Author's response to reviews

Title: Arthroscopic Partial Meniscectomy in Middle-Aged Patients with Mild or No Knee Osteoarthritis: A Double-Blind, Randomized Sham-Controlled Multi-Centre Trial

Authors:

Kristoffer B Hare (kbhr@regionsjaelland.dk)
Stefan Lohmander (stefan.lohmander@med.lu.se)
Robin Christensen (robin.christensen@frh.regionh.dk)
Ewa M Roos (eroos@health.sdu.dk)

Version: 3 Date: 12 October 2012

Author's response to reviews: see over
Cover letter

We thank the three reviewers for the thorough review of our design article. We have identified a number of comments noted by the reviewers. All these comments have been responded to below.

Reviewer - Gunter Spahn

1. **Reviewer Comment**
   Title: The title must more clearly enunciate that it is a study concept only. It suggest: Arthroscopic partial meniscectomy in middle-aged patients with mild or no knee osteoarthritis: Pre-study concept for a double-blind, randomized sham-controlled multi-centre study.

   **Author’s Response**
   As pointed out by the reviewer, this is a design paper. The title is already preceded with the heading; “Study Protocol”, which is similar to other study protocol articles published in BMC. We hope that the reviewer will agree that this particular Protocol premise for BMC Journals are acceptable explanation. Anticipating that the readers of a protocol article would not expect to see any real results except for the conceptual framework.

   **Author’s Action**
   None

2. **Reviewer Comment**
   Abstract: Background: The first sentence isn’t correct. Many studies did demonstrate the positive effect of partial meniscectomy in symptomatic (!) patients. It should be enough to say “The optimal treatment of a degenerative meniscus tear in patients with mild OA is unknown”.

   **Author’s Response**
   The sentence refers to patients with concomitant knee osteoarthritis and meniscus lesion and refers to the randomized clinical trials by Moseley, Kirkley and Herrlin [1-3] which failed to show any positive effect of “house cleaning” including meniscectomy, as compared to sham surgery, medical treatment ‘as usual’, or exercise. These publications represent the highest level of evidence in this area. Other studies are lower quality observational studies and prone to bias.

   **Author’s Action**
   None

3. **Reviewer Comment**
   Background: In this chapter the authors are very one-sided. You haven’t done your study! In this section it is urgently needed to discuss the difference between symptomatic meniscus-tears positive clinical tests) and the casually tear (e.g. MRI finding)! The authors cite Mosely et al. and Kirkley et al. Why don’t you cite Hubbard et al. with discordant results? [Hubbard MJ (1996) articular debridement versus washout for degeneration of the medial femoral condyle. A five-year study. J Bone Joint Surg Br 78:217-219] Please be correct!

   **Author’s Response**
   The matter of symptomatic meniscus tear is discussed below (paragraph 6). In our opinion the study by Hubbard does not contribute to the background. Firstly, it regards osteochondral lesions and not meniscus lesions. Our inclusion criteria are set to identify a sample with a degenerative meniscal tear but with no or at most mild cartilage damage. Secondly, the methodological quality of the Hubbard study is poor (lack of blinding, clinician assessed outcomes, randomization process not adequately described, etc.) and the study is thus prone to severe bias. Specifically, it seems like the primary outcome, defined as pain relief or not, was collected by the treating surgeon (worst
case scenario of observer bias). Neither the surgeon nor the patients were blinded to the intervention.

**Author's Action**
None

4. **Reviewer Comment**
Methods: The main default of your concept is the patient selection! This must be revised before this study concept is acceptable. How many patients you will include in the study. This must be calculated before you perform a randomized study. Did you perform a pre-study power analysis to calculate the number of patients? This is urgently needed from the ethical point of view to reduce the number of needless sham-operations.

**Author's Response**
We did perform an a priori sample size calculation and have carefully considered the design and methods part of the study to set up a study that meets the highest methodological and ethical criteria. The sample size calculation is described on page 8: “The sample-size calculation is based on the assumed superiority of the arthroscopic procedures over the sham procedure. For a two-sample pooled t-test of a normal mean difference with a two-sided significance level of 0.05, assuming a common standard deviation of 15 in the KOOS\textsubscript{5} score, a sample size estimation of the ITT population indicated that 36 individuals per group would be required to obtain a power of at least 80% to detect a difference between groups of 10 KOOS\textsubscript{5} score units. Following these estimations, it was decided to include 80 individuals in total (40 patients in each group), allowing for a 10% drop-out rate.”

**Author's Action**
None

5. **Reviewer Comment**
How do you recruit your patients? It is very important to describe the time-point of inclusion. I'm sure high-class sportsmen or patients with heavy work will decline the participation into the study. But these patients have to be listed also! Mosely performed his study in highly selected patients collective! This is the main bias of this study! He only included US-army veterans with a very high demand for financial benefits. Similar effects we know from our German patients in the “occupational insurance”! You have to register all potential candidates for this study before inclusion! This must clearly describe in the study protocol.

**Author's Response**
We agree with the reviewer that describing the screening stage and patient eligibility is an important part of any RCT and this is included in the Consort Statement which we will follow closely. Though not described in detail in the study protocol, the recruitment phase will be thoroughly described in our main article when the results are published. This article will also contain a flow chart describing patients screened, excluded and included.

**Author's Action**
None

6. **Reviewer Comment**
The preoperative evaluation isn’t sufficient for an arthroscopic meniscectomy! The prevalence of meniscus tears/damage in MRI is very high. Only few patients really are suffering from symptomatic” meniscus tears. This only can made by are careful clinical diagnostic. Before you perform a meniscectomy the patients should present positive clinical meniscus-signs (e.g. Steinmann, McMurray…). This must documented preoperatively as well as in your follow-ups.
Author’s Response
All patients in the study are referred from general practitioners on the suspicion of a meniscus tear and deemed eligible (i.e. with relevant symptoms consistent with a meniscal tear) for arthroscopy by the primary investigator, who is an orthopedic surgeon. Unfortunately, there is no consensus on what defines a symptomatic meniscus tear, and clinical tests (McMurray, Apley, etc.) have not been proven to diagnose a meniscus tear accurately [4]. It is not the purpose of this study to discuss what defines a symptomatic meniscus tear compared to a non-symptomatic meniscus tear, but to reflect routine practice. Therefore, we used MRI as a diagnostic tool.

Author’s Action
This is somewhat already described in the discussion but we have changed the wording on page 11 to: “There is no consensus on what defines a symptomatic meniscus tear or whether or not to perform an MRI before surgery. Clinical tests (McMurray, Apley, etc.) have not been proven to diagnose a meniscus tear accurately [4].”

7. Reviewer Comment
The KL-score alone has a high “interobserver variance”. I suggest to perform standardized 30°-flexion standing radiographs. Thus you are able to evaluate objective parameters for the “diagnosis OA”. The joint space narrowing, the varus or valgus angle, the extension of osteophytes has to describe.

Author’s Response
We perform standardized 20° flexion radiographs [5] as described in the article. The Kellgren & Lawrence score is often used in clinical settings for making a treatment decision. We only use the Kellgren & Lawrence classification to exclude moderate and severe knee osteoarthritis (K&L ≥ 3). A high inter-observer variance has primarily been shown to relate to mild stages of disease (K&L 0 vs. K&L 1) [6].

Author’s Action
We have included the reference regarding the inter-observer variance of the Kellgren-Lawrence score into the article on page 5: “Patients with grade 3 or 4 knee OA on the Kellgren & Lawrence classification [6, 7] or knee surgery within the previous 2 years will also be excluded.”

8. Reviewer Comment
The interventions also need some complements. When do you perform “meniscus surgery”? How do you deals with marginal fringes for example? What do you do in case of cartilage lesions? Do you remove loose bodies in both groups? Please clarify!

Author’s Response
The arthroscopic intervention is described as Intervention A in the article on page 5: “The arthroscopic partial meniscectomy will be performed on an outpatient basis by experienced surgeons who are at least in their final year of residency or are attending orthopedic surgeons. All arthroscopies will be performed with general anesthesia combined with local anesthesia (Bupivacain combined with Adrenalin) 20 + 20 ml extra- and intra-articularly, respectively. After general anesthesia is induced, the knee will be examined for stability. Thereafter, two standard portals on the lateral and medial sides of the ligamentum patella will be created but no outflow cannula inserted. An arthroscope will be used with a pressure-controlled irrigation system. Tourniquet use will be at the discretion of the surgeon. The strategy for the meniscectomy will be to preserve as much tissue as possible. A standard operation protocol will be used to document possible findings in cartilage, ligaments, synovium and the medial and lateral menisci. The type, and extent of meniscus lesion will be registered and changes in the articular cartilage will be classified and documented according to the ICRS classification [8].”
Any other findings by the surgeon other than meniscus lesion will be treated as indicated, with common arthroscopic procedures. These possible additional procedures will be noted in the perioperative report and described in the results paper.

**Author's Action**
None

9. **Reviewer Comment**
How many surgeons are involved? Are there well-experienced surgeons only?

**Author's Response**
We expect between 5-10 surgeons to be involved in the study who are all experienced as described under the same paragraph as above on page 5: “The arthroscopic partial meniscectomy will be performed on an outpatient basis by experienced surgeons who are at least in their final year of residency or are attending orthopedic surgeons”.

**Author's Action**
The number of surgeons involved in the study has been added to the sentence above on page 5 in the article: “The arthroscopic partial meniscectomy will be performed on an outpatient basis by experienced surgeons who are at least in their final year of residency or are attending orthopedic surgeons. We expect between 5-10 surgeons to be involved in the study.”

10. **Reviewer Comment**
Have you criteria for abortion? For my personal meaning it isn’t ethical to let well alone a “buckle-handle tear” in a young patient! How do you handle patients with insufferable knee pain at 3month control?

**Author's Response**
We have no criteria for dropping out of the study. Bucket-handle tears are already excluded. The study is voluntary and any patient can select to drop out of the study at any time and request an alternative treatment. Patient-reported outcomes will be collected at the time of drop-out, and the patient will continue to be monitored according to the Intention To Treat population.

**Author's Action**
None

11. **Reviewer Comment**
The postoperative evaluations must include indications about recovery to occupation and sports (job-title, sports-kinds, time-interval etc.).

**Author's Response**
The inclusion criteria are set to maximize external validity of the results. Patients ranging from 35-55 years are included, and no inclusion criteria are set with regard to physical activity level, yielding a very diverse group. When reviewing the literature re. measures of activity level it becomes clear that to adequately measure physical activity accelerometers are required. That was not possible for logistic reasons in this study. Scores available for assessing return to sport include the Tegner score and the Marx activity score. Tegner score is the most commonly used. However, it has been noted that in a sample of soccer players aged 18-35, the Tegner score is dependent on gender, age and prior level of activity. While younger men overestimate their activity level older women underestimate their activity level. The Marx score is not applicable in subjects who are not athletes. The activity scores available for middle-aged people are usually validated against METs, and where not considered appropriate in this context. We acknowledge the lack of valid measures of activity level for this patient group, and have thus not included any such measure.

**Author's Action**
We have added the following to the discussion on page 11: “The study does not include an activity score. A literature search revealed a lack of valid self-reported instruments of activity level for this diverse middle-aged population of varying physical activity levels. Since providing patients with accelerometers was not an option due to logistic reasons, we have not included any measure of activity level in this trial.”

Reviewer – Morgan Jones

1. Reviewer Comment
The authors mention the issue of arthroscopic surgery during the follow-up period, which represents either re-operation or crossover from non-operative to operative treatment. In some ways this represents another secondary outcome measure. However, it is unclear from reading the manuscript what criteria will be applied to determine whether or not patients undergo a repeat arthroscopy, and whether any of the subsequent care (such as repeat MRI) might unblind the patients or the treatment team.

Author’s Response
Patients are free to discontinue their participation in the study at any time, and without prejudice to further treatment. We don’t have any specific treatment failure criteria. If a patient reports “treatment failure” and suspect they have undergone sham surgery they are free to elect to be unblinded and to be offered a usual arthroscopy. We call these patients cross-overs.

Author’s Action
The following section has been added to the article on page 10: “Treatment failure. No a priori criteria for cross-over are given. Should a patient contact the department because of unbearable symptoms they will be un-blinded and in case of having had placebo surgery they will be offered a new arthroscopy. These patients will be treated as cross-overs and still be included in the study. In case of the patient having had arthroscopic surgery in the first place the patient will be referred to the responsible surgeon who will be in charge of referring to further surgical or non-surgical treatment and/or investigation. Both patient groups will be asked to fill out a KOOS questionnaire at the extra visit and will continue to be followed at the follow-ups determined by the study protocol.”

2. Reviewer Comment
While the sample size determination is very clearly presented based on the primary outcome, the number depends very much upon the standard deviation and clinically important difference chosen for this calculation. The authors should better explain and support their reasons for these numbers, as some investigators have suggested a clinically important difference of 8 points for individual scales of the KOOS, which would increase sample size to 124 if accounting for 10 percent dropout.

Author’s Response
Today it is increasingly recognized that there is no such thing as a single minimal important change (MIC) for an instrument. The MIC is dependent on context factors such as patient group, intervention, time to follow-up etc. This understanding is reflected in the updated KOOS users guide available at www.koos.nu. In this study, we have based our power calculation, MIC of 10 points and SD of 15 on findings from similar patient groups and interventions [9].

Author’s Action
We’ve changed the description of the sample size calculation to: “The sample-size calculation is based on the assumed superiority of the arthroscopic procedures over the sham procedure. For a two-sample pooled t-test of a normal mean difference with a two-sided significance level of 0.05, assuming a common standard deviation (SD) of 15 in the KOOS3 score, a sample size estimation of the ITT population indicated that 36 individuals per group would be required to obtain a power of at
least 80% to detect a minimal important change (MIC) of 10 KOOS5 score units. The MIC of 10 points and SD of 15 is based on findings from similar patient groups and interventions [9].”

3. **Reviewer Comment**
   The authors framed their study as a trial that may fail to demonstrate that arthroscopy is better than sham surgery for younger patients with meniscus tear and that the study has the potential to change practice. However, a study with 80 percent power and the somewhat larger clinically important difference leads to a smaller sample size and less potential to change practice if the groups are not statistically different.

   **Author’s Response**
   This is correct. The findings from our study have less potential to change practice if no statistically significant difference is found between groups. Whereas a significant improvement in the arthroscopy group compared to the placebo group would show that meniscectomy is beneficial, it will be harder to state that the treatments are equivalent since the sample size is based on a superiority assumption. We believe that not being able to show that meniscectomy has a greater benefit than a sham procedure should lead to a change in practice, though we accept the risk of a type 2 error.

   **Author’s Action**
   None

4. **Reviewer Comment**
   The authors describe joint space width narrowing as a secondary outcome. It would be helpful if they could present the difference they might expect to detect with the sample size that they have proposed.

   **Author’s Response**
   In RCT:s the sample size calculation is based on the primary outcome. Thus no formal sample size calculation has been performed for JSW. We will base expected difference on best available evidence at time of data analysis.

   **Author’s Action**
   None

**Reviewer – Lars Wiedenhielm**

**Reviewer Comment**
My only problem with this study as the authors discuss themselves on page 10 in the 3 paragraph is that the results of the study will heavily rely on a correct interpretation of the MRI-examinations and the ability to diagnose a degenerative tear in the meniscus. If, for instance, no meniscus lesion exists despite the interpretation of the preop MRI-examinations stating a degenerative meniscus lesion and the patient will be randomized to a sham operation, this error will never be discovered and if the patients symptoms resolve the false conclusion in this case will be that there was a degenerative tear in the meniscus that didn’t need treatment. I have not been able to find any published study about the number of false positive and false negative interpretations of meniscus lesions in MRI-examinations of the knee but I know that in our study (Herrlin at al. 2012) there were a couple of cases with positive MRI-findings and no meniscal tear on arthroscopic examination. An alternative approach to the sham operation without entering the knee might be to do an arthroscopic examination on all the cases to be sure that there actually is a tear in the meniscus and then only resect the torn meniscus in the examination group. However, the authors might feel that an arthroscopic examination will be too big to qualify as sham surgery.

**Author’s Response**
This is an obvious concern that we have debated intensively. We have three major reasons for not performing a diagnostic arthroscopy in the sham group: 1) When performing sham surgery one must make the risk as low as possible. By only making a skin incision we will fully avoid the risk of a deep knee infection from the sham procedure; 2) We wanted the sham group to be free of any accidental osteochondral lesions from the arthroscope; 3) We didn’t want to risk any unnecessary house cleaning performed by the surgeon.

**Author’s Action**

We have included these reasons in the discussion on page 11: “Another limitation of using MRI as a diagnostic tool is the risk of false positive result. If the patient will be randomized to a sham operation, this error will never be discovered. We chose not to perform a diagnostic arthroscopy in the sham group primarily to reduce the risk of deep infection which we find would be unacceptable for a sham intervention. Other reasons were to avoid any accidental osteochondral lesions from the arthroscope and unwanted intervention from the surgeons.”

**References**


