheduled for physical therapy (PT) following surgery. While there is no consensus on the need for or effectiveness of post-operative PT in this setting, the investigators felt that including PT in both arms would help to isolate the independent effects of surgery. The PT regimen was analogous to that provided in the nonoperative arm (see below).

##### 5.2. Nonoperative arm

The investigators opted for a standardized nonoperative protocol rather than ‘usual care.’ The latter makes the comparisons less incisive, as there are many variations of usual care for knee problems [19]. The physical therapy protocol was developed by a team of physical therapists from the various trial centers based on evidence in the literature supporting land-based, individualized physical therapy with concomitant

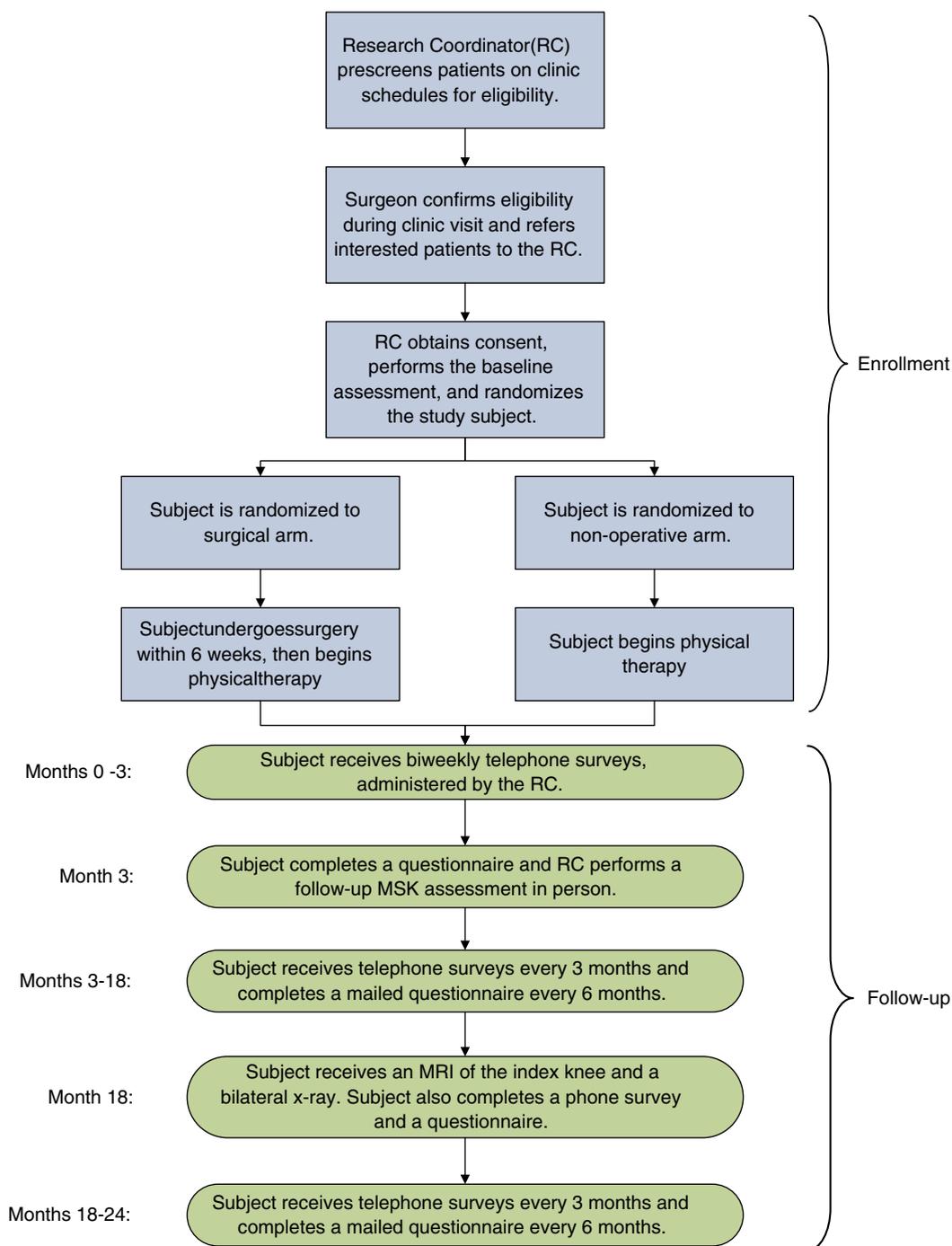


Fig. 1. Flow diagram depicting enrollment and follow-up procedures in MeTeOR trial.

progressive home exercise for patients with knee OA [20,21]. The program was designed to address range of motion, concentric/eccentric muscle strength, muscle length restrictions, aerobic conditioning (e.g. bicycling), functional mobility and proprioceptive/balance training. Patients followed a three-stage structured program. Criteria for advancing from stages I to II and II to III included the level of self-reported pain, observed strength, range of knee motion, knee effusion and functional mobility. In

each stage, the patient attended PT sessions approximately twice weekly and performed a follow-up home exercise program in between sessions.

In both the nonoperative and operative arms, subjects used acetaminophen and nonsteroidal anti-inflammatory agents as needed. Subjects with persistent pain or who were unable to tolerate these medications could receive codeine as needed. Intraarticular injections of corticosteroids were permitted at

the discretion of the surgeon. Use of these medications and injections was documented carefully, permitting analysis of the role of these concomitant treatments on outcome and differential use of these treatments across randomized groups.

## 6. Outcome measures

The WOMAC functional status scale [22], assessed at 6 months following randomization, is the primary outcome measure. We gathered preliminary data on the performance of this scale in a pilot study among patients with OA and meniscal tear. In these patients, the WOMAC function scale was highly reliable (Cronbach's alpha 0.97) and valid, had little missing data (just 1%) and no ceiling or floor effects (unpublished pilot data).

We measure secondary outcomes in several domains including pain, generic functional status, quality of life, and health care utilization. We measure pain and symptoms with the KOOS (Knee injury and Osteoarthritis Outcome Scale) pain and symptom scales [23,24]. Each has been validated extensively. In our pilot work in subjects receiving APM, the WOMAC pain scale had a ceiling effect, with the top 25% of patients scoring a perfect 100. The KOOS pain scale (comprised of the five WOMAC pain items and four other items that tap symptoms typical of internal derangements) did not have this limitation. Both scales had Cronbach's alpha greater than 0.85 in our pilot work. Therefore, we chose the KOOS scale as our pain measure. We included the Physical Activity Scale of the SF-36, a widely used, reliable, valid generic health status measure shown to capture the functional burden of patients following APM [25]. We also included the five-item mental health contained in the SF-36, which has been validated against clinical assessments of depression [26]. We administered the Euroqol (EQ-5D) for economic analyses. This five-item instrument assesses health-related quality of life in five domains and has been calibrated to measure utility in large national samples. It has been validated for use in samples with knee osteoarthritis [27].

## 7. Assessments

### 7.1. Timing (Fig. 1 and Table 2)

We chose times of assessment to capture key milestones in the management of patients with meniscal tears in the presence of OA. We assessed subjects prior to randomization and after 3 months to capture the early response to surgical or nonoperative therapy. The principal endpoint was at 6 months, at which point we hypothesize subjects typically reach maximum improvement. We also evaluated subjects with questionnaires at 12, 18 and 24 months to characterize the trajectory of pain and function over the first 2 years. We contacted subjects by phone every 2 weeks for the first 3 months after randomization and then quarterly until 2 years. (Of note, the reference point was the date of randomization.) These calls elicit data on health resource utilization, compliance with physical therapy, and pain and function to provide a detailed temporal profile of response to each intervention. Physical examinations were performed at enrollment and 3 months. The physical examination was intended to permit us to analyze the effect of baseline impairments on response to treatment and the extent that these impairments are influenced by treatment. We will follow this

cohort beyond 2 years to evaluate longer-term functional outcomes, subsequent surgeries and structural OA progression.

### 7.2. Blinding

The research team had to decide whether to blind the subjects, surgeons and assessors to treatment assignment. While blinding with a sham control would have averted potential observer bias by assessors and reporting bias by patients, our investigators were uncomfortable conducting sham surgery. They felt patients would not accept randomization to a sham procedure and that the hospitals would be troubled by the legal and billing aspects of the sham procedure. The primary outcome measures are obtained from validated self-report questionnaires, reducing the possibility of observer bias. However, we acknowledge that the subjects' knowledge of their randomization status poses potential bias. Analyses will be performed with the analysts blinded to randomization assignment.

## 8. Crossover, noncompliance and termination

Crossover is defined as switching from the randomly assigned treatment arm to the other arm of the study. In a surgical RCT such as MeTeOR, a crossover occurs when a study subject assigned to the nonoperative arm undergoes surgery or when a subject assigned to the surgical arm does not have surgery but rather is treated nonoperatively. Because the primary analysis is "intention to treat," crossovers dilute the estimates of efficacy, if indeed there is a true difference in the treatment arms. While "as treated" analyses can be performed to compare the treatments actually received, these analyses lose the unique benefits of randomization—the random allocation of all measured and unmeasured patient characteristics across the treatment groups. We attempted to minimize crossover in both directions. For subjects randomized to surgery, we attempted to deliver the procedure within 3 weeks of randomization. For those randomized to nonoperative therapy, we attempted to maintain the treatment assignment at least until the primary endpoint was assessed at 6 months. We also did a full assessment at 3 months in order to capture all outcome measures on subjects that did cross over before the 6-month mark. Crossover from nonoperative therapy to APM is informative because it likely reflects failure of nonoperative therapy to control symptoms and restore function. We anticipate that crossovers from surgery to nonoperative therapy will be less informative of the patient's clinical status and more reflective of logistical barriers in scheduling APM promptly.

## 9. Analytic approach

The primary analysis will be a comparison of changes in WOMAC function score between the surgical and the nonoperative groups, using the "intention to treat" approach, in which subjects are analyzed in the same group to which they were randomly assigned. We will account for potential center effects by including a random effect for center in multivariate analyses. We planned one subgroup analysis a priori of the effect of baseline Kellgren–Lawrence grade on outcome. This analysis will be performed by testing for an interaction between treatment assignment (APM vs. nonoperative) and baseline Kellgren–Lawrence grade (0–2 vs. 3).

**Table 2**  
Measures collected at each follow-up time point.

| Measures                       | Time point |                                |          |          |          |           |           |           |           |           |                                    |
|--------------------------------|------------|--------------------------------|----------|----------|----------|-----------|-----------|-----------|-----------|-----------|------------------------------------|
|                                | Baseline   | 0–10 week<br>biweekly<br>calls | 3 months | 6 months | 9 months | 12 months | 15 months | 18 months | 21 months | 24 months | 30–60 month<br>biannual<br>surveys |
| Pain,<br>functional<br>status* | X          |                                | X        | X        |          | X         |           | X         |           | X         | X                                  |
| Satisfaction                   |            |                                | X        | X        |          | X         |           | X         |           | X         | X                                  |
| Resource<br>utilization        | X          | X                              | X        | X        | X        | X         | X         | X         | X         | X         | X                                  |
| Work<br>disability**           | X          | X                              | X        | X        | X        | X         | X         | X         | X         | X         | X                                  |
| Physical<br>examination        | X          |                                | X        |          |          |           |           |           |           |           |                                    |
| Imaging***                     | X          |                                |          |          |          |           |           | X         |           |           |                                    |

\* WOMAC, KOOS, meniscal symptoms.

\*\* Work productivity and activity impairment.

\*\*\*MRI, fixed flexion radiographs.

Secondary analyses will use an “as treated” approach, in which patients are analyzed according to the treatments actually received. We will also perform a secondary “intention to treat” analysis in which the dependent variable is a failure indicator defined as failing to achieve a minimal clinically important improvement in WOMAC function score (8 points [28,29]), or electing to cross over to the other treatment arm. This analysis provides an estimate of efficacy at the level of the patient, rather than group.

## 10. Sample size and power

The principal trial outcome measure is the WOMAC function scale at 6 months. We powered the study to detect a 10-point difference on this scale between the operative and nonoperative arms. This was the difference we noted in observational pilot data. It was also in the range of the minimal clinically important difference in the WOMAC function scale among OA patients estimated by Angst et al. [28,29]. We adopted a type I error rate of 5% and power of 80%. The sample size calculation took into account two sources of sample degradation: losses to follow-up and crossover from the assigned arm to the other arm prior to the primary outcome assessment at 6 months. We powered the study for one pre-planned subgroup analysis in which patients with KL grade 3 (joint space narrowing) will be analyzed in one subgroup and those with KL grades 0–2 in the other. The subgroup analysis is a formal multiplicative interaction and is powered as such [30]. On the basis of these considerations, we set the target sample size at 340 patients.

## 11. Economic analysis

A formal economic analysis is planned to determine the cost-effectiveness of APM as compared with nonoperative therapy in patients with symptomatic meniscal tear and concomitant OA. The analysis will be based on recommendations by the Panel on Cost-Effectiveness in Health and Medicine [31]. The cost-effectiveness of APM will be determined by calculating the incremental cost-effectiveness ratio, defined as

“dC/dE,” where “dC” is the difference in direct medical and non-medical costs between APM and non-surgical management, and “dE” is the difference in effectiveness. We measure this ratio in dollars per quality-adjusted life year gained (\$/QALY). The choice between alternative strategies is best made via incremental analysis. We will perform the analysis from a societal perspective, meaning that we will include all costs and health outcomes that flow from APM or non-surgical management, as recommended by the Panel on Cost-Effectiveness in Health and Medicine. In all analyses, we will apply 3% per year discounting to both costs and effectiveness valuations, according to recommendation of the Panel [31].

The cost-effectiveness analysis will be performed in two stages. First we will conduct the cost-effectiveness analysis over the 2-year horizon of the trial. We then will follow with an analysis conducted over the remaining lifetime of the subjects. The lifetime analysis will involve projecting outcomes and costs beyond the course of the trial using modeling techniques. We will use the Osteoarthritis Policy (OAPol) Model for these analyses. The OAPol Model is a computer-based, state-transition, Monte Carlo simulation model of the progression and outcomes of knee OA [32,33].

Effectiveness will be measured in terms of quality-adjusted life years (QALYs). Quality adjustment is accomplished with a measure of utility, which measures capture patients' preferences for health states scored from 0.00 (essentially equivalent to death) to 1.00 (perfect health). Utility will be derived from EuroQol (EQ-5D) [27]. Direct medical costs include costs of inpatient stays and procedures (including total knee replacement), outpatient physician, physical therapy (PT) and emergency room visits, as well as costs of laboratory studies, medical devices and prescription and non-prescription medications. These data are derived from subject reports of resource utilization on the routinely administered questionnaires. We will apply costs to these utilization items using published sources. Patients receiving APM are assigned costs associated with the surgical procedure, including those related to laboratory, preoperative screening, hospital and surgeon. Indirect costs are not included, as suggested by the Panel [31]. To determine

how alternative management strategies compare with other uses of medical resources, we will compare the incremental cost-effectiveness ratios with those in other published studies.

## 12. Data management

Study data are managed by the Data Coordinating Center (DCC) at Boston University School of Public Health. The DCC developed a secure, password-protected MeTeOR website for this trial, which is used for randomization, direct entry of some data (including screening and eligibility information, terminations, crossovers and serious adverse events) and communication across centers. The phone interviews, questionnaires, physical examination and surgical forms are sent to the DCC in hard copy. All data are housed on a secure network drive and backed up nightly.

## 13. Imaging assessment

The research team received funding from the American Rehabilitation and Recovery Act to add an imaging component to MeTeOR. The principal goal of the imaging component is to determine whether either of the treatment arms is associated with more rapid progression of osteoarthritis. This is a critically important clinical policy question, as any effect of surgery on short-term outcomes should be viewed in the context of potential long-term risk of accelerated progression of osteoarthritis. We hypothesize that subjects with severe meniscal damage at baseline are more likely to experience incident or worsening of cartilage damage (semiquantitatively assessed using MOAKS - MRI Osteoarthritis Knee Score) [34], adjusting for other factors including baseline K-L grade, alignment and bone marrow lesions in the affected compartment. We also hypothesize that APM is associated with a higher likelihood of worsening cartilage status than nonoperative therapy in persons with severe meniscal damage at baseline.

All MeTeOR patients are invited to an 18-month visit in which they undergo a knee MRI and a plain radiograph and complete an additional questionnaire. The MRI is performed on 3 T scanners whenever possible and uses the same sequences as were used in each patient's baseline scan. Baseline and 18-month scans are de-identified and sent to the DCC for uploading to a secure imaging database.

The baseline and 18-month MRIs are scored using the MOAKS (MRI Osteoarthritis Knee Score) scoring system. MOAKS is a semiquantitative scoring system that incorporates elements from the WOMBS (Whole Organ MRI Score) and BLOKS (Boston Leeds Osteoarthritis Knee Score) scoring systems [34–37]. It has undergone preliminary reliability testing and validation. Scans are read paired (baseline and 18 months), but not blinded to order [38]. The readers are blinded to the subjects' treatment arm. The analyses of these data will examine whether the surgical and nonoperative groups differ with respect to loss of cartilage, change in the size or number of bone marrow lesions, synovitis/effusion, or size of osteophytes. The scans are obtained using standard clinical protocols and therefore do not allow quantitative cartilage morphometry, which requires high resolution 3D gradient echo sequences in addition.

The 18-month radiographic exams use a SynaFlexer™ (Synarc Inc., San Francisco) frame for standard fixed flexion views [39]. We will repeat the radiographs at 5 years using

the same technique, permitting an analysis of longitudinal progression in joint space narrowing from 18 months to 5 years. The imaging component presented several challenges, including the lack of standardized baseline scans at the seven clinical scanners, the limitations of routinely obtained knee MRIs for three-dimensional cartilage segmentation, and the possibility of secular change in equipment or sequences over the course of this longitudinal study. We note that routine clinical MRIs typically provide excellent assessment of cartilage damages [40].

We will conduct an analysis to establish whether subjects randomized to APM were more likely to progress by at least one grade on the MOAKS cartilage scale. This analysis will be conducted using a Chi-square test. We will also use multivariate models that include baseline values of BMI, alignment, K-L grade, MOAKS cartilage score and BML. We will include time between the baseline and APM to adjust for the fact that subjects may have crossed over to surgery at any time during the extended follow-up period. We anticipate that 250 subjects will complete the imaging assessment. With a sample of 250 subjects, we will have power of 80% to detect a 2.2-fold difference between operative and nonoperative arms in the likelihood of progressing by at least one grade in MOAKS cartilage score.

## 14. Summary

MeTeOR asks a fundamental question: whether arthroscopic partial meniscectomy combined with physical therapy leads to better pain and functional outcomes than physical therapy alone for middle-aged and older patients with a symptomatic meniscal tear in the setting of mild to moderate osteoarthritis. Study staff enrolled 351 subjects at seven clinical centers between June 2008 and August 2011. Key challenges have included the issues of patient and physician equipoise, blinding, crossover and noncompliance and examination of imaging studies that were done originally for clinical rather than research purposes. We have addressed each of these challenges in our work with the intention of limiting their potential to threaten validity. We acknowledge that no single trial can resolve an important question unequivocally, but hope that MeTeOR helps clinicians and patients to make wise, evidence-based decisions about the common problem of symptomatic meniscal tear in the setting of knee osteoarthritis.

## Acknowledgements

The authors appreciate the expert technical assistance of Julian Prokopetz, BA, and the tireless efforts of the entire MeTeOR team.

## References

- [1] Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* 2008;58:26–35.
- [2] Englund M, Guermazi A, Gale D, Hunter DJ, Aliabadi P, Clancy M, et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med* 2008;359:1108–15.
- [3] Bhattacharyya T, Gale D, Dewire P, Totterman S, Gale ME, McLaughlin S, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. *J Bone Joint Surg Am* 2003;85-A:4–9.

- [4] Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:81–8.
- [5] Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, et al. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008;359:1097–107.
- [6] Cullen KA, Golosinskiy A. *Ambulatory Surgery in the United States, 2006*. In: Hyattsville, editor. *Statistics NCHH*; 2009.
- [7] Herrlin SV, Wange PO, Lapidus G, Hallander M, Werner S, Weidenhielm L. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. *Knee Surg Sports Traumatol Arthrosc* March 23 2012 [Epub ahead of print].
- [8] Herrlin S, Hallander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007;15:393–401.
- [9] Freedman B. Equipose and the ethics of clinical research. *N Engl J Med* 1987;317:141–5.
- [10] Katz JN, Wright J, Levy BA, Baron JA, Losina E. Departures from community equipose may lead to incorrect inference in randomized trials. *J Clin Epidemiol* 2011;64:280–5.
- [11] Lyman S, Oh LS, Reinhardt KR, et al. Clinical factors that affect surgical decision making for arthroscopic partial meniscectomy in patients over age 40. *J Arthroscopy* 2012;28:492–501.
- [12] Creel AH, Losina E, Mandl LA, Marx RJ, Mahomed NN, Martin SD, et al. 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.
- [13] Weinstein JN, Lurie JD, Tosteson TD, Hanscomb B, Totoson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med* 2007;356:2257–70.
- [14] Weinstein JN, Tosteson TD, Lurie JD, Totoson AN, Blood E, Hanscomb B, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008;358:794–810.
- [15] Weinstein JN, Tosteson TD, Lurie JD, Totoson AN, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 2006;296:2441–50.
- [16] Wright RW, Boyce RH, Michener T, Shyr Y, McCarty EC, Spindler KP. Radiographs are not useful in detecting arthroscopically confirmed mild chondral damage. *Clin Orthop Relat Res* 2006;442:245–51.
- [17] Hunter DJ, Arden N, Conaghan PG, Eckstein F, Gold G, Grainger A, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. *Osteoarthritis Cartilage* 2011;19:963–9.
- [18] Niu NN, Losina E, Martin SD, Wright J, Solomon DH, Katz JN. Development and preliminary validation of a meniscal symptom index. *Arthritis Care Res (Hoboken)* 2011;63:208–15.
- [19] Roddy E, Zhang W, Doherty M, Arden NK, Barlow J, Birrell F, et al. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee—the MOVE consensus. *Rheumatology (Oxford)* 2005;44:67–73.
- [20] Franssen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2008;CD004376.
- [21] Franssen M, McConnell S. Land-based exercise for osteoarthritis of the knee: a meta-analysis of randomized controlled trials. *J Rheumatol* 2009;36:1109–17.
- [22] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
- [23] Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)—validation of a Swedish version. *Scand J Med Sci Sports* 1998;8:439–48.
- [24] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
- [25] Katz JN, Harris TM, Larson MG, Krushell RJ, Brown CH, Fossel AH, et al. Predictors of functional outcomes after arthroscopic partial meniscectomy. *J Rheumatol* 1992;19:1938–42.
- [26] Berwick DM, Murphy JM, Goldman PA, Ware Jr JE, Barsky AJ, Weinstein MC. Performance of a five-item mental health screening test. *Med Care* 1991;29:169–76.
- [27] Franssen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. *Rheumatology (Oxford)* 1999;38:807–13.
- [28] Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. *J Rheumatol* 2002;29:131–8.
- [29] Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis Rheum* 2001;45:384–91.
- [30] Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine—reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007;357:2189–94.
- [31] Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996;276:1339–41.
- [32] Holt HL, Katz JN, Reichmann WM, Gerlovin H, Wright EA, Hunter DJ, et al. Forecasting the burden of advanced knee osteoarthritis over a 10-year period in a cohort of 60–64 year-old US adults. *Osteoarthritis Cartilage* 2011;19:44–50.
- [33] Losina E, Walensky RP, Reichmann WM, Holt HL, Gerlovin H, Solomon DH, et al. Impact of obesity and knee osteoarthritis on morbidity and mortality in older Americans. *Ann Intern Med* 2011;154:217–26.
- [34] Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). *Osteoarthritis Cartilage* 2011;19:990–1002.
- [35] Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). *Ann Rheum Dis* 2008;67:206–11.
- [36] Guermazi A, Hunter DJ, Roemer FW. Plain radiography and magnetic resonance imaging diagnostics in osteoarthritis: validated staging and scoring. *J Bone Joint Surg Am* 2009;91(Suppl. 1):54–62.
- [37] Peterfy CG, Guermazi A, Zaim S, Tirman PF, Mieux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. *Osteoarthritis Cartilage* 2004;12:177–90.
- [38] Felton DT, Nevitt MC. Blinding images to sequence in osteoarthritis: evidence from other diseases. *Osteoarthritis Cartilage* 2009;17:281–3.
- [39] Kothari M, Guermazi A, von Ingersleben G, Mieux Y, Sieffert M, Block JE, et al. Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. *Eur Radiol* 2004;14:1568–73.
- [40] Roemer FW, Kwok CK, Hannon MJ, Crema MD, Moore CE, Jakicic JM, et al. Semiquantitative assessment of focal cartilage damage at 3T MRI: a comparative study of dual echo at steady state (DESS) and intermediate-weighted (IW) fat suppressed fast spin echo sequences. *Eur J Radiol* 2011;80:e126–31.