BMJ open
accessible medical research

Finnish Degenerative Meniscal Lesion Study (FIDELITY): a protocol for a randomised, placebo surgery controlled trial on the efficacy of arthroscopic partial meniscectomy for patients with degenerative meniscus injury with a novel 'RCT within-a-cohort' study design

Raine Sihvonen,[1] Mika Paavola,[2] Antti Malmivaara,[3] Teppo L N Järvinen[2]

To cite: Sihvonen R, Paavola M, Malmivaara A, *et al*. Finnish Degenerative Meniscal Lesion Study (FIDELITY): a protocol for a randomised, placebo surgery controlled trial on the efficacy of arthroscopic partial meniscectomy for patients with degenerative meniscus injury with a novel 'RCT within-a-cohort' study design. *BMJ Open* 2013;**3**:e002510. doi:10.1136/bmjopen-2012-002510

▸ Prepublication history for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-002510).

Received 17 December 2012
Revised 20 January 2013
Accepted 8 February 2013

For numbered affiliations see end of article.

Correspondence to
Dr Teppo L N Järvinen; teppo.jarvinen@uta.fi

ABSTRACT

Introduction: Arthroscopic partial meniscectomy (APM) to treat degenerative meniscus injury is the most common orthopaedic procedure. However, valid evidence of the efficacy of APM is lacking. Controlling for the placebo effect of any medical intervention is important, but seems particularly pertinent for the assessment of APM, as the symptoms commonly attributed to a degenerative meniscal injury (medial joint line symptoms and perceived disability) are subjective and display considerable fluctuation, and accordingly difficult to gauge objectively.

Methods and analysis: A multicentre, parallel randomised, placebo surgery controlled trial is being carried out to assess the efficacy of APM for patients from 35 to 65 years of age with a degenerative meniscus injury. Patients with degenerative medial meniscus tear and medial joint line symptoms, without clinical or radiographic osteoarthritis of the index knee, were enrolled and then randomly assigned (1 : 1) to either APM or diagnostic arthroscopy (placebo surgery). Patients are followed up for 12 months. According to the prior power calculation, 140 patients were randomised. The two randomised patient groups will be compared at 12 months with intention-to-treat analysis. To safeguard against bias, patients, healthcare providers, data collectors, data analysts, outcome adjudicators and the researchers interpreting the findings will be blind to the patients' interventions (APM/placebo). Primary outcomes are Lysholm knee score (a generic knee instrument), knee pain (using a numerical rating scale), and WOMET score (a disease-specific, health-related quality of life index). The secondary outcome is 15D (a generic quality of life instrument). Further, in one of the five centres recruiting patients for the randomised controlled trial (RCT), all patients scheduled for knee arthroscopy due to a degenerative meniscus injury are prospectively followed up using the same protocol as in the RCT to provide an external validation cohort. In this article, we present and discuss our study design, focusing particularly on the internal and external validity of our trial and the ethics of carrying out a placebo surgery controlled trial.

ARTICLE SUMMARY

Article focus

- This article describes a protocol for a multicentre, blinded, randomised placebo surgery controlled trial, using a parallel two-arm design to investigate the efficacy of arthroscopic partial meniscectomy (APM) in patients with degenerative meniscus injury.
- We will also describe a novel 'RCT within-a-cohort' study design that aims to improve the generalisability of randomised controlled trials (RCTs) in general.

Key messages

- The immediate goal of this trial is to provide evidence on the 'true' efficacy of APM, the most common orthopaedic procedure.
- Controlling for the placebo effect is a critical aspect of experimental design in any clinical research, particularly when assessing the success of treatment of degenerative orthopaedic complaints. Optimally, it requires the use of placebo surgery.
- The 'RCT within-a-cohort' study design will mark the first undertaking of an orthopaedic trial to address the inevitable trade-off between internal and external validity. The rigorous design, organisation and execution of this trial aim to set a new standard for future clinical trials in orthopaedics.

ARTICLE SUMMARY

Strengths and limitations of this study

- Our protocol paper will present a detailed ethical analysis of the use of placebo surgery, a framework that future placebo surgery controlled trials in orthopaedics are suggested to follow.
- The rationale supporting placebo surgery controlled design includes demonstrated considerable 'placebo effect' in surgical trials and optimal blinding of both the patient and the physician-researcher.
- The study is limited to degenerative meniscus injury. The definition of the concepts 'degenerative' or 'traumatic' in the context of meniscal injuries is arbitrary in nature.
- This is a protocol paper; the trial is underway, but has not been completed.

Ethics and dissemination: The protocol has been approved by the institutional review board of the Pirkanmaa Hospital District and the trial has been duly registered at ClinicalTrials.gov. The findings of this study will be disseminated widely through peer-reviewed publications and conference presentations.
Trial registration: ClinicalTrials.gov, number NCT00549172.

INTRODUCTION

Middle-aged men and women with knee pain attributed to degenerative knee disease (a degenerative meniscal injury accompanied by varying degrees of knee osteoarthritis (OA)) constitute the largest group of patients referred to orthopaedic surgeons.[1–4] The prevailing understanding regarding the aetiology and treatment of degenerative knee disease is quite mechanical in nature: knee pain perceived by the patient is attributed to a degenerative meniscus injury or chondral derangement. The symptoms related to degenerative meniscus injury are commonly chronic and fluctuating in nature. However, somewhat paradoxically, a recent population-based MRI study showed degenerative meniscal tears to be highly prevalent in middle-aged and elderly adults and they seem to be asymptomatic in the majority of individuals.[5 6] The current treatment strategy of early degenerative knee disease includes rest, oral non-steroidal anti-inflammatory drugs, glucocorticosteroid injections, physiotherapy and ultimately arthroscopic removal of mechanical derangement (arthroscopic partial meniscectomy (APM) or debridement).[7] In the USA alone, one million knee arthroscopies are performed annually to treat degenerative knee disease, making arthroscopic surgery of the knee by far the most common orthopaedic procedure.[8 9] Many patients report improvement (reduced knee pain, better function and improved quality of life) after this kind of surgery.[10–13] However, similar results have also been obtained with conservative treatment in randomised, non-placebo-controlled trials.[14–16] To date, there is a dire lack of high-quality evidence (from randomised controlled trials, RCT) on the efficacy of APM.[17] This issue cannot be addressed simply by evaluating the outcomes of patients who have undergone surgery, as the role of the underlying degenerative process and the surgical procedure cannot be unravelled in such a study design.[18 19]

It is widely agreed that controlling for the placebo effect is a critical aspect of experimental design in any clinical research.[20 21] A comprehensive meta-analysis of the 'placebo effect' indicated that subjective outcomes may be attributable to 'placebo effects' or bias.[20] As subjective measures or patient-reported outcomes—such as the amount of pain or perceived disability—are often used and also widely recommended as primary outcomes in assessing the success of treatment of degenerative orthopaedic complaints,[22 23] the issue of the possible 'placebo effects' of surgery is of utmost relevance. This is an obvious argument favouring the use of sham surgery to control for the imminent bias. Optimal blinding of both the patient and the physician-researcher is another clear asset of the sham-surgical model: patients are biased towards favourable outcomes after surgical intervention because they want to believe that they chose the correct option for their care. This 'leap of faith' is believed to be greater in surgery than in medical trials, in which the perceived and real risks of the intervention may be more subtle, less severe and do not involve the pain and risks of invasive procedures.[24] Equally importantly, sham surgery is ideal for minimising the investigator-bias through true blinding of the outcome assessor.[25]

Although the first-ever sham surgery controlled clinical trial was published as early as in 1959,[26] there has been a paucity of placebo surgery controlled trials in orthopaedics.[27 28] Even discussion regarding the use of the sham surgery model has been limited and generally condemning or censorious.[27 29 30] The essential question regarding the ethics of the use of placebo surgery concerns the dilemma between the provision of the highest standard of research design (the double-blind, randomised, placebo-controlled trial) and the highest standard of ethics (to do no harm to the patient).[29] When evaluating the merits of a study protocol, the institutional review board (IRB) should consider not only the risks and benefits that may result from the research but also the possible long-term effects of applying knowledge gained in the research as among the research risks that fall within the purview of its responsibility.[21 27] Dowrick and Bhandari recently proposed that there are four areas that deserve particular attention when deciding on the appropriateness of a sham surgery controlled RCT, namely the equipoise, risk minimisation, informed consent and deception. As our views on the last three issues are thoroughly expressed in the ethical annex (box 1), we will focus here on equipoise. In short, the statement by Dowrick and Bhandari[21] that "In orthopaedics, the reality is that many treatments are 'known,' 'accepted,' or 'standard' but may not yet be

Box 1 Ethical considerations of using placebo surgery control, as outlined by Miller[31]

I. Is there scientific and clinical value in conducting this study?

Despite being the most common orthopaedic procedure, there exists no randomised, placebo-controlled study on the efficacy (risk-benefit ratio) of arthroscopic partial meniscectomy (APM). We would like to emphasise that a properly conducted efficacy trial on APM could provide better evidence-base for this procedure and theoretically eliminate unnecessary procedures or at least provide the field with prognostic factors for good/poor outcome.

II. Is the use of placebo surgery methodologically necessary or desirable to achieve valid results?

Given on the one hand the subjective and somewhat vague nature of the symptoms related to degenerative meniscus injury (fluctuating course of symptoms that range from one individual to another from spontaneous disappearance of the symptoms to quite dramatic impairment in the quality of life), and on the other, the well-known strong placebo effect of surgical interventions,[27 32] we feel confident in stating that there is indeed no alternative to using placebo surgery as a control.

III–V. Risk–benefit assessment

A basic ethical requirement of clinical research is to minimise risks.[33] This does not mean that risks can be reduced to zero or that risks must be minimal.[31] The risks inherent in our placebo surgery procedure are related to spinal anaesthesia and arthroscopy. Spinal anaesthesia does carry a modest risk (mainly of postspinal headache), but similar risks have been considered acceptable for other medical investigators, such as spinal tap for neurological diseases or when used to carry out muscle biopsies for neurological or orthopaedic study designs or bronchoscopy.[34] As for the risks related to knee arthroscopy per se, the most common are wound infection, deep venous thrombosis and pulmonary embolism with respective incidences of 0.22%, 0.12% and 0.08%.[35] There is naturally also a risk for iatrogenic lesions (chondral surfaces) inside the knee during the procedure, but given the fact that all surgeons performing the operations in the RCT are highly skilled with experience of hundreds of previous knee arthroscopies, these are considered negligible. In addition, knee arthroscopy is commonly used as a diagnostic procedure for patients with knee pain, and accordingly the patients allocated to the placebo surgery arm can be considered to gain by being assured that no pathology other than a degenerative meniscus injury exists in their knee joint. The other ethical dilemma regarding our procedure actually lies at the heart of our study design, the potential harms and benefits related to the resection of the torn meniscus. The existing evidence shows quite convincingly no benefit of knee arthroscopy in the prevention of the degenerative process (knee osteoarthritis (OA)),[36–38] and accordingly few would argue that the goal of APM in patients with degenerative meniscus injury is solely to relieve symptoms. On this note, there is naturally a risk that patients allocated to the sham arm will suffer from the delay in getting the potentially beneficial procedure, a risk that we have attempted to minimise by allowing patients to request a reassessment of their symptoms (and potentially cross over to the APM arm) after 6 months of follow-up. It should also be noted that all the risks discussed above are related to the current treatment strategy of degenerative meniscus injury, the resection of the meniscus and the recent evidence actually implies that meniscectomy is associated with an increased risk of future knee OA.[39 40] Finally, we contend that there are no objective tools for measuring risk-benefit ratios in research. As Gillett has suggested: 'an acceptable procedure for the placebo arm of a surgical trial is one that carries no more risk than the investigation that would be used to diagnosis and treatment'.[41] We argue that the relatively minor risks related to placebo procedure were justifiable to answer the clinically important question of whether arthroscopic surgery is effective in the treatment of degenerative meniscus injury.

VI. Do the individuals give informed consent?

The ethical issues of the trial will be thoroughly explained and discussed with each recruited patient, both verbally and in writing. The basic principles laid down in the Declaration of Helsinki will be followed throughout the execution of the trial. Accordingly, each participant has the right to withdraw from the study at any given moment without having to explain this decision in any way. Finally, the possibility of being randomised to placebo surgery will be reiterated to each participant before the actual randomisation.

'proven' in the true sense of the word as it applies to scientific research" captures the essence of equipoise and APM.

RCT is considered the gold standard of research design in terms of methodological rigour (internal validity). Ideally, a well-designed RCT should not only have high internal validity but also preferably high external validity (generalisability).[25] However, realistically, such a 'wish' is an obvious paradox, as there is an almost inevitable trade-off between internal and external validity. In essence, the purpose of a true *efficacy* (or explanatory) trial is to demonstrate that an intervention can work theoretically under optimal conditions ('best-case scenario'). The patients recruited are thus carefully selected to obtain as homogeneous a sample as possible and they are also treated under ideal conditions (eg, by the most skilled/experienced physicians). An effectiveness (or pragmatic) trial, in turn, is aimed at testing how an intervention works under usual practice circumstances, and for this reason has a high external validity, but the internal validity is usually lower.[25 42 43] Relton *et al*[44] recently described a novel 'cohort multiple randomised controlled trial' design, in which the benefits of a large observational cohort and a pragmatic RCT can be combined. However, this pragmatic design cannot be used for trials with a placebo comparator (ie, 'true' efficacy design) as the patients may not receive placebo in routine healthcare.

In this paper, we describe the protocol of our ongoing placebo surgery controlled trial assessing the efficacy of APM. To address some of the problems related to both efficacy and effectiveness RCTs, we also briefly introduce a novel 'RCT within-a-cohort design'. This protocol paper has been written according to the recently published SPIRIT guidelines.[45] Hopefully, the publication of this article will make our study objectives and methods more transparent and consequently improve the eventual utility of our study.[46]

METHODS AND ANALYSIS

Study design

We are conducting a parallel (1 : 1), multicentre, randomised and placebo surgery controlled superiority trial to assess the efficacy of APM for patients with a degenerative medial meniscus injury (figure 1). Patients, outcome assessors and data analysts (before the final writing of the manuscript) will be unaware of treatment assignments (triple blind approach).

All patients have been given a written informed consent form and have been made to realise that on entering the study they may receive only placebo surgery, in which case the meniscus tear will be left untreated. They also know that participation in the study is entirely voluntary and the decision they make will not affect their possible future care (in case of refusal). In addition, every patient has been informed of their right to withdraw from the trial whenever they desire without giving the researchers any reason for their decision.

We have enrolled patients at the outpatient clinics of five secondary or tertiary hospitals in Finland. The study was first launched at one site, but to improve recruitment after 2 years (and to ensure multicentre design with its obvious benefits to the generalisability of the results), the study was expanded to four additional sites. The sites were selected on the basis of having an established practice of knee arthroscopy and also an experienced arthroscopic knee surgeon keen to participate in the study. All orthopaedic surgeons involved in the actual interventions of the study are members of the Finnish Arthroscopy Association and have experience of more than 1300 (range 1300–2500) knee arthroscopies with an annual rate of at least 200 (range 200–350) arthroscopies. Before the actual launch of the study at the four additional study centres, the protocol was discussed in the minutest details to ensure that it would be adhered to in a similar manner in all five centres. The principal clinical investigator and the study coordinator

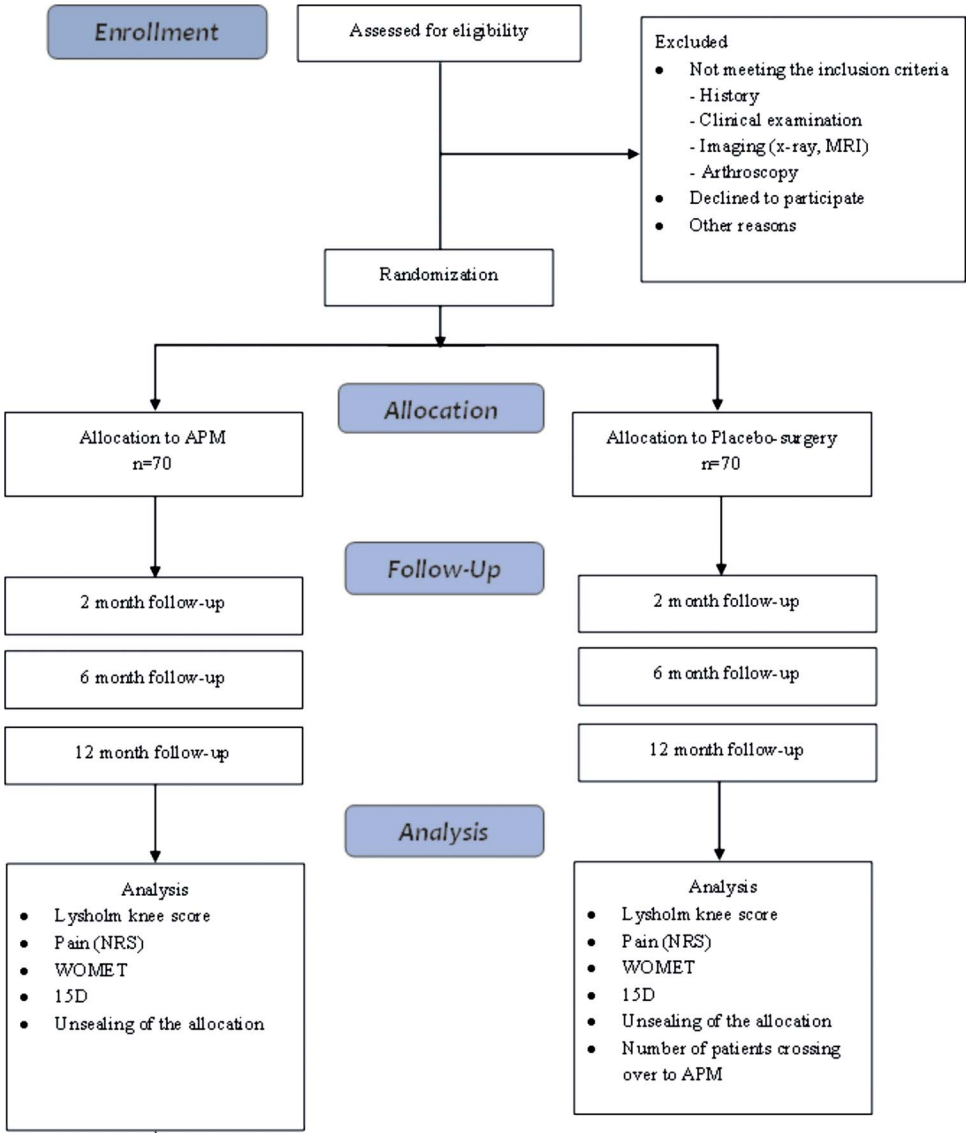

Figure 1 Flowchart of the efficacy trial: enrolment, assigned intervention and follow-up scheme.

also visited every centre at least once before the study was begun at these selected sites. Further, regular meetings have been held biannually among all study members to ensure that the protocol is carried out as planned.

Participants

Patients referred to any of the five orthopaedic surgeons and complaining of knee pain consistent with a meniscus injury were assessed for eligibility. Eligible patients were 35–65 years of age with persisting symptoms (more than 3 months) consistent with a degenerative medial meniscus injury and unresponsive to the appropriate conservative treatment. Surgeons performed a detailed examination of the knee, documenting the range of motion, the presence of an effusion and the results of the meniscal and stability tests. The meniscus injury had to be symptomatic (joint line tenderness, pain in forced flexion or a positive McMurray test). Patients with clinical (American College of Rheumatology) criteria[47] or radiographic knee OA were excluded. Radiographic exclusion criteria included knee OA (grade 2–4) according to the Kellgren and Lawrence scale[48] depicting definite joint space narrowing or definite osteophytes. Other exclusion criteria were a history of obvious traumatic onset of the symptoms (symptoms resulting from kneeling, twisting or slipping without an actual fall were not considered 'traumatic' despite a possible sudden onset), recent episodes of locked knee (the knee cannot be fully extended), previous surgical treatment of the knee, instability or restricted range of motion, previous (within 1 year) fracture of the index extremity and findings on MRI or arthroscopy of any other pathology than a degenerative meniscus injury (requiring an intervention other than APM; the final assessment of chondral lesions was carried out arthroscopically and, if deemed osteoarthritic, they were left untreated). Patients unable to provide informed consent were also excluded. All inclusion and exclusion criteria are listed in box 2.

Box 2 Inclusion and exclusion criteria used in the randomised controlled trial

Inclusion criteria

1. Age: 35–65 years of age
2. Persistent (>3 months) pain on the medial joint line of the knee
3. Pain provoked by palpation or compression of the joint line or a positive McMurray sign
4. MRI showing signals characteristic of medial meniscus injury
5. Degenerative injury to the medial meniscus confirmed at arthroscopy

Exclusion criteria

1. Trauma-induced onset of symptoms
2. Locked knee (that cannot be straightened normally)
3. Previous surgical procedure on the affected knee
4. Clinical osteoarthritis (OA) of the knee (American College of Rheumatology criteria)[47]
5. Radiological OA of the knee (Kellgren-Lawrence grade >1)[48]
6. Acute (within the previous year) fracture of the affected extremity
7. Decreased range of motion of the knee
8. Instability of the knee
9. MRI assessment shows pathology other than degenerative knee disease requiring treatment other than arthroscopic partial meniscectomy (APM)
10. Arthroscopic examination reveals pathology other than a degenerative injury to the medial meniscus requiring intervention other than APM

Baseline

Baseline assessment included sex, birth date, height and weight, mechanism of injury (onset of symptoms), time from the onset of symptoms and standard clinical examination. After being deemed eligible for inclusion, a standardised conservative treatment protocol (exercise programme) was administered by a physiotherapist to all patients who were also given written instructions to ensure the adequacy of conservative treatment before the operation.

Radiographic and MRI

Weight-bearing posteroanterior knee radiography with the use of a fixed-flexion protocol (knees in 20° of flexion and the beam oriented 10° above the horizontal axis)[49 50] and a whole leg radiograph for assessing the possible malalignment was performed for all individuals. MRI scans were obtained with the use of a 1–3-Tesla scanner with a phase-array knee coil. An increased meniscal signal communicating with the inferior, superior or free edge of the meniscal surface seen on two consecutive slices was considered indicative of a meniscal injury.[51] The x-rays and MRI were assessed by both a musculoskeletal radiologist and an orthopaedic surgeon and the final grading was carried out by consensus.

After the baseline visit, the eligible patients were placed on a waiting list for surgery to be operated on within the subsequent 3 months. If the symptoms subsided between the baseline assessment and the day of surgery, the patient was excluded from the trial. These patients, along with those who met the eligibility criteria but declined to participate in the RCT, were asked to give permission for follow-up with a protocol identical to the RCT participants.

Interventions

Arthroscopic examination of the knee was first performed for all patients using standard anterolateral and anteromedial portals with a 4 mm arthroscope (under spinal or general anaesthesia (one site) according to the normal practice of the respective hospitals). The orthopaedic surgeon evaluated the medial, lateral and patellofemoral joint compartments and graded the articular lesions according to the ICRS[52] classification and the meniscus injury(-ies) according to the classification by Cooper *et al*.[53] Patients with obvious chondral flaps or other chondral injuries indicative of trauma, loose

bodies or any finding other than a degenerative medial meniscus injury (requiring an arthroscopic procedure) were excluded from the trial. After a thorough arthroscopic examination of the knee (confirming the eligibility of the patient for the RCT), the patients were randomly assigned to receive APM or placebo surgery. To ensure the blinding of the patients at the four study sites using spinal anaesthesia, the blinding of the patient was further ensured by shielding the patients' view with a vertical drape and aiming the arthroscopy monitors away from the patient's line of vision.

Randomisation and concealment

To enter a patient into the study, a research/staff nurse opened an envelope containing the treatment assignment and revealed it to the surgeon by showing the paper, but the allocation was not expressed verbally. The sequentially numbered, opaque, sealed envelopes were prepared by a statistician with no clinical involvement in the execution of the trial using a computer-generated schedule and the envelopes were kept in a secure, agreed location at each centre. To minimise the risk of predicting the treatment assignment of the next eligible patient (to ensure concealment), randomisation was performed in unfixed blocks (block size known only to the statistician). A randomisation sequence of fourfold stratification, including the study site, the age (35–50 or 51–65 years of age), the gender (female or male) and the existence of radiographic knee OA (no (=K/L 0) or minor degenerative changes (=K/L 1))[48] was used.

Arthroscopic partial meniscectomy

The damaged and loose part of the meniscus tissue was removed with arthroscopic instruments (mechanised shaver and meniscal punches) until solid meniscus tissue was reached. The meniscus was then propped to ensure that all loose and weak fragments and unstable meniscus tissue had been successfully resected, preserving as much of the meniscus tissue as possible.[54] No other surgical procedure (synovectomy, debridement, excision of fragments of articular cartilage or chondral flaps, or abrasion and/or microfracture of chondral defects) was performed.

Placebo surgery

A standard APM procedure was simulated. The surgeon asked for all instruments and manipulated the knee as if APM was being performed. The mechanised shaver (without the blade) was also pushed firmly against the patella, outside of the knee, to mimic as closely as possible the feelings and sounds of the normal use of the arthroscopic shaver. Further, to simulate the sounds of normal APM, suction was also used to drain the joint and saline was splashed. The patient was kept in the operating room for the amount of time required to perform an actual APM.

All procedures were standardised across sites. Furthermore, all arthroscopies were video-recorded for later analysis and to ensure the comparability of all interventions.

In both APM and placebo surgery groups, the postoperative care was delivered according to a standard protocol specifying that all patients received the same walking aids and graduated exercise programme. Analgesia was given according to standard practice. All patients were discharged after full postoperative recovery. The staff delivering the care was blind to the treatment allocation. The notes specifying the actual procedure (APM or placebo) were sealed inside an envelope, but the arthroscopic findings and a phrase 'a procedure according to the degenerative meniscus trial was performed' were included in the patients' medical records.

Prospective cohort (for the 'RCT within-a-cohort' design)

Additionally, in one of the five study centres (the initial site in which over half of the participating patients were enrolled), all patients scheduled for a knee arthroscopy were followed up prospectively with a follow-up assessment protocol identical to that in the actual trial. Those with a degenerative meniscus injury (but who had somehow bypassed our recruitment, were treated by an orthopaedic surgeon not participating in the trial or were excluded from the trial) form a non-randomised, pragmatic, prospective cohort. This pragmatic cohort, together with the patients treated in the actual RCT (those randomised and also those who declined to be randomised), constitutes the 'RCT within-a-cohort' study population (figure 2A).

Outcome measures

The symptoms currently believed to be attributable to a torn meniscus are mostly subjective (pain, discomfort and possible functional deficits) and fluctuating in nature, and accordingly our outcomes are patient-administered outcome measures. The follow-up scheme of this trial consists of four questionnaires (one condition-specific outcome measure, one pain measure, one disease-specific quality of life index and one generic health-related quality of life (HRQoL) index), which are administered preoperatively (on the day of surgery), and then again at 2, 6 and 12 months postoperatively.

Postoperative questionnaires were/are mailed to the patients with an enclosed prepaid return envelope. In case of non-response, the study coordinator contacts these participants by phone, requesting the return of the follow-up questionnaire. At the 12-month follow-up point, the patients are also asked to respond to the following questions: (1) which procedure (APM or placebo surgery) do they think they have undergone, (2) is the knee better than before the intervention, (3) are they satisfied with their knee at present and (4) would they choose to be operated on again if they had to make the decision now. A flowchart of the follow-up scheme is presented in figure 3.

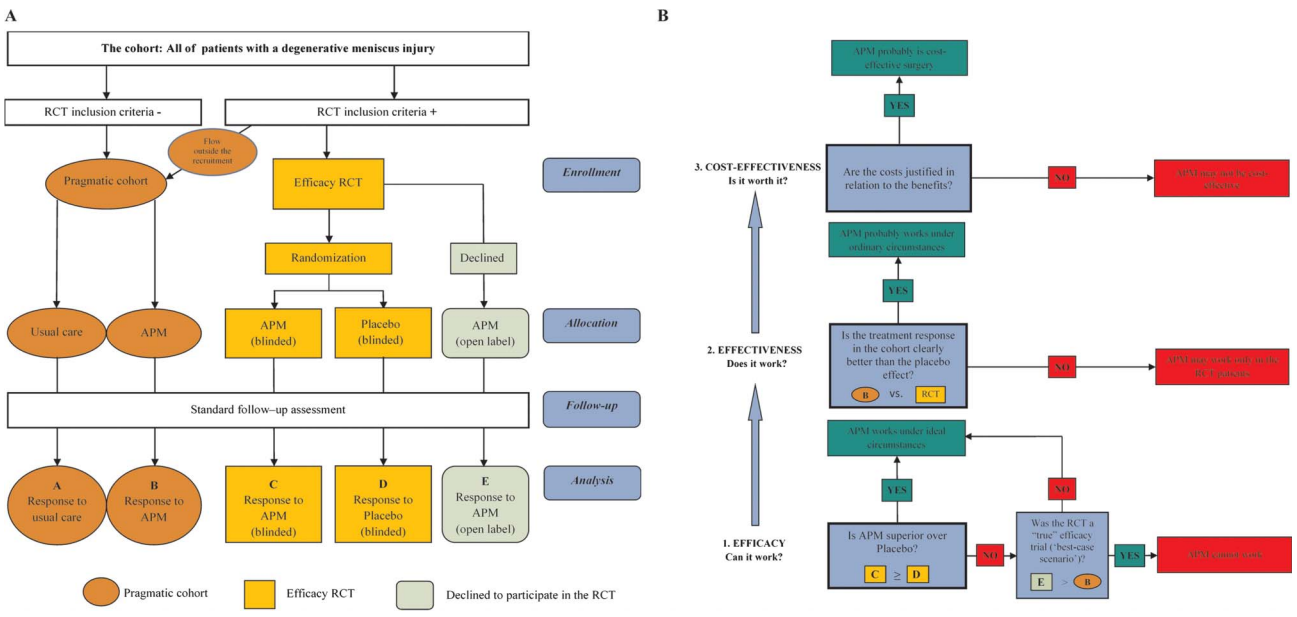

Figure 2 (A) The diagram illustrates the randomised controlled trial (RCT) within-a-cohort study design. By prospectively following up all patients undergoing arthroscopic surgery for their degenerative meniscus injury (and not just those randomised to the efficacy RCT), we are able to obtain several distinct groups of patients with distinct prognostic factors. These groups will enable us to address several interesting hypotheses regarding the efficacy, effectiveness and cost-effectiveness of arthroscopic partial meniscectomy (APM). (B) A theoretical scheme based on Cochrane's three-step hierarchy of evidence for addressing the efficacy, effectiveness and cost-effectiveness of APM. By comparing the treatment response of groups C and D, we can address the efficacy of APM: Is the APM superior to placebo? Then, to elucidate whether our RCT truly represents a 'best-case scenario' (a *true* efficacy trial), we will assess whether the outcome of group E is superior to any subgroup of patients within the pragmatic cohort (B). If these hypotheses hold true, we can continue to 'the second ladder of evidence', effectiveness. We feel that the treatment response observed in the pragmatic cohort (B) should be clearly more robust than the placebo effect of APM to claim that evidence of effectiveness exists. The placebo effect of APM, in turn, can be 'extracted' by subtracting the difference between the groups D and C (blinded APM minus blinded placebo) from the outcome of group E (open APM), a group that should entail the combined effects of placebo and APM. Finally, if evidence of effectiveness is obtained, we can proceed to the assessment of the possible cost-effectiveness ('cost-efficacy'—extrapolation) of APM.

Lysholm knee score

The Lysholm knee score was originally designed to evaluate knee function and symptoms in activities of daily living in patients with anterior cruciate ligament insufficiency.[55] It has later been adjusted and validated for the evaluation of meniscal injuries.[56 57] The Lysholm knee score is a condition-specific outcome measure with eight domains: limp, locking, pain, stair climbing, use of supports, instability, swelling and squatting.[58] An overall score of 0–100 points is calculated, 0 denoting the worst possible outcome and 100 the opposite.[58]

Numerical rating scale for pain

Knee pain at rest and after exercise (over the last week) will be assessed on a 0–10 numerical rating scale, (NRS) comprising 'no pain' at the left (0) and 'extreme pain' at the right (10). NRS is easy to administer and has been validated as a measure of pain intensity in populations with known pain.[59]

Western Ontario Meniscal Evaluation Tool

The Western Ontario Meniscal Evaluation Tool (WOMET) is a disease-specific tool designed to evaluate HRQoL in patients with meniscal pathology.[60] It has been developed to be the measure including the most important items for patients with meniscal tear[61] and was recently validated for patients with degenerative meniscus injury.[62] WOMET contains 16 items addressing three domains: nine items for domains of physical symptoms, four for domains of disabilities due to sports,

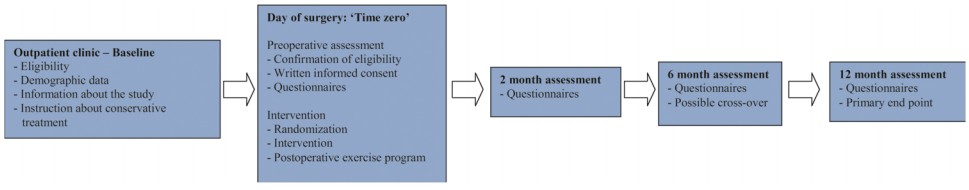

Figure 3 A diagram outlining the follow-up scheme used in the efficacy randomised controlled trial.

recreation, work and lifestyle and three for the emotions domain.[60] Each of the WOMET items is given a visual analogue scale (100 mm lines anchored at the ends). The best or least symptomatic situation ranks 0 and the most symptomatic rank possible is 1600.[60] For the sake of simplicity, the score is usually converted to a percentage of normal, in which zero (0) denotes the worst possible situation and 100 the best possible situation.

15D

The 15D instrument is a generic HRQoL instrument comprising 15 dimensions.[63] For each dimension, the respondent must choose one of the five levels that best describes his/her state of health at the moment (the best level being 1 and the worst level being 5). A set of utility or preference weights is used in an addition aggregate formula to generate a single index number, the utility or 15D score. The maximum 15D score is 1 (no problems on any dimension) and the minimum score is 0 (being dead). The responsiveness, reliability and validity of 15D have been thoroughly established, and this instrument has been used extensively in clinical and healthcare research.[64 65]

Initially, our two primary outcome measures were the Lysholm knee score[55] and the NRS for pain (VAS) at 12 months after surgery. The secondary outcome was the WOMET[60] score at the same time point and a cost-utility analysis (based on the patients' general quality of life using the 15D score and the utilisation of healthcare resources). At this preregistration stage (2007), the Lysholm score was conceived of as a prespecified primary outcome as this tool has been validated for meniscus injury[56] and is also the outcome instrument used most often for various knee conditions.[58 66] However, during the recruitment phase of the trial, we showed that WOMET has acceptable psychometric properties for the evaluation of patients with a degenerative meniscus injury and early OA,[62] and accordingly have now decided to use it as our third primary outcome.

Reoperations/treatment conversions

Patients were told at the time of giving their consent that they would be allowed to consider crossing over to the other procedure 6 months or later after the arthroscopic procedure if adequate relief of symptoms was not achieved (based on the existing evidence showing that it can take up to 6 months postoperatively to obtain the actual benefit from the APM).[67 68] No specific numerical thresholds of outcome measures were set for the option for crossover, but the patients will be asked in the 6-month follow-up questionnaire about their willingness to undergo reoperation based on their current symptoms. If a patient reports symptoms requiring attention at 6-month follow-up or becomes symptomatic after that, an appointment will be scheduled with an orthopaedic surgeon both blind to the initial treatment and not involved in the care of the patient before the appointment. If the clinical examination carried out at this point reveals clinical signs indicative of/consistent with a meniscal tear and the patient requests reoperation, the allocation will be unsealed and the patient will be offered APM. The number of patients both requiring additional follow-up appointments and unsealing of the allocation is registered in both the APM and placebo groups. MRI is not automatically performed for symptomatic patients because of its low specificity in differentiating residual or recurrent symptoms from the signal due to postresection (APM).[69] However, in case of a new trauma to the index knee or exceptionally severe symptoms, an MRI examination will also be carried out. In addition to the aforementioned predetermined assessment of the possible adverse events in the 6-month follow-up questionnaire, the patients are instructed to contact the study coordinator with any possible complaints related to the knee surgery at any given moment during the postoperative period. These contacts (events) will be registered.

Loss to follow-up

As in an earlier placebo surgery controlled trial,[70] we will document the number and proportion of individuals eligible for and compliant with each follow-up. Individuals who die during the study (from causes unrelated to the study or procedure) will also be tabulated. If the proportion of individuals withdrawing from either arm is greater than the anticipated 20% at 1 year, an analysis of the demographic and prognostic characteristics will be made between the individuals who withdraw and those who remain in the study. For continuous variables, parametric or non-parametric analysis of variance will be used. For categorical variables, χ^2 or Fisher's exact test will be applied.

Missing items

When individual items are missing from a scale, we will calculate the per cent of missing items.

If the missing items are less than 10%, we will impute values using the mean of the remaining items. If more than 10% of the scale score is unavailable for analysis, the patient will be excluded from the analysis at that time point. For the WOMET score, our approach to missing values is slightly different: If a patient fails to complete 1–3 items, we will substitute the missing value (s) with the average value for the answered items according to the protocol described previously for the Western Ontario and McMaster Universities Osteoarthritis Index, a similar outcome tool used for the assessment of established knee OA.[71 72]

Sample size

For the sample size calculation, we used our prospectively collected database of 377 consecutive patients with a degenerative meniscus injury who had undergone APM (data not published). Minimal clinically important improvement (MCII) was assessed with an anchor-based method (patient's global impression of change used as

an external anchor).[73] MCII was found to be 11.5 for the Lysholm knee score and 15.5 for the WOMET score. SDs for the response (postoperative score—preoperative score) of the same population were 21.1 and 24.5 for the Lysholm and WOMET scores, respectively. For the determination of the MCII for pain, we extrapolated our estimates from the existing literature on pain-VAS. Patients with knee OA report the MCII for pain to be 19.9 mm[74] and in our own cohort, the SD of response of pain after exercise was 31.4 mm. Using these estimates, the required sample sizes were 40, 54 and 40 participants per group for the Lysholm score, WOMET score and pain assessment, respectively, with 80% power to show a clinically meaningful advantage of APM over placebo, based on a two-sided type 1 error rate of 5%. Given an anticipated dropout level of 20% and the possible problems related to uneven randomisation (multicentre design with sites as one stratum), we decided to recruit 70 patients per group.

Safety analysis

For obvious ethical and safety reasons, an interim analysis by an independent data and safety monitoring board (the National Institute for Health and Welfare) was carried out after the enrolment of 70 (50%) patients to ensure that the rates of complications or rearthroscopies were within acceptable limits (within the normal rate of complications and/or reoperations related to knee arthroscopy). This analysis is planned to be carried out blind to the group assignment and unless a marked discrepancy is found in the incidence of complications/reoperations, the allocation concealment will be preserved. No other interim analysis is envisaged.

Data management

At each study site, the data are collected by a research nurse dedicated to the study. Forms are mailed to the coordinating centre (Hatanpää Hospital, Tampere), where all data are entered into a computerised database. After the completion of follow-up, all data are re-entered by an individual with no clinical involvement in the execution of the trial and then these two separate databases are compared for consistency. Participant files will be maintained in storage (both in electronic and paper format) at the coordinating centre for a period of 5 years after completion of the study.

Blinded data analysis

The Steering and Writing Committees (RS, MP, AM and TLNJ) will develop and record two interpretations of the results on the basis of a blinded review of the primary outcome data (treatment A compared with treatment B), with one assuming that A is the APM group and another assuming that A is the placebo surgery group. The Writing Committee will also deliberate before data analysis to determine the key analyses and presentation format for the primary publication. The minutes of this meeting are recorded as a statement of interpretation document, which will be signed by all members of the Steering and Writing Committees before the unsealing of the randomisation.[75]

Statistical methods

The trial was designed to ascertain the superiority of APM over placebo at 12 months with primary outcome measures by the means of statistically significant differences between the groups. Baseline characteristics will be analysed by descriptive statistics. For the primary analysis, the change in scores (Lysholm and WOMET) and pain (NRS) at 12 months will be compared between the two study groups. A secondary analysis of the primary endpoint will adjust for those prerandomisation variables which might reasonably be expected to influence the outcome (ie, sex, age, radiological OA and baseline score). A two-sided p value of 0.05 will be considered to indicate statistical significance. We will use the Bonferroni method to appropriately adjust the overall level of significance for multiple (three) primary outcomes. Analyses of the outcome measures score will also be performed at 2 and 6 months. Additionally, the number and percentage of satisfied patients, those subjectively improved, and those whose allocation was unsealed will be calculated and compared between the two groups. All statistical analyses will be performed on an intention-to-treat basis, meaning that data from patients will be analysed according to the initial study-group assignment regardless of compliance with randomisation. However, as the potential for treatment conversion exists, but only from placebo surgery to APM, per-protocol analysis will also be carried out. We plan to conduct a subgroup analysis to compare the result based on the degree of radiographic OA (K&L 0 vs I). The serious adverse event rates occurring within 12 months of surgery for both study arms will be evaluated descriptively in aggregate and separately. If the number of events is large enough to allow more sophisticated analysis, the rates between the APM and placebo arms will be compared. The data will be presented by the total number of events and the total number of individuals with at least one event.[70]

IBM SPSS Statistics (V.20 or later, IBM Corporation) is used for all statistical analyses.

ETHICS AND DISSEMINATION

Ethical issues related to the use of placebo surgery

Our application to the Pirkanmaa Hospital District IRB contained a specific, six-point ethical analysis focusing on the methodological rationale for use of placebo surgery, risk-benefit assessment and informed consent as outlined by Miller (box 1).[31]

Dissemination

The findings of this study will be disseminated widely through peer-reviewed publications and conference presentations.

DISCUSSION

In this protocol paper, we describe the execution of a randomised, placebo surgery controlled trial for the assessment of the efficacy of APM in patients with a degenerative meniscus injury. Given the imminent risk for bias from an unblended trial of surgery versus medical (conservative) treatment—particularly acknowledging the potential of surgery to produce powerful placebo effects[32 76 77]—it is doubtful that a rigorous trial of surgery can be conducted without a sham surgery control when the primary outcome is pain, patient-reported improvement or quality of life.[24 78 79] As the only previous placebo surgery controlled trial in knee arthroscopy[32] prompted significant criticism from the orthopaedic discipline at large,[15 80–91] we decided to use this debate as a template in rationalising our methodological choices (table 1).

GENERALISABILITY

Alleged poor generalisability was the most prevalent concern regarding the study by Moseley *et al*.[15 80–83 85 87–91] Explanatory trials are designed to measure the efficacy of an intervention—to find out if the treatment exerts a biological effect in a research setting *under ideal and controlled conditions*. The factors most consistently found to predict a poor outcome after APM in patients with a degenerative meniscus injury are advanced knee OA and chondral damage.[39 94 95] Also, lateral meniscectomy has been identified as a predictor of poorer prognosis rather than medial meniscectomy.[39 96] Accordingly, to obtain a population with 'best-possible response to APM', we chose to recruit patients with a stable knee, no or minimal OA and an isolated, degenerative injury of the medial meniscus. Although age per se is not a predictor for poor outcome, it has been shown to be associated with increased

Table 1 Concerns expressed regarding the Moseley *et al*[32] trial and their related pilot study[92] and our solution to these concerns

Concern	Author (year)	Our solution
Generalisibility		
▶ All men	Felson (2002)[80]	
▶ Age	Jackson (2002)[81]	
▶ Lack of well-defined exclusion criteria	Ewing (2002)[82]	
▶ Suboptimal study population	Chambers (2002)[83]	
– Stage of OA	Johnson (2002)[85]	Study population represents 'best-case scenario'
– Mechanical symptoms	Gillespie (2003)[87]	RCT within-a-cohort design (pragmatic cohort included)
– Meniscal tears	Fowler (2003)[88]	
– Range of motion	Marx (2008)[89]	
– Mechanical malalignment	Kirkley (2008)[15]	
	Ilahi (2010)[90]	
	Felson (2010)[91]	
Number of denial (selection bias)	Chambers (2002)[83]	Study population represents 'best-case scenario'
Only one surgeons	Kirkley (2008)[15]	Multicentre study, five experienced knee surgeons
	Felson (2010)[91]	
Non-validated outcome measurement	Felson (2002)[80]	Validated outcomes
	Chambers (2002)[83]	
	Kirkley (2008)[15]	
	Felson (2010)[91]	
No intention-to-treat analysis	Felson (2002)[80]	Intention-to-treat analysis
Power calculations not based on primary outcome	Chambers (2002)[83]	Prior power calculations based on primary outcomes
Post hoc equivalency bounds	Chambers (2002)[83]	
	Gillespie (2003)[87]	
Superiority trial design, but equivalency trial conclusions	Chambers (2002)[83]	Superiority design and conclusions
Inappropriate publication forum (non-orthopaedic)	Johnson (2002)[85]	
Insufficient socioeconomic data	Johnson (2002)[85]	
Insufficient data on anatomy and surgical procedures performed	Johnson (2002)[85]	Standardised data collecting form and video recordings of all surgical procedures
Data reported as group average	Chambers (2002)[83]	Possibility to report the results as proportion of those who benefited/satisfied
Investigator bias	Ilahi (2010)[90]	
Concerns regarding possible existence of placebo effect	Hartman (1997)[93]	RCT within-a-cohort design

OA, osteoarthritis; RCT, randomised controlled trial.

prevalence of OA and (consequently) inferior functional recovery after APM.[39] Similarly, women have also been found to have a less favourable prognosis after APM.[94] These two factors were thus used as stratifications.

Acknowledging the obvious risk of high internal but low external validity due to the strict eligibility criteria, we decided to also keep track of the patients eliminated from the RCT by prospectively following these individuals with the same outcome assessment arsenal used in the RCT. One of the reasons for collecting this pragmatic cohort was to safeguard against bias attributable to patients declining to take part in a placebo surgery controlled trial. However, unlike in the trial by Moseley *et al*[32] (only about 40% of the eligible patients agreed to participate), we have been very successful in convincing patients to participate (15% of eligible patients declined). This pragmatic cohort will also be used to corroborate or refute whether our efficacy (RCT) population represents the best-case scenario (those with optimal response to APM) (figure 2B).

SINGLE VERSUS MULTIPLE SURGEONS/CENTRES

The study by Moseley *et al*[32] has also been criticised for having only one surgeon performing all operations.[15 32 91] Similar to Moseley *et al*,[32] we also began our trial in only one centre/surgeon, but once the study protocol had been thoroughly refined, it was expanded to four additional sites. The inclusion of five surgeons/study sites entails one problem: we chose to use 'study site' as our fourth stratification (separate randomisation sequences for each site), and accordingly, there is an imminent risk of uneven randomisation as the number of patients allocated at each of the four additional sites is relatively small (projected to be 10–20/site) and there are three stratifications in the randomisation. However, the use of separate randomisation sequences was a conscious choice, as we wanted to retain the possibility of analysing the data separately from each site.

RANDOMISATION

By postponing the randomisation into the operation suite, we have managed to completely eliminate the chance that any eligible patient who gives informed consent would decline to participate in the trial after being randomised. Even though our 'RCT within-a-cohort' design allows us to also follow up those declining, the elimination of postrandomisation declining obviously decreases the risk of bias.

OUTCOME ASSESSMENT AND STATISTICAL ANALYSIS

According to the CONSORT statement,[97] authors should not only define the prespecified outcome considered to be of the greatest importance to relevant stakeholders, but they should also indicate the prespecified time point of primary interest when outcomes are assessed at several time points after randomisation. Few would argue that the goal of treatment in patients with degenerative meniscus injury is anything but to relieve symptoms, particularly as the existing evidence quite convincingly shows that degenerative meniscus injury per se and APM both inevitably lead to the development of knee OA.[36–38] As noted previously, it can take up to 6 months postoperatively to obtain the full benefit from the APM,[67 68] but a more lasting relief of symptoms seems to be confounded by eventual progression of the underlying degenerative process. Accordingly, to be able to showcase the potential efficacy of APM on pain and the quality of life while minimising various types of confounding (eg, non-retention/loss to follow-up and progression of knee OA), we chose a 12-month time point as our time point of primary interest. However, the potential beneficial or detrimental short-term effects of either of the studied procedures (eg, possible transient accelerated recovery) can also be assessed using the data from the 2-month and 6-month follow-up assessments. Further, as Moseley *et al*[32] were also criticised for reporting only average change of a group,[86] we decided to ask our patients about their global impression of change with their knee status and their satisfaction with their knee. These two questions enable us to also report the outcome of our population as a proportion of patients who consider they have benefited from the treatment.

As for the prespecified outcome of the greatest importance to patients, Moseley *et al*[32] developed a patient-administered outcome instrument of their own, a methodological choice that was criticised as the tool was not properly validated.[15 80 83 86] We have chosen three validated outcome instruments in our trial. To safeguard against potential problems associated with multiplicity of analyses,[97] our solution is twofold. We will use the Bonferroni method to appropriately adjust the overall level of significance and also take this issue into account in interpreting our findings: to be deemed hypothesis-proving, the intervention-induced changes have to be well aligned and clinically meaningful in all three validated primary outcomes.

LIMITATIONS OF THE STUDY

The definition of the concept 'degenerative' or 'traumatic' in the context of meniscal injuries is arbitrary in nature. Traditionally, meniscal injuries or tears have been classified as traumatic or degenerative based on the aetiology (injury mechanism) or morphology (tear pattern observed in MRI or at arthroscopy), but no validated and generally agreed criteria exist to make the distinction between the two entities. Even the definition of the word 'trauma' represents a challenge. Although a 'traumatic' onset of symptoms is indeed an exclusion criterion in our RCT, the majority of patients with a 'degenerative' meniscus injury do experience some kind of twisting movement or other relatively modest injury prior to the onset of their knee symptoms.[98] Accordingly, all patients with sudden injuries related to their own voluntary muscle

activities (such as kneeling, bending or kicking) and patients with a minor twisting of the knee were included in our trial. In essence, our criteria labelling a tear as 'traumatic' describe a more substantial event, such as falling from a chair, stairs or bicycle or slipping on ice. To increase the validity of the classification of the meniscus injury, our assessment includes not only history/clinical examination but also MRI and arthroscopy. Finally, our RCT within-a-cohort design allows us to examine whether the response to APM of those excluded from the RCT because of 'traumatic' meniscal injury differs from those included in the RCT.

A similarly controversial issue is the differentiation of chondral injuries that require surgical procedure (debridement, microfracture or any other procedure) from those that should be considered merely osteoarthritic.[99] To the best of our knowledge, there is no objective classification system (either MRI or arthroscopy based) for this distinction. Our rationale was to use preoperative MRI to verify that the patients' symptoms could be due to degenerative meniscus injury (and to exclude possibly rare situations as knee tumours). No cartilage lesions detected in MRI were considered an exclusion, but the final decision whether or not to include the patient in the trial was made at arthroscopy, where the meniscus injury (and commonly associated chondral lesions) were visually and physically (probing) verified and assessed. As there really is no valid (RCT -based) evidence to show that surgery would provide any clinically meaningful improvement in patients with degenerative chondral lesions,[15 32] we instructed our surgeons to generally forbear from surgical intervention. However, if a single large chondral flap or local deep osteochondral lesion was found, the surgeons were given permission to carry out the surgical procedure they deemed appropriate and the patient was excluded. Finally, the only difference between our two allocation groups is whether a meniscus resection was carried out or not. Accordingly, if we made a wrong decision concerning chondral lesions, they should still be evenly divided into the two study arms, thus not compromising our primary objective.

INTERPRETATION OF THE RESULTS: INTRODUCING VARIOUS POTENTIAL SCENARIOS

To illustrate the various potential findings of the study, we have fitted different scenarios within a larger context of Cochrane's three-step hierarchy of evidence (figure 2B). This scheme outlines the requirements for evidence to justify claims of proven efficacy, effectiveness and, ultimately, the cost-effectiveness of APM.

Quantifying the placebo effect

The placebo effect refers to a positive change in symptoms related to a participant's perceptions of the treatment rather than to the mechanisms of the treatment itself.[21] After the publication of their pilot study,[92] Moseley *et al* were questioned about the possibility that patients who enrolled in their placebo surgery controlled study would be more 'suggestible' and therefore more susceptible to the placebo effect.[93] Moseley responded by proposing two alternative scenarios: either the patients in their RCT would be more susceptible to the placebo effect than the patients who declined to be in the study (suggesting that the characteristics and personality traits of patients who choose to enrol in the studies are different from those declining) or patients who enrol in the RCT are more susceptible to the placebo effect than patients who undergo surgery outside a study. For the sake of brevity, we will not enter into a lengthy elaboration of these two alternative interpretations (which can be found in the detailed response by Moseley[100]).

Translating this line of thought into our study design, potential differences in the responses to treatment between the RCT study groups (C, D and E, figure 2A) are attributable to either the effect of resection (groups C vs D, figure 2A) or the magnitude of the placebo effect (groups C vs E, figure 2A). What makes our design unique is the opportunity to actually disentangle (quantify) the respective effects of 'resection' and 'placebo'. Further, by comparing (with appropriate baseline adjustment) these estimates with the treatment responses observed in the patients who have undergone APM outside our study (group B, figure 2A) and the more heterogeneous pragmatic cohort (group A, figure 2A), we should also be able to extrapolate the potential effectiveness of the procedure (figure 2B). Theoretically, the placebo effect is more robust in the 'open label' (E) and pragmatic groups (A and B) than in the efficacy trial groups (C and D) due to the doubt inherent in any 'blinded' RCT (figure 2A). To summarise, there is a placebo effect in all five study groups, not just the placebo surgery group (D).

CONCLUSION

In this article, we present a protocol for a randomised placebo surgery controlled trial to assess the efficacy of APM in patients with degenerative meniscus tear. We have also discussed some of the methodological issues that we feel are most pertinent to the successful execution of a controlled surgical trial, particularly with respect to minimising bias and maximising the internal and external validity of the study.

Author affiliations
[1]Department of Orthopaedics and Traumatology, Hatanpää Hospital, Tampere, Finland
[2]Department of Orthopaedics and Traumatology, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland
[3]National Institute for Health and Welfare, Centre for Health and Social Economics, Helsinki, Finland

Acknowledgements We would like to thank Timo Järvelä, MD, PhD for his help and expertise during the design phase of the trial, Heini Huhtala, MSc (Tampere School of Public Health/Biostatistics) for her help in planning the

statistical analyses, and research coordinator Pirjo Toivonen for her vital role in implementation. We would also like to acknowledge Kari Tikkinen, MD, PhD for his critical comments on the manuscript and Virginia Mattila, MA for linguistic expertise and language revisions.

Contributors TLNJ, RS, MP and AM designed the efficacy study. TLNJ and RS designed the 'RCT within-a-cohort' extension and wrote the manuscript. MP and AM provided substantive input on drafts of the manuscript. All authors contributed to the refinement of the study protocol and approved the final manuscript. TLNJ is the principal investigator of the study.

Funding This project is supported by the Sigrid Juselius Foundation, the Competitive Research Fund of Pirkanmaa Hospital District grant number: 9M121, the Social Insurance Institution of Finland grant number: 50/26/2012, and the Academy of Finland grant number: 259503. The funding sources had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data or decision to submit results.

Competing interests None.

Patient consent Obtained.

Ethics approval The institutional review board of the Pirkanmaa Hospital District (R06157).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Jinks C, Jordan K, Ong BN, *et al*. A brief screening tool for knee pain in primary care (KNEST). 2. Results from a survey in the general population aged 50 and over. *Rheumatology (Oxford)* 2004;43:55–61.
2. McAlindon TE, Cooper C, Kirwan JR, *et al*. Knee pain and disability in the community. *Br J Rheumatol* 1992;31:189–92.
3. O'Reilly SC, Muir KR, Doherty M. Knee pain and disability in the Nottingham community: association with poor health status and psychological distress. *Br J Rheumatol* 1998;37:870–3.
4. Mantyselka P, Kumpusalo E, Ahonen R, *et al*. Pain as a reason to visit the doctor: a study in Finnish primary health care. *Pain* 2001;89:175–80.
5. Englund M, Guermazi A, Gale D, *et al*. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med* 2008;359:1108–15.
6. Englund M, Niu J, Guermazi A, *et al*. Effect of meniscal damage on the development of frequent knee pain, aching, or stiffness. *Arthritis Rheum* 2007;56:4048–54.
7. Weber KL. The AAOS clinical practice guidelines. *J Am Acad Orthop Surg* 2009;17:335–6.
8. Cullen KA, Hall MJ, Golosinskiy A. Ambulatory surgery in the United States, 2006. *Natl Health Stat Report* 2009;11:1–25.
9. Kim S, Bosque J, Meehan JP, *et al*. Increase in outpatient knee arthroscopy in the United States: a comparison of national surveys of ambulatory surgery, 1996 and 2006. *J Bone Joint Surg Am* 2011;93:994–1000.
10. Baumgaertner MR, Cannon WD Jr, Vittori JM, *et al*. Arthroscopic debridement of the arthritic knee. *Clin Orthop Relat Res* 1990;253:197–202.
11. McLaren AC, Blokker CP, Fowler PJ, *et al*. Arthroscopic debridement of the knee for osteoarthrosis. *Can J Surg* 1991;34:595–8.
12. Ogilvie-Harris DJ, Fitsialos DP. Arthroscopic management of the degenerative knee. *Arthroscopy* 1991;7:151–7.
13. Yang SS, Nisonson B. Arthroscopic surgery of the knee in the geriatric patient. *Clin Orthop Relat Res* 1995;316:50–8.
14. Herrlin SV, Wange PO, Lapidus G, *et al*. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. *Knee Surg Sports Traumatol Arthrosc* 2013;21:358–64.
15. Kirkley A, Birmingham TB, Litchfield RB, *et al*. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008;359:1097–107.
16. Katz JN, Chaisson C, Chaisson C, *et al*. The meteor trial: preliminary results of an RCT of arthroscopic partial meniscectomy vs. physical therapy in patients greater than 45. *Arthritis Rheum* 2012;64:SS1123.
17. Howell J, Handoll H. WITHDRAWN: surgical treatment for meniscal injuries of the knee in adults. *Cochrane Database Syst Rev* 2009;(1):CD001353.
18. Katz JN, Martin SD. Meniscus—friend or foe: epidemiologic observations and surgical implications. *Arthritis Rheum* 2009;60:633–5.
19. Englund M, Guermazi A, Lohmander LS. The meniscus in knee osteoarthritis. *Rheum Dis Clin North Am* 2009;35:579–90.
20. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2004;(3):CD003974.
21. Dowrick AS, Bhandari M. Ethical issues in the design of randomized trials: to sham or not to sham. *J Bone Joint Surg Am* 2012;94(Suppl 1):7–10.
22. Revicki DA, Osoba D, Fairclough D, *et al*. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. *Qual Life Res* 2000;9:887–900.
23. Willke RJ, Burke LB, Erickson P. Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved product labels. *Control Clin Trials* 2004;25:535–52.
24. Zhang W, Robertson J, Jones AC, *et al*. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2008;67:1716–23.
25. Farrokhyar F, Karanicolas PJ, Thoma A, *et al*. Randomized controlled trials of surgical interventions. *Ann Surg* 2010;251:409–16.
26. Cobb LA, Thomas GI, Dillard DH, *et al*. An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med* 1959;260:1115–18.
27. Mehta S, Myers TG, Lonner JH, *et al*. The ethics of sham surgery in clinical orthopaedic research. *J Bone Joint Surg Am* 2007;89:1650–3.
28. Boutron I, Ravaud P, Nizard R. The design and assessment of prospective randomised, controlled trials in orthopaedic surgery. *J Bone Joint Surg Br* 2007;89:858–63.
29. Macklin R. The ethical problems with sham surgery in clinical research. *N Engl J Med* 1999;341:992–6.
30. Tenery R, Rakatansky H, Riddick FA Jr, *et al*. Surgical 'placebo' controls. *Ann Surg* 2002;235:303–7.
31. Miller FG. Sham surgery: an ethical analysis. *Sci Eng Ethics* 2004;10:157–66.
32. Moseley JB, O'Malley K, Petersen NJ, *et al*. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:81–8.
33. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* 2000;283:2701–11.
34. Horng S, Miller FG. Is placebo surgery unethical? *N Engl J Med* 2002;347:137–9.
35. Jameson SS, Reed MR, Dowen D, *et al*. The burden of arthroscopy of the knee: a contemporary analysis of data from the English NHS. *J Bone Joint Surg Br* 2011;93:1327–33.
36. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. *Arthritis Rheum* 2003;48:2178–87.
37. McDermott ID, Amis AA. The consequences of meniscectomy. *J Bone Joint Surg Br* 2006;88:1549–56.
38. Petty CA, Lubowitz JH. Does arthroscopic partial meniscectomy result in knee osteoarthritis? A systematic review with a minimum of 8 years' follow-up. *Arthroscopy* 2010;27:419–24.
39. Salata MJ, Gibbs AE, Sekiya JK. A systematic review of clinical outcomes in patients undergoing meniscectomy. *Am J Sports Med* 2010;38:1907–16.
40. Roos EM, Ostenberg A, Roos H, *et al*. Long-term outcome of meniscectomy: symptoms, function, and performance tests in patients with or without radiographic osteoarthritis compared to matched controls. *Osteoarthritis Cartilage* 2001;9:316–24.
41. Gillett GR. Unnecessary holes in the head. *IRB* 2001;23:1–6.
42. Roland M, Torgerson DJ. What are pragmatic trials? *BMJ* 1998;316:285.
43. Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheum Dis* 1999;58(Suppl 1):I82–5.
44. Relton C, Torgerson D, O'Cathain A, *et al*. Rethinking pragmatic randomised controlled trials: introducing the 'cohort multiple randomised controlled trial' design. *BMJ* 2010;340:c1066.
45. Chan AW, Tetzlaff JM, Gotzsche PC, *et al*. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
46. Godlee F. Publishing study protocols: making them more visible will improve registration, reporting and recruitment. *BMC News Views* 2001;2:4.
47. Altman R, Asch E, Bloch D, *et al*. Development of criteria for the classification and reporting of osteoarthritis. Classification of

osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039–49.
48. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16:494–502.
49. Nevitt MC, Peterfy C, Guermazi A, *et al*. Longitudinal performance evaluation and validation of fixed-flexion radiography of the knee for detection of joint space loss. *Arthritis Rheum* 2007;56:1512–20.
50. Kothari M, Guermazi A, von Ingersleben G, *et al*. Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. *Eur Radiol* 2004;14:1568–73.
51. Fox MG. MR imaging of the meniscus: review, current trends, and clinical implications. *Radiol Clin North Am* 2007;45:1033–53, vii.
52. Brittberg M, Winalski CS. Evaluation of cartilage injuries and repair. *J Bone Joint Surg Am* 2003;85-A(Suppl 2):58–69.
53. Cooper DE, Arnoczky SP, Warren RF. Meniscal repair. *Clin Sports Med* 1991;10:529–48.
54. Metcalf R, Burks R, Metcalf M, *et al*. *Athroscopic meniscectomy*. 2nd edn. Philadelphia, PA: Lippincott-Raven, 1996.
55. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med* 1982;10:150–4.
56. Briggs KK, Kocher MS, Rodkey WG, *et al*. Reliability, validity, and responsiveness of the Lysholm knee score and Tegner activity scale for patients with meniscal injury of the knee. *J Bone Joint Surg Am* 2006;88:698–705.
57. Kocher MS, Steadman JR, Briggs KK, *et al*. Reliability, validity, and responsiveness of the Lysholm knee scale for various chondral disorders of the knee. *J Bone Joint Surg Am* 2004;86-A:1139–45.
58. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res* 1985:43–9.
59. Jensen M, Karoly P. *Handbook of pain assessment*. 2nd edn. New York: Giuldford Press, 2001.
60. Kirkley A, Griffin S, Whelan D. The development and validation of a quality of life-measurement tool for patients with meniscal pathology: the Western Ontario Meniscal Evaluation Tool (WOMET). *Clin J Sport Med* 2007;17:349–56.
61. Tanner SM, Dainty KN, Marx RG, *et al*. Knee-specific quality-of-life instruments: which ones measure symptoms and disabilities most important to patients? *Am J Sports Med* 2007;35:1450–8.
62. Sihvonen R, Jarvela T, Aho H, *et al*. Validation of the Western Ontario Meniscal Evaluation Tool (WOMET), a meniscal pathology-specific Quality-of-Life index, for patients with a degenerative meniscus tear. *J Bone Joint Surg Am* 2012;94:e65-1–8.
63. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33:328–36.
64. Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res* 2008;17:485–95.
65. Bowling A. *Measuring health. A review of quality of life measurement scores*. 3rd edn. Berkshire: Open University Press, 2004.
66. Garratt AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford)* 2004;43:1414–23.
67. Roos EM, Roos HP, Ryd L, *et al*. Substantial disability 3 months after arthroscopic partial meniscectomy: a prospective study of patient-relevant outcomes. *Arthroscopy* 2000;16:619–26.
68. Herrlin S, Hallander M, Wange P, *et al*. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007;15:393–401.
69. Rosas HG, De Smet AA. Magnetic resonance imaging of the meniscus. *Top Magn Reson Imaging* 2009;20:151–73.
70. Gray LA, Jarvik JG, Heagerty PJ, *et al*. INvestigational Vertebroplasty Efficacy and Safety Trial (INVEST): a randomized controlled trial of percutaneous vertebroplasty. *BMC Musculoskelet Disord* 2007;8:126.
71. Bellamy N, Buchanan WW, Goldsmith CH, *et al*. 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.
72. Bellamy N. *Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index: users guide*. London, Ontario: University of Western Ontario, 1996.
73. Copay AG, Subach BR, Glassman SD, *et al*. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7:541–6.
74. Tubach F, Ravaud P, Baron G, *et al*. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005;64:29–33.
75. Bhandari M, Guyatt G, Tornetta P 3rd, *et al*. Study to prospectively evaluate reamed intramedually nails in patients with tibial fractures (S.P.R.I.N.T.): study rationale and design. *BMC Musculoskelet Disord* 2008;9:91.
76. Kallmes DF, Comstock BA, Heagerty PJ, *et al*. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *N Engl J Med* 2009;361:569–79.
77. Buchbinder R, Osborne RH, Ebeling PR, *et al*. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med* 2009;361:557–68.
78. Doherty M, Dieppe P. The 'placebo' response in osteoarthritis and its implications for clinical practice. *Osteoarthritis Cartilage* 2009;17:1255–62.
79. Moerman DE, Jonas WB. Deconstructing the placebo effect and finding the meaning response. *Ann Intern Med* 2002;136: 471–6.
80. Felson DT, Buckwalter J. Debridement and lavage for osteoarthritis of the knee. *N Engl J Med* 2002;347:132–3.
81. Jackson RW. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:1717–19; author reply 17–9.
82. Ewing W, Ewing JW. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:1717–19; author reply 17–9.
83. Chambers KG, Schulzer M. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:1717–19; author reply 17–9.
84. Poehling GG. Degenerative arthritis arthroscopy and research. *Arthroscopy* 2002;18:683–7.
85. Johnson LL. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *Arthroscopy* 2002;18:683–7.
86. Chambers K, Schulzer M, Sobolev B. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *Arthroscopy* 2002;18:683–7.
87. Gillespie WJ. Arthroscopic surgery was not effective for relieving pain or improving function in osteoarthritis of the knee. *ACP J Club* 2003;138:49.
88. Fowler P. Arthroscopic lavage or debridement did not reduce pain more than placebo did in patients with osteoarthritis. *J Bone Joint Surg Am* 2003;85-A:387.
89. Marx RG. Arthroscopic surgery for osteoarthritis of the knee? *N Engl J Med* 2008;359:1169–70.
90. Ilahi OA. Selection bias results in misinterpretation of randomized controlled trials on arthroscopic treatment of patients with knee osteoarthritis. *Arthroscopy* 2010;26:144–6.
91. Felson DT. Arthroscopy as a treatment for knee osteoarthritis. *Best Pract Res Clin Rheumatol* 2010;24:47–50.
92. Moseley JB Jr, Wray NP, Kuykendall D, *et al*. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo-controlled trial. Results of a pilot study. *Am J Sports Med* 1996;24:28–34.
93. Hartman GP. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo controlled trial. *Am J Sports Med* 1997;25:724; author reply 24–5.
94. Meredith DS, Losina E, Mahomed NN, *et al*. Factors predicting functional and radiographic outcomes after arthroscopic partial meniscectomy: a review of the literature. *Arthroscopy* 2005;21:211–23.
95. Fabricant PD, Rosenberger PH, Jokl P, *et al*. Predictors of short-term recovery differ from those of long-term outcome after arthroscopic partial meniscectomy. *Arthroscopy* 2008;24: 769–78.
96. Chatain F, Adeleine P, Chambat P, *et al*. A comparative study of medial versus lateral arthroscopic partial meniscectomy on stable knees: 10-year minimum follow-up. *Arthroscopy* 2003;19:842–9.
97. Moher D, Hopewell S, Schulz KF, *et al*. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.
98. Camanho GL, Hernandez AJ, Bitar AC, *et al*. Results of meniscectomy for treatment of isolated meniscal injuries: correlation between results and etiology of injury. *Clinics (Sao Paulo)* 2006;61:133–8.
99. Katz JN, Chaisson CE, Cole B, *et al*. The MeTeOR trial (meniscal tear in osteoarthritis research): rationale and design features. *Contemp Clin Trials* 2012;33:1189–96.
100. Fleiss DJ. Arthroscopic treatment of osteoarthritis of the knee: a prospective randomised, placebo controlled trial. *Am J Sports Med* 1997;25:724; author reply 24–5.