CHANGING OUR TREATMENT OF DEGENERATIVE MENISCAL TEARS

Victor A. van de Graaf

Changing our treatment of degenerative meniscal tears

Victor Adrianus van de Graaf

Publication of this thesis was kindly supported by:

Joint Research OLVG – Nederlandse Orthopaedische Vereniging – Stichting Wetenschap OLVG – Nederlandse Vereniging voor Arthroscopie – D.H. Heijne Stichting / Basko Healthcare – Spronken Orthopedie NV – Zimmer Biomet Nederland BV – LM Orthopedie – Mathys Orthopaedics BV – Centrum Orthopedie Rotterdam – Smith & Nephew – Bauerfeind Benelux BV – Össur – Heraeus Medical – Castor EDC – Chipsoft.

ISBN: 978-94-6182-997-9

Layout and printing: Offpage, Amsterdam

Cover design: Malou F. Schoevaars

Changing our treatment of degenerative meniscal tears

Herziening van onze behandeling van degeneratieve meniscusscheuren
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 13 maart 2020 des middags te 2.30 uur

door

Victor Adrianus van de Graaf

geboren op 7 maart 1986
te Fijnaart en Heijningen

Promotoren:

Prof. dr. D.B.F. Saris
Prof. dr. R.W. Poolman

Copromotoren:

Dr. A. de Gast
Dr. N.W. Willigenburg

This thesis is based on the following papers

van de Graaf VA, Bloembergen CH, Willigenburg NW, Noorduyn JCA, Saris DBF, Harris IA, Poolman RW; for the Escape research group. Can even experienced orthopaedic surgeons predict who will benefit from surgery when patients present with degenerative meniscal tears? A survey of 194 orthopaedic surgeons who made 3880 predictions. Br J Sports Med. 2019 Aug 1. pii: bjsports-2019-100567.

van de Graaf VA, van Dongen JM, Willigenburg NW, Noorduyn JCA, Butter IK, de Gast A, Saris DBF, van Tulder MW, Poolman RW; for the Escape research group. How do the costs of physical therapy and arthroscopic partial meniscectomy compare? A trial-based economic evaluation of two treatments in patients with meniscal tears alongside the Escape study. Br J Sports Med. 2019 Jun 21. pii: bjsports-2018-100065.

Noorduyn JCA, **van de Graaf VA**, Mokkink LB, Willigenburg NW, Poolman RW; for the Escape research group. Responsiveness and minimal important change of the IKDC in middle-aged and older patients with a meniscal tear. Am J Sports Med. 2019 Feb;47(2):364-371.

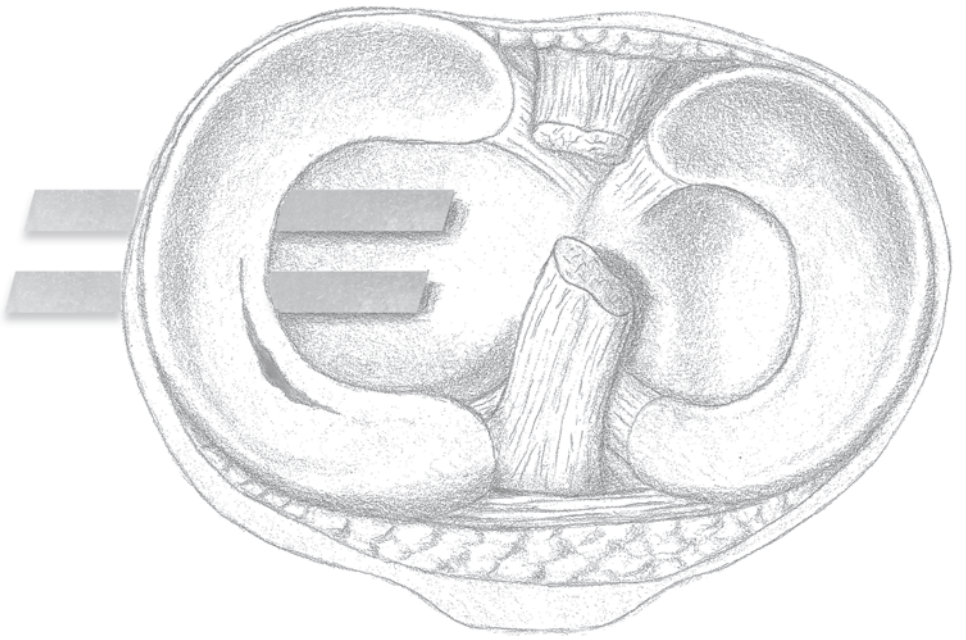
van de Graaf VA, Noorduyn JCA, Willigenburg NW, Butter IK, de Gast A, Mol BW, Saris DBF, Twisk JWR, Poolman RW; for the Escape research group. Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears: the Escape randomized clinical trial. JAMA. 2018 Oct 2;320(13):1328-1337.

van de Graaf VA, Scholtes VA, Wolterbeek N, Noorduyn JC, Neeter C, van Tulder MW, Saris DBF, de Gast A, Poolman RW; for the Escape research group. Cost-effectiveness of early surgery versus conservative treatment with optional delayed meniscectomy for patients over 45 years with non-obstructive meniscal tears (Escape study): protocol of a randomised controlled trial. BMJ Open. 2016 Dec 21;6(12):e014381.

van de Graaf VA, Wolterbeek N, Mutsaerts ELAR, Scholtes VAB, Saris DBF, de Gast A, Poolman RW. Arthroscopic partial meniscectomy or conservative treatment for nonobstructive meniscal tears: a systematic review and meta-analysis of randomized controlled trials. Arthroscopy. 2016 Sep;32(9):1855-1865.e4.

van de Graaf VA, Wolterbeek N, Scholtes VAB, Mutsaerts ELAR, Poolman RW. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. Am J Sports Med. 2014 Jun;42(6):1408-16.

Table of Contents



1

Introduction and outline of this thesis

Introduction

The menisci are essential for retaining normal function of the knee joint. Their semilunar and wedge profile with attachment to the joint capsule at the peripheral rim make that they transmit tibiofemoral joint load forces. With ageing, there is a gradual loss of cellular elements resulting in increased water content and decreased elasticity.[1-4] As a result, the menisci become more vulnerable to damage and meniscal tears may occur spontaneously as part of a degeneration process.[1,2]

When the torn or degenerative meniscus is surgically removed, the joint contact area decreases and as a consequence the joint peak load increases (Figure 1).[5] The magnitude of this increase in joint peak load depends on the amount of meniscal tissue removed.[5,6] A higher joint peak load is believed to contribute to a faster progression of osteoarthritis.[5,7]

In the early 2000s, arthroscopic lavage and debridement for knee osteoarthritis was one the most frequently performed orthopaedic surgical procedures. Surgeons moved away from this procedure after a placebo-controlled randomised clinical trial (RCT) demonstrated that the improvement experienced by patients from this procedure could be attributed to a placebo effect.[8] Surgeons had been performing arthroscopic surgery for osteoarthritis on a large scale for years without ever rigorously studying its effectiveness.

After the publication of this placebo-controlled trial by Mosely and colleagues, the number of knee arthroscopies performed for osteoarthritis declined by 18% between 1996 and 2006 in the United States.[9] It seemed that orthopaedic surgeons changed their practice to arthroscopic partial meniscectomies (APM) as the number of meniscectomies increased by 25% in the same period.[9] In England, the most significant change was seen

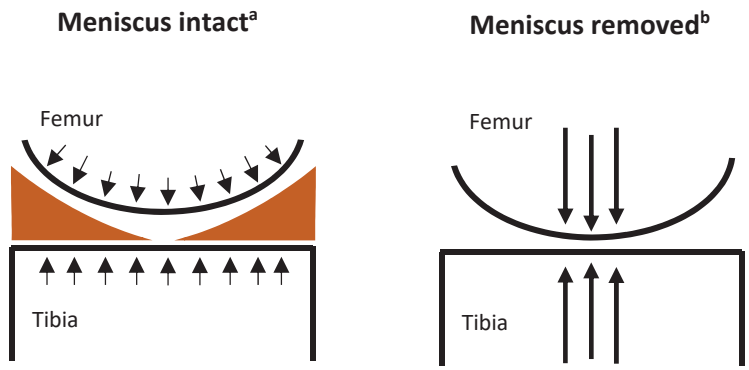


Figure 1. The biomechanical load distribution of the meniscus.
[a] The menisci increase the joint contact area and reduce the contact pressure on the articular cartilage.
[b] After a meniscectomy, the joint contact area decreases. This causes an increase in the peak contact pressure on the cartilage and in the risk of osteoarthritis, depending on the amount of meniscus removed. Illustration adapted from Mc Dermott and colleagues [6]

after the introduction of the National Institute for Health and Care Excellence (NICE) guidance in 2007. This guidance resulted in a decrease of 80% in the number of knee arthroscopies for osteoarthritis but an increase of 230% in the number of meniscectomies from 2000 to 2012.[10] Whether this change could entirely be explained by policy changes or also partly by rebranding (using a different code to perform the same procedure) is unclear.[11-13] Interestingly, this shift to partial meniscectomy lacked any evidence-based grounds as well.[14-16]

In 2008, shortly after this shift to partial meniscectomy, Englund and colleagues described that meniscal tears are common incidental findings in older patients. They found meniscal tears in up to 60% of asymptomatic persons with knee osteoarthritis (Kellgren-Lawrence grade 2 or higher).[17,18] They were the first to describe that no causality exists between having such a degenerative meniscal tear and knee symptoms. The authors hypothesised that meniscal tears in knees with osteoarthritis should be considered as part of the degenerative process rather than a separate disorder.[17] A meniscectomy in a degenerative knee further increases the joint peak load and accelerates the osteoarthritis. The authors therefore questioned whether a meniscectomy in these patients would reduce the knee symptoms.[15]

Herrlin and colleagues published the first RCT assessing the effectiveness of partial meniscectomy in patients with degenerative meniscal tears in 2007.[19] The authors found no difference in knee function and knee pain between partial meniscectomy followed by supervised exercise therapy or supervised exercise therapy alone after 2 years follow-up.[19] These results did not stir the orthopaedic community as the number of partial meniscectomies continued to grow after this publication.[10,14,20]

It took until 2013 for more evidence to become available.[3,4,21,22] Again, the effects of arthroscopic surgery in knees with osteoarthritis could be attributed to a placebo effect.[4] The number of partial meniscectomies decreased in the following years. However, this decrease was smaller than expected.[20]

Given the results of the published trials, the use of arthroscopic surgery in the degenerative knee appears hard to justify. However, orthopaedic surgeons appeared unconvinced by the evidence as partial meniscectomies continued to be performed in large numbers. Reasons for this include:[23]

- Published trials demonstrating that meniscectomy was effective for improving knee function and knee pain;
- Trials using outcome instruments less likely to pick up clinically relevant differences in treatment;
- A 30% non-responder rate to conservative treatment;
- Patients in the RCTs not representing the patients that orthopaedic surgeons select for surgery;
- Having concerns about the quality and generalizability of the existing evidence;

- Having difficulty in implementing the conclusions from scientific evidence into clinical practice;
- Being convinced that the patient in the outpatient clinic will benefit more from surgery;
- High demanding patents with higher expectations from surgery;
- Sending patients to a physical therapist feels like doing nothing;
- Being convinced that patients will recover faster from surgery;
- General resistance to move away from generally accepted treatments among clinicians;
- Believing in being able to predict who will benefit more from surgery.

Demonstrating that the benefit of partial meniscectomy could be attributed to a placebo effect appeared insufficient to move away from this procedure as common practice. This is mainly the result of our healthcare system in which demonstrated (cost-)effectiveness is not a strict condition for financial coverage. The fact that surgeons continued to perform meniscectomies despite this lack of effectiveness is not a complete surprise as the outcome of 65% of current medical treatments is unknown or ineffective.[24,25] In times of an ongoing rise of healthcare costs, physicians should aim to deliver only those interventions that add value for our patients (Figure 2). Ineffective care will not increase the value for our patients and should be banned.

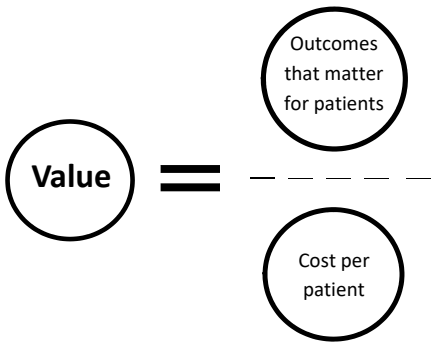


Figure 2. The concept of value-based healthcare. Value-Based Healthcare aims to maximise the value of healthcare for patients by providing the best care at the lowest cost.

Given the significant clinical and economic implications of partial meniscectomies, research should focus on the societal impact in order to create a more efficient healthcare system. The difference between effectiveness and efficiency can best be explained as:

"Being effective is about doing the right things, while being efficient is about doing things right."
- Peter Drucker -

The main goal of healthcare efficiency is to provide the greatest benefit per unit of cost. In other words, efficient healthcare is about finding the optimal balance between the highest quality of care at the most affordable costs.[26] Therefore, healthcare efficiency

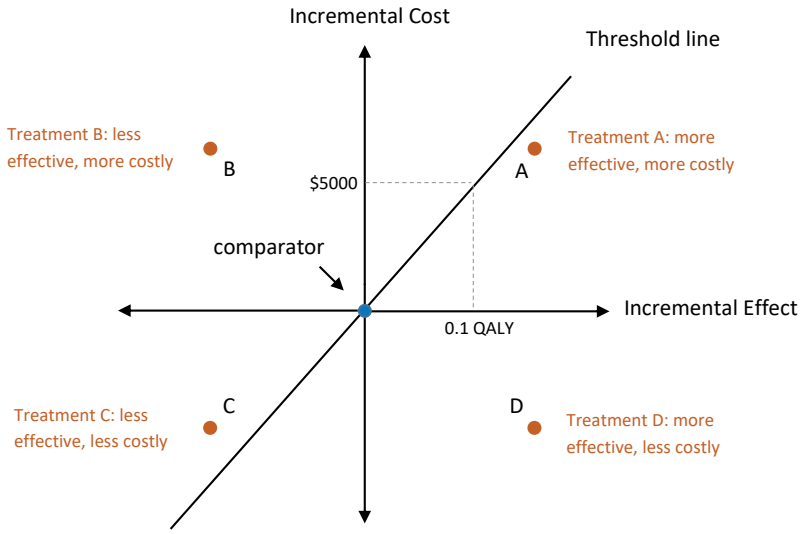


Figure 3. Incremental Cost-Effectiveness Plane. The horizontal axis represents the difference in effects and the vertical axis represents the difference in costs between the treatment of interest and the comparator (incremental cost-effectiveness ratio (ICER)). The diagonal line represents the threshold line and represents the maximum acceptable costs-effectiveness ratio (i.e. the trade-off between efficient and non-efficient care).
Treatment A is more effective and more expensive, but under the trade-off and therefore accepted.
Treatment B is less effective and more costly and therefore rejected.
Treatment C is less effective and less costly and above the trade-off and therefore rejected.
Treatment D is more effective and less costly and therefore accepted.
If a treatment is in the northwest (B) or southeast (D) quadrant, the decision is clear (accept). However, if a treatment is in the north-east (A) or south-west (C) quadrant, the decision depends on the maximum acceptable costs-effectiveness ratio (threshold line). The slope of this threshold line depends on the treatment. [Illustration adapted from [27]]

studies add important information for decision-makers in practice and policy. When comparing 2 treatments, both the change in cost and the change in effect should be considered and weighed, for example in a cost-effectiveness analysis (Figure 3).

Economic evaluations provide information on the relative efficiency of at least 2 interventions.[28] If a treatment is both more expensive and more effective or less expensive and less effective than the comparator, the criterion for efficiency depends on the willingness to pay. If this is below the maximum acceptable cost-effectiveness ratio one is willing to accept, the treatment will be accepted.[26] Economic evaluation studies nowadays are of particular interest as we are currently in the transition phase towards a more value-based healthcare system (Figure 2).

Aims

Our approach to the treatment of degenerative meniscal tears needs serious reconsideration. The first publication by Herrlin and colleagues that found no difference in effects between partial meniscectomy and conservative treatment was the main incentive for the work in this thesis.[19] Being aware of an ongoing placebo-controlled trial [4] and a large RCT [3], our research group focused on the cost-effectiveness of partial meniscectomy compared to conservative treatment. We aimed to determine whether arthroscopic partial meniscectomy or conservative treatment provides better value for money in patients over 45 years of age with degenerative meniscal tears.

We hypothesised that supervised exercise therapy provides better value for money compared to arthroscopic partial meniscectomy in these patients.

The following research goals were formulated for this thesis:

1. To review and evaluate the available literature to support the use of partial meniscectomy and to expose the lack of evidence in the treatment of patients with degenerative meniscal tears (*Chapter 2*)
2. To assess the reliability and validity of different measurement instruments used for evaluating treatment in patients with meniscal tears (*Chapter 3*)
3. To describe a well-designed feasible and methodologically sound RCT to study the effectiveness and cost-effectiveness in patients with degenerative meniscal tears (*Chapter 4*)
4. To assess whether supervised exercise therapy is non-inferior to partial meniscectomy for improving patient-reported knee function in patients with meniscal tears (*Chapter 5*)
5. To determine whether partial meniscectomy or supervised exercise therapy is more cost-effective in the treatment of degenerative meniscal tears (*Chapter 6*)
6. To assess the responsiveness and determine the minimal important change of the IKDC questionnaire (*Chapter 7*)
7. To determine the ability of orthopaedic surgeons to predict the outcome in patients treated for meniscal tears (*Chapter 8*)

Outline

Chapter 2 systematically reviews and appraises the literature up to 2016 after the publication of several RCTs in a meta-analysis. The published RCTs included different measurement instruments for the (primary) outcome knee function, leading to difficulty interpreting these results. The quality of a measurement instrument depends on the quality of its measurement properties. Therefore, *Chapter 3* describes the measurement properties (reliability and validity) for the 3 most used measurement instruments, the International Knee Documentation Committee (IKDC), Knee Injury and Osteoarthritis Outcome Score (KOOS), and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). These results formed the basis for the primary outcome as described in our trial protocol (*Chapter 4*).

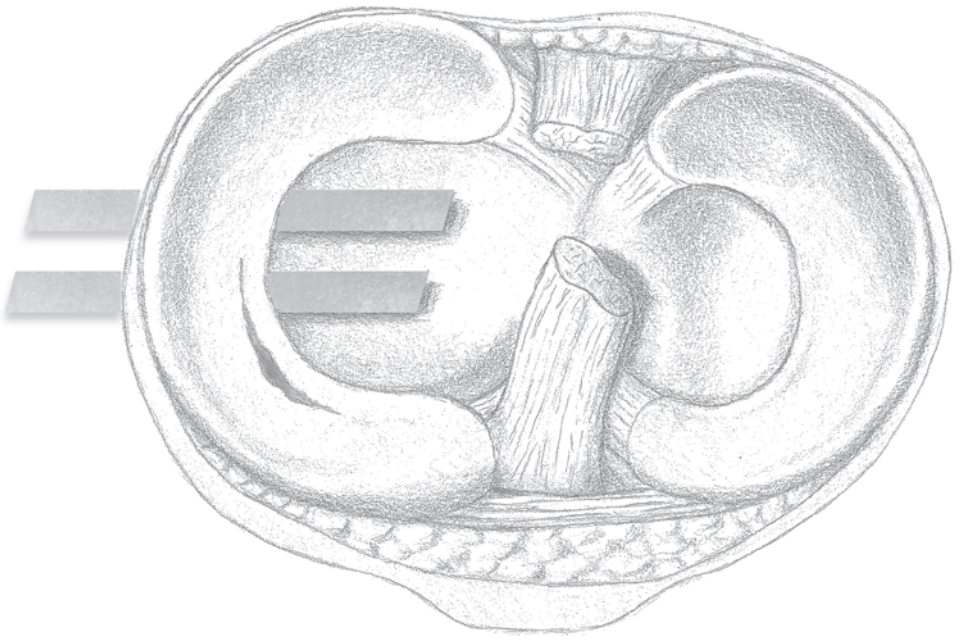
With the exposed limited available evidence for the treatment of degenerative meniscal tears, we aim to improve and contribute to the body of evidence on healthcare efficiency (i.e. both effectivity and cost-effectiveness). Therefore, *Chapter 4* describes the methods of a non-inferiority multicentre RCT with economic evaluation. In this trial, we compare partial meniscectomy to conservative treatment in patients with degenerative meniscal tears. The 2 main objectives are to determine whether 1) supervised exercise therapy is non-inferior to partial meniscectomy for improving self-reported knee function (C*hapter 5*) and 2) to determine whether supervised exercise therapy or partial meniscectomy is more cost-effective, from a societal perspective (*Chapter 6*), over a 24-month follow-up period in patients with non-obstructive meniscal tears.

Chapter 7 describes the results of the measurement properties reliability and the minimal important change of the IKDC, to improve the interpretability of the results from *Chapters 5 and 6*.

Finally, we search for reasons among orthopaedic surgeons why partial meniscectomies are still being performed on a large scale. One of the most frequently heard arguments is the ability to be able to predict who will (and who will not) benefit from surgery. Therefore, we examine this ability of orthopaedic surgeons to predict the outcome of treatment for meniscal tears by partial meniscectomy and supervised exercise therapy in middle-aged patients (*Chapter 8*).

We hypothesise that supervised exercise therapy provides better value for money compared to arthroscopic partial meniscectomy in patients with degenerative meniscal tears.

If our hypothesis holds and is accepted by clinicians, this will improve the quality of care for patients, payers and providers alike.



2

Arthroscopic partial meniscectomy or conservative treatment for nonobstructive meniscal tears: A systematic review and meta-analysis of randomized controlled trials

Victor A van de Graaf
Nienke Wolterbeek
Eduard LAR Mutsaerts
Vanessa AB Scholtes
Daniël BF Saris
Arthur de Gast
Rudolf W Poolman

Arthroscopy. 2016 Sep;32(9):1855-1865.e4

Abstract

Purpose

To conduct a meta-analysis of randomised controlled trials comparing the outcome of arthroscopic partial meniscectomy (APM) with conservative treatment in adults with non-obstructive meniscal tears and to recommend a treatment of choice.

Methods

We systematically searched the databases of MEDLINE, Excerpta Medica Database, Cochrane, the National Health Service Centre for Reviews and Dissemination, and Physiotherapy Evidence Database from inception to May 2, 2016. Two authors independently searched the literature and selected eligible studies. The meta-analyses used a random-effects model. The primary outcome was physical function, measured by knee-specific patient-reported outcomes. Secondary outcomes included knee pain, activity level, the progression of osteoarthritis, adverse events, general health, and quality of life.

Results

We included 6 randomised controlled trials, with a total of 773 patients, of whom 378 were randomised to APM and 395 were randomised to the control treatment. After pooling the data of 5 studies, we found small significant differences in favour of the APM group for physical function at 2 to 3 months (mean difference [MD], 3.31; 95% confidence interval [CI], 0.69–5.93; P=.01; I^2, 0% [Lysholm knee score]), and at 6 months (MD, 3.56; 95% CI, 0.24–6.88; P=.04; I^2, 0% (Knee injury and Osteoarthritis Outcome Score [KOOS] and Western Ontario and McMaster Universities Osteoarthritis Index); standardised MD, 0.17; 95% CI, 0.01–0.32; P=.03; I^2, 0% [Lysholm knee score, KOOS, and Western Ontario and McMaster Universities Osteoarthritis Index]). We also found small significant differences for pain at 6 months (MD, 3.56; 95% CI, 0.18–6.95; P=.04; I^2, 0% [KOOS] and MD, 0.56; 95% CI, 0.28–0.83; P=.0001; I^2, 0% [visual analogue scale and numeric rating scale]). We found no significant differences after 12 and 24 months.

Conclusions

We found small, although statistically significant, favourable results of APM up to 6 months for physical function and pain. However, we found no differences at longer follow-up.

Level of Evidence

Level I, systematic review and meta-analysis of Level I studies.

Introduction

Arthroscopic partial meniscectomy (APM) is the most performed procedure in orthopaedic surgery.[14] However, whether APM is superior to conservative treatment in patients with non-obstructive meniscal tears is controversial.[3,4]

The quality of the menisci decreases with ageing: the water content increases, whereas the cellularity, collagen content, and total amount of glycosaminoglycans decrease.[2,29,30] This results in a meniscus that is more vulnerable to degenerative damage and injuries.

Not surprisingly, meniscal tears are the most common type of knee injury in middle-aged and older patients.[9,31]

Meniscal tears can occur with or without mechanical obstruction. Although APM for the obstructive meniscal tear is widely accepted, APM for the symptomatic non-obstructive meniscal tear has come under scrutiny. Knee symptoms, such as pain, in patients with non-obstructive meniscal tears may not be triggered by the meniscus, but by early stages of osteoarthritis (OA). Knee pain and meniscal function are therefore not always directly related. This is strongly supported by Englund and colleagues,[17] who identified meniscal tears on magnetic resonance imaging in 61% of asymptomatic volunteers more than 50 years old.

Meniscal tears can be asymptomatic, as shown by Englund and colleagues.[17] The challenge is to determine who are and who are not likely to benefit from a meniscectomy, because surgery might not be beneficial in the asymptomatic group.

Still, APM is the most frequently performed orthopaedic surgical procedure and the numbers continue to rise. Kim and colleagues showed that the number of APMs increased by 49% to approximately 500 000 between 1996 and 2006 in the United States, two-thirds of which were more than 45 years old.[9] This increase was partially explained by population growth, patient demand, and the practice of defensive medicine.[9]

Recently, 2 meta-analyses were published on the outcome after arthroscopy for degenerative knee complaints (including meniscal injuries).[32,33] Both meta-analyses included studies that did not primarily focus on meniscal injuries. In this meta-analysis, we therefore aimed to summarise all available Level I studies focusing primarily on meniscal injuries.

There is currently no consensus for an evidence-based treatment of choice, being surgical or conservative, for middle-aged patients with non-obstructive meniscal tears. The purpose of this study was to conduct a meta-analysis of randomised controlled trials (RCTs) comparing the outcome of APM with conservative treatment in adults with non-obstructive meniscal tears and to recommend a treatment of choice.

We hypothesised that surgery would be equally effective as conservative treatment in the recovery of physical function in older patients with non-obstructive meniscal tears.

Methods

This meta-analysis followed the preferred reporting items for systematic reviews and meta-analyses guidelines,[34,35] and was performed in accordance with its protocol (Prospero registration number: CRD42012002870).

Eligibility Criteria

We included only RCTs in which at least 1 group of adults with primarily a meniscal injury received either APM or conservative treatment, including all types of non-operative approach. No restrictions on publication status were imposed. Language restrictions were set to English, German, or Dutch.

We excluded studies on discoid menisci, anterior cruciate ligament injuries, or meniscal repair.

Type of Outcome Measures

The primary outcome was physical function measured by knee-specific patient-reported outcome measures (PROMs). We searched the PROMs for subscales on physical function and presented these as our primary outcome. If PROMs had no subscales, the total score was used for our primary outcome.

Our secondary outcomes included knee pain, change of activity level, the development or progression of OA, the occurrence of complications and adverse events, general health, quality of life (QoL), and return to work.

Literature Search and Information Sources

An independent medical librarian searched the following databases twice from inception to May 2, 2016: the Cochrane Central Register of Controlled Trials, the National Library of Medicine (MEDLINE), the Excerpta Medica Database, the Physiotherapy Evidence Database, and the National Health Service Centre for Reviews and Dissemination. The search strings can be found in Appendix 1. We also searched for cross-references and "cited by" articles of the included articles to ensure that no relevant studies were missed. Finally, we searched for any ongoing and unpublished trials by searching for study protocols.

Study Selection

Two members of the project group (VAG and NW) independently assessed the eligibility of the search results with the criteria mentioned above by screening all titles and abstract. Any discrepancies were resolved by consensus.

Data Collection Process

One project member extracted the data into a modified Cochrane Collaboration data extraction form. Another project member checked the extraction forms for accuracy and completeness. Follow-up publications of the same study were included as one. Any disagreements were resolved by consensus.

For continuous outcomes, we extracted the means with standard deviations or the means with 95% confidence intervals (CI). We contacted the corresponding authors for any additionally required data.

Assessment of Risk of Bias and Methodological Quality

The same project members independently assessed the risk of bias on study level using the Cochrane Collaboration's Risk of bias tool (Higgins and Green, Chapter 8),[36] containing 6 domains, each of which addresses a potential source of bias relating to internal validity. No studies were removed from inclusion based on their risk of bias assessment.

The quality of the evidence was independently assessed on outcome level using the Grading of Recommendations, Assessments, Development and Evaluation (GRADE) tool,[37] containing 5 domains, each of which can downgrade the quality of evidence: limitations in study design (including the risk of bias across studies), inconsistency, indirectness, imprecision, and publication bias. For the latter, we searched for unpublished studies in the following trial registries: Clinicaltrials.gov, Current Controlled Trials, and the Dutch Trial Registry.

Planned Methods of Analysis

We pooled data between studies by using a random-effects model to increase the generalizability of the results. The inverse variance method was used to determine the weight of a study in a pooled analysis. We pooled data of similar measurement instruments and presented the pooled estimate as the mean difference (MD). The significance level was set at $P<.05$.

Deviating from the study protocol, we also pooled the data of similar outcomes measured with different measurement instruments and presented these as the standardised mean difference (SMD).[38]

We assessed statistical heterogeneity as proposed by Higgins and Thompson with the I^2-statistic, which describes the percentage of variability that is caused by heterogeneity rather than by chance. Values above 50% represent substantial heterogeneity and those above 75% represent considerable heterogeneity.[39]

We used the software of Review Manager (RevMan version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012) and Grade Profiler (GRADEpro version 3.6 for Windows. Hamilton, Ontario: McMaster University, J. Brozek, A. Oxman, H. Schünemann, 2008).

Results

Study Selection

The search resulted in 1 997 potentially eligible studies, of which 7 RCTs were included (Figure 1). Two RCTs reported the results of the same trial at different follow-ups.[19,22] We included the results of both though presented them as one;[22] the results of one other study (Stensrud and colleagues[40]) could not be pooled and therefore were excluded from our meta-analysis, leaving a total of 5 RCTs for data analysis.[3,4,21,22,41] No relevant completed unpublished RCTs were found.

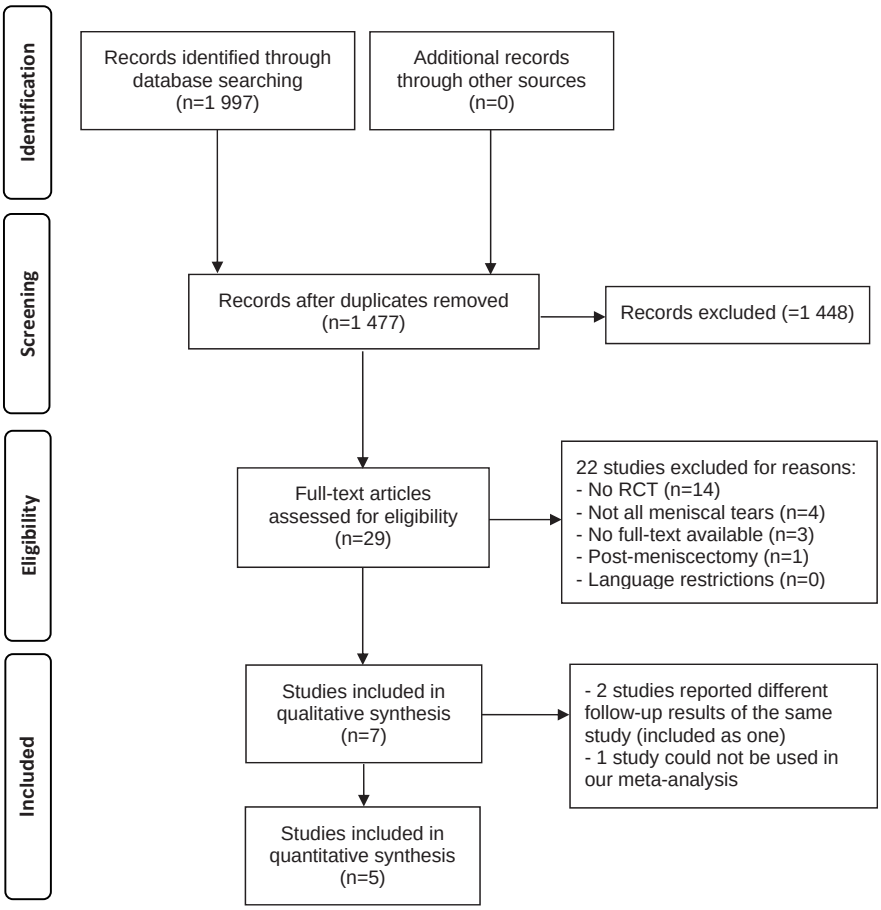


Figure 1. PRISMA Flow diagram of study selection process.
From the originally 1997 found studies, only 5 were eligible for the quantitative analysis. Abbreviations: n, number; PRISMA, Preferred Reporting Items for Systematic reviews and Meta-Analyses.

Characteristics of Included Studies

Data were available for 773 patients (47.6% male) with a mean age between 47 and 59 years. In total, 378 patients were randomised for APM and 395 patients were randomised for the control group.

Herrlin and colleagues described the results of 97 patients with a symptomatic degenerative meniscal tear.[22] The intervention group, 47 patients with a mean age of 54 years, received APM followed by exercise therapy. The control group, 50 patients with a mean age of 56 years, received only exercise therapy. The exercise therapy program consisted of 16 supervised sessions in 8 weeks. The length of follow-up was 60 months, during which 2 patients in each group were lost (4.1%). After 24 months, 27.7% of the control group had crossed over and received a meniscectomy.

Katz and colleagues described the results of 330 patients with a meniscal tear and mild-to-moderate OA.[3] The intervention group, 161 patients with a mean age of 59 years, received APM followed by physical therapy. The control group, 177 patients with a mean age of 58 years, received the same physical therapy with optional delayed APM. The physical therapy program consisted of a 6-week program that patients attended once or twice a week. The length of follow-up was 12 months, during which 31 patients (of 351 patients randomised; 9.3%) were lost, leaving 156 patients in the intervention group and 164 patients in the control group. After 12 months, 33.3% of the control group had crossed over and received a meniscectomy.

Østeras and colleagues described the results of 17 patients with a degenerative meniscal tear.[41] The intervention group, 8 patients with a mean age of 53 years, received APM. The control group, 9 patients with a mean age of 47 years, received medical exercise therapy. The exercise therapy program consisted of 36 sessions in 12 weeks. The length of follow-up was 3 months, during which none were lost.

Sihvonen and colleagues described the results of 146 patients with a medial meniscal tear and no knee OA, although 80% were found to have some level of chondral wear during arthroscopy.[4] The intervention group, 70 patients with a mean age of 52 years, received APM. The control group, 76 patients with a mean age of 52 years, received a sham procedure, an arthroscopy without the meniscectomy. All patients received the same home exercise instructions, consisting of some simple exercises. The length of follow-up was 12 months, during which none were lost.

Stensrud and colleagues described the short-term results of 82 patients with a medial meniscal tear and mild to no knee OA.[40] The intervention group, 42 patients with a mean age of 49 years, received APM. The control group, 40 patients with a mean age of 49 years, received an individualised and supervised neuromuscular and strength exercise program over 12 weeks (24 to 36 sessions). The length of follow-up was 3 months, during which 8 patients were lost. The results could not be embedded in the meta-analysis because of different chosen outcomes such as isokinetic knee muscle strength, lower

extremity performance, and self-reported global rating of change. The authors found a significantly better improvement in knee extension strength in the exercise group (16% difference; 95% CI, 7.1-24.0; P<.01).

Yim and colleagues described the results of 108 patients with a degenerative medial meniscal tear.[21] The intervention group (54 patients) received APM followed by a home exercise program. The control group (54 patients) received supervised exercise therapy followed by the same home exercise program. The exercise therapy consisted of 9 sessions in 3 weeks and the home exercise program of daily exercises for 8 weeks. The length of follow-up was 24 months, during which 6 patients (5.5%) were lost, leading to 50 patients with a mean age of 55 years in the intervention group, and 52 patients with a mean age of 58 years in the control group.

Risk of Bias Assessment

Figure 2 presents the results of the risk of bias assessment. Two studies provided insufficient data on both the process of sequence generation and the concealment of allocation (high risk of selection bias).[31]

Blinding was impossible in 5 of the 6 included studies.[3,21,22,40,41] However, we are unaware to what extent the lack of blinding influenced the patient-reported outcome assessments (unknown risk of detection bias).

Herrlin and colleagues' follow-up study contained 6 additional patients, all of whom were in the control group, without any clarification (high risk of attrition bias).[22] Furthermore, pre-specified data on sick leave, medication use, and differences in type of work were not presented (high risk of reporting bias).

Study protocols were not available for 4 studies (unknown risk of reporting bias).[21,22,40,41]

Katz and colleagues found large differences between participating centres in cross-over to APM.[3] Variation in the physical therapy protocols could have resulted in differences in the provided physical therapy treatment between centres (unknown risk of other bias).

Østeras and colleagues reported on the outcomes of a pilot study with only 17 patients. This study is likely to be under-powered (unknown risk of other bias).[41]

Results of Included Studies

Table 1 presents the summary of findings, including the results on outcome level per follow-up moment, the GRADE assessment, and the results of the pooled estimates. Assessments for primary and secondary outcomes were performed at 2 to 3, 6, 12, and 24 months. We present only the significant results in the text. The forest plots for our primary outcome are presented in Figure 3. The forest plots for our secondary outcomes can be found in Appendix 2 and 3. The pooled estimates are also presented in Table 1.

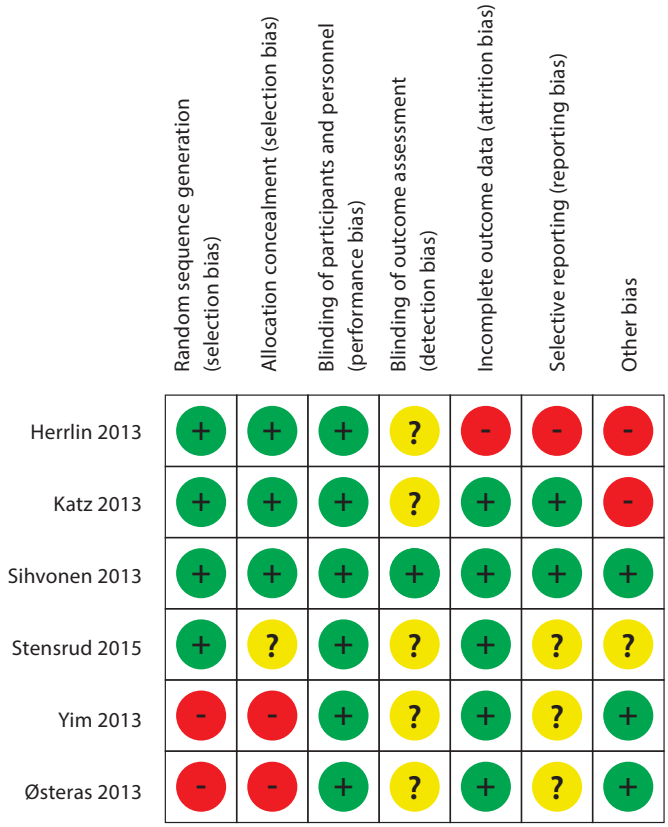


Figure 2. Risk of bias assessment.
Review authors' judgements about each risk of bias item for each included study.

Primary Outcome

Both groups in all of the included studies significantly improved in physical function from baseline to 6-month follow-up (30.3% to 39.4% for the APM group; 20.1% to 37.4% for the control group). None of the studies found significant differences between the groups for knee-related physical function at baseline or at any of the follow-ups, measured with knee-specific PROMs. However, different PROMs were used in the included studies: the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC),[3] the Knee injury and Osteoarthritis Outcome Score (KOOS),[22,41] the Lysholm Knee Scoring Scale (LKSS),[4,21,22] and the Western Ontario Meniscal Evaluation Tool.[4] The WOMAC physical function dimension and the KOOS function in daily living dimension of activities of daily living consist of the same items and could directly be pooled.

When we pooled these results, we found a significantly better outcome in the APM group at 2 to 3 months measured with the LKSS (MD, 3.31; 95% CI, 0.69-5.93; P=.01;

Table 1. Summary of findings
Population: patients with symptomatic non-obstructive meniscal tears
Intervention: Arthroscopic partial meniscectomy
Comparison: conservative treatment

Outcome	Follow-up	Quality of the evidence (GRADE)	No. of Participants (studies)	Effect size[d] (95% CI)
Physical function				
WOMAC – KOOS	2–3 months	High	426 (2 studies)[3,22]	MD 3.21 (-0.88–7.29)
LKSS	2–3 months	Very low[a,b,c]	344 (3 studies)[4,21,22]	MD 3.31 (0.69–5.93)[e]
WOMAC – KOOS	6 months	High	426 (2 studies)[3,22]	MD 3.56 (0.24–6.88)[e]
LKSS–WOMAC–KOOS	6 months	low[a,b]	764 (4 studies)[3,4,21,22]	SMD 0.17 (0.01–0.32)[a]
LKSS	6 months	Very low[a,b,c]	344 (3 studies)[4,21,22]	MD 0.77 (-1.61–3.15)
LKSS – WOMAC	6 months	Low[a,b]	674 (4 studies)[3,4,21,22]	SMD 0.11 (-0.07–0.30)
LKSS–WOMAC–KOOS	12 months	Low[a,b]	662 (4 studies)[3,4,21,22]	SMD 0.01 (-0.14–0.16)
WOMAC – KOOS	12 months	High	414 (2 studies)[3,22]	MD 1.14 (-2.01–4.30)
LKSS	12 months	Very low[a,b,c]	342 (3 studies)[4,21,22]	MD -0.24 (-2.65–2.17)
LKSS – WOMAC	12 months	Low[a,b]	662 (4 studies)[3,4,21,22]	SMD 0.02 (-0.13–0.17)
LKSS	24 months	Low[a,c]	194 (2 studies)[3,22]	MD -1.14 (-3.72–1.45)
Pain				
KOOS	2–3 months	Moderate[d]	426 (2 studies)[3,22]	MD 4.54 (-0.89–9.96)
VAS and NRS activity	2–3 months	Very low[a,b,c]	344 (3 studies)[4,21,22]	MD 0.42 (-0.06–0.89)
VAS and NRS rest	2–3 months	Very low[a,b,c,d]	58 (2 studies)[22,41]	MD -0.04 (-0.99–0.91)
KOOS	6 months	High	426 (2 studies)[3,22]	MD 3.56 (0.18–6.95)[e]
VAS and NRS activity	6 months	Very low[a,b,c]	344 (3 studies)[4,21,22]	MD 0.56 (0.28–0.83)[e]
KOOS	12 months	High	414 (2 studies)[3,22]	MD 0.83 (-2.53–4.19)
VAS and NRS activity	12 months	Very low[a,b,c]	342 (3 studies)[4,21,22]	MD 0.12 (-0.15–0.38)
VAS and NRS rest	12 months	Low[b,c]	240 (2 studies)[4,22]	MD 0.18 (-0.34–0.70)
VAS and NRS activity	24 months	Low[a,c]	194 (2 studies)[21,22]	MD -0.10 (-0.51–0.31)
Activity level				
TAS	2–3 months	Low[a,c]	101 (2 studies)[21,22]	MD -0.06 (-0.42–0.31)
TAS	12 months	Low[a,c]	100 (2 studies)[21,22]	MD 0.31 (-0.06–0.68)
TAS	24 months	Low[a,c]	98 (2 studies)[21,22]	MD 0.14 (-0.26–0.54)

[a] Risk of bias: limitations in study design (see Figure 2).
[b] Indirectness: large differences between the conservative treatments.
[c] Imprecision: small sample / large standard deviation.
[d] Inconsistency: large heterogeneity.
[e] positive values indicate results in favour of surgery.
[f] significant difference.
Abbreviations: CI, Confidence interval; GRADE, Grading of Recommendations, Assessments, Development and Evaluation; MD, mean difference; No, numbers; KOOS, Knee injury and Osteoarthritis Outcome Score; LKSS, Lysholm Knee Scoring Scale; NRS, Numeric Rating Scale; SMD, standardised mean difference; TAS, Tegner Activity Score; VAS, Visual analogue Scale; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Table 1. (continued)

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

I^2, 0%),[4,21,22] and at 6 months measured with the WOMAC physical function dimension and KOOS dimension of activities of daily living (MD, 3.56; 95% CI, 0.24-6.88; P=.04; I^2, 0%).[3,22] The quality of evidence of the LKSS at 2 to 3 months was graded as very low, and the quality of evidence of the KOOS at 6 months was graded as high. We calculated the SMD for our primary outcome from the results of different PROMs (Appendix 4).[3,4,21,22] Herrlin and colleagues used both the KOOS and the LKSS. We compared both outcomes separately with the LKSS,[3,21] and the WOMAC2 in the pooled analyses. When pooling the KOOS results, we found a significantly better outcome in the APM group at 6 months (SMD, 0.17; 95% CI, 0.01-0.32; P=.03; I^2, 0%). No significant differences between groups were found at 12 months, or when pooling the LKSS data from Herrlin and colleagues with the other studies on both 6 and 12 months, we found no differences between groups at 6 (n=674) and 12 months (n=662) for physical function (combined LKSS and WOMAC).

Secondary Outcomes

None of the studies found significant differences between the groups for knee pain. Two studies used the KOOS Pain dimension,[3,22] 2 studies used the visual analogue scale (VAS),[21,22] and 1 study used the numeric rating scale (NRS).[4] The data of the VAS and NRS during activity were reported in the same unit of measurement and could therefore be pooled.

When we pooled the results (Appendix 2), we found significantly better outcomes in the APM group at 6 months measured with the VAS and NRS activity score (MD, 0.56; 95% CI, 0.28-0.83; P<.001; I^2, 0%),[4,21,22] and measured with the KOOS Pain dimension (MD, 3.56; 95% CI, 0.18-6.95; P=.04; I^2, 0%).[3,22] The quality of evidence of the VAS and NRS activity at 6 months was graded as very low, and the quality of evidence of the KOOS Pain dimension at 6 months was graded as high.

Two studies reported on the change of the activity level.[21,22] Herrlin and colleagues found that after 6 months, only 40% of the control group had returned to their pre-injury activity level, compared with 51% of the APM group. However, when we pooled the data of the Tegner activity score, we found no significant differences between the groups (Appendix 3).[21,22]

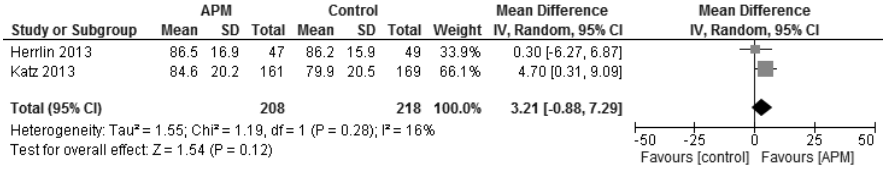


Figure 3.1 Forest plot of comparison WOMAC physical function and KOOS ADL – 2-3 months

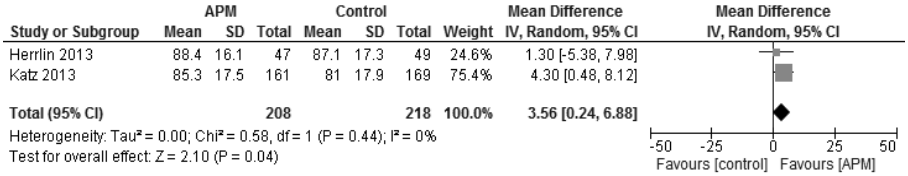


Figure 3.2 Forest plot of comparison WOMAC physical function and KOOS ADL – 6 months

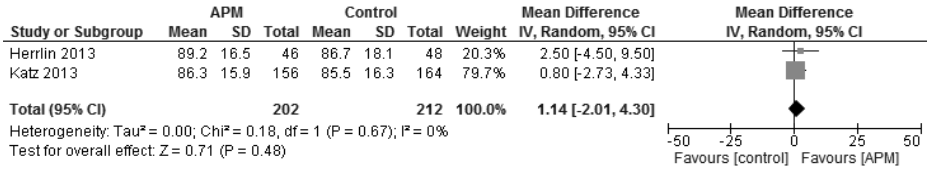


Figure 3.3 Forest plot of comparison WOMAC physical function and KOOS ADL – 12 months

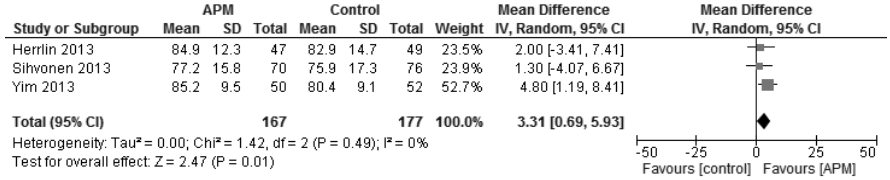


Figure 3.4 Forest plot of comparison LKSS – 2-3 months

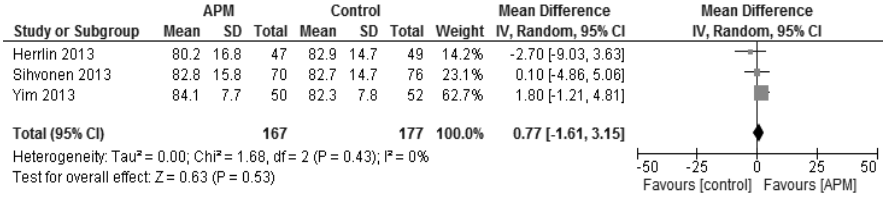


Figure 3.5 Forest plot of comparison LKSS – 6 months

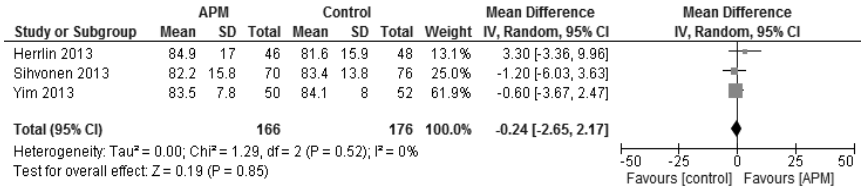


Figure 3.6 Forest plot of comparison LKSS – 12 months

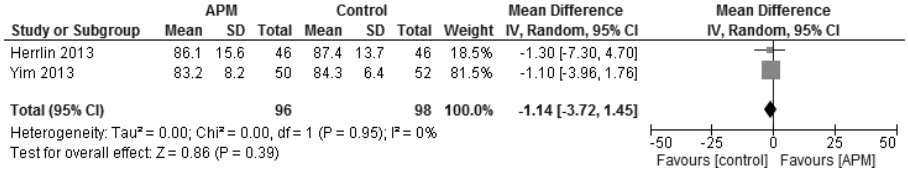


Figure 3.7 Forest plot of comparison LKSS – 24 months

Figure 3. Forest plots of primary outcome physical function.
Statistically significant favourable results for surgery were found at 2-3 months (pooling KOOS and WOMAC data) and 6 months (pooling of LKSS data). Abbreviations: ADL, Activities of Daily Living; APM, Arthroscopic Partial Meniscectomy; CI, Confidence Interval; df, Degrees of Freedom; I^2, level of heterogeneity (50-75% substantial, >75% considerable); KOOS, Knee injury and Osteoarthritis Outcome Score; LKSS, Lysholm Knee Scoring Scale; p, probability level; SD, standard deviation; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Two studies reported on the progression of OA.[21,22] Yim and colleagues found that after 24 months 2 patients in the APM group (8.5%) and 3 patients in the conservative group (6.7%) showed a progression of more than 1 grade on the Kellgren-Lawrence score.[21] Herrlin and colleagues found that after 60 months, 2 patients in each group (4.3% in the APM group; 4.1% in the conservative group) showed a progression of at least 1 grade on the Ahlbäck classification.[22] These data could not be pooled.

Two studies reported on the occurrence of adverse events after 12 months.[3,4] Katz and colleagues found 3 serious and 15 non-serious adverse events in the APM group, compared with 2 serious and 13 non-serious adverse events in the physical therapy group.[3] Sihvonen and colleagues described only the occurrence of serious adverse events.[4] The authors found 1 serious adverse event in the APM group and none in the sham surgery group.

One study reported on general health.[3] Both groups improved equally, measured with the Short Form 36, after 6 and 12 months.

One study reported QoL as an outcome measure.[4] The authors found no differences in QoL at 12 months, measured with the 15D questionnaire.

Although mentioned in the methods section of one study, we found no data reporting on the outcome of return to work.[19]

Discussion

Key Findings

None of the individual studies found significant differences between the groups for any of the outcomes. When we pooled the data, we found small, although statistically significant, favourable results of APM up to 6 months for physical function (LKSS at 2 to 3 months, KOOS and WOMAC at 6 months, and WOMAC, KOOS, and LKSS at 6 months) and pain (VAS and NRS, and KOOS both at 6 months). However, we found no differences at longer follow-up.

Comparison with Literature

We could only identify small significant differences at 3 and 6 months for some of the outcomes. Although we are unaware of any publications on the minimal clinically important difference in patients with meniscal injuries, we expect the minimal clinically important difference to be similar to the results in other patient groups, which is much larger than the effect sizes found in this meta-analysis.[42-45] With just small significant differences in favour of surgery at 3 and 6 months, and the absence of any differences in outcome between the groups at longer follow-up, we are unconvinced whether these differences are of clinical importance. It remains to be defined what clinical benefit in the short-term is relevant to allow surgical intervention for this indication, given the comparable outcomes after 6 months. Furthermore, it is not always possible to relate knee pain to meniscal function. Because general degenerative changes in the knee need to be taken into account as well, the results of this study should be interpreted carefully.

Future studies will increase the numbers of patients and thereby may narrow down the CI while not affecting the effect size. These results are therefore likely to increase the number of significant differences, such as for (1) pain on both the VAS and NRS, and on the KOOS at 2 to 3 months, and (2) physical function on the WOMAC at 2 to 3 months. However, because the effect size is unlikely to be affected, increasing the patient numbers is unlikely to exceed the threshold value for clinically relevant differences between both treatments.

In the literature, APM has been suggested as a risk factor for the development and progression of OA.[18,46-53] How this influences physical function after surgery remains unclear.[46,54-56] One study described the progression of OA after both treatments in the long-term (60 months) and found no differences between the groups on radiological examination.[22]

Two of the included studies reported subgroup analyses for the influence of the level of OA on the outcome. Although it is generally believed that APM in patients with mild-to-moderate OA results in better outcome than in patients with severe OA,[57] this was contradicted by the subgroup analyses in which the severity of OA was of no influence on the outcome. Given the low patient numbers on this outcome, more studies are needed before any firm conclusions can be drawn on this topic.

Recently, 2 meta-analyses were published on the outcome after arthroscopy for several degenerative knee disorders.[32,33] In the meta-analyses of Khan and colleagues,[32] small short-term beneficial effects were found for the surgically treated group for functional outcome (SMD, 0.25; 95% CI, 0.02-0.48; I^2, 56%). Thorlund and colleagues found small significant beneficial effects for pain measured with the VAS (2.4 mm; 95% CI, 0.03-0.26; I^2, 20.6%).[33] No significant differences were found at longer follow-up. These results are similar to our findings, because we found small significant favourable short-term outcomes for both pain and function for surgery. However, the clinical relevance of the small differences is questionable and neither of the meta-analyses found any significant differences at longer follow-up.

Despite the wide use of APM for these injuries, high-quality studies on this subject are sparse. We found only 6 published RCTs that compared APM with conservative treatment (n=5), or that compared APM with sham surgery (n=1).

This meta-analysis was conducted following the preferred reporting items for systematic reviews and meta-analyses guidelines. We systematically assessed the risk of bias using the Cochrane methods and the quality of the evidence using the GRADE methods.

Despite the use of different outcome measures in the included studies, we were able to pool the data from the included studies increasing the quality and strength of the conclusions. Most of the pooled results of this study are therefore suggestive towards a certain outcome, rather than a firm conclusion can be drawn.

However, none of the results and/or outcomes are contradictive, which increase the quality of the evidence. Compared with other published meta-analyses, we were able to obtain more (unpublished) data from the included studies by contacting the authors and the current meta-analysis focused specifically on patients with primarily a meniscal injury instead of several degenerative disorders.

Implications of this Meta-analysis

With the findings of this meta-analysis, we aim to support healthcare providers (orthopaedic surgeons, physical therapists, and general practitioners) in their clinical decision making and to provide a complete overview of the best available evidence for guideline panels. For clinical decision making, however, more knowledge is needed of those patients who will and who will not benefit from APM. In 28% and 33% of the conservatively treated patients in the studies of Herrlin and colleagues and Katz and colleagues, respectively,[3,22] delayed surgery was necessary. Furthermore, Sihvonen and colleagues showed in a recently published post hoc analysis that APM has no beneficial effect over sham surgery in patients with occasional knee catching or locking symptoms.[58] Future studies should therefore focus on the development of the true indication for surgery and on specific patient profiles, based on baseline characteristics (e.g., body mass index, leg axis), who are likely to benefit from surgery, as guidance for clinical decision making.[59]

In general, the indirect costs represent the most of the total costs associated with treatment of medical conditions. There are no economic studies on this topic yet. However, such studies can add very valuable information, particularly to see whether the small short-term favourable outcomes of surgery will lead to lower indirect (both medical and nonmedical) costs. An economic decision analysis recently showed the economic benefits of improving the reliability of patient history and physical examination over the magnetic resonance imaging scan in diagnosing meniscal injuries.[60]

We are also insufficiently informed about the impact of APM on the progression of OA in the long-term. Future studies should focus more on this aspect of intervention.

Of the 4 significant findings, 2 were graded as very low quality of evidence. Currently, several RCTs are ongoing of which some focus on cost-effectiveness,[61,62] and some on the progression of OA[63,64] after APM and conservative treatment in patients with degenerative meniscal tears. These studies may further close the gap of knowledge, result in stronger conclusions, and support a more definite change in medical guidelines on the treatment of choice for non-obstructive degenerative meniscal tears.

Limitations

This meta-analysis has some limitations. First, even though all included studies were Level I studies, the quality of evidence of the outcomes using GRADE varied from high, when combining only 2 studies, to low or even very low, when combining more than 2 studies. The latter was mainly explained due to limitations in study design, indirectness, and imprecision. With GRADE, it is difficult to be classified as high-level quality of evidence when comparing different measurement instruments. None of the outcomes are contradictive, increasing the strength of the evidence.

Second, our primary outcome was measured with different knee-specific PROMs. PROMs should be chosen with knowledge of their measurement properties because only the instrument with the best properties most adequately reflects a patient's function.

Third, because of the use of different PROMs and the small number of RCTs (only 5 of 1 997 found studies), we could only pool a few studies per outcome. For none of the outcomes, we were able to use all included studies. To increase the generalizability of our findings, we decided, after completion of the protocol, to pool different measurement instruments used for our primary outcome. We found no significant differences for the pooled estimates of the LKSS and the WOMAC. Results from the SMD should be interpreted with care because the size of the effect cannot be derived from this. The direction of the effect, on the other hand, can be derived from the SMD and indicates similar outcome for the physical function of APM and conservative treatment.

Fourth, the exercise programs of the studies varied from only home exercise instructions to 36 supervised physical therapy sessions, and 1 study did not describe the exercise program or physical therapy at all. These differences could have had influenced the outcome of the control group. Because all studies found similar results

and we found low heterogeneity, it is unlikely that differences in conservative treatment largely influenced the results.

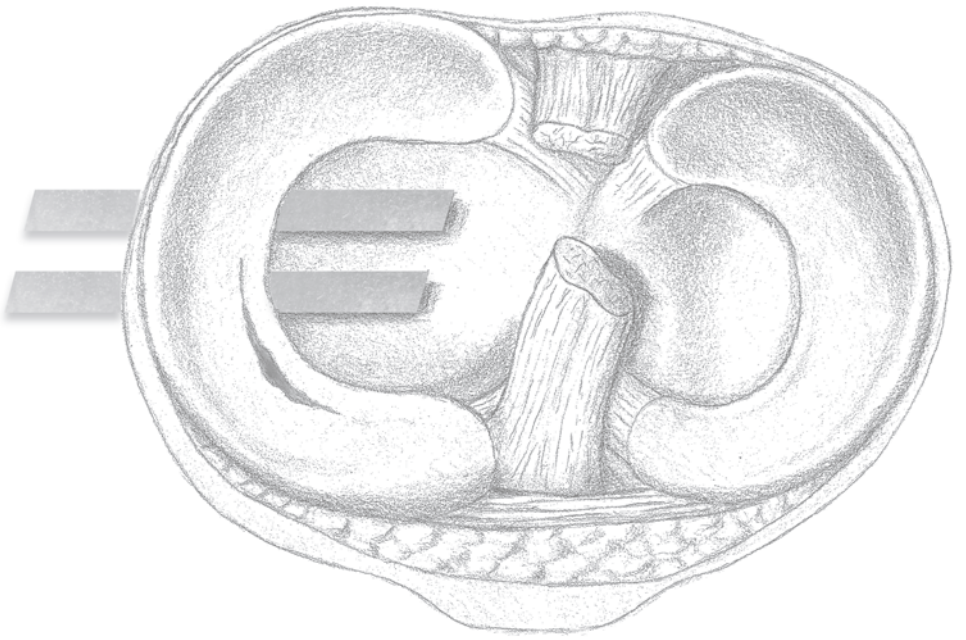
Unfortunately, all these limitations have an effect on the strength of this meta-analysis and weaken the outcome of this study.

Conclusions

We found small, although statistically significant, favourable results of APM up to 6 months for physical function and pain. However, we found no differences at longer follow-up.

Acknowledgements

The authors thank Prof. Dr L. Weidenhielm, Prof. Dr J. N. Katz, Dr J. K. Seon, and Dr J. H. Yim for providing us with the requested data needed for the meta-analysis. Furthermore, they thank Bert Berenschot, information specialist at the Onze Lieve Vrouwe Gasthuis Amsterdam, for the literature search.



3

Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries

Victor A van de Graaf
Nienke Wolterbeek
Vanessa AB Scholtes
Eduard LAR Mutsaerts
Rudolf W Poolman

Am J Sports Med. 2014 Jun;42(6):1408-16

Abstract

Background

Several patient-reported outcome measurements are used to measure functional outcome after treatment of meniscal injuries. However, for comparison of study results, there is a need for a uniform and standardised approach of measuring functional outcome. Selection of the instrument should be based on the quality of its measurement properties, and only the best instrument can be justified to be used.

Purpose

This study aimed to determine and compare the measurement properties of the Dutch-language versions of the International Knee Documentation Committee (IKDC) Subjective Knee Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), and Western Ontario and McMaster Universities Arthritis Index (WOMAC) in a homogeneous group of patients with meniscal tears.

Study Design

Cohort study (design); Level of evidence, 2.

Methods

Patients on the waiting list for meniscal surgery and patients between 6 weeks and 6 months after meniscal surgery were included (n=75). Patients were excluded if they received an arthroplasty or had surgery on the anterior cruciate ligament. Internal consistency (Cronbach alpha), test-retest reliability (intraclass correlation coefficient [ICC]), measurement error (SEM), smallest detectable difference (SDD), content validity, construct validity (factor analysis and hypothesis testing), and floor and ceiling effects were determined.

Results

Results for the IKDC, KOOS dimensions, and WOMAC dimensions, respectively, were as follows: Cronbach alpha, 0.90, 0.72–0.95, and 0.84–0.95; ICC, 0.93, 0.84–0.89, and 0.77–0.89; SEM, 5.3, 7.0–12.6, and 7.3–12.2; SDD, 14.6, 19.4–35.0, and 20.2–33.9; hypotheses testing confirmation, 100%, 86%, and 85%. Floor effects within the SDD from the minimum score were found for the KOOS Sports/Recreation and Quality of Life dimensions. Ceiling effects within the SDD from the maximum score were found for the KOOS Activities of Daily Living and for all WOMAC dimensions.

Conclusion

The IKDC showed the best performance on all measurement properties, implying that the IKDC, rather than the KOOS or WOMAC, should be used to assess functional outcome in patients with meniscal tears.

Introduction

Arthroscopic meniscal surgery, in which the damaged part of the meniscus is removed, is the most common procedure in orthopaedic surgery, with nearly 500 000 surgeries in the United States[9] and more than 41 000 surgeries in the Netherlands (Centraal Bureau voor Statistiek [CBS]) each year. Despite these high numbers, this treatment has gained a lot of attention in the current literature,[3,21,22] as its efficacy has been questioned.

With the shift from objective outcome measures to the patients' perspective on health, it is essential that high-quality patient-reported outcome measurements (PROMs) are used.

According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN),[65] the information retrieved from a PROM is only as useful and reliable as the quality of its measurement properties. Measurement properties refer to the different features of a PROM reflecting its quality.[65] The COSMIN group developed standardised terminologies and definitions for measurement properties and established a checklist for systematically evaluating the measurement properties of an instrument. The checklist is divided into 3 domains: reliability, validity, and responsiveness.[65]

In the literature, several PROMs are used to measure functional outcome after treatment of meniscal injuries. The most commonly used are the International Knee Documentation Committee (IKDC) Subjective Knee Form,[66-69] the Knee Injury and Osteoarthritis Outcome Score (KOOS),[22,41,63,70] the Lysholm Knee Score,[21,67,69-71] and the Western Ontario and McMaster Universities Arthritis Index (WOMAC) Score.[3] The use of different PROMs makes it difficult to compare and interpret conclusions from these studies. With the effectiveness of meniscal surgery being questioned, only the best PROM is justified for its use in evaluating the outcome of this treatment. Similar validation studies have been performed for other knee conditions, such as anterior cruciate ligament (ACL) injuries[72] and cartilage repair.[73-75]

In this study, we determined and compared the reliability and validity of the Dutch-language versions of the IKDC, KOOS, and WOMAC in a homogeneous group of patients with meniscal injuries. The Lysholm score was not included, since Briggs and colleagues[76] already described the Lysholm score as a suboptimal PROM for meniscal injuries. We aimed to explore which of the PROMs should preferably be used in this specific patient group in future research.

Materials and Methods

We conducted an institutional review board approved cross-sectional cohort study at the Department of Orthopaedic Surgery at the Onze Lieve Vrouwe Gasthuis between May 2011 and March 2012.

Study Participants

Between May 2011 and March 2012, adult patients (age ≥18 years) with proper knowledge of the Dutch language and either on the waiting list for meniscal surgery or between 6 weeks and 6 months after meniscal surgery were approached for participation. Patients were excluded if they received an arthroplasty on either knee or if they had previous ACL surgery on the knee of interest.

Patient-Reported Outcome Measurements

IKDC Subjective Knee Form

The International Knee Documentation Committee developed the Subjective Knee Form in 2001 for knee-specific measurement of symptoms, function, and sports activities in patients with a variety of knee conditions, including meniscal injuries.[77] The IKDC has a total of 19 items and takes approximately 5 minutes to complete. All but 1 of the items are converted to a score, with a maximum of 100 indicating no restrictions in daily and sports activities and the absence of symptoms. In 2006, the IKDC was translated to Dutch by Haverkamp and colleagues, and validated for a heterogeneous population with osteoarthritis, meniscal injuries, and ligament problems.[78] The authors reported that sub-analyses were performed for the different conditions with excellent values of reliability and validity, however, these numbers were not provided.

Knee Injury and Osteoarthritis Outcome Score

Roos and colleagues developed the KOOS in 1998 for any type of knee injury in patients at risk of developing osteoarthritis (OA).[79] It consists of 5 dimensions: Pain, Symptoms, Activities of Daily Living (ADL), Function in Sports and Recreation (Sport/Rec), and Knee-Related Quality of Life (Qol). It has a total of 42 items and takes approximately 15 minutes to complete. The scores are calculated for each dimension, with a maximum score of 100 indicating no knee symptoms. De Groot and colleagues translated the KOOS to Dutch in 2008 and validated it for patients with mild to moderate OA and a primary total knee arthroplasty.[80]

Western Ontario and McMaster Universities Arthritis Index

The Western Ontario and McMaster Universities developed the WOMAC score in 1982 for use among patients with knee and/or hip OA.[81] The WOMAC has 3 dimensions: Pain, Stiffness, and Physical Function. It has a total of 24 items and takes approximately 7 minutes to complete. The scores are calculated for each dimension and for the total score, with a maximum score of 100 indicating no knee symptoms. The WOMAC score can also be calculated from the KOOS.

RAND-36

The RAND-36 is a widely used general health measurement instrument. It consists of 8 dimensions: Physical Functioning, Bodily Pain, Role Limitations due to Physical Health Problems, Role Limitations due to Personal or Emotional Problems, General Mental Health, Social Functioning, Vitality, and General Health Perceptions. It has a total of 36 questions, and the overall score varies between 0 and 100, a higher score indicating better general health status. Moreover, 2 aggregated scores, the Physical Component and Mental Component Score, can be calculated based on the average scores of the Dutch population (set at 50). The RAND-36 takes approximately 10 minutes to complete.

Study Procedures

A sample size of at least 50 patients was recommended for the analysis of measurement properties.[82] Participants completed an online questionnaire twice within 1 to 2 weeks. A hard copy was sent if patients had no internet access. All questionnaires were filled out individually by the patients at home.

The first measurement comprised the IKDC, KOOS, and RAND-36, and the second measurement comprised the IKDC and KOOS. The WOMAC scores were derived from the KOOS. An experienced independent musculoskeletal radiologist determined the degree of OA on each participant's pre-operative radiographs according to the well-accepted Kellgren-Lawrence classification.[83]

Measurement Properties

Reliability

This domain refers to the degree to which the measurement is free from measurement error and to the ability of an instrument to measure the same outcome when repeated measurements are performed in a patient under various conditions.[65] It consists of internal consistency, test-retest reliability, and measurement error.[65,84]

Internal consistency was measured with Cronbach alpha. A Cronbach alpha of 0.70–0.95 was defined as good, with values >0.95 indicating a redundancy of items in a construct.[82,84] Test-retest reliability was calculated with the intraclass correlation coefficient (ICC), using a 2-way mixed-effects model for absolute agreement. Both measurements were obtained within an interval of 1 to 2 weeks. An ICC of >0.7 was considered acceptable, >0.8 good, and >0.9 excellent.[65,84] Measurement error was calculated using the standard error of measurement (SEM), defined as √(variance of measurements 1 variance of the error of the ICC), and the smallest detectable difference (SDD), defined as 1.96 * √(2) * SEM.[86]

Validity

Validity is the degree to which the instrument measures the construct(s) that it intends to measure. It consists of content and construct validity.[65]

Content validity is a relatively subjective judgment, which explains why a standard approach for measuring this property is lacking. We explored the PROMs for resemblance with 6 relevant domains of the International Classification of Functioning, Disability and Health (ICF): Mobility, Self-Care, Domestic Life, Interpersonal Interactions and Relations, Major Life Areas and Community, and Social and Civic Life.[87] Furthermore, we searched the content of the PROMs for resemblance with items found important by this specific group of patients, as described by Tanner and colleagues (Table 1).[85]

Construct validity is determined by structural validity and hypothesis testing. Structural validity can be assessed with factor analysis. In a factor analysis, the number of dimensions (subscales) is determined with the eigenvalue, which should be >1.0, and by judgment of the contribution of each dimension to the total variance, which should be >5%. In addition, the correlation and contribution of single items in a dimension can be assessed with the loading factor, which should be >0.5 to sufficiently contribute to the dimension. If not, the item might be removed from the PROM. Items that substantially

Table 1. Symptoms and Disabilities found most important by patients according to Tanner and colleagues[a]

Condition	Preoperatively	Postoperatively
Meniscal tear	Sports and recreation performance expectations have changed	Fear of reinjuring knee
	Amount of time participating at preinjury level is affected	Often aware and conscious of knee problem
	Often aware and conscious of knee problem	Knee hurts
	Difficult to participate in favourite sport or recreational activity	Squatting is difficult
	Squatting hurts	Frustrating to consider knee with respect to sports and recreation
	Fear of reinjuring knee	Worried what will happen to knee in the future
	Worried what will happen to knee in the future	Knee pain makes it difficult to perform heavy physical labour
	Difficult to quickly change direction	Apprehensive about knee
	Frustrating to consider knee with respect to sports and recreation	Difficult to quickly change direction
	Knee makes it difficult to participate in second most favourite sport or activity	Modified lifestyle to avoid activities that are potentially damaging the knee

[a]Adapted from Tanner and colleagues.[85] Reproduced with permission.

load on more than 1 dimension, >0.3, indicate an inadequate distribution of dimensions and again might be removed from the PROM.

Hypothesis testing was done in relation to the RAND-36.9,[48] It was defined as good if ≥75% was confirmed, as moderate if 50% to 75% was confirmed, and as poor if <50% was confirmed.

1. Strong correlations (r >0.5) were expected between:
   a. RAND-36 Bodily Pain with IKDC Pain (items 1, 2, and 3), KOOS Pain, and WOMAC Pain;
   b. RAND-36 Physical Function with IKDC Activity (items 8 and 9), KOOS ADL, and WOMAC Physical Function; and
   c. RAND-36 Physical Component score with IKDC total score, KOOS dimensions and total score, and WOMAC dimensions and total score.
2. The correlations mentioned in part 1A should be at least 0.1 higher than the correlations between the IKDC Pain, KOOS Pain, and WOMAC Pain with the other dimensions of the RAND-36.
3. The correlations mentioned in part 1B should be at least 0.1 higher than the correlations between the IDKC Activity, KOOS ADL, and WOMAC Physical Function with the other dimensions of the RAND-36.
4. Weak correlations (r<0.3) were expected between:
   a. RAND-36 General Health with the IKDC total, KOOS dimensions and total score, and WOMAC dimensions and total score; and
   b. RAND-36 Mental Component score with the IKDC total, KOOS dimensions and total score, and WOMAC dimensions and total score.

Discriminative hypotheses were assessed for subgroup analyses as follows:

   a. Less cartilage damage (Outerbridge classification 0–2) should indicate a better outcome on IKDC, KOOS, and WOMAC; and
   b. A body mass index <25 kg/m2 should indicate a better outcome on IKDC, KOOS, and WOMAC

Interpretability

The distribution of scores was determined for floor and ceiling effects. More than 15% of the minimum or maximum scores reflected a floor or ceiling effect, respectively. Furthermore, we determined the numbers within the SDD from the minimum or maximum scores. Any score within the SDD from the minimum or maximum score indicated that no deterioration or improvement could be accomplished, respectively. This was determined only in the patients on the waiting list, since only this group should have been able to demonstrate a true change.

Data Management

Data obtained from the first measurement were used for calculation of internal consistency, construct validity, and floor and ceiling effects. Data obtained from both measurements were used for test-retest reliability and measurement error.

All hypotheses were tested for correlations as mentioned above. For normally distributed scores (P>.05, Kolmogorov-Smirnov), Pearson correlation coefficients were used; otherwise, Spearman correlation coefficients were used. Discriminative hypotheses were tested with independent samples t-tests.

Data collection was done in Excel (Microsoft Inc., Redmond, Washington, USA), and statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) v18 (SPSS Inc., Chicago, Illinois, USA).

Results

Figure 1 presents the patient flow. After initial screening, 222 patients were approached for participation, of which 75 responded. Four patients were excluded, and another 4 did not complete the second measurement. Table 2 contains the baseline characteristics of the participants. Since all items in the online questionnaires were mandatory, these had no missing data. Eight patients received paper versions, and 4 of these questionnaires contained missing data. The patients were contacted and asked for the missing answers. As a result, there were no missing data.

Reliability

The results of the internal consistency, test-retest reliability, and measurement error are presented in Table 3. The internal consistency was good for all PROMs, both in the total scores and in the various dimensions, with all Cronbach alphas >0.72. However, there may have been some redundancy of items in the KOOS and WOMAC total scores, since the Cronbach alpha was >0.95.

The test-retest reliability was acceptable to good (ICCs, 0.77–0.89) for WOMAC total scores and for WOMAC and KOOS dimensions. The test-retest reliability was excellent (ICCs, 0.93) for both IKDC and KOOS total scores.

The amount of measurement error was highest in the KOOS Sport/Rec, WOMAC Stiffness, and WOMAC Pain (SEM >10) and smallest in the IKDC and KOOS total scores (SEM<6). The calculated SDD seemed high for all PROM scores, with values ranging from 15, which still might be acceptable, to 35, which seems inadequate.

Validity

The content of all PROMs mainly focuses on the ICF domains Mobility and Domestic Life, whereas the other domains are not covered by the PROMs.

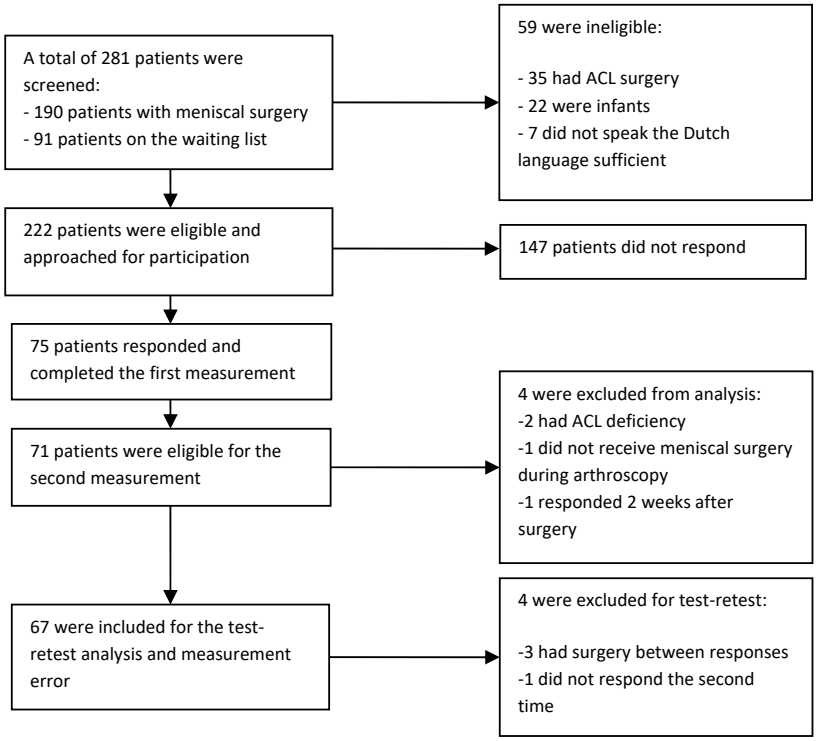


Figure 1. Flowchart of patient recruitment.
Abbreviations: ACL, anterior cruciate ligament; TKA, total knee arthroplasty.

Compared with the list of Tanner and colleagues,[85] none of the PROMs covered topics such as expectations regarding sports and recreation performance, difficulty in participating in favourite or second-favourite sport, fear of reinjuring the knee, worries towards the future, frustration regarding the knee, or influence of knee pain on performing heavy physical labour. The only major difference between the IKDC and the KOOS was the coverage of apprehensiveness and an item regarding the modification of lifestyle, in favour of the KOOS. Most of the items of the KOOS Symptoms dimension, however, did not fit in either the ICF or the list of Tanner and colleagues, and moreover, the KOOS contains multiple similar items. The WOMAC covered only the topics of difficulty in quickly changing direction, knee pain, rising from a chair, and difficulty performing heavy physical labour. The factor analysis of the IKDC revealed 4 dimensions with an eigenvalue >1.0, of which all contributed for >5% to the total variance. However, with 4 dimensions, 11 items loaded on more than 1 dimension. As a result, we performed a single-dimension factor analysis, similar to the current version of the IKDC. One dimension accounted for 41.5% of the total variance, with only 1 item (10a) that did not fit in this dimension (loading factor <0.5).

Table 2. Baseline characteristics

Characteristic	Baseline Characteristics[a]	
	Patients on Waiting List (n=28)	**Postoperative Patients (n=43)**
Age (years)	48.8 (± 13.7)	57.7 (± 9.7)
Sex (male)	14 (50.0%)	21 (48.8%)
BMI (kg/m^2)	24.5 (± 2.7)	26.1 (± 4.0)
Time interval questionnaires (days)	9.9 (± 4.5)[b]	9.3 (± 4.1)[c]
Time after surgery (weeks)	NA	19.4 (± 6.7)
Affected knee (left)	11 (39.3%)	24 (56.8%)
Affected meniscus (medial)	76.7%	74.5%
Level of osteoarthritis[d]	1.0 (± 0.0)[e]	1.1 (± 0.27)[f]
IKDC total score	50.9 (± 14.2)	59.8 (± 20.2)
KOOS total score	45.2 (± 13.1)	60.5 (± 22.3)
KOOS Symptoms	58.0 (± 15.5)	71.7 (± 19.1)
KOOS Pain	45.5 (± 16.4)	61.1 (± 27.0)
KOOS ADL	61.8 (± 16.3)	74.1 (± 21.9)
KOOS SP/R	26.4 (± 17.6)	45.8 (± 31.0)
KOOS QOL	34.4 (± 19.9)	49.9 (± 26.5)
WOMAC Total	60.2 (± 15.5)	72.0 (± 22.5)
WOMAC Pain	54.1 (± 21.4)	65.6 (± 28.8)
WOMAC Stiffness	61.6 (± 21.0)	70.1 (± 23.8)
WOMAC Physical function	61.8 (± 16.3)	74.1 (± 21.9)
RAND-36	61.8 (± 16.2)	64.9 (± 17.7)
Age (years)	48.8 (± 13.7)	57.7 (± 9.7)

[a] Values are expressed as mean ± standard deviation unless otherwise indicated.
[b] n=25 (3 exclusions for second questionnaire).
[c] n=42 (1 did not answer second questionnaire).
[d] According to Kellgren-Lawrence classification.
[e] n=18 (10 missing).
[f] n=35 (8 missing).
Abbreviations: ADL, Activities of Daily Living; BMI, body mass index; IKDC, International Knee Documentation Committee; KOOS, Knee Injury and Osteoarthritis Outcome Score; NA, not applicable; Qol, Quality of Life; Sport/Rec, Function in Sports and Recreation; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

The factor analysis of the KOOS revealed 8 dimensions with an eigenvalue >1.0, whereas only 3 dimensions contributed for >5% to the total variance. We therefore performed a 3–dimension factor analysis. With 3 dimensions, 27 of the 42 items loaded on more than 1 dimension. Therefore, we ultimately performed a single–dimension factor analysis. One dimension accounted for 53.5% of the total variance, with only 4 items that did not fit into this dimension (loading factor <0.5). We also performed a 5–dimension factor analysis, as in the current version of the KOOS. We found that 25 items loaded on more than 1 dimension, and 5 items did not fit in any of these dimensions at all.

Table 3. Measurement properties of IKDC, KOOS and WOMAC

Measurement instrument	Cronbach's alpha[a]	ICC (95% CI)	SEM (%)[b]	SDD (%)[b]	Floor effects (%)	Ceiling effects (%)	Floor effects (%)[c] related to the SDD	Ceiling effects (%)[c] related to the SDD
IKDC total score	0.90	0.93 (0.89–0.96)	5.3 (9.3)	14.6 (25.8)	0	0	0	0
KOOS total score	0.97	0.93 (0.89–0.96)	5.4 (9.8)	15.0 (27.3)	0	2 (3.0)	0	0
KOOS Symptoms	0.72	0.85 (0.77–0.91)	7.0 (10.4)	19.4 (28.8)	0	4 (6.0)	0	(3 (10.7)
KOOS Pain	0.89	0.84 (0.76–0.90)	9.3 (16.9)	25.7 (46.8)	1 (1.5)	5 (7.5)	0	0
KOOS ADL	0.95	0.89 (0.83–0.93)	7.3 (10.5)	20.2 (29.2)	0	4 (6.0)	0	5 (17.9)
KOOS Sport/Rec	0.91	0.85 (0.76–0.90)	12.6 (31.5)	35.0 (87.2)	6 (9.0)	2 (3.0)	18 (64.3)	0
KOOS Qol	0.87	0.88 (0.81–0.92)	9.5 (21.7)	26.2 (60.3)	6 (9.0)	2 (3.0)	11 (39.3)	1 (3.6)
WOMAC Total	0.96	0.89 (0.83–0.93)	7.2 (10.8)	20.1 (29.8)	0	4 (6.0)	0	4 (14.3)
WOMAC Stiffness	0.84	0.77 (0.65–0.85)	11.3 (16.9)	31.4 (46.9)	0	12 (17.9)	3 (10.7)	7 (25.0)
WOMAC Pain	0.87	0.78 (0.66–0.86	12.2 (20.0)	33.9 (55.4)	0	8 (11.9)	1 (3.6)	10 (35.7)
WOMAC Physical	0.95	0.89 (0.83–0.93)	7.3 (10.5)	20.2 (29.2)	0	4 (6.0)	0	5 (17.9)

[a] Based on first measurement (n=71).

[c] Percentages are based on the mean scores.

[d] Based on first measurement of patients on the waiting list (n=28).

Abbreviations: ADL, Activities of Daily Living; CI, confidence interval; ICC, intraclass correlation coefficient; IKDC, International Knee Documentation Committee; KOOS, Knee Injury and Osteoarthritis Outcome Score; Qol, Quality of Life; SEM, standard error of measurement; SDD, smallest detectable difference; Sport/Rec, Function in Sports and Recreation; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

The factor analysis of the WOMAC revealed 4 dimensions with an eigenvalue >1, of which all contributed for >5% to the total variance. With 4 dimensions, 15 items loaded on more than 1 dimension, and 1 item did not fit in any dimension. A 3–dimension factor analysis, which represents the current number of dimensions of the WOMAC, revealed that 17 items loaded on more than 1 dimension. We therefore performed a single-dimension factor analysis as well, which accounted for 55% of total variance. All items fit into this single dimension.

Hypothesis testing confirmed 100% of the predefined hypotheses of the IKDC, 88% of the KOOS, and 86% of the WOMAC, as shown in Tables 4 and 5.

Interpretability

No maximum or minimum scores were found in >15% of the IKDC, KOOS, or WOMAC scores, except for WOMAC Stiffness (17.9%). In relation to the SDD, floor and ceiling effects were found for the KOOS ADL, Sport/Rec, and Qol dimensions and for all dimensions of the WOMAC, as presented in Table 3.

Discussion

Key Findings

We found good to excellent measurement properties for the IKDC. The KOOS and WOMAC dimensions performed suboptimally on internal consistency, measurement error, the ability to measure a true change, and content validity. Moreover, there was a large redundancy in items, and the constitution of the 5 dimensions of the KOOS and the 3 dimensions of the WOMAC could not be confirmed.

Strengths and Limitations

To our knowledge, this is the first study that directly compared the measurement properties of 3 of the most commonly used PROMs in a homogeneous group of patients with meniscal injuries. Furthermore, this is the first study that adequately assessed the measurement properties of the KOOS in its current form in this group of patients.

A major issue in comparing PROMs on their measurement properties is that researchers use different definitions and methods to assess these properties. The COSMIN provided useful guidelines and instructions for the assessment of the measurement properties. We believe that in this study we have given a complete overview of the reliability and validity of the Dutch IKDC, KOOS, and WOMAC in following the COSMIN guidelines.[84]

With the recently published studies of Yim and colleagues,[21] Katz and colleagues,[3] and the current ongoing studies,[63,71] the evaluation of meniscal injuries is gaining interest. It is surprising that 3 different PROMs are used (KOOS,[63] WOMAC,[62] and Lysholm[21,71]) as the primary outcome measure in these 4 studies. This indicates the persistent inconsistency regarding the PROM of first choice, and it perfectly illustrates the need for

Table 4. Results of hypotheses testing[a]

Correlation[b]	IKDC Total	IKDC Pain	IKDC Activity	KOOS Total	KOOS Symptoms	KOOS Pain	KOOS Physical	KOOS SP/R	KOOS QOL	WOMAC Total	WOMAC Stiffness	WOMAC Pain	WOMAC physical
RAND-36 Physical function		**0.59**	**0.78**			0.66	**0.72**					0.66	**0.72**
RAND-36 Bodily pain		**0.70**	**0.62**			**0.61**	0.67					**0.56**	0.67
RAND-36 Role Physical		**0.52**	**0.55**			0.52	**0.57**					0.50	**0.57**
RAND-36 Role Emotional		**0.31**	**0.35**			**0.19**	**0.27**					**0.19**	**0.27**
RAND-36 Mental health		**0.19**	**0.25**			**0.29**	**0.29**					**0.31**	**0.29**
RAND-36 Social function		**0.35**	**0.41**			**0.33**	**0.33**					**0.34**	**0.33**
RAND-36 Vitality		**0.28**	**0.35**			**0.36**	**0.39**					**0.39**	**0.39**
RAND-36 General health	**0.26**	**0.12**	**0.27**	**0.16**	**0.03**	**0.14**	**0.22**	**0.15**	**0.27**	**.23**	**0.14**	**0.20**	**0.22**
Physical component score	**0.78**			**0.71**	0.47	**0.67**	**0.74**	**0.64**	**0.65**	**.74**	0.48	**0.65**	**0.74**
Mental component score	**0.11**			**0.16**	**0.12**	**0.09**	**0.09**	**0.13**	**0.24**	**.13**	**0.14**	**0.11**	**0.09**

[a] Calculated with Spearman correlation. Predefined correlations were defined as strong if r>0.5 and poor if r<0.3. Bolded values indicate that the predefined hypothesis is confirmed. Abbreviations: ADL, Activities of Daily Living; IKDC, International Knee Documentation Committee; KOOS, Knee Injury and Osteoarthritis Outcome Score; QoI, Quality of Life; Sport/Rec, Function in Sports and Recreation; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

Table 5. Discriminative hypotheses

		IKDC			KOOS			WOMAC		
Outcome	**n**	**Mean (SD)**	**Sign.**	**95% CI**	**Mean (SD)**	**Sign.**	**95% CI**	**Mean (SD)**	**Sign.**	**95% CI**
BMI <25 km/m^2	37	60.6 (18.7)			59.1 (20.3)			72.4 (19.4)		
BMI >25 km/m^2	34	51.6 (17.3)	.04	0.53–17.61	49.4 (19.7)	0.05	0.24–19.24	61.9 (21.1)	.03	0.92–20.07
Cartilage injury 0-2[a]	25	64.7 (19.6)			66.6 (21.5)			78.9 (19.6)		
Cartilage injury 3-4[a]	18	53.0 (19.4)	.06	-0.41–23.96	52.0 (21.0)	0.03	1.31–27.89	62.4 (23.2)	.2	3.31–29.75

[a] According to the International Cartilage Repair Society (ICRS) classification, only obtained from postoperative patients.

Abbreviations: BMI, body mass index; CI, confidence interval; IKDC, International Knee Documentation Committee; KOOS, Knee Injury and Osteoarthritis Outcome Score; SD, standard deviation; sign, significance, presented as p-value; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

this study. This study could help authors in their decision making and therefore lead to a uniform approach in the evaluation of meniscal injuries.

This study has limitations. First, the low response rate could be a source of sampling bias. Since waiting lists for meniscal injuries are very short, most patients had undergone surgery when a reminder for participation was sent. Second, the conclusions drawn from the factor analysis should be taken with great care due to the relatively small sample size. The minimum sample size for a factor analysis should be at least 100 patients.[65] We still performed the factor analysis to gain insight into the composition of the dimensions in the PROMs. We searched the literature for the means in similar preoperative patient groups. Since we found similar means for the IKDC in 1 090 patients,[67] for the KOOS in 96 patients,[22] and for the WOMAC in a group of 330 patients,[3] we consider our results generalizable for a larger population.

Relation to Prior Work

The KOOS is even more commonly used than the IKDC or WOMAC for meniscal injuries. Roos and colleagues found an ICC of 0.78–0.86 and a Cronbach alpha of 0.93 for 2 dimensions (Sport/Rec and QoL) added to the WOMAC in patients with OA 21 years after meniscectomy.[88] In another study, Roos and colleagues described an ICC of 0.78–0.91 and a Cronbach alpha of 0.71–0.95 for the 5 dimensions of the Swedish KOOS version in a heterogeneous group of patients consisting of ACL deficiencies, meniscal injuries, and cartilage damage.[89] Based on these study results, it cannot be stated that the KOOS is adequately validated for a homogeneous population with meniscal injuries, unlike what some studies suggest.[80,84,90,91]

Reliability

For internal consistency, we found a Cronbach alpha of 0.90 for the IKDC, compared with 0.77[92] and 0.92[77,78] described in the literature. The high Cronbach alpha implies that the IKDC has a very high level of inter-relatedness among the items. For the KOOS, we found values of 0.72–0.95 for the dimensions, somewhat similar to the 0.56–0.98 that we found in the literature for the different dimensions.[80,88] For the total score, we found a value of 0.97, indicating a redundancy of items for this group of patients. For the WOMAC, no values were found for comparison. The Cronbach alpha for the WOMAC total score (0.96) indicates a redundancy of items.

For test-retest reliability, we found an excellent ICC of 0.93 for the IKDC, comparable with the results described by Irrgang and colleagues,[77] Haverkamp and colleagues,[78] and Crawford and colleagues,[92] who all found values ≥0.94 in patients with a variety of knee problems[65,77,78] and meniscal injuries.[92] For the KOOS dimensions, we found good ICC values of 0.84–0.89 and excellent (0.93) for the total score. Compared with the values described in the literature, Roos and colleagues,[88,89,93] and de Groot and colleagues[80]

found values of 0.45–0.97 in patients with posttraumatic OA,[88] arthroscopic surgery on the knee,[89] total knee arthroplasty,[93] and different stages of OA.[80] We found acceptable to good ICC values for the WOMAC of 0.78–0.89; again, no values were found in the literature.

These results indicate that patient-reported scores with the IKDC remain most consistent with outcomes when functional outcome has not changed.

The test-retest reliability depends on the variation between subjects. When reliability is measured in a heterogeneous group of patients, such as our study population, this automatically results in a high ICC.[84] On the other hand, the measurement error is not affected by the heterogeneity of the group and should therefore always be assessed.

We found an SEM of 5.3 and an SDD of 14.6 for the IKDC, slightly larger than the 4.6 and 9.0, respectively, described by Irrgang and colleagues[77] in the developing study and the 3.2 and 8.8, respectively, found by Crawford and colleagues[92] for meniscal injuries. For the KOOS dimensions, we found values between 7.0 and 12.6 for the SEM with SDD values between 19.4 and 35.0. For the KOOS total score, we found an SEM of 5.3 and an SDD of 15.0. Only de Groot and colleagues[80] reported the SEM for KOOS, varying between 5.2 for ADL and 24.6 for Sport/Rec. For the WOMAC dimensions, we found SDD values up to 33.9 on the Pain dimension. It is unfortunate that no values were found in the literature for comparison.

This means that on some dimensions of the KOOS and WOMAC, changes up to 35 have to be achieved to measure a true change, instead of measurement error.

Validity

We compared the PROMs with the domains of the ICF and the list of Tanner and colleagues.[85] The IKDC, KOOS, and WOMAC focused mainly on physical function and therefore missed several of the domains of the ICF. However, the PROMs all covered the same domains of the ICF. With the addition of a general health questionnaire, such as the RAND-36, the missing domains were covered. As compared with symptoms and disabilities found most important to patients, the WOMAC showed the least resemblance. Furthermore, for the WOMAC, recently used in a large study,[3] the redundancy of items had been demonstrated years ago.[94,95] One study even validated a reduced version.[96] It is not known why this version is not used.

The KOOS, the longest of the PROMs, contains multiple similar items, which gives a higher degree of patient burden. The notion of patient burden is important since PROMs are inseparable from current research. We therefore believe it is of great importance to minimise the amount of patient burden by selecting the PROM that provides no more than the necessary information. The factor of the KOOS could not confirm the 5 current dimensions, similar to the results reported by de Groot and colleagues.[80] This indicates that the current composition of dimensions does not measure what they are supposed to measure.

Interpretability

To measure a real change, the change in score should exceed the SDD. Besides the number of minimum and maximum scores, we calculated the occurrence of floor and ceiling effects in relation to the SDD. In this manner, several dimensions of the KOOS and WOMAC were unable to measure a deterioration (KOOS Sport/Rec and KOOS Qol) or an improvement (KOOS ADL and all the WOMAC dimensions) when the PROM would be reassessed. Since PROMs are designed to measure functional change, the occurrence of floor and ceiling effects are highly undesirable.

Other Knee Conditions

With hundreds of PROMs available, it is almost impossible to know all their measurement properties. Ideally, only 1 or 2 PROMs should be used for each joint and only 1 PROM for each condition. As stated earlier, for other knee conditions, the measurement properties of the same PROMs have been determined as well. For ACL injuries, van Meer and colleagues[72] concluded that the IKDC was more useful than the KOOS in the first year after ACL reconstruction. For cartilage repair, the IKDC and WOMAC have been shown to perform equally on their measurement properties.[75] The properties of the KOOS have individually been assessed in several studies, suggesting the KOOS as the PROM of choice.[73,74] However, a direct comparison between the KOOS and both other PROMs has not yet been performed.

Implications

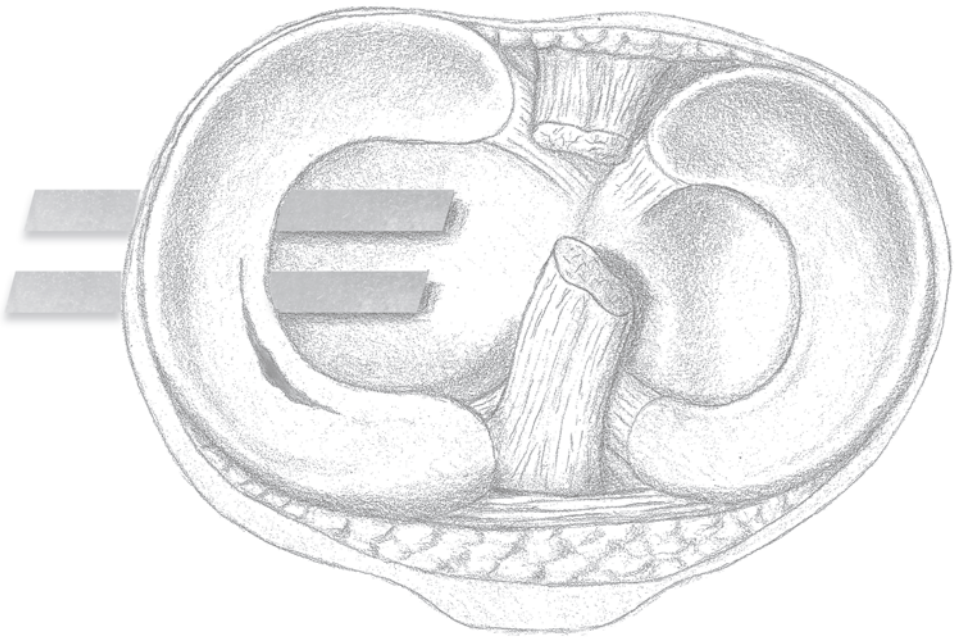
The current distribution of all the dimensions of the KOOS is suboptimal, similar to the findings of de Groot.[80] Therefore, in a larger population, the composition of the dimensions should be adequately reassessed. Furthermore, the responsiveness should be assessed to get a complete overview of the measurement properties of PROMs for meniscal injuries.

Tanner and colleagues[85] concluded that the Western Ontario Meniscal Evaluation Tool (WOMET) is much more specific as a meniscus evaluation tool compared with the IKDC and KOOS, which are rather general knee PROMs and contain too many questions. It is unfortunate that a Dutch version of the WOMET evaluation tool is lacking and we were unable to assess its measurement properties in this study. For future research, this could be of great interest.

Finally, a study directly comparing the measurement properties of the KOOS, IKDC, and optionally the WOMAC for cartilage repair could provide further information and conclude the IKDC as the PROM of choice for all knee conditions.

Conclusions

This cross-sectional cohort study showed favourable results for reliability and validity of the Dutch IKDC compared with the Dutch KOOS and WOMAC for patients with meniscal injuries. Despite a tendency towards the KOOS as the PROM for meniscal injuries, our findings imply that the IKDC Subjective Knee Form is the best applicable PROM for patients with meniscal injuries. We therefore advise the use of the IKDC in future research on meniscal injuries.



4

Cost-effectiveness of early surgery versus conservative treatment with optional delayed meniscectomy for patients over 45 years with non-obstructive meniscal tears (Escape study): protocol of a randomised controlled trial

Victor A van de Graaf
Vanessa AB Scholtes
Nienke Wolterbeek
Julia CA Noorduyn
Camille Neeter
Maurits W van Tulder
Daniël BF Saris
Arthur de Gast
Rudolf W Poolman
for the Escape Research Group

BMJ Open. 2016 Dec 21;6(12):e014381

Abstract

Introduction

Recent studies show similar outcomes between surgery and conservative treatment in patients with non-obstructive meniscal tears. However, surgery is still often preferred over conservative treatment. When conservative treatment is non-inferior to surgery, shifting the current standard treatment choice to conservative treatment alone could save over €30 millions of direct medical costs on an annual basis. Economic evaluation studies comparing surgery to conservative treatment are lacking.

Methods and analysis

A multicentre randomised controlled trial (RCT) with an economic evaluation alongside was performed to assess the (cost-)effectiveness of surgery and conservative treatment for meniscal tears. We will include 402 participants between 45 and 70 years with MRI-confirmed symptomatic, non-obstructive meniscal tears to prove non-inferiority of conservative treatment. Block randomisation will be web-based. The primary outcome measure is a physical function, measured by the International Knee Documentation Committee 'Subjective Knee Form'. Furthermore, we will perform a cost-effectiveness and cost-utility analysis from a societal perspective and a budget impact analysis from a societal, government and insurer perspective. Secondary outcomes include general health, quality of life, activity level, knee pain, physical examination, the progression of osteoarthritis and the occurrence of adverse events.

Ethics and dissemination

This RCT will be performed in accordance with the Declaration of Helsinki and has been approved by the Ethics Committee (number NL44188.100.13). The results of this study will be reported in peer-reviewed journals and at international conferences. We further aim to disseminate our results to guideline committees.

Trial registration number

NCT01850719.

Introduction

Meniscal surgery is the most performed orthopaedic surgical intervention with over 41 000 procedures annually in the Netherlands.[31] In the USA, an increase of 49% was seen in arthroscopic partial meniscectomies (APMs) between 1996 and 2006.[9] Half of these were performed in patients over 45 years old and these numbers continue to rise since the proportion of population over 60 years will double from 11% to 22% between 2000 and 2050 (WHO). APM therefore contributes significantly to the costs of our healthcare system.

The quality of the menisci decreases with age and they become more vulnerable to damage and tears.[2,29,30] Both surgery and conservative treatment do not prevent the development of osteoarthritis (OA). APM in degenerative knees may even accelerate the process of OA more than a non-operative approach since more of the meniscus tissue is removed. However, to the best of our knowledge, no properly designed studies have been published investigating this hypothesis. The expected accelerated progression of OA after APM may influence the number of knee arthroplasties subsequently needed. Faster progression to OA will lead to more patients on waiting lists for knee replacement and subsequently raise costs. In 2003, the National Hospital Discharge Survey in the USA described a total of 402 100 knee arthroplasties in that year and predicted this to grow by 673% to 3.48 million by 2030.[97] Preventing the accelerated progression of OA may result in stagnation of these numbers. Therefore, it could accomplish a substantial reduction of costs of healthcare usage.

Although arthroscopy for obstructive meniscal tears is widely accepted,[19,31] non-obstructive symptoms may not be triggered by the meniscal tear, but by early-onset OA in middle-aged and older patients. Englund and colleagues identified a meniscal tear on MRI in 61% of nearly 1000 asymptomatic volunteers over 50 years old.[3] APM in the non-obstructive meniscal tear group could therefore be seen as overtreatment since many are asymptomatic. Despite the wide use of APM for treatment of non-obstructive meniscal lesions, randomised controlled trials (RCTs) on this subject are sparse. Three recently published meta-analyses of 6–9 RCTs all found a small short-term benefit of surgery over conservative treatment, disappearing over time.[32,33,98] With these data and the lack of economic data, no recommendations can be made on a treatment of choice.

A meniscal tear could lead to knee OA, but knee OA could also lead to a meniscal tear.[99]

The main objective of this study is to evaluate the effectiveness and cost-effectiveness of surgical and conservative treatment, consisting of physical therapy (PT), of non-obstructive meniscal injuries in patients older than 45 years.

We hypothesise that meniscal tears are not a predominant factor causing knee symptoms in patients over 45 years and assume equal improvement of physical function in both groups and reduced costs with PT.

Methods

Study design

We will perform a multicentre RCT with an economic evaluation in the Netherlands. This trial was registered at clinicaltrials.gov (NCT01850719) and the Dutch Trial Registry (the Nederlands Trial Register; NTR3908) prior to the start of inclusion.

Setting

We included the first patient on 3 July 2013. We recruited patients at the orthopaedic outpatient clinic of 9 hospitals, of which 1 was an academic hospital, in the Netherlands (Academic Medical Centre Amsterdam, Diakonessenhuis Utrecht, OLVG Amsterdam, Medisch Centrum Alkmaar, Medisch Centrum Haaglanden Den Haag, Medisch Centrum Jan van Goyen Amsterdam, Sint Elisabeth hospital Tilburg, Slotervaart hospital Amsterdam, Tergooi hospital Hilversum). Eligible participants are randomised into 2 equal groups receiving either APM at the hospital of inclusion or PT. PT is performed at several preselected PT clinics in the area of the hospitals. These PT clinics are selected according to their qualifications and specific instructions regarding the protocol are provided prior to the start of the trial. Participants may prefer receiving treatment at another PT clinic. In these cases, the researcher will contact these clinics prior to the start of the treatment to inform them about the study and provide them with the PT protocol.

Participants

Participants between 45 and 70 years old with a symptomatic, non-obstructive, MRI-confirmed meniscal tear are being recruited at the outpatient clinic of the participating medical centres. Participants will be excluded when meeting one or more of the following exclusion criteria:

- Knee locking or trauma leading to acute surgery;
- Associated injuries on the index knee consisting of:
  - Symptomatic partial or total tear of the anterior cruciate ligament (ACL),
  - Posterior cruciate ligament tear,
  - OA of the knee, grade 4 on the Kellgren and Lawrence Grading Scale,
  - An injury to the lateral or posterolateral ligament complex with significant laxity;
- Previous knee surgery on the index knee (with the exception of diagnostic arthroscopy);
- Tumour that is suspected of malignancy, detectable on MRI;
- Obesity with a body mass index >35;
- American Society of Anaesthesiologists (ASA) class 4 or 5 patients;
- General disease that affects physical function or systemic medication/abuse of steroids;

- Any other medical condition or treatment interfering with the completion or assessment of the trial, for example, contraindications to MRI or surgery;
- Drugs or alcohol abuse;
- Patients who are unable to fill out the Dutch questionnaires.

Participant recruitment

We will screen all patients with knee symptoms who visit the orthopaedic outpatient clinic for eligibility. Patients are informed verbally and in writing about the trial during their first visit. MRI will be conducted for confirmation of the diagnosis meniscal tear. Informed consent is signed when patients agree on participating in the trial at the second visit (on average after 7–14 days) for the result of the MRI.

Randomisation and blinding

After informed consent has been signed, patients are randomly assigned to the treatment group (APM) or control group (PT). The randomisation is performed online using a computerised software program (TENALEA Clinical Trial Data Management system) in a 1:1 ratio using random blocks with a maximum block size of 6. Patients are stratified for centre and age (45– 57 and 58–70 years old).

Interventions

Treatment group

APM is performed within 4 weeks after randomisation by the orthopaedic surgeons experienced in arthroscopic surgery, or orthopaedic residents skilled in arthroscopic surgery under the supervision of an orthopaedic surgeon. Standardised surgery forms for this study are used including assessment of the lateral and medial menisci, the ACL, the level of chondropathy, and a general classification of the level of degeneration. After general or spinal anaesthesia, standard anteromedial and anterolateral portals were introduced for inspection of the knee joint. The affected meniscus is partially removed until a stable and solid meniscus is reached and all unstable and loose fragments are removed. All patients receive an information letter with perioperative instructions. Eight weeks after surgery (about 3 months after randomisation), patients visit the outpatient orthopaedic department for a post-surgery check-up. Considering that standard PT after APM has not been proven effective, patients are referred for PT in case of swelling or signs of atrophy, as advised by the Dutch Orthopaedic Association Guidelines.[31]

Control group

After randomisations, participants are referred to a PT clinic and the treatment on average starts within 1–2 weeks. The treatment protocol consists of a total of 16 sessions of 30

min (Appendix 5). Patients will visit the PT twice a week for 8 weeks. The PT programme consists of a progressive exercise programme and is based on the PT programme developed by Herrlin and colleagues.[22] Three months after randomisation, the patients of the PT group visit the outpatient department to check for function and persistence of symptoms. Additionally, both groups receive the same home exercise instructions (Appendix 5).

Cross-over

Based on the persistence of the symptoms, physical examination (PE) and the level of pain, the physician and participant will decide whether conservative treatment has been successful. When conservative treatment has failed, a delayed APM can be performed. This can be done during the entire follow-up time of the study.

Outcomes

Table 1 provides an overview of the outcomes at the different measurement moments.

Primary outcome

The primary outcome to evaluate the clinical effectiveness is the change in physical function from baseline to 2 years measured by the International Knee Documentation Committee (IKDC) 'Subjective Knee Form'. The IKDC is developed for knee-specific measurements of symptoms, function and sports activities in patients with ligament and meniscal injuries.[77] This self-administered questionnaire is validated for meniscal injuries and consists of 19 questions.[92,100] All items, except item 10a, are converted to a score with a maximum of 100 points, indicating no restrictions in daily and sports activities and the absence of symptoms. A difference of more than 8.8 points in IKDC score is considered clinically relevant.[100]

Secondary outcomes

The secondary outcomes to evaluate clinical effectiveness will be:
Change in:

1. General health, measured by RAND-36;[101]
    a. Quality of life, measured by EuroQol 5 Dimensions 5 Level Survey (EQ-5D-5L);[102]
    b. Pain, measured with the visual analogue scale in rest and during weight-bearing;
    c. Level of activity, measured by Tegner Activity Scale (TAS);[103]
    d. Patient-specific complaints measured by the Patient Specific Complaints (PSC) questionnaire;[104]
    e. Percentage of cross-overs; the number of patients initially treated conservatively, treated secondarily by APM;
2. The progression of OA of the knee using the Kellgren and Lawrence score on X-rays;[83]

3. The relation between a participant's expectation of treatment and their satisfaction;
4. PE at baseline and 3 months, consisting of performance on physical tests (squatting with Duckwalk, Thessaly test, McMurray), the range of motion, joint line tenderness and existence of joint effusion in the knee;
5. Adverse events including:
    a. Minor: prolonged synovial fluid leakage from arthroscopy portals and bleeding;
    b. Moderate: surgical site infection, vascular and neurological damage;
    c. Severe: septic arthritis, cardiac events, pulmonary embolism and death.

Surgical instrument malfunction will be recorded, as well as reoperations including knee arthroplasties and rehospitalisation.

Sample size

Prior to the start of this trial, we calculated the initial sample size based on a power of 90%, an α of .05 and SD of 20 points (retrieved from the study of Crawford and colleagues[92]). We used the previously mentioned clinically relevant difference of 8.8 points on the IKDC 'Subjective Knee Form', and to increase the power of our results, we rounded this down to a non-inferiority threshold of 8 points. We calculated that with 10% loss to follow-up after 24 months and 25% delayed APM in the PT group, 201 patients were needed per group in this non-inferiority trial. This meant a total of 402 patients. The sample size was calculated for the intention-to-treat analysis.

In order to avoid unnecessary inclusions and unnecessary delay, we recalculated our SD halfway through the study. This interim analysis was performed by an independent committee consisting of an orthopaedic surgeon/expert in the field and an orthopaedic research coordinator/statistical expert. Only the SD was recalculated, all other outcome data remained blinded and no analyses were performed for any of the outcomes with different sample sizes. With an SD of 18 points (compared to the SD of 20 in our initial calculation), the committee recalculated the sample size. We agreed on a sample size reduction to a total number of 320 patients (160 per group). The Ethics Review Board granted approval for this on October 27, 2015. The change of sample size has been updated in the trial registries.

Data analysis

Effectiveness analysis

To investigate the clinical effectiveness of both treatment groups, we will use linear mixed-model analysis for continuous outcomes. Logistic generalised estimation equation analysis will be used for dichotomous outcomes. This method takes into account the dependency of observations within a patient and the fact that not all patients may be assessed at each time point (missing data). All analyses will be carried out on an intention-to-treat and per/protocol basis, as well as cross-over analysis.

Table 1. Measurement moments

Baseline (t0)	3 months (t2)	6 months (t3)	9 months (t4)	12 months (t5)	18 months (t6)	24 months (t7)
CRF-1	CRF-2	CRF-3	CRF-4	CRF-5	CRF-6	CRF-7
Visit						
IKDC	IKDC	IKDC		IKDC		IKDC
RAND-36	RAND-36	RAND-36		RAND-36		RAND-36
EQ-5D-5L	EQ-5D-5L	EQ-5D-5L	EQ-5D-5L	EQ-5D-5L	EQ-5D-5L	EQ-5D-5L
VAS	VAS	VAS		VAS		VAS
TAS	TAS	TAS		TAS		TAS
PSC	PSC	PSC		PSC		PSC
TiC-P	TiC-P	TiC-P	TiC-P	TiC-P	TiC-P	TiC-P
PE	PE					
X-ray						X-ray
Expectation	Satisfaction	Satisfaction		Satisfaction		Satisfaction
MRI						

Abbreviations: CRF, Case Report Form; EQ-5D-5L, EuroQol 5-Dimensions 5-Level Survey; IKDC, International Knee Documentation Committee; PE, physical examination; PSC, Patient Specific Complaints questionnaire; TAS, Tegner Activity Scale; TiC-P, Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness; VAS, visual analogue scale.

In the primary linear mixed model, the outcome variable studied (e.g., physical function on the IKDC) will be analysed as a dependent variable. To investigate the effect at the different time points, we will analyse the model, according to a 4-level structure (treatment group, centre, patient and time, in which time will be treated as a categorical variable to assess the treatment effects at the different time points). Time will be included as a dummy variable (reference is baseline T0), and 4 interaction terms will be analysed (T2Xgroup; T3Xgroup; T5Xgroup; T7Xgroup). To investigate the overall effect of both treatments (irrespective of time), we will also analyse the model according to a 3-level structure (treatment group, centre, patient). The baseline outcome will be included as a covariate in all models.

Besides analysing the basic model (e.g., analysis of main effects for treatment group and time and a time-by-treatment interaction), we will also control for possible confounders, by adding them as covariates (e.g., body mass index, gender, profession, ASA classification, the affected meniscus, the type of tear and the status of OA according to Kellgren and Lawrence Grading Scale for Osteoarthritis). Covariates are defined as resulting in more than 10% change in the parameter estimate of time-by-treatment interaction.

In the secondary linear mixed models, the outcome variables studied (e.g., general health on the RAND-36, quality of life on the EQ-5D-5L, level of activity on the TAS, knee

pain on the question 10 of IKDC, the correlation between a patient's expectation and satisfaction, productivity losses on the Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness (TiC-P), muscle strength, range of motion and squatting) will be analysed in a similar way.

The estimated main effects for treatment at different assessment points under these different models are reported as in differences in means with 95% CIs for continuous outcomes, and ORs with 95% CIs for dichotomous outcomes

At the time points 3 months (T2), 6 months (T3), 12 months (T5) and 2 years (T7), we will describe the incidence of revisions (intervention group) or treatment failures (=delayed APM, control group) using descriptives. After 2 years (T7), we will compare the incidence of development or progression of OA between groups using a $\chi 2$ test (or Fisher's exact as appropriate).

For all analyses, a 2-tailed value of $p<.05$ is considered to be significant.

We will consult a statistician for all longitudinal analysis.

Cost-effectiveness analysis

General considerations

The economic evaluation will be conducted from a societal perspective. The aim of the economic evaluation is to measure, value and analyse total costs of patients in both groups and to relate the difference in costs between the 2 treatment groups to the difference in clinical effects. We will perform both a cost-effectiveness and cost-utility analysis. The time horizon of the economic evaluation is 24 months, so discounting will be used. Sensitivity analysis will be performed to assess the robustness of the results using different assumptions regarding costs and effects.

Patient outcome analysis

Effect measures in the economic evaluation are physical function, pain intensity and general health. Quality-adjusted life-years (QALYs) based on the Dutch tariff for the EuroQol will also be measured.[105,106]

The analysis will be carried out according to the intention-to-treat principle. Missing cost and effect data will be imputed using multiple imputations, according to the National Institute for Health and Care Excellence (NICE) algorithm developed by van Buuren and colleagues.[107]

We will perform a full cost-effectiveness and cost-utility analysis. Incremental cost-effectiveness ratios (ICERs) will be calculated by dividing the difference in mean total costs between the treatment groups by the difference in mean effects.

Bias-corrected and accelerated bootstrapping with 5000 replications will be used to estimate 95% CIs around cost differences and the uncertainty surrounding the ICERs. Rubin's

rules will be used to pool the results from the different multiple imputed data sets. Uncertainty surrounding the ICERs will be graphically presented on cost-effectiveness planes.

Cost-effectiveness acceptability curves will also be estimated using the net benefit framework.[108] Cost-effectiveness acceptability curves show the probability that APM is cost-effective compared with PT for a range of different ceiling ratios thereby showing decision uncertainty.[109]

Cost-analysis

Costs will be measured using a web-based questionnaire, which is a modified version of the TiC-P.[110] Direct costs include costs of APM surgery and costs of PT, but also other healthcare expenses for knee problems such as general practitioner care, costs of visits to other primary care providers, ambulatory and inpatient hospital care, medication and home care. Indirect costs include absenteeism from paid and unpaid work and presenteeism. The friction cost approach will be used in the primary analysis to estimate indirect costs.[111] We will use standard prices published in the Dutch costing guidelines for the valuation of healthcare usage.[112] Medication use will be valued using prices of the Royal Dutch Society for Pharmacy.

Cost-Effectiveness analysis

Effect measures in the economic evaluation are physical function based on the IKDC 'Subjective Knee Form' and general health based on the EuroQol. QALYs based on the Dutch tariff for the EuroQol will also be measured.[105,106]

The analysis will be carried out according to the intention-to-treat principle. Missing cost and effect data will be imputed using multiple imputations according to the NICE algorithm developed by van Buuren and colleagues.[107]

We will perform a full cost-effectiveness and cost-utility analysis. ICERs will be calculated by dividing the difference in mean total costs between the treatment groups by the difference in mean effects. Bias-corrected and accelerated bootstrapping with 5000 replications will be used to estimate 95% CIs around cost differences and the uncertainty surrounding the ICERs. Rubin's rules will be used to pool the results from the different multiple imputed datasets. Uncertainty surrounding the ICERs will be graphically presented on cost-effectiveness planes. Cost-effectiveness acceptability curves will also be estimated using the net benefit framework.[108] Cost-effectiveness acceptability curves show the probability that APM is cost-effective compared with PT for a range of different ceiling ratios thereby showing decision uncertainty.[109]

Budget impact analysis

General considerations

In the budget impact analysis, the results of the economic evaluation will be linearly extrapolated over a period of 5 years to estimate the financial consequences of implementation of the study results. An estimate of the long-term financial consequences will also be given to quantify the impact of the expected decrease of the progression of OA and therefore the number of knee arthroplasties. The intervention will be offered to patients aged 45–70 years who were diagnosed with symptomatic, non-obstructive, MRI-confirmed meniscal tears. Perspectives that will be considered are the societal, government (Budget Kader Zorg) and insurer. Different scenarios will be evaluated including the following: (1) all patients will receive APM; (2) all patients will receive PT; (3) PT will replace APM gradually over a period of 4 years (25% change per year).

One-way sensitivity analysis will be performed in which the change rate per year and the reduction of the number of knee arthroplasties will be varied.

Cost-analysis

The total number of patients aged 45–70 years who were diagnosed with symptomatic, non-obstructive, MRI-confirmed meniscal tears will be estimated based on Dutch incidence and prevalence rates. Resource usage is calculated by multiplying the number of eligible patients with the resource usage rates obtained from the cost-effectiveness analysis.

We will use different prices to value resource use depending on the perspective of the analysis: Dutch standard costs for the societal perspective, actual Nederlandse Zorgautoriteit (in English: Dutch Healthcare Authority) (NZA) tariffs for the government perspective, and average tariffs NZA for the insurer perspective.

Both resource use and annual costs will be presented over a 5-year period for all perspectives. Aggregated and disaggregated total costs per year will be presented for the different perspectives and scenarios. For the long-term analysis, total costs over the whole time horizon will be estimated.

Data analysts are blinded to the type of treatment by numerical coding of the performed intervention. After finalising data analysis, this code will be broken for publication purposes.

Data handling and confidentiality

Data will be collected using online questionnaires. All participant data will be anonymised by assigning study numbers to each participant. The study numbers will not be based on the patient initials or birth date. The key to these study numbers is only available to the researchers (JCAN and on demand by the principal investigators). Outcome data, anonymised, is only accessible for the coordinating investigator (VAG), principal investigators (RWP and AG), research assistant (JCAN), statistical analysers (NW

and VABS) and authorised research personnel of the Joint Research Group at the OLVG Amsterdam. Data will be collected and stored for a period of 15 years. Paper and original questionnaires will be kept in a database at the initiating hospital (OLVG). Data will be processed and stored in SPSS, password protected.

Security requirements: Data input capabilities are limited to the coordinating investigator (VAG) and the research assistant (JCAN). Data processing capabilities are limited to the coordinating investigator, statistical analysers (NW and VABS), the principal investigators, and authorised research staff.

The handling of personal data will comply with the Dutch Personal Data Protection Act (de Wet Bescherming Persoonsgegevens, Wbp).

Steering and data monitoring committee

There is no official steering committee for this study. The following representatives from the participating organisations are involved in the project oversight and control: RWP, MD PhD (principal investigator and sponsor); VAG, MD; NW, PhD; VABS PhD; MWT, PhD; and JCAN, Msc.

All study-related problems or (serious) adverse events (SAEs) will be discussed with the principal investigator RWP, and researchers VAG, VABS and JCAN. SAEs will be officially reported to the ethical committee. The ethical committee judges will decide whether the safety of the patients is jeopardised and whether the trial can be continued or not.

There is no official data monitoring committee. Data entry will be performed by one of the researchers (JCAN) and checked and cleaned according to the quality handbook of the EMGO+ institute for health and care research (http://www.emgo.nl/kc). In addition, a random sample of 5% of the data will be re-entered by another researcher to check for inconsistencies. A third researcher will be involved with the data processing and analysis, which will be performed without having knowledge of the allocation key. All data analyses will be discussed with the researchers (RWP, VAG and JCAN) before the final presentation of the results. A professor (MWT) specialised in cost-effectiveness will perform the economic evaluation in association with one of the researchers (VAG).

Ethics and dissemination

This study will be conducted in accordance with the Declaration of Helsinki and the Medical Research Involving Human Subjects Act (WMO). Also, all institutional review boards have approved the start of the study. All substantial amendments to the protocol will be notified to the ethics committee and to the competent authority. Non-substantial amendments will not be notified to the accredited Medisch Ethische ToetsingsCommissie (in English Medical Ethical Committee) (METC) and the competent authority, but will be recorded and filed by the sponsor. Written informed consent will be obtained from all participating patients. The research coordinator will report all SAEs within 24 hours

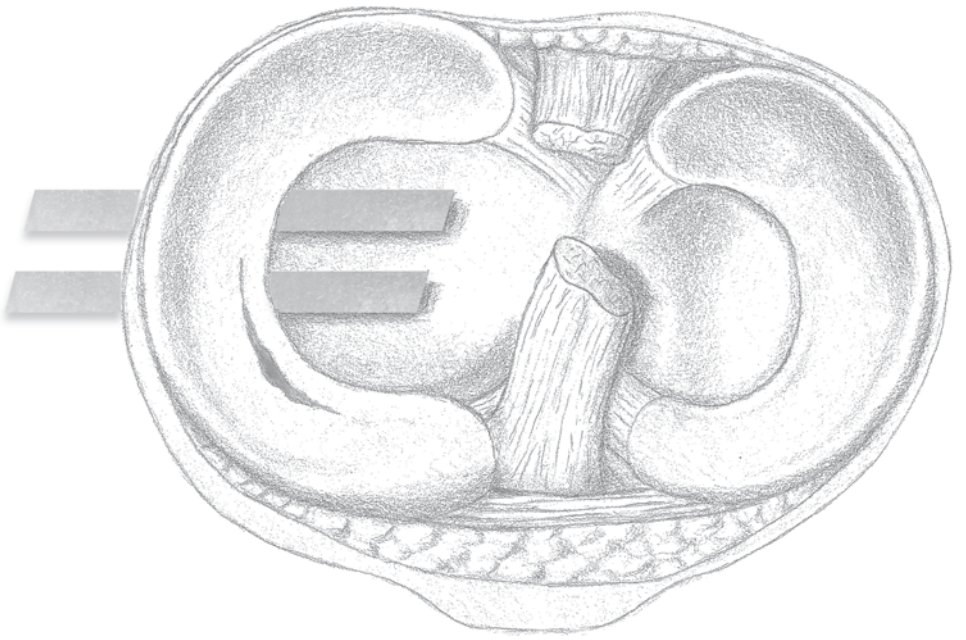
of noticing, using the online submission system of the ethics committee. The ethical committee judges will decide whether the safety of the patients is jeopardised and whether the trial can be continued or not. We will submit our study results for publication in peer-reviewed journals and present at international conferences. Furthermore, we aim to disseminate our results to guideline committees.

Discussion

In this protocol paper, we propose the protocol of an economic evaluation study for the assessment of (cost-)effectiveness of early APM versus conservative treatment with optional delayed meniscectomy for patients between 45 and 70 years old with a meniscal tear. Previous RCTs found no difference in outcome between surgical and conservative treatment.[3,19,21,22,41,66]

Since we were unaware of the exact SD of the IKDC in this patient group, we decided to calculate the SD in our own group. Subsequently, we could use this for a recalculation of our sample size in order to avoid unnecessary inclusions and any further (unnecessary) delay. The SD in our own group was found to be 18, compared with the SD of 20 used for our initial sample size calculation. This resulted in a reduction of 82 patients. As previously mentioned, an independent committee consisting of an orthopaedic surgeon/expert in the field and an orthopaedic research coordinator/statistical expert were appointed for this recalculation. During this process, all other data remained blinded and no analyses were performed for any of the outcomes with different sample sizes. The Ethics Review Board approved this recalculation.

This RCT will be the first to investigate and publish data on the cost-effectiveness of both treatment groups in this specific group of patients. Therefore, this trial adds to the clinical evidence of treatment of meniscal tears which contributes to the ongoing debate to reduce healthcare costs in the western world.

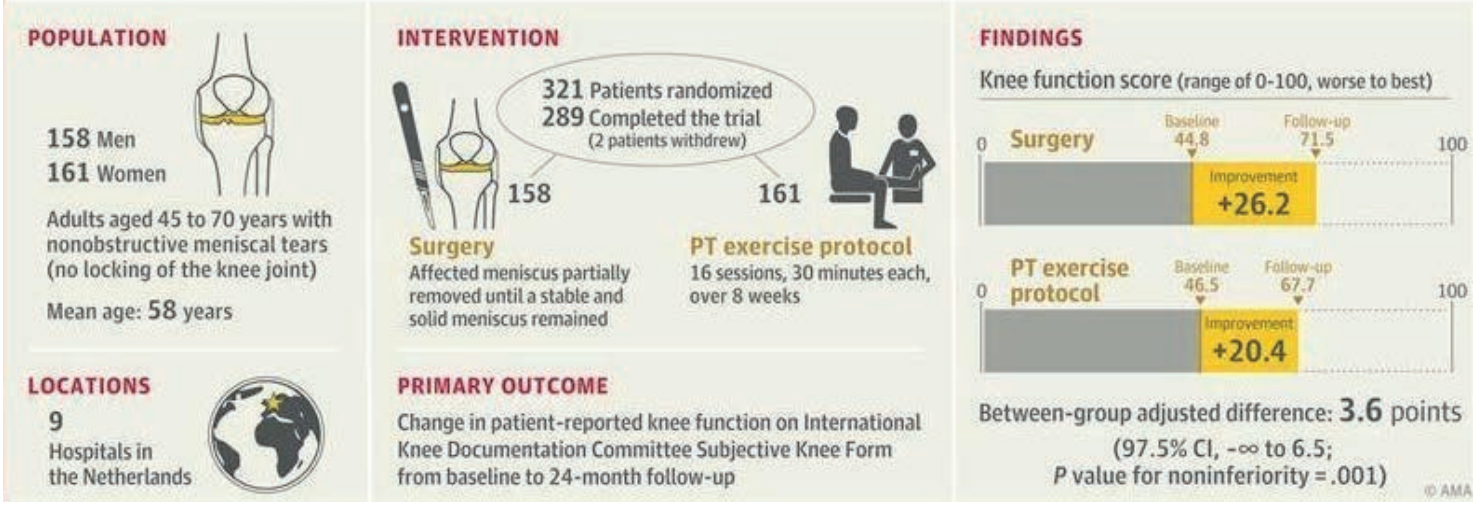


5

Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears: the Escape Randomized Clinical Trial

Victor A van de Graaf
Julia CA Noorduyn
Nienke W Willigenburg
Ise K Butter
Arthur de Gast
Ben W Mol
Daniël BF Saris
Jos WR Twisk
Rudolf W Poolman
for the Escape Research Group

JAMA. 2018 Oct 2;320(13):1328-1337



Visual abstract

Abstract

Importance

Despite recent studies suggesting arthroscopic partial meniscectomy (APM) is not more effective than physical therapy (PT), the procedure is still frequently performed in patients with meniscal tears.

Objective

To assess whether PT is non-inferior to APM for improving patient-reported knee function in patients with meniscal tears.

Design, Setting, and Participants

Non-inferiority, multicentre, randomised clinical trial conducted in 9 hospitals in the Netherlands. Participants were aged 45 to 70 years with non-obstructive meniscal tears (i.e., no locking of the knee joint). Patients with knee instability, severe osteoarthritis, and body mass index greater than 35 were excluded. Recruitment took place between July 17, 2013, and November 4, 2015. Participants were followed up for 24 months (final participant follow-up, October 11, 2017).

Interventions

Three hundred twenty-one participants were randomly assigned to APM (n=159) or a predefined PT protocol (n=162). The PT protocol consisted of 16 sessions of exercise therapy over 8 weeks focused on coordination and closed kinetic chain strength exercises.

Main Outcomes and Measures

The primary outcome was change in patient-reported knee function on the International Knee Documentation Committee Subjective Knee Form (range, 0 to 100; from worse to best) from baseline over a 24-month follow-up period. The non-inferiority margin was defined as a difference between treatment groups of 8 points and was assessed with a 1-sided α of .025. The primary analysis followed the intention-to-treat principle.

Results

Among 321 patients who were randomised (mean [SD] age, 58 [6.6] years; 161 women [50%]), 289 (90%) completed the trial (161 women and 158 men). In the PT group, 47 participants (29%) had APM during the 24-month follow-up period, and 8 participants randomised to APM (5%) did not have APM. Over a 24-month follow-up period, knee function improved in the APM group by 26.2 points (from 44.8 to 71.5) and in the PT group by 20.4 points (from 46.5 to 67.7). The overall between-group difference was 3.6 points (97.5% CI, −∞ to 6.5; p-value for non-inferiority = .001). Adverse events occurred

in 18 participants in the APM group and 12 in the PT group. Repeat surgery (3 in the APM group and 1 in the PT group) and additional outpatient visits for knee pain (6 in the APM group and 2 in the PT group) were the most frequent adverse events.

Conclusions

And Relevance Among patients with non-obstructive meniscal tears, PT was non-inferior to APM for improving patient-reported knee function over a 24-month follow-up period. Based on these results, PT may be considered an alternative to surgery for patients with non-obstructive meniscal tears.

Trial Registration

ClinicalTrials.gov Identifier: NCT01850719

Introduction

Arthroscopic partial meniscectomy (APM) is among the most frequently performed procedures in orthopaedic surgery. It was estimated that in 2014, 516 800 meniscectomies were performed in the United States, and the global annual cost was estimated at $4 billion in 2006.[4,113]

Meniscal tears may occur as part of a degenerative process of the knee joint and occur in up to 60% of persons older than 50 years of age without knee pain.[17] Because physical therapy (PT) has positive short-term effects on knee pain and function in patients with knee osteoarthritis,[114] the benefit of surgical resection of the degenerative meniscal tear compared with PT is unclear.[115]

To date, 6 randomised clinical trials (RCTs) have assessed superiority of APM compared with either PT[3,21,22,41,116] or sham surgery[4] in patients with a confirmed meniscal tear. These trials reported no significant differences between treatment groups for knee function. A meta-analysis that included data from 5 RCTs found a statistically significant benefit of APM for knee function and pain at up to 6-month follow-up, but this benefit did not persist at 1- or 2-year follow-up.[98,115] Evidence published to date has not led to a major decline in APM for managing meniscal tears.[115,117]

Therefore, the objective of this study was to determine whether PT is non-inferior to APM for improving self-reported knee function over a 24-month follow-up period in patients with non-obstructive meniscal tears.

Methods

Trial Oversight and Design

This trial was a non-inferiority, multicentre RCT performed in 9 hospitals in the Netherlands, comparing APM with PT in patients with meniscal tears. The study protocol has been published (*Chapter 4*). The study was conducted in accordance with the Declaration of Helsinki and was approved by the Medical Research Ethics Committees United (MEC-U; NL44188.100.13) and by the board of directors of each of the participating hospitals. All participants provided written informed consent prior to randomisation. The number of patients screened for eligibility was not documented.

Patient Population

Participants aged 45 to 70 years who were referred to 1 of 9 participating hospitals with knee pain and a non-obstructive (i.e., no locking of the knee joint) meniscal tear confirmed by magnetic resonance imaging (MRI) were eligible for inclusion. Exclusion criteria were locking of the knee, prior knee surgery, instability caused by an anterior or posterior cruciate ligament rupture, severe osteoarthritis (Kellgren-Lawrence score of 4, indicating large osteophytes, marked joint-space narrowing, severe sclerosis, and definite bone ends deformity),[83] and a body mass index (BMI, calculated as weight in kilograms

divided by height in meters squared) greater than 35. No distinction was made between traumatic and degenerative tears because in older patients, even traumatic tears may be related to degenerative changes in the knee. Recruitment was carried out between July 17, 2013, and November 4, 2015. Follow-up testing was completed on October 11, 2017.

Intervention, Randomisation, and Blinding

Patients with a meniscal tear were informed about the study by their treating orthopaedic surgeon at their first outpatient visit. After written informed consent was obtained, the research coordinator randomised study participants to APM or PT. Randomisation was concealed and performed using a central computer-generated randomisation scheme in a 1:1 ratio with variable block size (minimum block size of 2 and maximum block size of 6). Randomisation was stratified by hospital and by age (45–57 and 58–70 years). Participants, physicians, and physical therapists were not blinded. Investigators who performed the statistical analysis were blinded. After the analysis was completed, data were unblinded for the final interpretation of the results.

Participants randomised to APM were scheduled for surgery within 4 weeks of randomisation. Surgery was performed in an outpatient clinic under general or spinal anaesthesia. During surgery, standard anteromedial and anterolateral portals were introduced for inspection of the knee joint. The affected meniscus was partially removed until a stable and solid meniscus remained. All participants received perioperative instructions and a home exercise program. Participants were only referred to PT after APM if they did not recover as anticipated (i.e., they did not adequately improve or experienced a decrease in knee function, ability to participate in daily activities, and/or had an increase in knee pain), as defined by the Dutch Orthopaedic Association guideline.[31]

Participants randomised to PT were referred to PT clinics directly after randomisation and their initial PT session was scheduled within 2 weeks after randomisation. Participating PT clinics were instructed about the exercise protocol by a knee-specialised physical therapist or the primary investigator, prior to the first participant's referral. The PT exercise protocol was developed by a knee-specialised physical therapist and consisted of 16 sessions of 30 minutes each conducted over 8 weeks (Appendix 5). The PT protocol included cardio-vascular, coordination/balance, and closed kinetic chain strength exercises (in which the distal part of the extremity is fixed to an object that is stationary). If PT failed (e.g., knee pain or limitations in daily activities persisted or locking occurred), the participant could attend additional PT sessions or have APM, depending on their preference.

Outcomes

The primary outcome was the change in patient-reported knee function on the Subjective Knee Form of the International Knee Documentation Committee

(IKDC) from baseline over 24 months. The IKDC is a validated and self-administered questionnaire designed for patients with a variety of knee disorders that assesses knee function, symptoms, and ability to engage in sports activities.[77,92,100] IKDC scores range from 0 to 100, in which 100 indicates no knee symptoms or limitations in daily or sporting activities. Normative values for the IKDC in the United States are 88 (SD, 14) for men aged 51 to 65 years and 85 (SD, 16) for women aged 51 to 65 years in a population without current knee problems or a history of knee surgery. In a population of people with and without current knee problems, the normative scores are 77 (SD, 23) for men aged 51 to 65 years and 71 (SD, 26) for women aged 51 to 65 years.[118]

Because a minimal clinically important difference (MCID) for the IKDC has not been defined in a population consisting only of patients with meniscal tears, the non-inferiority margin was defined as the smallest detectable change of 8.8 points,[92] rounded down to a margin of 8 points.

Secondary outcomes included knee pain during weight-bearing, general health, the progression of osteoarthritis, and activity level. Knee pain during weight-bearing was measured on a visual analogue scale (VAS[119]; range, 0-100, with 0 anchored as "no pain" and 100 as "worst pain imaginable"). General health was measured with the RAND-36 Physical Component Score, derived from the RAND-36 questionnaire[101] (range, 0-100, with higher scores indicating better health). The mean (SD) score in the general population is 50 (10).[101] Progression of osteoarthritis was measured using the Kellgren-Lawrence classification (range, 0-4, in which grade 0 [no osteophytes or joint-space narrowing] indicates no osteoarthritis and grade 4 [>50% joint-space narrowing] indicates severe osteoarthritis).[83] Activity level was measured with the Tegner Activity Scale, which measures the level of working activities and sport activities on a scale from 0 to 10, with higher scores indicating a higher level of activity.[103] Although these outcomes were originally intended to test for non-inferiority, no non-inferiority margins were defined in advance. Therefore, the secondary outcomes were tested for superiority. After data analyses, MCIDs were identified in the literature to guide the interpretation of observed differences between treatment groups. Adverse events were categorised as serious and non-serious.

Other pre-specified outcomes included resource utilisation, health-related quality of life, patient-specific complaints, participant expectations, and participant satisfaction. These outcomes will be analysed and reported separately. Participants completed study questionnaires at baseline, 3, 6, 12, and 24 months after randomisation. If participants did not respond, up to 3 reminders were sent and, if needed, the participant was contacted by telephone.

MRIs and X-rays were taken at the time of enrolment and X-rays were performed 24 months after randomisation. One radiologist reviewed all X-rays to grade osteoarthritis severity on the Kellgren-Lawrence classification,[83] while another radiologist reviewed all MRIs for classification of the meniscal tears according to the Modified International

Cartilage Repair Society classification.[120] Both radiologists were unable to assess whether a participant had surgery and were blinded to treatment allocation.

Sample Size Calculation and Statistical Analysis

The sample size was based on an SD of 18 points on the IKDC, a power of 90%, a 2-sided α of .05, and a non-inferiority margin of 8 points on the IKDC. With an anticipated 20% loss to follow-up and a 25% delayed APM rate after 24 months, 160 participants per treatment group were needed.

Mixed models were used for longitudinal data analyses, with a 3-level structure, i.e., repeated measurements were clustered within participants and participants were clustered within the participating centres, to calculate the overall between-group differences. Unadjusted between-group differences were calculated based on a model with the baseline score and treatment group as independent variables. To define the between-group differences per follow-up period, time was added to the model as a categorical variable, as well as a time-by-treatment interaction. Adjusted between-group differences were calculated based on similar models including potential confounders as independent variables. These confounders were sex, age, BMI, education level,[121] Kellgren-Lawrence classification,[83] location of the tear (medial, lateral, both), mechanical complaints, and baseline knee pain during weight-bearing.

Progression of osteoarthritis was analysed using a mixed model with Kellgren-Lawrence score at 24 months as the dependent variable and intervention group and baseline Kellgren-Lawrence score as independent variables. Analyses followed the intention-to-treat principle, in which patients were analysed according to their randomised treatment allocation, regardless of any deviations from the protocol.

To test the robustness of the results, analyses that followed the as-treated principle were performed for the unadjusted and adjusted between-group differences for the outcomes of knee function and pain. In these analyses, participants were analysed based on their adherence to randomised treatment allocation in 3 groups: (1) participants randomised to the APM group who received APM, (2) participants randomised to the PT group who completed the PT protocol without having APM during the follow-up period, and (3) participants randomised to the PT group who had APM during follow-up (delayed APM group). Patients who were randomised to the APM treatment group but did not have surgery and patients who were randomised to PT but did not complete the PT protocol and did not have delayed APM were not included in the as-treated analysis.

Adverse events were reported descriptively. Post hoc exploratory analyses were performed to assess effect modification on the primary outcome by evaluating the interaction term between each of the potential confounders listed above.

A 97.5% CI (i.e., a 1-sided α of .025) was used for the knee function outcome, whereas a 95% CI (i.e., a 2-sided α of .05) was used for the other outcomes. Because the analyses

for the secondary outcomes were not corrected for type 1 error, they should be considered exploratory.

Non-inferiority was demonstrated when the 97.5% CI did not include the non-inferiority margin. Missing data were handled using full maximum likelihood estimation. If participants withdrew from the trial, the data collected prior to withdrawal were used in the analyses, with the participant's approval. All analyses were conducted using IBM SPSS version 22 (IBM).

Deviations from the Original Trial Protocol

Four inconsistencies with the original protocol need to be addressed (Appendix 6). First, the power calculation was adjusted after an interim analysis of the first 100 participants' IKDC scores at 1-year follow-up, which demonstrated that the SD was 18 points compared with the anticipated 20 points derived from Crawford and colleagues[92] used in the sample size calculation. This allowed for a reduction of the sample size from 402 to 320.

Second, the description of the primary outcome in the original protocol was the change from baseline to 24-month follow-up. However, the trial was designed to include all follow-up time points in the primary outcome measure and, therefore, the originally intended primary outcome measure was the change from baseline over the 24-month follow-up. This was not explicitly specified in the protocol.

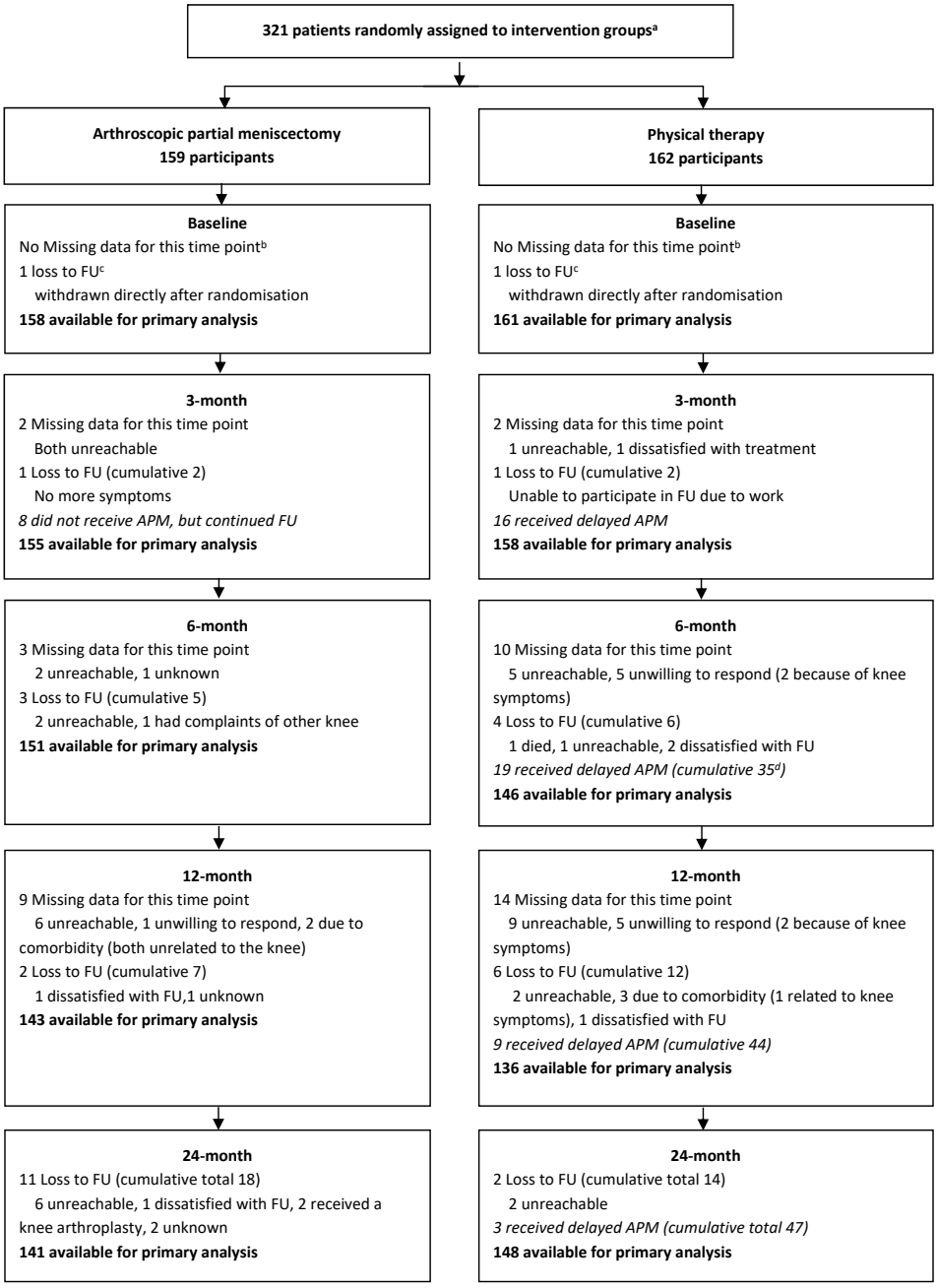
Third, the original protocol described generalised estimating equations for longitudinal analyses. During the study, and more than a year prior to data analysis, statistical consultation indicated that a mixed-model approach was more suitable for our data set. The advantages of a mixed-model analysis over generalised estimating equations are the ability to correct for recruitment centre without significant loss of power and the ability to handle missing data using maximum likelihood estimation without the need for imputation.

Fourth, an error was made in the protocol as it mentions a loss to follow-up of 10%, while the anticipated percentage was 20%.

Results

Participants

Between July 17, 2013, and November 4, 2015, 321 participants were enrolled and randomly assigned to APM (n=159) or PT (n=162) (Figure 1). Two participants (1 in each group) withdrew consent immediately after randomisation without providing a reason. After 24 months, 289 participants (90%) completed follow-up, with the final participant's follow-up visit occurring on October 11, 2017. The baseline characteristics in the 2 treatment groups were comparable (mean [SD] age, 58 [6.6] years; 161 women (50%); Table 1). The distributions of baseline knee function and knee pain during weight-bearing are presented in Appendix 7 and 8.


321 patients randomly assigned to intervention groups[a]
Arthroscopic partial meniscectomy
159 participants
Physical therapy
162 participants
Baseline
No Missing data for this time point[b]
1 loss to FU[c]
withdrawn directly after randomisation
158 available for primary analysis
Baseline
No Missing data for this time point[b]
1 loss to FU[c]
withdrawn directly after randomisation
161 available for primary analysis
3-month
2 Missing data for this time point
Both unreachable
1 Loss to FU (cumulative 2)
No more symptoms
8 did not receive APM, but continued FU
155 available for primary analysis
3-month
2 Missing data for this time point
1 unreachable, 1 dissatisfied with treatment
1 Loss to FU (cumulative 2)
Unable to participate in FU due to work
16 received delayed APM
158 available for primary analysis
6-month
3 Missing data for this time point
2 unreachable, 1 unknown
3 Loss to FU (cumulative 5)
2 unreachable, 1 had complaints of other knee
151 available for primary analysis
6-month
10 Missing data for this time point
5 unreachable, 5 unwilling to respond (2 because of knee symptoms)
4 Loss to FU (cumulative 6)
1 died, 1 unreachable, 2 dissatisfied with FU
19 received delayed APM (cumulative 35[d])
146 available for primary analysis
12-month
9 Missing data for this time point
6 unreachable, 1 unwilling to respond, 2 due to comorbidity (both unrelated to the knee)
2 Loss to FU (cumulative 7)
1 dissatisfied with FU,1 unknown
143 available for primary analysis
12-month
14 Missing data for this time point
9 unreachable, 5 unwilling to respond (2 because of knee symptoms)
6 Loss to FU (cumulative 12)
2 unreachable, 3 due to comorbidity (1 related to knee symptoms), 1 dissatisfied with FU
9 received delayed APM (cumulative 44)
136 available for primary analysis
24-month
11 Loss to FU (cumulative total 18)
6 unreachable, 1 dissatisfied with FU, 2 received a knee arthroplasty, 2 unknown
141 available for primary analysis
24-month
2 Loss to FU (cumulative total 14)
2 unreachable
3 received delayed APM (cumulative total 47)
148 available for primary analysis

Eight participants (5%) of the APM group chose not to have surgery because they determined that their symptoms were not severe enough for surgical management. Four of these participants attended a mean of 14.5 PT sessions (range, 6-23), while the other 4 did not receive PT. Two participants in the APM group had a total knee arthroplasty within 2 years of their APM.

Participants in the PT group attended a mean of 17 PT sessions (range, 0–40). In the PT group, 47 participants (29%) had APM due to the persistence of symptoms, and 35 (75%) had APM within 6 months of randomisation. Three participants in the PT group had a total knee arthroplasty within the 24-month follow-up period. Seventeen participants did not complete the PT protocol (<16 sessions).

Primary Outcome

Figure 2 shows the improvement in knee function from baseline over the 24-month follow-up period for the APM and PT groups. In the APM group, knee function improved from 44.8 points at baseline to 71.5 points at 24 months (mean difference [MD], 26.2 points [95% CI, 23.2 to 29.3]). In the PT group, knee function improved from 46.5 points at baseline to 67.7 points at 24 months (MD, 20.4 points [95% CI, 17.5 to 23.2]). The primary mixed model analysis of the overall effects found a between-group difference of 3.6 points (97.5% CI, −∞ to 6.5; p-value for non-inferiority = .001) in favour of the APM group, indicating non-inferiority of PT compared with APM (Table 2).

The between-group differences at 3 (0.78 points [97.5% CI, −∞ to 4.3]) and 6 (3.4 points [97.5% CI, −∞ to 7.0]) months after randomisation also demonstrated non-inferiority of PT; however, the effects at 12 (5.7 points [97.5% CI, −∞ to 9.4]) and 24 (4.8 points [97.5% CI, −∞ to 8.5]) months after randomisation did not demonstrate non-inferiority. The adjusted between-group differences and the individual change scores from baseline to 24 months are presented in Appendix 9 and 10, respectively.

Exploratory Outcomes

Knee pain during weight-bearing improved in the APM group from 61.1 mm at baseline to 19.6 mm at 24 months (MD, 39.2 mm [95% CI, 33.8 to 44.6]) and in the PT group, knee

Figure 1. Flow of patients through the trial.

[a] The number of patients screened for eligibility was not available.

[b] Missing data refer to data that were missing at a specific time point, while patients remained available for the remaining follow-up times.

[c] Loss to follow-up refers to actual dropout from the study; e.g., patients who did not participate at any of the remaining time points (cumulative numbers are total number of dropouts).

[d] Cumulative number of delayed APM refers to total number of participants from the physical therapy group who received delayed APM from baseline until that follow-up.

Abbreviations: APM, arthroscopic partial meniscectomy, FU, follow-up.

Table 1. Baseline characteristics of the intention-to-treat population

	APM group	PT group
Demographics	N=158	N=161
Age, mean (SD), y	57.6 (6.5)	57.3 (6.8)
Women	80 (50.6)	81 (50.3)
Right knee	88 (55.7)	81 (50.3)
Education level, beyond high school[a]	67 (42.4)	86 (53.4)
BMI, mean (SD)	26.7 (3.8)	27.2 (4.0)
18.5<BMI<25	56 (35.4)	53 (32.9)
25≤BMI<30	72 (45.6)	67 (41.6)
30≤BMI<35	30 (19.0)	41 (25.5)
Mechanical complaints[b]	56 (35.4)	67 (41.6)
Imaging[c]		
Affected meniscus		
Medial	126 (79.7)	136 (84.5)
Lateral	30 (19.0)	25 (15.5)
Both	2 (1.3)	0 (0)
Type of tear on MRI (ISAKOS[120])	N=151	N=152
Longitudinal-vertical	5 (3.3)	5 (3.3)
Horizontal	80 (53.0)	69 (45.4)
Complex degenerative	47 (31.1)	58 (38.1)
Radial	13 (8.6)	10 (6.6)
Vertical flap	2 (1.3)	5 (3.3)
Unclassifiable	1 (0.7)	5 (3.3)
Horizontal flap	3 (2.0)	0 (0)
Osteoarthritis score on radiographs (KL classification[83])[d]	N=150	N=149
0 – No OA	18 (12.0)	15 (10.1)
1 – Doubtful	81 (54.0)	74 (49.7)
2 – Minimal OA	45 (30.0)	55 (36.9)
3 – Moderate OA	6 (4.0)	5 (3.3)
Knee Function and Pain		
Knee function	N=158	N=161
IKDC score (0 (most limitations) to 100 (no limitations))[f]	44.8 ± 16.6	46.5 ± 14.6
Knee pain	N=145	N=151
VAS during weight-bearing (0 [no pain] to 100 [worst pain imaginable])[g]	61.1 (44.9–83.4)	59.3 (44.9–77.4)

[a] Education level was measured according to the International Standard Classification of Education (ISCED) and dichotomized to low (ISCED level 0–3; e.g., early childhood education, primary education, or high school) or high (ISCED level 4–8; e.g., any education beyond high school, including bachelor's, master's, or doctoral degree).[121]

[b] In contrast to locking of the knee joint, which was an exclusion criterion, mechanical complaints, such as clicking or catching, were allowed for inclusion.

[c] All participants underwent MRI prior to inclusion, and information on the affected meniscus was based on clinical readings by different radiologists and orthopaedic surgeons in participating centres. The type of tear was based on post hoc readings of the MRIs by 1 radiologist, and osteoarthritis scores were based on study readings of the MRIs by 1 other radiologist. Some radiographs and MRIs were unavailable (8 and 7 for APM and 12 and 9 for PT, respectively) for study readings.

Table 1. (continued)

[d] KL grade 0 (no osteophytes or joint-space narrowing) indicates no osteoarthritis, grade 1 (questionable osteophytes) indicates early-onset osteoarthritis, grade 2 (definite osteophytes, possible joint-space narrowing) indicates mild osteoarthritis, grade 3 (moderate osteophytes, definite joint-space narrowing, some sclerosis, possible bone-end deformity) indicates moderate osteoarthritis, and grade 4 (large osteophytes, marked joint-space narrowing, severe sclerosis, definite bone ends deformity) indicates severe osteoarthritis.[83] KL grade 4 was an exclusion criterion.

[e] IKDC scores range from 0 to 100, in which 0 indicates the highest level of knee symptoms and lowest level of function in daily or sporting activities, and 100 indicates the lowest knee symptoms and highest level of function in daily or sporting activities. Normative mean (SD) values for the IKDC in the United States are 88 (14) for men aged 51 to 65 years and 85 (16) for women aged 51 to 65 years in a population without current knee problems or a history of knee surgery. In a mixed population of people with and without current knee problems, the normative mean (SD) scores are 77 (23) for men aged 51 to 65 years and 71 (26) for women aged 51 to 65 years.[118] For example, a patient scores 50.6 on the IKDC if the highest level of activities and the effect of the knee on activities are graded as "moderate" for all asked activities, the level of knee pain is "moderate" (5 of 11), mechanical complaints are absent, and the current knee function is graded as "moderate" (5 of 11).

[f] VAS score ranged from 0 to 100, and was anchored as 0 indicating no pain and 100 indicating maximum pain.

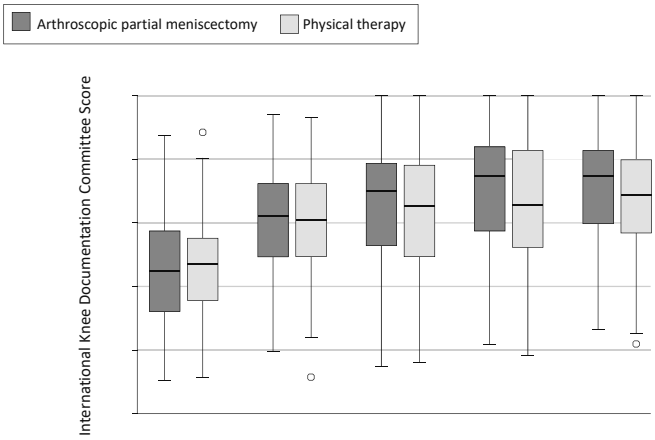
Abbreviations: APM, arthroscopic partial meniscectomy; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); IKDC, International Knee Documentation Committee; IQR, interquartile range; ISAKOS, International Society of Arthroscopy, Knee Surgery, and Orthopaedic Sports Medicine; KL, Kellgren-Lawrence; MRI, magnetic resonance imaging; OA, osteoarthritis; PT, physical therapy; VAS, visual analogue scale; y, year.

pain during weight-bearing improved from 59.3 mm at baseline to 25.5 mm at 24 months (MD, 32.5 mm [95% CI, 26.7 to 38.3]) (Table 3 and Appendix 12). The mixed-model analysis of the overall effects found a between-group difference of 5.9 mm (95% CI, 1.4 to 10.3; P=.01) in favour of APM (Table 3). The adjusted between-group differences are presented in Appendix 13.

General health in the APM group improved from 37.6 points at baseline to 51.1 points at 24 months (MD, 13.1 points [95% CI, 11.6 to 14.6]) and in the PT group, general health improved from 37.9 points at baseline to 48.7 points at 24 months (MD, 10.5 points [95% CI, 8.9 to 12.1]) (Appendix 14). The mixed-model analysis of the overall effects found an overall between-group difference of 1.3 points (95% CI, −0.2 to 2.7; P=.08) in favour of APM.

The activity level in the APM group improved from 2.6 points at baseline to 2.9 points at 24 months (MD, 0.34 points [95% CI, −0.00 to 0.69]) and in the PT group, the activity level improved from 2.5 points at baseline to 3.0 points at 24 months (MD, 0.38 points [95% CI, 0.08 to 0.68]) (Appendix 15). The mixed-model analysis of the overall effects found no significant between-group difference (0.04 points [95% CI, −0.3 to 0.2; P=.73]).

Osteoarthritis severity in the APM group progressed from 1.3 points at baseline to 1.6 points at 24 months (MD, 0.37 points [95% CI, 0.25 to 0.49]) and in the PT group, osteoarthritis severity progressed from 1.3 points at baseline to 1.5 points at 24 months (MD, 0.18 points [95% CI, 0.04 to 0.31]). The mixed-model analysis found no significant between-group difference (0.10 points more progression in the APM group [95% CI, −0.05 to 0.26; P=.18]).



Time	Baseline	3 mo	6 mo	12 mo	24 mo
No. of patients	158 161	155 158	151 146	143 136	141 148

Figure 2. knee function on the IKDC at different follow-ups. The figure represents the results of the primary outcome knee function on the IKDC (range, 0 to 100; from worse to best) for intention-to-treat analysis. The data represent actual patient data at each time. In each comparison, the box indicates the range between the 25th and 75th percentile, with the median indicated as a horizontal line within the box. The whiskers extend to the upper and lower adjacent values, the most extreme values that are within 1.5 × interquartile range (IQR) beyond the 26th and 75th percentiles. Circles indicate points beyond these values. The median IKDC data are in Appendix 11.

As-Treated Analysis

In the as-treated analysis, 150 participants were analysed in the APM group, 97 participants in the PT group, and 47 in the delayed APM group. All differences in knee function and pain during weight-bearing between the APM and PT groups were smaller in the as-treated analysis compared with the intention-to-treat analysis (Table 2 and Table 3).

Knee function improved in all 3 treatment groups (in the APM group from 43.9 points at baseline to 71.1 points at 24 months [MD, 26.8 points (95% CI, 23.6 to 29.9)]; in the PT group, from 48.6 points at baseline to 69.2 points at 24 months [MD, 20.2 points (95% CI, 16.5 to 23.8)]; and in the delayed APM group, from 40.8 points at baseline to 63.0 points at 24 months [MD, 21.5 points (95% CI, 15.8 to 27.3)]) (Table 2). The mixed-model analysis of the overall effects found a between-group difference of 1.7 points in favour of APM compared with PT (97.5% CI, −∞ to 5.1; p-value for non-inferiority <.001), indicating non-inferiority of PT compared with APM (Table 2). The between-group differences at 3 (−2.9 points [97.5% CI, −∞ to 1.3]) and 6 (1.6 points [97.5% CI, −∞ to 5.7]) months after randomisation also demonstrated non-inferiority; however, the effects at 12 (4.9 points

Table 2. Unadjusted Intervention Effects for the IKDC Score of Knee Function

	Mean IKDC Score						APM vs PT		Delayed APM vs PT	
Outcome	APM	N	PT	N	Delayed APM	N	Between-group difference[a] (97.5% CI)	p-value for non-inferiority[b]	Between-group difference[a] (97.5% CI)	p-value for non-inferiority[b]
Primary outcome										
Intention-to-treat										
Baseline	44.8	158	46.5	161	-	-	-	-	-	-
Overall[c]	66.5	590	64.2	588	-	-	3.6 (-∞ – 6.5)	.001	-	-
3 months	59.9	155	60.0	158	-	-	0.78 (-∞ – 4.3)	<.001	-	-
6 months	64.7	151	63.2	146	-	-	3.4 (-∞ – 7.0)	.006	-	-
12 months	70.7	143	66.4	136	-	-	5.7 (-∞ – 9.4)	.11	-	-
24 months	71.5	141	67.7	148	-	-	4.8 (-∞ – 8.5)	.04	-	-
Secondary outcome										
As-treated[d]										
Baseline	43.9	150	48.6	97	40.8	47	-	-	-	-
Overall[c]	66.2	561	66.6	372	57.6	163	1.7 (-∞ – 5.1)	<.001	-5.8 (-∞, -1.2)	<.001
3 months	59.3	147	64.3	97	49.8	47	-2.9 (-∞ – 1.3)	<.001	-10.8 (-∞, -5.2)	<.001
6 months	64.5	144	65.3	93	55.7	40	1.6 (-∞ – 5.7)	.001	-6.9 (-∞, -1.0)	<.001
12 months	70.5	137	67.5	90	63.8	35	4.9 (-∞ – 9.1)	.07	-0.78 (-∞ – 5.3)	.002
24 months	71.1	133	69.2	92	63.0	41	4.1 (-∞ – 8.3)	.04	-2.1 (-∞ – 3.8)	<.001

[a] The between-group difference for different follow-up points and as overall effect corrected only for IKDC score at baseline. Positive values signify that patients did better with (delayed) APM.

[b] P-values for non-inferiority based on a 1-sample Z test with respect to the non-inferiority margin. One-sided α of .025.

[c] Overall estimate over 24 months refers to the overall IKDC score within each group and between groups including all time points. The overall effect is based on a model including intervention group and baseline IKDC score. The No. represents the number of observations entered into the model.

[d] In the as-treated model, participants were analysed who adhered to their randomised treatment in 3 groups: (1) participants allocated to the APM group who received APM, (2) participants allocated to the PT group who completed the PT protocol without having APM during the follow-up period (e.g., ≥16 PT sessions; <16 sessions was considered a protocol violation), and (3) participants randomised to the PT group who had APM during follow-up (delayed APM group). Patients who were randomised to APM but did not have surgery and patients who were randomised to PT but did not complete the PT protocol and did not have delayed APM were removed from the as-treated analysis.

Abbreviations: APM, arthroscopic partial meniscectomy; IKDC, International Knee Documentation Committee (range, 0 [most limitations] to 100 [no limitations]); PT, physical therapy.

Table 3. Unadjusted Intervention Effects for the VAS for Weight-Bearing Pain

	VAS weight-bearing (0 (no pain) to 100 (worst pain imaginable))						APM vs PT		Delayed APM vs PT	
Outcome	APM	N	PT	N	Delayed APM	N	Between-group difference[a] (95% CI)	p-value[b]	Between-group difference[a] (95% CI)	p-value[b]
Secondary outcome										
Intention-to-treat										
Baseline	61.1	145	59.3	150	-	-	-	-	-	-
Overall[c]	24.5	559	28.8	555	-	-	-5.9 (-10.3, -1.4)	.01	-	-
3 months	30.4	154	33.4	151	-	-	-3.1 (-8.8–2.7)	.30	-	-
6 months	25.4	151	31.0	145	-	-	-8.2 (-14.1, -2.3)	.007	-	-
12 months	21.0	139	24.4	134	-	-	-5.3 (-11.4–0.73)	.08	-	-
24 months	19.6	115	25.5	125	-	-	-7.7 (-14.0, -1.3)	.02	-	-
As-treated[d]										
Baseline	62.5	137	56.0	91	66.4	43	-	-	-	-
Overall[c]	24.9	531	24.7	349	39.8	154	-1.7 (-6.8–3.3)	.50	13.8 (6.8–20.8)	<.001
3 months	30.8	146	27.5	91	48.0	46	2.4 (-4.3–9.1)	.49	18.8 (9.7–27.9)	<.001
6 months	26.0	144	28.0	92	40.8	40	-4.5 (-11.2–2.3)	.20	12.7 (3.2–22.2)	.009
12 months	21.0	133	21.0	89	31.4	34	-2.4 (-9.3–4.5)	.49	6.8 (-3.2–16.8)	.18
24 months	20.5	108	21.6	77	36.0	34	-3.6 (-10.9–3.7)	.33	13.4 (3.2–23.5)	.01

[a] The between-group difference for different follow-up points and as overall effect corrected only for VAS score at baseline. Negative values signify that patients did better with (delayed) APM.

[b] P-values with respect to zero (superiority testing). Standard p-values resulting from the mixed-model analysis in SPSS. One-sided α of .05.

[c] Overall estimate over 24 months refers to the overall VAS score within each group and between groups including all time points. The overall effect is based on a model including intervention group and baseline VAS score. The No. represents the number of observations entered into the model.

[d] In the as-treated model, participants were analysed who adhered to their randomised treatment in 3 groups: (1) participants randomised to the APM group who received APM, (2) participants randomised to the PT group who completed the PT protocol without having APM during the follow-up period (e.g., ≥16 PT sessions; <16 sessions was considered a protocol violation), and (3) participants randomised to the PT group who had APM during follow-up (delayed APM group). Patients who were randomised for APM but did not have surgery and patients who were randomised to PT but did not complete the PT protocol and did not have delayed APM were removed from the as-treated analysis.

Abbreviations: APM, arthroscopic partial meniscectomy; PT, physical therapy; VAS, visual analogue scale (range, 0 [no pain] to 100 [worst pain imaginable]).

[97.5% CI, −∞ to 9.1]) and 24 (4.1 points [97.5% CI, −∞ to 8.3]) months after randomisation did not demonstrate non-inferiority.

When comparing PT with delayed APM, all between-group differences in knee function favoured PT, demonstrating non-inferiority of PT compared with delayed APM overall and at all time points (P<.002). The adjusted between-group differences are presented in Appendix 9.

Knee pain during weight-bearing improved in all 3 groups (in the APM group, from 62.5 mm at baseline to 20.5 mm at 24 months [MD, 39.8 mm (95% CI, 34.1 to 45.5)]; in the PT group, from 56.0 mm at baseline to 21.6 mm at 24 months [MD, 33.6 mm (95% CI, 26.0 to 41.2)]; and in the delayed APM group, from 66.4 mm at baseline to 36.0 mm at 24 months [MD, 27.5 mm (95% CI, 16.0 to 39.1)]) (Table 3). The mixed-model analysis of the overall effects found no statistically significant between-group differences when comparing APM with PT (all P≥.20), and an overall between-group difference of 13.8 mm (95% CI, 6.8 to 20.8; P<.001) in favour of the PT group when comparing delayed APM with PT (Table 3). The adjusted between-group differences are presented in Appendix 13.

Adverse Events

Serious adverse events (e.g., cardiovascular, neurological or internal medicine conditions, venous thromboembolism, or repeat knee surgery) occurred in 9 participants in the APM group and 8 in the PT group. Non-serious adverse events (e.g., reactive arthritis, joint paint resulting in extra consultation or surgical site infection) occurred in 9 participants in the APM group and 4 in the PT group. All adverse events are reported in Appendix 16.

Post Hoc Exploratory Analysis

Post hoc exploratory analyses evaluated effect modification by the predefined potential confounders on the primary outcome, the IKDC score for knee function, and identified 2 statistically significant effect modifiers. First, there was interaction between baseline pain during weight-bearing and the treatment effect (regression coefficient, 0.14 [95% CI, 0.01 to 0.27; p-value for interaction = .03]), indicating that the effect of APM was 0.14 points larger compared with PT on the IKDC score for each millimetre increase in baseline pain. Second, the effect of the intervention differed between BMI categories (P for interaction = .02 for obesity vs normal; P for interaction = .01 for obesity vs overweight). Specifically, obese participants in the APM group scored on average 10.7 IKDC points (95% CI, 4.7 to 16.8; P=.001) higher than obese participants in the PT group, while the difference between treatments was not statistically significant in the other groups (normal: 1.4 [95% CI, −3.4 to 6.2; P=.57] and overweight: 1.2 [95% CI, −3.1 to 5.4; P=.60]). There were no statistically significant interaction effects observed for location of the tear (P=.12), education level (P=.15), osteoarthritis severity (P=.74), mechanical complaints (P=.81), sex (P=.60), age (P=.53), and baseline IKDC score (P=.21) (Appendix 17).

Discussion

This multicentre RCT showed that, in patients older than 45 years old with knee pain and non-obstructive meniscal tears, PT was non-inferior to APM for knee function over a 24-month follow-up period. The results of this trial support the recommendations from the current guidelines that PT may be considered an appropriate alternative to APM as first-line therapy for patients with meniscal tears.[122,123]

Although non-inferiority was demonstrated for the overall between-group difference in patient-reported knee function, and for the 3- and 6-month follow-up time points, results did not demonstrate non-inferiority at the 12- and 24-month time points. Longer follow-up will provide more details on the effect of time on the between-group differences. To date, only 1 trial has reported 5-year outcomes in a similar population with meniscal tears and reported no statistically significant differences in knee function between APM combined with PT (n=45) and PT alone (n=47).[22]

The previously reported clinical trials included a combined 838 patients and each trial reported no statistically significant differences between groups for knee function. These trials, however, were designed to assess superiority and had smaller sample sizes or shorter follow-up. Pooling these trials resulted in small statistically significant benefits of APM at up to 6-month follow-up.[98,115] The current trial had a larger sample size and longer follow-up and demonstrated a small benefit of APM, consistent with the pooled results from earlier trials.[98,115] Because this benefit from APM as compared with PT was smaller than the predefined non-inferiority margin, and within the previously reported MCID in patients with different knee pathologies,[124] the results of this trial demonstrated non-inferiority of PT as compared with APM, and are consistent with current consensus that APM should not be the first treatment in middle-aged and older patients with meniscal tears.[122,123]

In this trial, 29% of participants from the PT group received delayed APM, demonstrating that not all patients initially treated with PT were satisfied with their results. The post hoc exploratory findings on effect modification could guide future research on the characteristics of individuals who may be less likely to respond to PT to improve their treatment options and functional outcome.

Limitations

This study has several limitations. First, participants were not blinded to group assignment. Second, a screening log was not kept of patients who were eligible but not randomised, which limits the ability to assess the generalizability of the results. Third, the predefined non-inferiority margin was a conservative estimate of potentially relevant differences, based on the smallest detectable change of 8.8 points.[92] The smallest detectable change quantifies the amount of change that can be reliably detected by a measurement instrument, while non-inferiority margins should be based on the maximum clinically

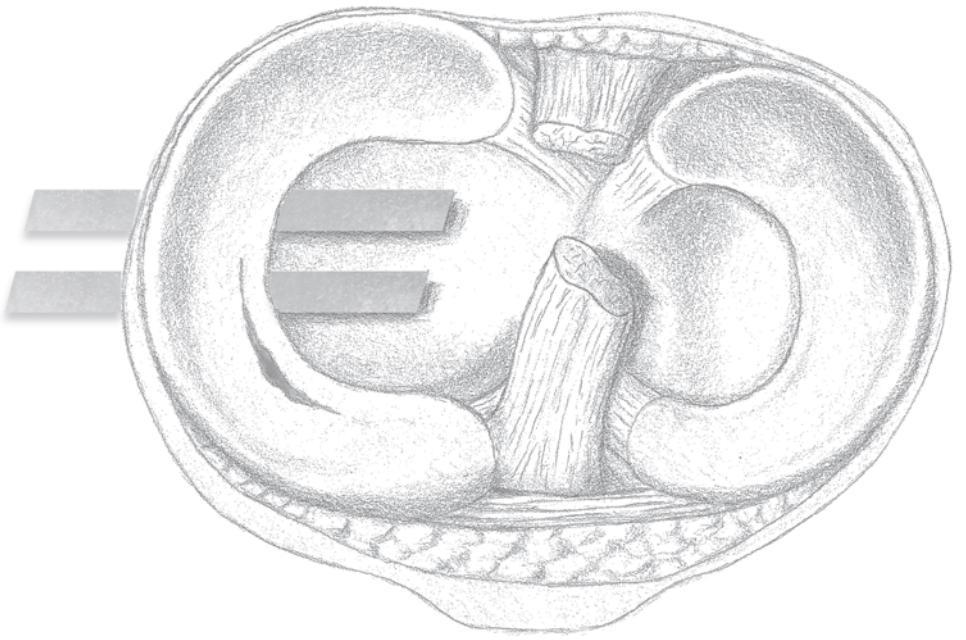
acceptable difference that a patient is willing to give up in exchange for the secondary benefits of the alternative therapy. While the exact threshold of clinical relevance is unknown in a population with meniscal tears, Irrgang and colleagues[124] reported an MCID for the IKDC of 11.5 points in patients with different types of knee disorders. If an MCID of 11.5 points was applied, PT would have been non-inferior to APM both overall and at all individual time points. Fourth, non-inferiority testing was intended for the secondary analyses, but no non-inferiority margins were specified in the protocol. Therefore, the comparisons between the groups for the secondary outcomes were treated as standard 2-sided superiority hypotheses. Fifth, MCIDs for the secondary outcomes were not defined until after data analyses (identified in different populations; 13.7 mm for pain on the VAS[125] and 2.0 points for RAND-36 PCS[126]). None of the observed between-group differences in these secondary outcomes exceeded MCID values, indicating that the clinical relevance of these findings is likely limited. Sixth, X-rays were interpreted by a single radiologist. Having 2 or more radiologists interpret X-rays may have resulted in more valid interpretations of osteoarthritis progression. Seventh, the combination of APM and PT may be more effective than APM alone. However, the authors followed the national guideline for generalizability of the study results and therefore APM was not always followed by PT.[31]

Conclusions

Among patients with non-obstructive meniscal tears, PT was non-inferior to APM for improving patient-reported knee function over a 24-month follow-up period. Based on these results, PT may be considered as an alternative to surgery for patients with non-obstructive meniscal tears.

Acknowledgements

We thank the study participants for their involvement as well as the participating centres for their help in this study. We thank Mirjam Schavemaker, MD (OLVG Amsterdam, the Netherlands), and Rogier van Dijk, MD (Meander Medical Centre, the Netherlands), for their post hoc reading of the magnetic resonance images and X-rays, respectively. We also thank Cees Verheyen, MD, PhD (Isala Clinics, the Netherlands), and Walter van der Weegen, PhD (St Anna Ziekenhuis, the Netherlands), for their work as independent committee members in the adjustment of the sample size. None of these individuals received any compensation for their involvement in this study.



6

How do the costs of physical therapy and arthroscopic partial meniscectomy compare? A trial-based economic evaluation of two treatments in patients with meniscal tears alongside the Escape study

Victor A van de Graaf
Johanna M van Dongen
Nienke W Willigenburg
Julia CA Noorduyn
Ise K Butter
Arthur de Gast
Daniël BF Saris
Maurits W van Tulder
Rudolf W Poolman
for the Escape Research Group

Br J Sports Med. 2019 Jun 21. pii: bjsports-2018-100065

Abstract

Objectives

To examine whether physical therapy (PT) is cost-effective compared with arthroscopic partial meniscectomy (APM) in patients with a non-obstructive meniscal tear, we performed a full trial-based economic evaluation from a societal perspective. In a secondary analysis - this paper - we examined whether PT is non-inferior to APM.

Methods

We recruited patients aged 45–70 years with a non-obstructive meniscal tear in 9 Dutch hospitals. Resource use was measured using web-based questionnaires. Measures of effectiveness included knee function using the International Knee Documentation Committee (IKDC) and quality-adjusted life-years (QALYs). Follow-up was 24 months. Uncertainty was assessed using bootstrapping techniques. The non-inferiority margins for societal costs, the IKDC and QALYs, were €670, 8 points and 0.057 points, respectively.

Results

We randomly assigned 321 patients to PT (n=162) or APM (n=159). PT was associated with significantly lower costs after 24 months compared with APM (–€1 803; 95% CI, –€3 008 to –€838). The probability of PT being cost-effective compared with APM was 1.00 at a willingness to pay of €0/unit of effect for the IKDC (knee function) and QALYs (quality of life) and decreased with increasing values of willingness to pay. The probability that PT is non-inferior to APM was 0.97 for all non-inferiority margins for the IKDC and 0.89 for QALYs.

Conclusions

The probability of PT being cost-effective compared with APM was relatively high at reasonable values of willingness to pay for the IKDC and QALYs. Also, PT had a relatively high probability of being non-inferior to APM for both outcomes. This warrants further de-implementation of APM in patients with non-obstructive meniscal tears.

Trial registration numbers

NCT01850719 and NTR3908.

Introduction

Each year, approximately 2 million arthroscopic knee surgeries are performed in the world, associated with $4 billion of direct medical costs.[127] Even though a clinically important benefit of surgery over conservative treatment has not been demonstrated,[115] the number of arthroscopic surgeries is decreasing slower than expected.[11]

Therefore, an economic evaluation, comparing conservative treatment with surgery could confirm the findings of prior research and support implementation of changes in clinical care. A recent model-based economic evaluation found that arthroscopic partial meniscectomy (APM) was not cost-effective in patients with or at risk for osteoarthritis compared with a group of matched controls receiving no treatment.[128] As no treatment at all is not a common alternative for surgical treatment in clinical practice, this model should be interpreted with caution. With treatment alternatives such as physical therapy (PT), pain medication or injections, the actual difference in costs compared with surgery is likely smaller.

To address this gap in the literature, we conducted an economic evaluation alongside a multicentre randomised controlled trial (RCT) comparing PT and APM in patients between 45 years and 70 years of age with a non-obstructive meniscal tear (i.e., no locking of the knee joint). In this study, we aimed to determine whether PT is cost-effective to APM, from a societal perspective, in patients with a non-obstructive meniscal tear. Since both PT and APM are considered standard and effective treatments, the multicentre RCT was set up as a non-inferiority trial. We performed a secondary analysis in which we explored whether PT (which is related to fewer side effects) is at least as cost-effective as APM (i.e., non-inferior).[129]

Methods

Participants and settings

We conducted an economic evaluation from a societal perspective alongside a multicentre RCT with a 2-year follow-up in which 321 participants (45–75 years) with an MRI-confirmed non-obstructive meniscal tear entered the trial between 3 July 2013 and 4 November 2015 in 9 Dutch hospitals.[130] We excluded patients with a locked knee, an anterior cruciate ligament rupture, severe osteoarthritis (Kellgren-Lawrence 4)[83] and a body mass index (BMI) >35 kg/m2.

The study was conducted in accordance with the Declaration of Helsinki. The board of directors of each of the participating hospitals approved the study. We registered the trial at clinicaltrials.gov and the Dutch Trial Register. We did not keep a log of patients who were screened for eligibility. Further details of the study are published elsewhere.[130,131]

Interventions

Physical therapy

After randomisation, we referred participants to one of the participating primary care PT clinics, and treatment was started within 2 weeks. The PT protocol was developed by a knee-specialised physical therapist and consisted of 16 sessions of 30 min each in 8 weeks (Appendix 5). Participating PT clinics were instructed about the protocol prior to the first study participant referral. Additionally, participants completed a home exercise programme (Appendix 5). Participants who were not satisfied with PT were allowed to receive delayed APM during follow-up.

Arthroscopic partial meniscectomy

APM was generally performed within 4 weeks after randomisation under general or spinal anaesthesia in an outpatient clinic. Standard anteromedial and anterolateral portals were introduced for inspection of the knee joint and partial removal of the affected meniscus until a stable and solid meniscus was reached. All participants received an information letter with perioperative instructions and the same home exercise programme as the PT group (Appendix 5). Participants were only referred for PT in case of swelling or signs of atrophy, as advised by the Dutch Orthopaedic Association Guidelines.[31]

Measures and outcomes

We collected effect and cost data using web-based questionnaires at baseline, 3, 6, 9, 12, 18 and 24 months.

Effect measures

The International Knee Documentation Committee (IKDC) 'Subjective Knee Form' was completed at baseline, 3, 6, 12 and 24 months. The IKDC is a validated and self-administered questionnaire designed for patients with a variety of knee disorders that assesses knee function, symptoms and ability to engage in sports activities,[77,92,100] with a range from 0 to 100, in which 100 indicated no limitations in daily or sporting activities.

The EuroQol 5-dimensional 5-level questionnaire (EQ-5D-5L) was used to measure health-related quality of life (HR-QoL).[105] The patients' health states were converted into utilities, anchored at 0.0 (death) and 1.0 (full health), using the Dutch EQ-5D-5L tariff.[132] Quality-adjusted life years (QALYs) were calculated by multiplying the utility of a patient's health state by the duration of time spent in that health state. Transitions between health states were linearly interpolated. Effects occurring after 12 months were discounted at a rate of 1.5%.[133]

Cost measures

Costs included intervention and other healthcare costs, paid help at home, informal care, work absenteeism and presenteeism and unpaid productivity costs.

For estimating *intervention costs*, we collected data on the participants' number of PT sessions using questionnaires and on the number and type of surgery from hospital records. For valuing the costs of PT, we used Dutch standard costs,[133] and for surgeries, we used the average costs from all hospitals in the Netherlands, derived from the Dutch Healthcare Authority.[134]

Other healthcare costs included costs related to the use of primary healthcare (e.g., general practitioner), secondary healthcare (e.g., hospital visits other than the initial APM) and prescribed and over-the-counter medication. For valuing these costs, we used Dutch standard costs, prices according to professional organisations and those of the Dutch Society of Pharmacy.[133]

Paid home care costs were assessed by asking participants to report the number of hours they received paid home care, which were valued using Dutch standard costs.[133]

For estimating *informal care costs*, we asked participants to report the total number of hours they received help from family, friends and other volunteers, which were valued using a Dutch recommended shadow price.[133]

For estimating *absenteeism and presenteeism costs*, we used the Productivity Cost Questionnaire.[135] We valued the patients' number of sickness absence days in accordance with the Friction Cost Approach (FCA; friction period=12 weeks) using gender-specific price weights.[133] For *presenteeism costs*, we asked participants to report the total number of days that they went to work while experiencing health complaints and their performance level on these days on a scale ranging from 0 (not able to do anything) to 10 (able to do everything). Subsequently, we calculated the total number of presenteeism days using the following formula:

Presenteeism days = ((10 – performance level)/10) * number of days with health complaints.

Presenteeism days were valued using gender-specific price weights.[133]

For estimating *unpaid productivity costs*, we asked participants to report the total number of hours they were unable to perform unpaid tasks (e.g., chores, volunteer work and educational activities), which were valued using a Dutch recommended shadow price.[133]

We converted all costs to Euros 2016 using consumer price indices and discounted costs occurring after 12 months at a rate of 4%.[133]

Other pre-specified outcomes included participant expectations and participant satisfaction. These outcomes will be analysed and reported separately.

Sample size, randomisation and blinding

Patients referred to one of the participating hospitals with symptomatic knee pain and suspected for a meniscal tear were informed about the study by the orthopaedic surgeon.

At the second outpatient visit, after written informed consent, we randomised eligible patients to either PT or APM using a central computer-generated randomisation scheme in a 1:1 ratio with random blocks (maximum block size of 6). We stratified for hospital and age (45–57 and 58–70 years). Participants, physicians and physical therapists were not blinded.

The sample size was based on an SD of 18 points on the IKDC, a power of 90%, a 2-sided α of .05 and a non-inferiority margin of 8 points on the IKDC. With an anticipated 20% loss to follow-up and a 25% delayed APM rate after 24 months, 160 participants per treatment group were needed.

Statistical analysis

We present all outcomes based on intention-to-treat principles. Missing data were multiply imputed, stratified by treatment group. Using Multiple Imputation by Chained Equations, we created 5 complete datasets (loss of efficiency <5%).[136] We analysed each dataset separately as specified below. Pooled estimates were calculated using Rubin's rules.[136]

We performed linear regression analyses to compare crude and adjusted aggregated and disaggregated costs between groups. To estimate total cost and effect differences, we performed seemingly unrelated regression (SUR) analyses in order to simultaneously correct for their possible correlation. We adjusted these total cost and effect differences for their baseline values, if available, level of osteoarthritis on the Kellgren-Lawrence scale,[83] mechanical complaints (IKDC question 6), the affected meniscus (medial, lateral or both), BMI (in 3 categories: <25, 25–30 or ≥30 kg/m2), age, gender and education level (high vs low).[137] Subsequently, we calculated 95% CIs surrounding all cost differences using Bias Corrected and Accelerated (BCA) bootstrapping (5 000 replications).

We calculated incremental cost-effectiveness ratios (ICERs) by dividing the adjusted difference in total costs by the adjusted difference in effects. Uncertainty surrounding the ICERs was estimated using BCA bootstrapping (5 000 replications) and graphically illustrated by plotting bootstrapped incremental cost-effect pairs (CE pairs) on cost-effectiveness planes (CE planes). We constructed Cost-Effectiveness Acceptability Curves (CEACs) indicating the probability of PT being cost-effective compared with APM for different values of willingness to pay. Data were analysed in STATA (v14) with a level of significance of $p<.05$. The unadjusted cost and effect differences and ICERs were calculated and presented in online supplementary appendix 18.

The deviations from the original trial protocol can be found in online supplementary Appendix 6.

Sensitivity analyses (SAs)

We performed 4 SAs to test the robustness of the results:

(1) only including participants with complete cost and effect data (SA1), (2) absenteeism costs estimated using the Human Capital Approach (SA2), (3) applying the healthcare perspective (SA3) and (4) an as-treated analysis in which we analysed 3 groups: (1) participants assigned to APM who received APM, (2) participants assigned to PT who completed the PT protocol (e.g., ≥16 PT sessions) and (3) participants assigned to PT but who received APM during follow-up (delayed APM group).

Secondary analysis: non-inferiority

We explored whether PT is non-inferior to APM according to the recommendations of Bosmans and colleagues.[129] For this, we defined a non-inferiority margin of 8 points for the IKDC, which is consistent with estimates of the smallest detectable change of this outcome.[92] For QALYs, a non-inferiority margin of 0.057 was chosen,[138,139] which is based on the assumption that a minimal clinically important difference in utility is sustained for 1 year (i.e., 1*0.057). As universally accepted non-inferiority margins for societal costs are currently lacking, we used the margin suggested by Bosmans and colleagues of €670 (i.e., €500 converted to Euros 2016).[129] Bosmans and colleagues based this margin on 2 visits to a primary healthcare provider, 1 outpatient visit and 3 days of absenteeism,[129] which we deemed appropriate for the condition under study as well. We estimated the proportion of CE pairs within these margins (i.e., the non-inferiority region) to explore the probability of PT being non-inferior to APM. As non-inferiority margins for total costs may vary greatly across countries, we constructed non-inferiority curves. These curves indicate the probability of PT being non-inferior to APM for various values of the non-inferiority margin for costs while the non-inferiority margin for effects is kept constant.[129] For PT being considered non-inferior to APM in terms of its cost-effectiveness, we assumed that the percentage of CE pairs in the non-inferiority region should be above 95% and the probability of non-inferiority above 0.95.

Patient involvement

No patients were involved in designing the study, nor were they involved in developing plans for recruitment, design or implementation of the study. No patients were asked to advise on interpretation or writing up of results.

Results

Participants

Between 3 July 2013 and 4 November 2015, we randomly assigned 321 patients to either APM (n=159) or PT (n=162) (Figure 1). The baseline characteristics can be found in Table 1. Participants with complete and incomplete data differed in terms of their

education level (highly educated; 55.9% vs 38.9%), smoking (yes; 12.4% vs 20.1%), the hospital of inclusion (recruited at OLVG; 43.4% vs 49.7%) and the level of pain on the VAS in rest (33.3 vs 42.1).

Clinical outcomes

Full details on the clinical outcomes, including the intervention effects per measurement point and over time, are described in a separate paper.[130] As for the economic evaluation (for which missing data were imputed), PT group patients' baseline and 24-month follow-up IKDC scores were 46.5 points and 62.6 points, respectively. For AMP group patients, these scores were 44.8 points and 64.6 points, respectively. During follow-up, PT group patients gained 1.65 QALYs and AMP group patients gained 1.68 QALYs. The corresponding adjusted effect differences were not statistically significant (IKDC, -4.0; 95% CI, -8.3 to 0.2; QALYs, -0.029; 95% CI, -0.074 to 0.016) (Table 2).

Costs

After 24 months, the mean intervention costs were statistically significantly lower in the PT group (€408) than in the APM group (€1 964) (€1 468; 95% CI, €1 347 to €1 680). Mean total societal costs were also statistically significantly lower in the PT group (€3 935) than in the APM group (€5 991) (€1 803; 95% CI, €838 to €3 008). The costs for paid help, absenteeism, informal care and unpaid productivity were lower in the PT group than in the APM group, whereas other healthcare and presenteeism costs were higher in the PT group than in the APM group. Of the disaggregate cost differences, only the differences in primary care, paid help and informal care costs were statistically significant (Table 3).

Cost-effectiveness

For the IKDC, we found an ICER of 449, indicating that 1 point decrease on the IKDC in the PT group as compared with the APM group was associated with a societal cost saving of €449 (i.e., PT was less costly and less effective) (Figure 2, Table 2). The CEAC indicated that the probability of PT being cost-effective compared with APM was 1.00 at a willingness to pay of €0/point improvement on the IKDC, decreasing to 0.07 at a willingness to pay of €2 500/point improvement (Appendix 19).

For QALYs, we found an ICER of 61 584, indicating that 1 QALY lost in the PT group as compared with the APM group was associated with a societal cost saving of €61 584 (i.e., PT was less costly and less effective) (Figure 2, Table 2). The CEAC indicated that the probability of PT being cost-effective compared with APM was 1.00, 0.99, and 0.40 at a willingness to pay of €0, €10 000 and €80 000/QALY, respectively (Appendix 19).

Sensitivity analyses

The overall conclusions of the present study would not change when only using data of patients with complete data (SA1), when using the HCA instead of the FCA for estimating absenteeism costs (SA2), and when applying the healthcare perspective instead of the societal perspective (SA3). When we excluded protocol violators and the group who received delayed APM from the PT group (SA4), the probability of PT being cost-effective compared with APM decreased much slower with increasing values of willingness to pay compared with the main analysis. For QALYs, for example, the probability of PT being cost-effective compared with APM was still 1.00 at a willingness to pay of €10 000/QALY, only decreasing to 0.98 at a willingness to pay of €80 000/QALY. Results of the group who received delayed APM were less favourable, with lower probabilities of cost-effectiveness for both the IKDC and QALYs (Table 2).

Secondary analysis: non-inferiority

We found the probability that PT is non-inferior to APM to be 0.97 for all non-inferiority margins for the IKDC and 0.89 for QALYs. SA2 and SA3 resulted in similar results. When we only included participants with complete data (SA1) non-inferiority of PT in comparison with APM was not demonstrated for both the IKDC and QALYs. As differences were observed between participants with complete and incomplete data, this was likely due to selective dropout of participants making the results of the main analysis more valid. In SA4, non-inferiority of PT in comparison with APM was demonstrated for both the IKDC and QALYs, whereas we found the group who received delayed APM to be inferior to APM for both the IKDC and QALYs (Table 2).

Discussion

In this first trial-based economic evaluation in patients with non-obstructive meniscal tears, the total societal costs of PT were statistically significantly lower to those of APM. The probability of PT being cost-effective compared with APM was 1.00 at a willingness to pay of €0/unit of effect for the IKDC (knee function) and QALYs (quality of life) and decreased with increasing values of willingness to pay. In a secondary analysis, the probability that PT is non-inferior to APM was 0.97 for all non-inferiority margins for the IKDC and 0.89 for QALYs. When we excluded patients who: (1) did not complete all 16 PT sessions and (2) received delayed APM, the probability of PT being cost-effective compared with APM decreased much slower with increasing values of willingness to pay compared with the main analysis and the probability that PT is non-inferior in comparison with APM was 0.99 for the IKDC and 1.00 for QALYs. The latter illustrates the need for further studies to focus on the characteristics of the non-responders to PT, that is, the patients who received delayed APM.

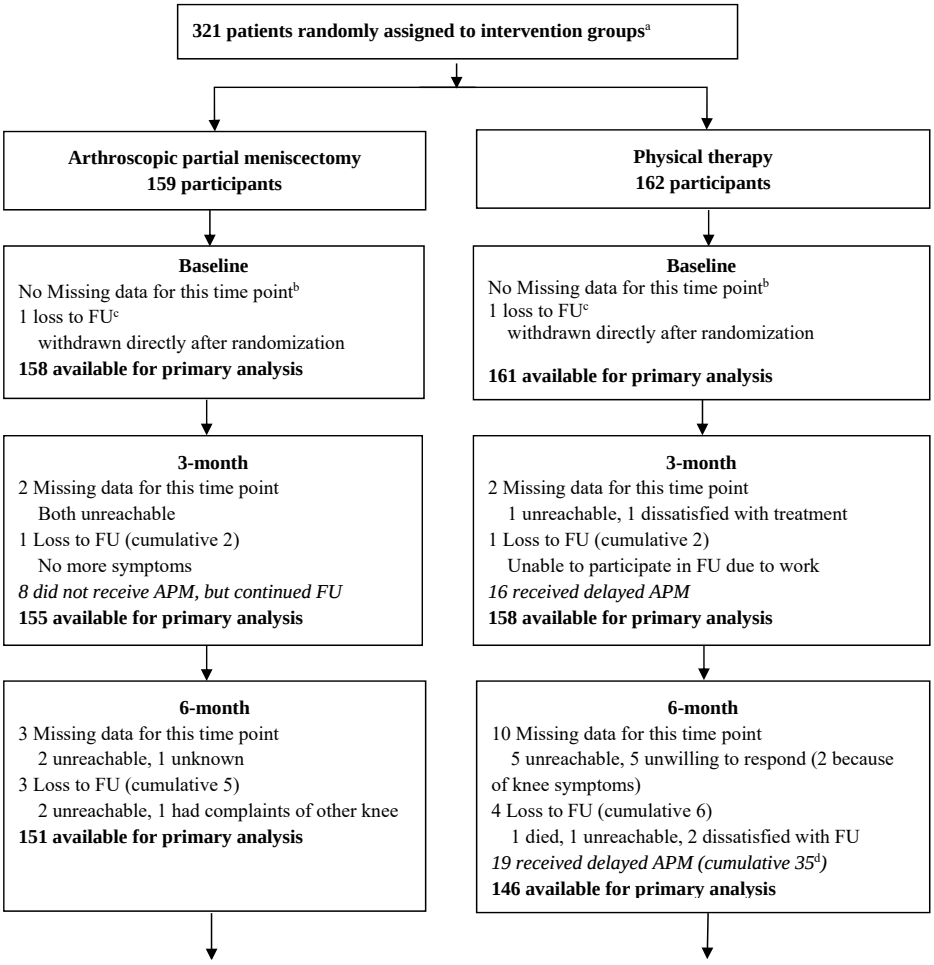


Figure 1. Flow of patients through the trial. [a] The number of patients screened for eligibility was not available. [b] Missing data refers to data that was missing at a specific time point, while patients remained available for the remaining follow-up moments. [c] Loss to follow-up refers to actual drop-out from the study; e.g. patients who did not participate at any of the remaining time points (cumulative numbers are total number of drop-outs). [d] cumulative number of delayed APM refers to total number of participants from the PT group that have received delayed APM from baseline until that follow-up Abbreviations: APM, arthroscopic partial meniscectomy; FU, follow-up.

Comparison with other studies

The literature on the economic aspects of APM for patients with meniscal tears is scarce. Although debate persists on the additional value of an economic evaluation in case of no difference in effectiveness, differences in costs could be missed if an economic analysis is not performed, nor can non-inferiority be investigated.[140] Our data will further assist clinicians and healthcare decision-makers in efficiently allocating already scarce

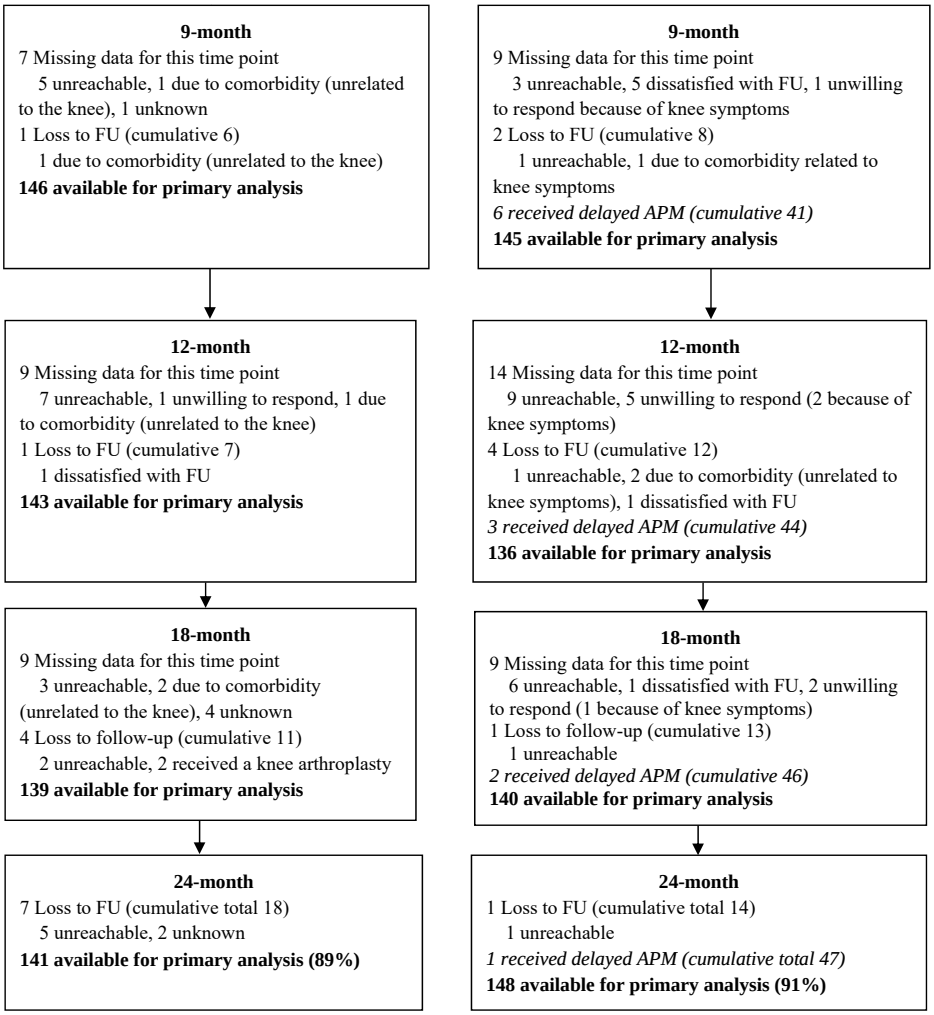


Figure 1. (continued)

healthcare resources[141] and will likely contribute to reducing healthcare costs.[122] Rongen and colleagues[128] reported the results of a model-based economic evaluation in which they compared APM with matched controls. APM was associated with a cost of €150 754 per QALY gained, which highly exceeds the generally accepted willingness to pay in the Netherlands (i.e., between €10 000 and €80 000 per QALY).[128] That study[128] has several limitations, as was illustrated previously.[142] First, since this model-based economic evaluation did not randomly assign patients to treatment groups, selection bias lures. Second, model-based economic evaluations involve making multiple assumptions and are less rigorous than trial-based economic evaluations in which individual patient data

Table 1. Baseline characteristics of the intention-to-treat population

	APM group	PT group
Demographics	N=158	N=161
Age, years	57.6±6.5	57.3±6.8
Women	80 (50.6)	81 (50.3)
Right knee	88 (55.7)	81 (50.3)
Education level, beyond high school	67 (42.4)	86 (53.4)
BMI (kg/m^2)	26.7±3.8	27.2±4.0
18.5<BMI<25	56 (35.4)	53 (32.9)
25≤BMI<30	72 (45.6)	67 (41.6)
30≤BMI<35	30 (19.0)	41 (25.5)
Mechanical complaints[a]	56 (35.4)	67 (41.6)
Imaging[b]		
Affected meniscus	N=158	N=161
Medial	126 (79.7)	136 (84.5)
Lateral	30 (19.0)	25 (15.5)
Both	2 (1.3)	0 (0)
Type of tear on MRI[120]	N=151	N=152
Longitudinal vertical	5 (3.3)	5 (3.3)
Horizontal	80 (53.0)	69 (45.4)
Complex degenerative	47 (31.1)	58 (38.1)
Radial	13 (8.6)	10 (6.6)
Vertical flap	2 (1.3)	5 (3.3)
Unclassifiable	1 (0.7)	5 (3.3)
Horizontal flap	3 (2.0)	0 (0)
OA level[c]	N=150	N=149
0 – None OA	18 (12.0)	15 (10.1)
1 – Doubtful	81 (54.0)	74 (49.7)
2 – Minimal	45 (30.0)	55 (36.9)
3 – Moderate	6 (4.0)	5 (3.3)
Knee function	N=158	N=161
IKDC score (0–100, worse to best)	44.8±16.6	46.5±14.6
EQ-5D-5L Index value	0.72±0.2	0.74±0.1
	N=146	N=158
EQ-5D-5L Quality of life scale	74.9±18.4	73.6±19.5

Data are n (%) or mean ± SD. [a] In contrast to locking of the knee joint, which was an exclusion criterion, mechanical complaints were allowed for inclusion. [b] Although inclusion was based on clinical readings by different radiologists and orthopaedic surgeons, 1 radiologist read all radiographs post hoc and 1 radiologist read all MRIs post hoc. Some of the radiographs (6.3%) and MRIs (5.0%) were unavailable to the viewing radiologist. [c] Kellgren-Lawrence grade 0 (no osteophytes or joint-space narrowing) indicates no osteoarthritis, grade 1 (questionable osteophytes) indicates early onset osteoarthritis, grade 2 (definite osteophytes, no joint-space narrowing) indicates mild osteoarthritis, grade 3 (50% joint-space narrowing) indicates moderate osteoarthritis, and grade 4 (>50% joint-space narrowing) indicates severe osteoarthritis.[83] Kellgren-Lawrence grade 4 was an exclusion criterion.

Table 1. (continued)

Abbreviations: APM, arthroscopic partial meniscectomy; BMI, body mass index; EQ-5D-5L, EuroQol 5-dimensional 5-level questionnaire; IKDC, International Knee Documentation Committee; ISAKOS, International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine; KL, Kellgren-Lawrence classification; MRI, magnetic resonance imaging; N, number; OA, osteoarthritis; PT, physical therapy.

are prospectively collected and few assumptions are made.[140,141] Third, the population in the control group was based on their probability of undergoing APM without being diagnosed with a meniscal tear and without receiving any treatment. This group does not adequately represent clinical practice in which conservative treatment (such as PT) is typically prescribed, which may increase the risk of bias since the patients in the intervention group are likely to have more complaints. Fourth, Rongen and colleagues determined the costs for APM (€4 407) based on their hospital records, whereas we determined these costs (€1 935) based on the average costs from all hospitals in the Netherlands. Finally, the authors used a superiority design compared with the non-inferiority design in the current study, which is preferred when surgical and non-surgical treatments are compared.[129]

During our 2-year follow-up, only 5 patients progressed to having a knee arthroplasty (3 in the PT group and 2 in the APM group). Therefore, our follow-up is insufficient to draw any conclusions on differences in the progression of OA between PT and APM. Rongen and colleagues estimated a 3-fold increase in the risk for future knee arthroplasty after APM.[143] Since the control group was not diagnosed with a meniscal tear and did not receive any treatment, this risk is likely to be overestimated.

The IKDC point estimate of the current trial-based economic evaluation slightly differs from that of the effect analyses[130] due to differences in the applied analytical methods. These different methods include: (1) multiple imputation, which is recommended for economic evaluations,[144] versus full maximum likelihood estimation, which is often used in longitudinal data analyses; (2) correcting for the possible correlation between costs and effects (e.g., by using SUR analyses), which is recommended in economic evaluations[137]; (3) discounting for cost and effect data, which is recommended in economic evaluations[145]; and (4) using longitudinal techniques in effect data, which is not applicable to cost data, since they require an estimate of the mean total cost difference during the entire follow-up, instead of an estimate of the mean cost difference per time period.

Strengths and limitations

The current study is the first trial-based economic evaluation in patients with meniscal tears. During 24 months, we prospectively collected cost and effect data with a response rate of 90% and performed a full economic evaluation from a societal perspective. The trial-based approach increases the generalisability of the results into clinical

Table 2. Adjusted differences in pooled mean costs and effects.

Analysis	Sample size PT	Sample size APM	Outcome	ΔC (95% CI) €	ΔE (95% CI) Points	ICER €/point	Distribution CE plane (%) NE[a]	SE[b]	SW[c]	NW[d]	non-inferiority region
Main analysis – imputed dataset	161	158	IKDC (range: 0–100)	-1 803 (-3 008, -838)	-4.0 (-8.3–0.2)	449	0.0	2.5	97.5	0.0	97.3
	161	158	QALYs (range: 0–1)	-1 803 (-3 008, -838)	-0.029 (-0.074–0.016)	61 584	0.0	9.8	90.1	0.0	89.0
sA1 – complete cases	89	81	IKDC (range: 0–100)	-976 (-1 365, -589)	-7.0 (-11.5–2.5)	139	0.0	0.1	99.9	0.0	66.9
	89	81	QALYs (range: 0–1)	-976 (-1 365, -589)	-0.024 (-0.077–0.030)	40 667	0.0	18.9	81.1	0.0	88.8
sA2 – human capital approach	161	158	IKDC (range: 0–100)	-1 866 (-3 129, -871)	-4.0 (-8.3–0.2)	465	0.0	2.5	97.4	0.0	97.3
	161	158	QALYs (range: 0–1)	-1 866 (-3 129, -871)	-0.029 (-0.074–0.016)	63 741	0.0	9.8	90.1	0.0	89.0
sA3 – healthcare perspective	161	158	IKDC (range: 0–100)	-1 120 (-1 767, -707)	-4.0 (-8.3–0.2)	279	0.0	2.5	97.4	0.0	97.5
	161	158	QALYs (range: 0–1)	-1 120 (-1 767, -707)	-0.029 (-0.074–0.016)	38 269	0.0	9.8	90.2	0.0	89.0
sA4 – as-treated analysis PT (≥16 sessions)	97	150	IKDC (range: 0–100)	-3 073 (-4 280, -2 150)	-3.0 (-7.5–1.4)	1 010	0.0	8.4	91.5	0.0	98.6
	97	150	QALYs (range: 0–1)	-3 073 (-4 280, -2 150)	0.021 (-0.024–0.065)	-148 866	0.0	81.6	18.4	0.0	99.9
Delayed APM	47	150	IKDC (range: 0–100)	525 (-1 312–2 272)	-6.0 (-13.1–1.1)	-88	1.8	2.0	25.4	70.8	43.4
	47	150	QALYs (range: 0–1)	525 (-1 312–2 272)	-0.108 (-0.185, -0.032)	-4 850	0.0	0.1	27.7	72.1	7.5

Differences in pooled mean costs and effects (95% CI), incremental cost-effectiveness ratios, distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes and percentage of bootstrapped cost-effectiveness pairs located in the non-inferiority region of the cost-effectiveness planes

[a] Refers to the NE of the CE plane, indicating that PT is more effective and more costly than APM.

[b] Refers to the SE of the CE plane, indicating that PT is more effective and less costly than APM.

[c] Refers to the SW of the CE plane, indicating that PT is less effective and less costly than APM.

[d] Refers to the NW of the CE plane, indicating that PT is less effective and more costly than APM.

Abbreviations: APM, arthroscopic partial meniscectomy; C, costs; CE plane, cost-effectiveness plane; E, effects; ICER, incremental cost-effectiveness ratio; IKDC, International Knee Documentation Committee; NE, northeast; NW, northwest; PT, physical therapy; QALYs, quality-adjusted life years; SA, sensitivity analysis; SE, southeast; SW, southwest.

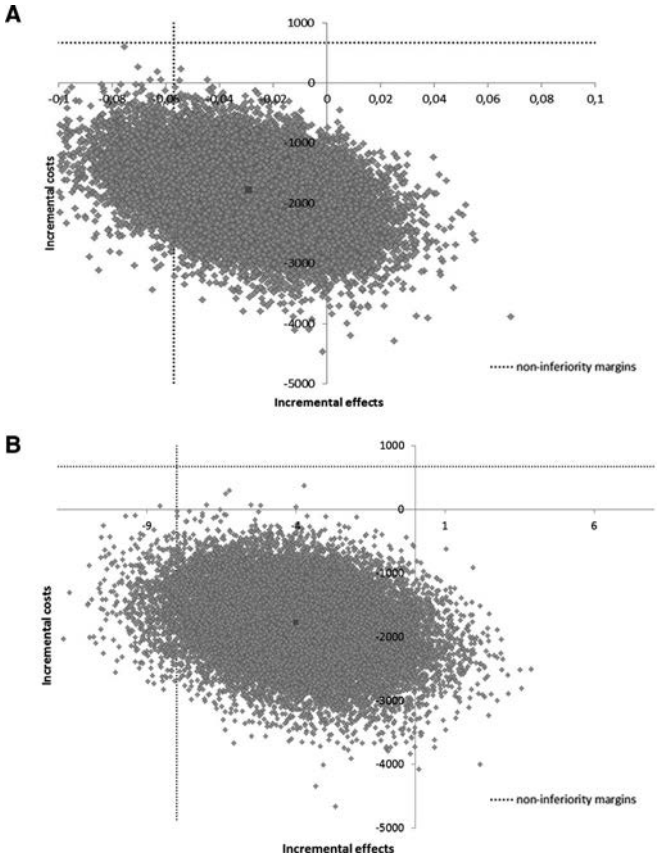


Figure 2. Cost-effectiveness planes.

Cost-effectiveness planes, including non-inferiority margins, for quality-adjusted life-years (A) and the IKDC (B).

Abbreviations: IKDC, International Knee Documentation Committee; QALY, quality-adjusted life years.

practice while simultaneously reducing the risk of selection bias and results in the most reliable estimates of costs and effects[146,147]; this is considered the most valid method for estimating the clinical and financial implications of a healthcare intervention.[140,141] The societal approach is recommended by the Dutch guidelines for costing research and is required by governmental funding agencies such as the Netherlands Organisation for Health Research and Development.[148] Second, we conducted our analyses using the SUR technique. The advantage of this technique is that it allows for the correction of a possible correlation between costs and effects.[137] Third, we had a relatively high rate of complete cases, that is, 91%, 81% and 71% for the IKDC, QALYs and costs data, respectively. We used Multiple Imputation by Chained Equations,[107] which is considered

Table 3. Mean cost in € per participant in the PT and APM group and mean cost differences between groups during the 2-year follow-up

Cost category	PT (n=161) mean (SEM)	APM (n=158) mean (SEM)	Cost difference crude, mean (95% CI)	Cost difference adjusted[a], mean (95% CI)
Intervention costs	488 (10)	1 964 (73)	-1 476 (-1 682, -1 370)	-1 468 (-1 680, -1 347)
Other healthcare costs	1 527 (145)	1 238 (205)	289 (-301–689)	347 (-276–726)
Primary care	407 (49)	734 (185)	-326 (-950, -81)	-309 (-954, -1 347)
Secondary care	1 114 (126)	499 (51)	615 (393–928)	655 (436–935)
Medication	6 (1)	5 (1)	1 (-2–4)	1 (-2–4)
Paid help costs	29 (12)	151 (60)	-122 (-333, -42)	-134 (-358, -49)
Informal care costs	290 (58)	573 (140)	-282 (-648, -62)	-216 (-489, -8)
Absenteeism costs	225 (48)	337 (51)	-112 (-238–12)	-83 (-200–35)
Presenteeism costs	424 (73)	328 (60)	96 (-77–265)	118 (-44–285)
Unpaid productivity costs	952 (169)	1 402 (218)	-449 (-988–49)	-369 (-845–79)
Total	3 935 (334)	5 991 (504)	-2 056 (-3 343, -1 002)	-1 803 (-3 008, -838)

[a] Adjusted for level of osteoarthritis on the Kellgren-Lawrence scale, mechanical complaints, the affected meniscus (medial, lateral or both), body mass index, age, gender and education level.
Abbreviations: APM, arthroscopic partial meniscectomy; PT, physical therapy; n, number of, 95% CI; Confidence Interval.

the most appropriate method for dealing with missing data in economic evaluations, since this accounts for uncertainties around the imputation of missing data by creating several imputed data sets.[136] Fourth, in this study, we included productivity-related costs due to reduced-on-the-job productivity, for example, presenteeism, which are often not collected in other economic evaluations.[149]

Some limitations warrant discussion. First, our study is vulnerable to performance bias due to the unblinded study design. However, we would expect this to result in an overestimation of the effect of APM as most patients would probably expect surgery to be more effective. Because of the small difference in effect, we believe that the risk for this bias is probably low. Second, we did not register the patients who were eligible but did not participate, leading to potentially reduced generalisability. Third, although cost and effect data were collected prospectively, this was done using self-report, which may have caused social desirable answers and/or recall bias. However, due to the randomisation, we do not expect this to systematically differ between treatment groups. Fourth, due to the follow-up of 24 months, conclusions on long-term effects of both groups, such as the numbers of knee arthroplasties, could not be drawn. Fifth, economic evaluation trials

often require large sample sizes. Since these numbers are not feasible in clinical trials, these trials risk being underpowered. Fifth, for the secondary analysis, non-inferiority margins of 8 points for the IKDC, 0.057 for QALYs and €670 for societal costs were used. These margins, however, are either based on narrative evidence or an established minimally clinically important difference, but it remains unclear whether they are appropriately justified in the context of trial-based economic evaluations. As such, the non-inferiority results should be interpreted in combination with the cost-effectiveness results only and further research into this topic is warranted.

Implications of this study

The probability of PT being cost-effective compared with APM was 1.00 at a willingness to pay of €0/unit of effect for the IKDC and QALYs and PT to be non-inferior to APM for the IKDC. Nonetheless, the probability of cost-effectiveness decreased with increasing values of willingness to pay for both outcomes and non-inferiority of PT to APM could not be unequivocally demonstrated for QALYs. It is therefore up to decision-makers whether they perceive the probability of PT being cost-effective compared with APM to be high enough at a reasonable value of willingness to pay and whether a probability of 0.89 is high enough to consider PT non-inferior as compared with APM for QALYs.

In the as-treated analysis, we removed the protocol violators and analysed those from the PT group who received delayed APM as a separate group. Then, cost-effectiveness results were more favourable than those of the main analysis and PT was non-inferior to APM for the IKDC and QALYs. The participants who received delayed APM were inferior to APM for both the IKDC and QALYs. Future research on the characteristics of these non-compliers to PT may help clinicians to recognise which patients are unlikely to benefit from a standardised PT programme.

The results of this trial-based economic evaluation support the results from previous RCTs[3,4,19,21,22,116,150] that all failed to demonstrate a clinically important benefit of APM, suggesting that APM should not be the first treatment choice in this population.

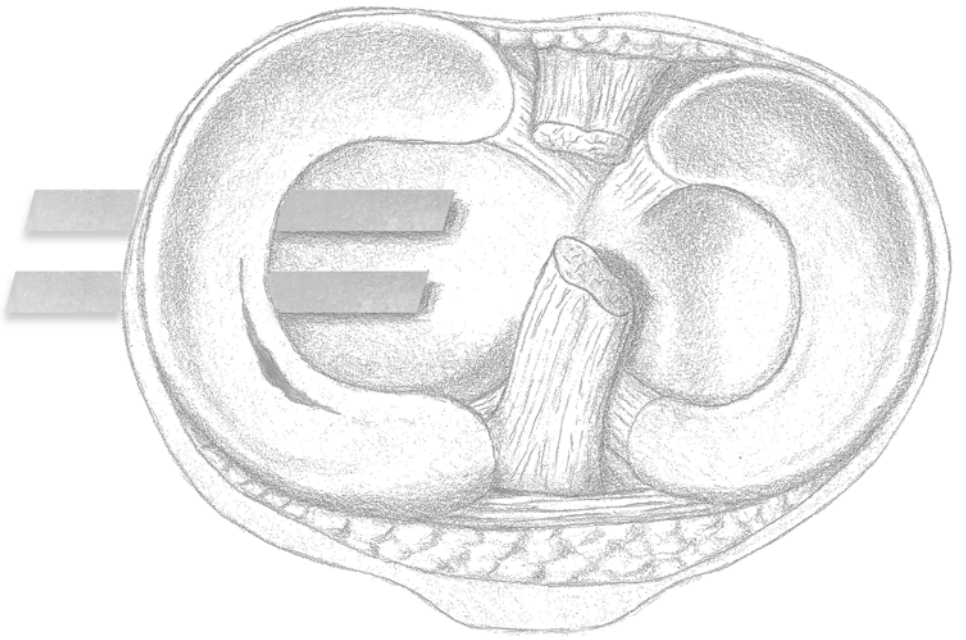
However, with the slower than expected decrease in the number of arthroscopies for meniscal tears,[20] studies identifying barriers to change practice for orthopaedic surgeons are important to further reduce the number of unnecessary arthroscopies.

Conclusion and Policy Implications

In this trial-based economic evaluation, the probability of PT being cost-effective compared with APM to be relatively high at reasonable values of willingness to pay for the IKDC and QALYs. Also, PT had a relatively high probability of being non-inferior to APM for both outcomes. These results support the results of previous RCTs and warrant further de-implementation of APM in patients with non-obstructive meniscal tears.

Acknowledgements

The authors would like to thank the study participants for their involvement in the study as well as the participating centres for their help in this study. The authors would like to thank Mirjam Schavemaker and Rogier van Dijk for their post hoc reading of the MRIs and X-rays, respectively. Finally, the authors would like to thank Cees Verheyen and Walter van der Weegen for their work as independent committee members in the adjustment of the sample size.



7

Responsiveness and minimal important change of the IKDC of middle-aged and older patients with a meniscal tear

Julia CA Noorduyn
Victor A van de Graaf
Lidwine B Mokkink
Nienke W Willigenburg
Rudolf W Poolman
for the Escape Research Group

Am J Sports Med. 2019 Feb;47(2):364-371

Abstract

Background

Responsiveness and the minimal important change (MIC) are important measurement properties to evaluate treatment effects and to interpret clinical trial results. The International Knee Documentation Committee (IKDC) Subjective Knee Form is a reliable and valid instrument for measuring patient-reported knee-specific symptoms, functioning, and sports activities in a population with meniscal tears. However, evidence on responsiveness is of limited methodological quality, and the MIC has not yet been established for patients with symptomatic meniscal tears.

Purpose

To evaluate the responsiveness and determine the MIC of the IKDC for patients with meniscal tears.

Study Design

Cohort study (design); Level of evidence 2.

Methods

This study was part of the Escape trial: a non-inferiority multicentre randomised controlled trial comparing arthroscopic partial meniscectomy with physical therapy. Patients aged 45 to 70 years who were treated for a meniscal tear by arthroscopic partial meniscectomy or physical therapy completed the IKDC and 3 other questionnaires (RAND 36-Item Health Survey, EuroQol-5D-5L, and visual analogue scales for pain) at baseline and 6-month follow-up. Responsiveness was evaluated by testing predefined hypotheses about the relation of the change in IKDC with regard to the change in the other self-reported outcomes. An external anchor question was used to distinguish patients reporting improvement versus no change in daily functioning. The MIC was determined by the optimal cut-off point in the receiver operating characteristic curve, which quantifies the IKDC score that best discriminated between patients with and without improvement in daily function.

Results

Data from all 298 patients who completed baseline and 6-month follow-up questionnaires were analysed. Responsiveness of the IKDC was confirmed in 7 of 10 predefined hypotheses about the change in IKDC score with regard to other patient-reported outcome measures. One hypothesis differed in the expected direction, while 2 hypotheses failed to meet the expected magnitude by 0.02 and 0.01 points. An MIC of 10.9 points was calculated for the IKDC of middle-aged and older patients with meniscal tears.

Conclusion

This study showed that the IKDC is responsive to change among patients aged 45 to 70 years with meniscal tears, with an MIC of 10.9 points. This strengthens the value of the IKDC in quantifying treatment effects in this population.

Introduction

Different patient-reported outcome measures (PROMs) have been developed and validated for patients with meniscal injuries. Many reflect the patients' perception of knee-specific symptoms, functioning, and sports activities, such as the KOOS (Knee injury and Osteoarthritis Outcome Score), the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), the Lysholm Knee Scoring Scale, the Western Ontario Meniscal Evaluation Tool, and the International Knee Documentation Committee (IKDC) Subjective Knee Form.[151] It is important to use high-quality PROMs to obtain outcomes that are trustworthy.[152] The quality of PROMs mainly depends on their reliability, validity, and responsiveness as described by the Consensus-Based Standards for the Selection of Health Measurements Instruments (COSMIN).[65] The IKDC, a 1-dimensional questionnaire as proven by confirmatory factor analysis, was previously shown to have the highest reliability and validity in measuring the functional outcome after treatment of meniscal injuries as compared with the KOOS and WOMAC.[100]

The responsiveness, however, of most PROMs, including the IKDC, is not well documented, and limited evidence is available in the specific population of patients with a meniscal tear.[151] The COSMIN initiative defines responsiveness as "the ability of a health-related PROM to detect change over time in the construct to be measured."[65] Adequate responsiveness of a PROM is important to properly assess intervention effects in clinical trials.

Aside from the responsiveness of a PROM, the interpretation of the changed score is essential in clinical practice. When changed scores are interpreted, the emphasis should be on the important change as perceived by the patient, represented by the minimal important change (MIC).[153] MIC is a measure that quantifies the smallest change score that patients perceive to be important.[153-155] It is relevant to know whether a measurement instrument is able to detect changes as small as the MIC value. This depends on the reliability and measurement error, often quantified as the smallest detectable change (SDC). When the SDC exceeds the MIC, an instrument cannot detect the MIC at the individual level on the basis of single measurements; when the SDC is smaller than the MIC, an instrument may detect statistically significant changes that lack clinical relevance. To ensure that observed changes are both statistically significant and clinically relevant, the change values have to exceed both the SDC and the MIC.[156]

Devji and colleagues[155] acknowledged the importance of the MIC in the interpretation of a treatment effect. The MIC for the IKDC is not yet determined for patients with an isolated meniscal tear.[151] Knowledge of both the responsiveness and the MIC in this patient population is important for designing clinical trials and to discriminate between responders and non-responders with regard to the treatment. Unknown responsiveness and MIC severely hamper the interpretation of clinical trial results and might explain why the preferred choice of treatment for meniscal tears is still a topic of debate, despite

several randomised controlled trials, systematic reviews, and meta-analyses comparing arthroscopic partial meniscectomy with physical therapy.[3,116,130,131]

Because the IKDC has high reliability and validity for patients with a meniscal tear, this study focuses on the other main measurement property, responsiveness, and the measure of interpretability, the MIC.[100] Specifically, we evaluated the responsiveness and MIC of the IKDC among middle-aged and older patients with meniscal tears.

Methods

Population

This study was part of the Escape trial, a non-inferiority multicentre randomised controlled trial comparing arthroscopic partial meniscectomy with a non-operatively treated control group receiving physical therapy.[130,131] Between July 2013 and October 2015, 321 patients between 45 and 70 years of age with a symptomatic, non-obstructive, degenerative meniscal tear (confirmed per magnetic resonance imaging) were included. Exclusion criteria consisted of severe osteoarthritis (Kellgren-Lawrence 4), body mass index >35 kg/m2, locking of the knee, prior knee surgery, and knee instability attributed to anterior or posterior cruciate ligament rupture. Previous knee injuries (e.g., anterior cruciate ligament rupture) that can interfere with the treatment outcome were assessed on magnetic resonance imaging and excluded from the trial. Further details can be found in the study protocol.[131] The Escape trial was approved by the Medical Ethical Committee (NL44188.100.13). All patients provided written informed consent for participation.

Treatment

Patients randomised to arthroscopic partial meniscectomy underwent surgery within 4 weeks after enrolment. The arthroscopic partial meniscectomy procedure started with a general assessment of the joint, whereupon the affected meniscus was partially removed, resulting in a stable and solid meniscus. Patients received standard written postoperative instructions. Participants were referred to physical therapy after arthroscopic partial meniscectomy if rehabilitation was not going according to the guideline of the Dutch Orthopaedic Association.[31]

Physical therapy started 1 to 2 weeks after randomisation. Patients in the physical therapy group participated in a supervised progressive exercise program consisting of 16 sessions of 30 minutes each (Appendix 5).[131]

Data Collection

Patients received self-administered questionnaires at baseline and 6 months after enrolment. Patients completed the questionnaires at home, either online or on paper. In the online questionnaires, no data were missing, as completion of each item was required to move on to the next item. When an item was missing in the paper-based

questionnaires, the missing item was obtained by telephone. To enhance the response rate, up to 3 response reminders were sent to the patients. Details on patient inclusions, randomisation, and follow-up are available in Appendix 20.

Outcome Measures

Four PROMs that were evaluated were all translated and validated for the Dutch population.[78,100,101,157] Sociodemographic information (age, sex, and body mass index) were collected at baseline. At follow-up, the same PROMs were administered, and an anchor question was added about the patients' assessment of change of functioning in daily activities.

The IKDC was developed to measure knee-specific symptoms, function, and sports activity for patients with ligament or meniscal injuries.[77] The IKDC consists of 19 items, of which 18 are converted into a total score. The answer to question 10a is not used for the overall score. Factor analysis confirmed the single dimension in a similar population.[100] The sum of these 18 items is converted into an IKDC score, ranging from 0 to 100 points. The minimum score of 0 points indicates that the patient is very limited in daily and sports activities, and the maximum score of 100 points indicates no restriction in functioning.[77] The IKDC was validated for patients with meniscal tears.[92,100]

The RAND 36-Item Health Survey (RAND-36) is a general health questionnaire that consists of 8 dimensions with a total of 36 questions.[101] From these 8 dimensions, 2 aggregated scores are calculated: the physical and mental component scores. These scores can be compared with the Dutch population with an average score of 50 points, in which higher scores represent better health. A study on its psychometric qualities concluded sufficient reliability and validity.[101]

The EuroQol–5 Dimension–5 Level (EQ-5D-5L) is a generic measure of health often used to assess the quality of life.[102] The questionnaire consists of 5 questions on mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Additionally, patients were asked to rate their general health on a visual analogue scale (EQ-5D-VAS) for a score between 0 and 100, with 0 indicating the worst possible health status as perceived by the patient and 100 indicating the best. The EQ-5D-VAS is responsive for patients undergoing knee arthroscopy[158] and was the only EQ-5D item that we used for further analysis.

Pain was assessed through 2 visual analogue scales of 100 mm. Patients were asked to rate their pain at rest and during weight-bearing activities in the previous week. The amount of pain was scored by marking on a line of 100 mm, with 0 indicating no pain and 100 indicating severe pain.

The external anchor question 'How did your function in daily activities change since the surgery/treatment of your knee?" was administered at 6 months after enrolment to determine the patient's perception of change in knee function after the treatment.[159]

The question was scored on a 7-point Likert scale, ranging from very much worsened to very much improved.

Responsiveness

Responsiveness of the IKDC was assessed with hypothesis testing based on the correlations of absolute changed scores, as recommended by the COSMIN panel.[82] Ten hypotheses were formulated (see Table 1): 5 before data collection (hypotheses 1,

Table 1. Hypotheses with expected and calculated correlations.

	Hypothesis	Expected r	Calculated r (95% CI)
1	The change in total IKDC score shows at least a very strong positive correlation with the change on the PCS of the RAND-36	(r≥0.7)	0.74 (0.67–0.81)
2	The change in the items for activity of the IKDC (questions 8 and 9) shows a very strong positive correlation with the change on the dimension for PCS of the RAND-36	(r ≥0.7)	0.70 (0.61–0.78)
3	The change in the items for activity of the IKDC (questions 8 and 9) shows a very strong positive correlation with the change on the dimension for physical function of the RAND-36	(r≥0.7)	0.72 (0.63–0.79)
4	The change in the items for pain of the IKDC (questions 1, 2 and 3) shows a very strong negative correlation with the change in VAS for pain during weight-bearing.	(r≤-0.7)	-0.68[a] (-0.76, -0.59)
5	The change in the items for pain of the IKDC (questions 1, 2 and 3) shows a moderate to strong positive correlation with the change on dimension for bodily pain of the RAND-36	(0.3≤r<0.7)	0.59 (0.51–0.69)
6	The change in VAS for pain at rest shows at least a moderate to strong negative correlation with the change in IKDC	(-0.3≥r>-0.7)	-0.55 (-0.60, -0.40)
7	The change in VAS for pain during weight-bearing shows a moderate to strong negative correlation with the change in IKDC	(-0.3≥r>-0.7)	-0.70[a] (-0.77, -0.60)
8	The change in EQ-VAS shows moderate to strong moderate positive correlation with change in IKDC.	(0.3≤r<0.7)	0.35 (0.21–0.43)
9	The change in total IKDC score shows a poor positive correlation with the change on the dimension for general health of the RAND-36	(r<0.3)	0.04 (-0.06–0.17)
10	The change in total IKDC score shows a poor positive correlation with the change on the MCS of the RAND-36	(r<0.3)	-0.11[a] (-0.123–0.11)

[a] Hypothesis was not confirmed Abbreviations: IKDC, International Knee Documentation Committee; VAS, Visual Analogue Scale; EQ-VAS, EuroQoL-Visual Analogue Scale; PCS physical component scale; MCS, mental component scale.

3, 5, 9, 10) and 5 after data collection but before data analysis (hypotheses 2, 4, 6-8). The expected correlations were predetermined per current literature, clinical experience, and consensus among the authors.

Correlations were categorized as very strong ($r \geq 0.7$), strong ($0.5 \geq r < 0.7$), moderate ($0.3 \geq r < 0.5$), and weak ($r < 0.3$). The hypotheses were tested with the Pearson correlation coefficient for normally distributed data and the Spearman rank correlation coefficient for non-normally distributed data. To demonstrate good responsiveness, 75% of the hypotheses should be confirmed.[65]

Minimal Important Change

The MIC was defined as the smallest change in outcome in the domain of interest as perceived beneficial by the patient.[154] The MIC value was established with an "anchor-based MIC distribution method", a blending of 2 methodologies: Specifically, an anchor-based method uses an external criterion to determine what patients consider important,[160] which is especially helpful in a study based on score distribution, given that distribution-based methods lack information on whether the observed changes are minimally important.[154]

First, we analysed the correlation between the changes in IKDC scores and the external anchor question. Next, if this correlation was >0.5, the study population was divided into changed and unchanged based on the external anchor question. The changed group comprised patients who reported to be very much, much, and slightly improved. The unchanged group included patients who reported to be unchanged. Patients who reported very much, much, or slight deterioration in daily functioning were excluded since we were comparing patients with and without important improvement.[154]

The receiver operating characteristic (ROC) curve was used because it searches for the optimal cut-off points, irrespective of how much misclassification occurs. A graphic display of the anchor-based MIC distribution was plotted, as well as the ROC curve.[160] Sensitivity and specificity were determined for all potential cut-off points. The MIC value was determined by the optimal cut-off point—that is, with the smallest value of the sum of the proportions of misclassifications: (1 – sensitivity) 1 (1–specificity).[160] In other words, the MIC was quantified by the IKDC score that best discriminated between patients with and without clinically relevant improvement.

Statistical Analyses

We used descriptive statistics to analyse the patients' demographics and tested all data for normality with the Kolmogorov-Smirnov test. The mean and SD were calculated for continuous normally distributed data ($P > .05$, Kolmogorov-Smirnov) and the median and interquartile range for continuous non-normally distributed data ($P < .05$, Kolmogorov-Smirnov). Frequencies and percentages were used for categorical data. We calculated

Table 2. Scores at baseline and 6-month follow-up and the changed scores[a]

PROM: Subscale	Baseline	6-mo Follow-up	Changed Scores	Percentage Changed Scores
IKDC total RAND-36	45.7 ± 15.1	66.7 (50.6–78.2)	19.5 (3.5–31.3)	44.6 (7.1–82.8)
PCS	37.7 ± 8.4	49.5 (41.8–54.2)	9.4 ± 9.6	25.8 (4.9–49.9)
MCS	52.9 (47.3–60.4)	55.3 (48.6–58.5)	-0.4 (-4.6–4.2)	-0.4 (-7.4–8.1)
PF	60 (45.0–75.0)	80.0 (60.0–90.0)	15.0 (0–30)	22.6 (0–70)
BP	42.9 (32.7–44.9)	77.6 (67.4–89.8)	32.7 (13.8–46.9)	77.3 (33.3–120)
GH	70.0 (60.0–80.0)	72.5 (65.0–85.0)	5 (-5–15)	6.5 (-6.7–25)
VAS for pain				
Rest	30.1 (15.8–56.1)	6 (0.0–24.1)	-18.9 (–36.9–21.9)	-82.0 (-100–217.5)
Weight-bearing	60.9 (42.0–78.1)	16.5 (4.6–51.4)	-30.2 ± 32.8	-61.9 (-90.2–217.4)
EQ-5D-VAS	78.1 (64.3–88.1)	82.6 (69.3–90.4)	3.1 (-7.6–11.6)	-3.8 (-8.9–15.7)

[a] Data are reported as median (interquartile range). For normally distributed data, values are reported as mean ± SD.

Abbreviations: BP, bodily pain; EQ-5D-VAS, EuroQol–5 Dimension–visual analogue scale; GH, general health; IKDC, International Knee Documentation Committee; MCS, Mental Component Scale; PCS Physical Component Scale; PF, physical functioning; PROM, patient-reported outcome measure; RAND-36, 36-Item Health Survey; VAS, visual analogue scale.

the changed scores by subtracting the baseline scores from the follow-up questionnaire scores. The percentage change scores are reported in Table 2, as it takes into account the scores at baseline. All analyses were performed with SPSS (v22; IBM Corporation).

Results

In total, 321 patients were randomised in the Escape trial; however, 2 patients (1 in each treatment group) withdrew immediately after randomisation. Of the remaining 319 patients, 298 (93.4%) returned the baseline and 6-month follow-up questionnaires. Baseline data of the 21 patients who did not complete the 6-month follow-up questionnaires were discarded. At baseline, the questionnaires (n=298) contained 0.4% missing items. At follow-up (n=298), 0.06% of the items were missing. Most patients (n=279; 94%) completed both questionnaires online. Fifteen patients completed both questionnaires on paper, and 4 patients completed the first questionnaire online and the second on paper. Patient characteristics are shown in Table 3, with the mean and changed scores of the PROMs in Table 2.

Responsiveness

Of 10 hypotheses, 7 (70%) were confirmed. The hypothesised and calculated correlation coefficients with the 95% CIs are shown in Table 1. For 2 unconfirmed hypotheses (hypotheses 4 and 7), the correlation coefficients deviated only slightly (≤0.02) from

Table 3. Baseline characteristics

Characteristic	n (%) or Mean ± SD
Patients	298
Sex	
Male	148 (49.7)
Female	150 (50.3)
Age, y	57.5 ± 6.7
Body mass index, kg/m^2	26.9 ± 3.9
Treatment	
APM	151 (50.7)
PT	147 (49.3)
Affected knee	
Left	136 (45.6)
Right	162 (54.4)
MRI: Affected meniscus	
Medial	245 (82.3)
Lateral	52 (17.4)
Both	1 (0.3)
Radiograph: Kellgren-Lawrence, n	281
0 – no OA	29 (9.7)
1 – Doubtful	147 (49.3)
2 – Minimal	95 (31.9)
3 – Moderate	10 (3.4)
4 – Severe[a]	0 (0)

[a] Kellgren Lawrence grade 4 was an exclusion criterion.
Abbreviations: APM, arthroscopic partial meniscectomy; kg, kilogram; MRI, Magnetic Resonance Imaging; m, meter; n, number; OA, osteoarthritis; PT, physical therapy; SD, standard deviation; y, year.

the predetermined threshold. Only hypothesis 10 differed from the predetermined direction, with a poor negative correlation while a poor positive correlation was expected.

Minimal Important Change

A strong correlation was found between the changed IKDC scores and the external anchor question (r=0.64; P<.001). On the basis of the external anchor question, 217 patients (72.8%) reported to be changed and 48 (16.1%) unchanged. Patients who reported slight (n=21, 7%), much (n=7; 2.3%), or very much (n=3; 1%) deterioration were excluded from the MIC analysis. Figure 1 shows the ROC curve. The optimal cut-off point was set at a sensitivity value of 79.7% and a specificity of 72.9%, resulting in an MIC of 10.9 points on the IKDC (range, 0–100 points). The anchor-based MIC distribution is displayed in Figure 2.

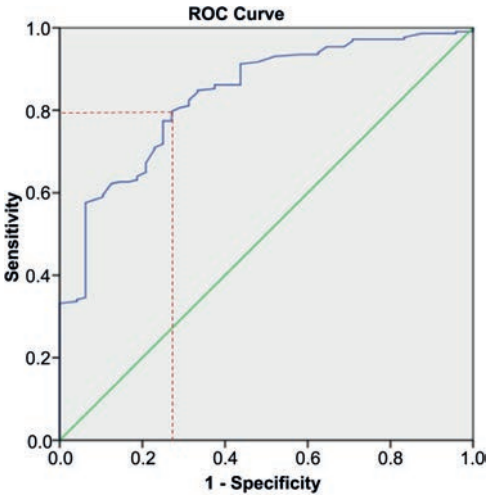


Figure 1. Receiver operating characteristic (ROC) curve, with optimal cut-off point.

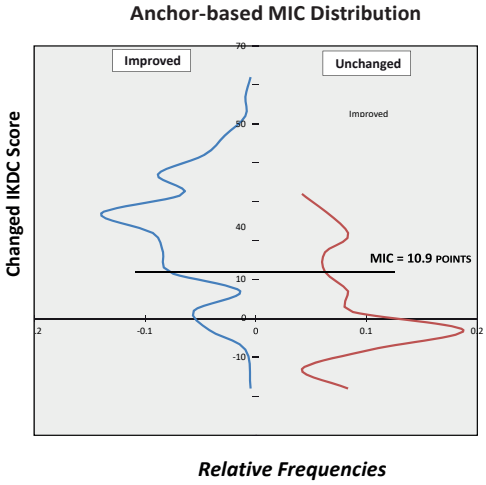


Figure 2. Anchor-based minimal important change (MIC) distribution. IKDC, International Knee Documentation Committee.

Discussion

Responsiveness of the IKDC among patients 45 to 70 years old with symptomatic meniscal tears was confirmed in 7 of the 10 predefined hypotheses. One unconfirmed hypothesis demonstrated a weak negative correlation while a weak positive correlation was expected—namely, between change in IKDC score and the Mental Component Scale of the RAND-36. Two unconfirmed hypotheses (4 and 7) deviated only slightly

in magnitude from the expected correlation. Therefore, we concluded that the IKDC was responsive in our population. Furthermore, we calculated an MIC of 10.9 points, reflecting the minimal change in IKDC score that a patient considers important. This value contributes to the interpretation of change scores as a result of the treatment of patients with meniscal tears.

Comparison with the Literature

Irrgang and colleagues[124] established the MIC for the IKDC at 11.5 points and 20.5 points in a study population with various knee injuries, using the point on the ROC curve closest to the upper left corner. These values are both higher as compared with the MIC in our study. However, we determined the MIC as the optimal cut-off point, using the smallest value of the sum of the proportions of misclassifications. Furthermore, we found that the MIC exceeded the SDC of 8.8 points that was reported by Crawford and colleagues.[92] Based on this SDC, there is 98% certainty that a change of 10.9 points was not due to measurement error.[161]

Responsiveness of the IKDC was previously reported by 2 studies. Crawford and colleagues[92] analysed responsiveness among 100 patients with meniscal injuries, and Irrgang and colleagues[124] analysed the responsiveness of 207 patients with a variety of knee disorders. Both studies concluded adequate responsiveness, using the effect size without predefined hypothesis as a measure of responsiveness. This is considered a less suitable method, since it measures the magnitude of change rather than the quality of the measurement.[151,156]Our results confirm that the IKDC is responsive to change based on recommended methodology.[162]

Strengths and Limitations

To our knowledge, this is the first study that determined the responsiveness and MIC of the IKDC among patients 45 to 70 years old with symptomatic meniscal tears, using predefined hypotheses with the expected magnitude and direction of the correlations. While previous studies investigating responsiveness with hypotheses testing used a general cut-off criterion of 0.5 for the expected correlations,[163-165] we defined more specific criteria to enhance the quality of our hypotheses. Another strength is that we utilised a large sample (n=298) with >90% complete data. Third, with a relatively short interval (6 months), we are confident that patients could adequately recall any changes in physical functioning and that these changes were largely related to the treatment that they received. Fourth, we used the anchor-based MIC distribution for the calculation of the MIC to give more insight into the interpretation of the MIC.

There were also limitations to this study. First, the data were retrieved from a randomised controlled trial, which could have led to selection bias. Second, the anchor question was not a true reflection of the construct measured by the IKDC. The anchor

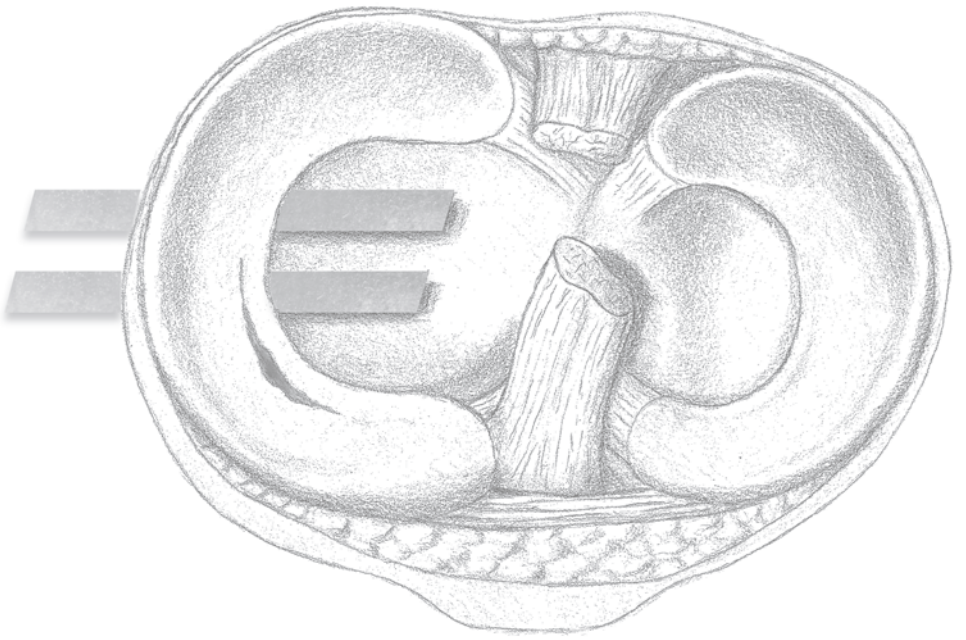
question focused on functioning in daily living, and the IKDC measures knee-specific symptoms, functioning, and activities. However, we found a strong correlation (r=0.64) between the anchor question and change in IKDC score. The results of our study apply specifically to patients 45 to 70 years old with degenerative meniscal tears and can be different for patients with traumatic meniscal tears or other knee pathologies.

Implications of the Study

The results of this study contribute to the evidence regarding the measurement properties of the IKDC among patients with meniscal tears; the IKDC is also responsive to change in this population and is valid and reliable. An MIC of 10.9 was established, which strengthened the value of the IKDC for assessing patient-reported knee function. The MIC of 10.9 points was determined on a group level. These results can therefore be used on a group level, whether by policymakers to determine treatment per recipient or by researchers to compare different treatments.[153-155] The distinctive character of the MIC between "changed" and "unchanged," on a group level, makes it highly relevant for developing clinical prediction models. Furthermore, based on the sensitivity and specificity levels (79.7% and 72.9%, respectively) and the probability of the measurement error (2%), the MIC of 10.9 can also be applied to individual patients.[153-155] However, one should take the patient's characteristics into account when applying the MIC on an individual level.[166]

Conclusion

The IKDC was responsive to change, with an MIC of 10.9 points for middle-aged and older patients with a meniscal tear. This study has shown that the IKDC has good measurement properties to evaluate the treatment effect on meniscal injuries. Therefore, we recommend the use of the IKDC for middle-aged and older patients with degenerative meniscal tears.

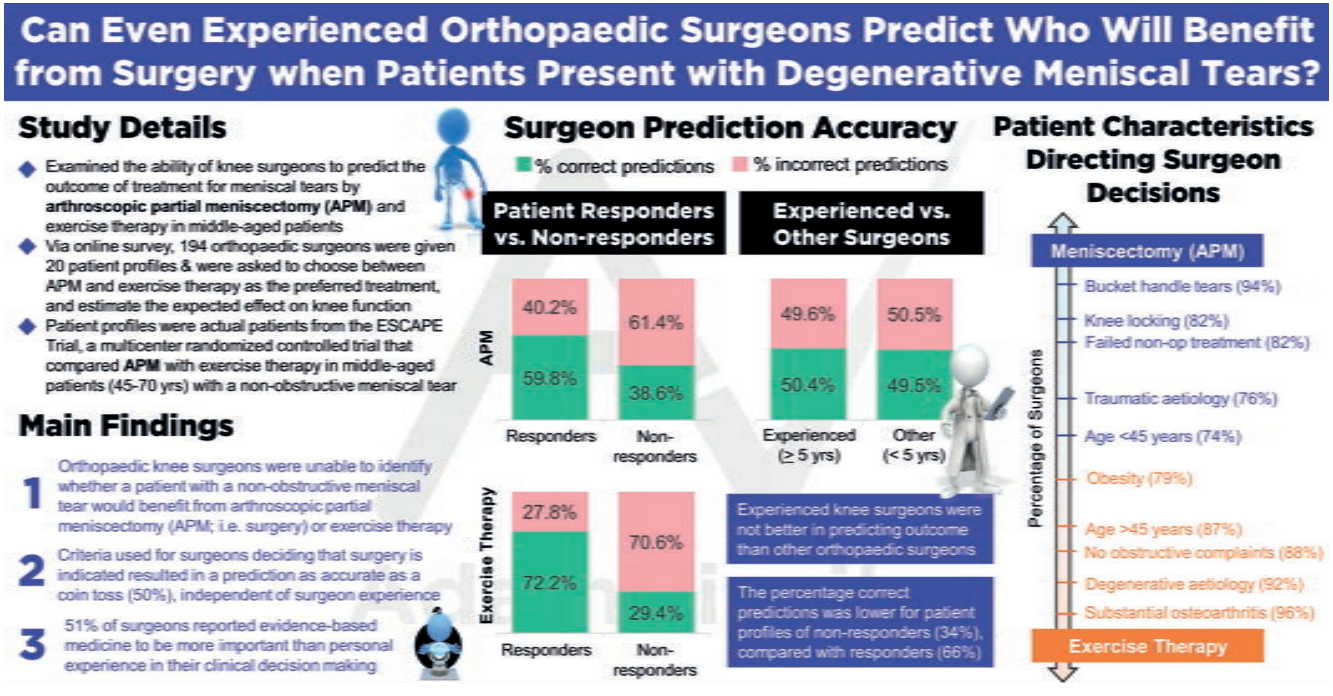


8

Can even experienced orthopaedic surgeons predict who will benefit from surgery when patients present with degenerative meniscal tears? A survey of 194 orthopaedic surgeons who made 3880 predictions

Victor A van de Graaf
Coen H Bloembergen
Nienke W Willigenburg
Julia CA Noorduyn
Daniël BF Saris
Ian A Harris
Rudolf W Poolman
for the Escape Research Group

Br J Sports Med. 2019 Aug 1. pii: bjsports-2019-100567



Visual abstract. Bloembergen CH, van de Graaf VA, Virgile A, Willigenburg NW, Noorduyn JCA, Saris DBF, Harris IA, Poolman RW. Infographic. Can even experienced orthopaedic surgeons predict who will benefit from surgery when patients present with degenerative meniscal tears? A survey of 194 orthopaedic surgeons who made 3880 predictions. Br J Sports Med. 2019 Oct 25. pii: bjsports-2019-101502.

Abstract

Objectives

To examine the ability of surgeons to predict the outcome of treatment for meniscal tears by arthroscopic partial meniscectomy (APM) and exercise therapy in middle-aged patients.

Design and setting

Electronic survey. Orthopaedic surgeon survey participants were presented 20 patient profiles. These profiles were derived from a randomised clinical trial comparing APM with exercise therapy in middle-aged patients with symptomatic non-obstructive meniscal tears. From each treatment group (APM and exercise therapy), we selected 5 patients with the best (responders) and 5 patients with the worst (non-responders) knee function after treatment. One thousand one hundred eleven orthopaedic surgeons and residents in the Netherlands and Australia were invited to participate in the survey. Interventions For each of the 20 patient profiles, surgeons (unaware of treatment allocation) had to choose between APM and exercise therapy as preferred treatment and subsequently had to estimate the expected change in knee function for both treatments on a 5-point Likert Scale. Finally, surgeons were asked which patient characteristics affected their treatment choice.

Main outcomes

The primary outcome was the surgeons' percentage correct predictions. We also compared this percentage between experienced knee surgeons and other orthopaedic surgeons, and between treatment responders and non-responders.

Results

We received 194 (17%) complete responses for all 20 patient profiles, resulting in 3880 predictions. Overall, 50.0% (95% CI, 39.6% to 60.4%) of the predictions were correct, which equals the proportion expected by chance. Experienced knee surgeons were not better in predicting outcome than other orthopaedic surgeons (50.4% vs 49.5%, respectively; p=.29). The percentage correct predictions was lower for patient profiles of non-responders (34%; 95% CI, 21.3% to 46.6%) compared with responders (66.0%; 95% CI, 57.0% to 75.0%; p=.01).

In general, bucket handle tears, knee locking and failed non-operative treatment directed the surgeons' choice towards APM, while a higher level of osteoarthritis, degenerative aetiology and the absence of locking complaints directed the surgeons' choice towards exercise therapy.

Conclusions

Surgeons' criteria for deciding that surgery was indicated did not pass statistical examination. This was true regardless of a surgeon's experience. These results suggest that non-surgical management is appropriate as first-line therapy in middle-aged patients with symptomatic non-obstructive meniscal tears.

Clinical trial registration

ClinicalTrials.gov Identifier: NCT03462134.

Introduction

The indication for arthroscopic partial meniscectomy (APM) is one of the most commonly made decisions in orthopaedic practice,[113] and 75% of APMs are performed in patients older than 40 years of age.[20] However, meniscal tears are common incidental findings in the general population. Incidental meniscal tears are found on MRI in 60% of asymptomatic adults older than 50 years with radiographic evidence of osteoarthritis.[17] Therefore, meniscal tears can be seen as part of a degenerative process of the knee.

Although several randomised controlled trials failed to demonstrate a clinically important benefit of APM over non-operative alternatives[3,21,22,40,41,116,130] or sham surgery[4] in middle-aged and older patients with symptomatic meniscal tears, these results have not led to a consistent decline in the number of APMs performed in daily practice.[20,167] Common arguments for performing APM include being a difficult habit to break, being influenced by personal experiences (observational evidence), criticism of the experimental evidence (e.g., low external validity) and a surgeon's belief in being capable to identify which patient may still benefit more from surgery.[127,168-172] Therefore, it is suggested to be up to the judgement of the treating surgeon to decide what is best for the individual patient.[171]

In this survey, we examined the ability of orthopaedic surgeons to predict the outcome in patients treated for meniscal tears. We also determined differences between surgeons with and without expertise in managing patients with knee pain, and how predictions differed between responders and non-responders to treatment. Finally, we evaluated which patient characteristics directed orthopaedic surgeons towards APM or non-surgical treatment.

Methods

Participants and setting

Between December 2017 and March 2018, an online survey was conducted among orthopaedic surgeons and orthopaedic surgery residents. The survey was sent to 1111 orthopaedic surgeons and residents active in the Netherlands (950 orthopaedic surgeons and residents) and Australia (161 orthopaedic surgeons). The Dutch participants were invited by the Dutch Orthopaedic Association (Nederlandse Orthopaedische Vereniging) and the Australian participants were invited by one of the authors (IH). The survey was constructed and distributed using Castor Electronic Data Capture 2019, Ciwit BV, Amsterdam, the Netherlands.[173] In December 2017, the first invitation was sent, and in January and February 2018 a maximum of 2 reminders were sent to all participants. Only 100% completed surveys were used for data analysis. We registered the study at clinicaltrials.gov (NCT03462134). Ethics approval was not required.

Patient profiles

Each participating surgeon was presented with 20 patient profiles. These profiles represented participants from the Escape trial,[174] a multicentre randomised controlled trial that compared APM with exercise therapy under the supervision of a physical therapist in middle-aged patients (45–70 years) with a non-obstructive meniscal tear.[130] A case description of each patient was presented in the survey, as shown in Figure 1. The profiles consisted of demographics, a description of symptoms, baseline knee function, baseline pain score, the results of physical examination, type of meniscal tear on MRI (on the Modified International Cartilage Repair Society classification)[120] and osteoarthritis level (Kellgren-Lawrence classification).[83]

Knee function was quantified as a score on the International Knee Documentation Committee (IKDC) Subjective Knee Form, which is a patient-reported knee function with a score ranging from 0 to 100 points, with a score of 0 corresponding to maximum knee symptoms and limitations in daily or sporting activities and a score of 100 reflecting no knee symptoms or limitations in daily or sporting activities.[77] Knee pain was quantified using a visual analogue scale, ranging from 0 mm to 100 mm, with a score of 0 mm corresponding to no pain and a score of 100 mm reflecting the worst possible pain.[119] Pain scores during rest and weight-bearing were both presented.

These selected patient profiles represented the top-5 and bottom-5 participants from the Escape trial with complete baseline data from each treatment group, discarding patients who were allocated to exercise therapy but received delayed APM. The top-5 were the patients with the most improvement on the IKDC Score after 24 months follow-up, therefore categorised as 'responders', with a mean improvement in IKDC Score of 64.6 (SD 4.6) points in the patients allocated to APM and 54.0 (SD 5.9) points in the patients allocated to exercise therapy. The bottom-5 per treatment group were the patients who deteriorated or had the least improvement on the IKDC Score after 24 months follow-up, therefore categorised as 'non-responders', with a mean deterioration in IKDC Score of -11.5 (SD 6.0) points in the patients allocated to APM and -13.1 (SD 6.6) points in the patients allocated to exercise therapy. All 20 patient profiles are presented in Appendix 21.

Survey

The participating surgeons, who were unaware of the treatment received, were asked to choose between APM and exercise therapy as the preferred treatment per profile. Subsequently, the surgeons had to estimate the expected effect on the patients' knee function after 2 years, twice: first for their preferred treatment and second for the other (non-preferred) treatment. The treatment effect on knee function was scored on a 5-point Likert Scale (strong deterioration, mild deterioration, no relevant change, mild improvement, strong improvement, see Figure 1).

Profile 1

Unemployed

50 y

BMI 23

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	1
VAS weight bearing (0-100)	7
IKDC score (0-100)	85

Physical exam	
Flexion	140°
Extension	+5°
Joint effusion	-
Joint line tenderness	-
McMurray	pain medial
Thessaly	-

Physical therapy in last 3 months
No

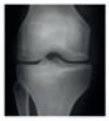
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Would you prefer meniscectomy or physical therapy as treatment in this particular patient?

- ☐ Arthroscopic partial meniscectomy (APM)
- ☐ Physical therapy (PT)

What would you think that will be the effect of your treatment of choice on knee function after two years?

- ☐ Strong deterioration (at least 20 points on IKDC)
- ☐ Mild deterioration (10-20 points on IKDC)
- ☐ No relevant difference (-10 to +10 points on IKDC)
- ☐ Some improvement (10-20 points on IKDC)
- ☐ Strong improvement (at least 20 points on IKDC)

What will the outcome be if the other treatment would be applied?

- ☐ Strong deterioration (at least 20 points on IKDC)
- ☐ Mild deterioration (10-20 points on IKDC)
- ☐ No relevant difference (-10 to +10 points on IKDC)
- ☐ Some improvement (10-20 points on IKDC)
- ☐ Strong improvement (at least 20 points on IKDC)

Figure 1. Example of a patient profile in the survey.

Each patient profile consisted of demographics, a description of symptoms, baseline knee function on the IKDC, baseline pain score on the VAS, the results of physical examination, type of meniscal tear on MRI, and osteoarthritis level.

The information in the figure above corresponds to patient profile 1. A clarification of the terms and clinical tests is presented in Supplement 22, 'Patient profiles 1–20; explanation of terms and abbreviations used'. For each profile, surgeons were asked to choose between meniscectomy and exercise therapy (in the survey referred to as physical therapy) as the preferred treatment. Subsequently, the surgeons had to estimate the expected effect on the patients' knee function after 2 years, twice: first for their preferred treatment and second for the other (non-preferred) treatment on a 5-point Likert Scale.

Abbreviations: BMI, body mass index; IKDC, International Knee Documentation Committee; MRI, magnetic resonance imaging; OA, osteoarthritis; VAS, visual analogue scale; y, years.

Furthermore, the surgeons were asked for their years of experience, field of expertise and opinion regarding the quality of the literature. In addition, they were presented a list of patients' characteristics, and were asked whether these typically affect their choice towards APM or exercise therapy, or do not affect their choice of preferred treatment. The complete survey content is attached in Supplement 22.

Outcome measures

The primary outcome was the percentage of correct predictions of treatment outcome. We assessed differences between surgeons with and without knee expertise, and how predictions differed between responders and non-responders to treatment. Secondary outcomes included the ratio of treatment choice between APM and exercise therapy, the surgeons' opinion towards the literature, and an overview of patient characteristics that direct the surgeons' choice towards APM or non-surgical treatment.

Data analysis

The predictions on the 5-point Likert Scale were dichotomised to discriminate between identified non-responders (Likert Scores 1, 2 and 3) and identified responders (Likert Scores 4 and 5) to treatment. The overall percentage of correct predictions (correct identification as either responder or non-responder) was first calculated over the profiles per surgeon and then averaged over the surgeons, with the 95% CI representing the reliability of the average estimate over all surgeons.

Surgeons were then divided into two groups based on expertise. The criterion for the group 'experienced knee surgeons' was a minimum of 5 years of experience in knee surgery. We used the $\chi 2$ test to compare the percentage of correct predictions between the surgeon groups (experienced knee surgeons vs other surgeons) and to compare the percentages of correct predictions in the responders and non-responders to treatment.

All other outcomes were analysed descriptively. Level of significance was set at .05. All analyses were performed using SPSS v22 (IBM Corporation, Armonk, New York, USA).

Patient and public involvement

No patients or the public were involved in designing the study, nor were they involved in developing plans for recruitment, design or implementation of the study. No patients were asked to advise on interpretation or writing up of results.

Results

Participant demographics

Of the 1111 invitations sent, we received 194 (17%) complete responses, 139 from the Netherlands and 55 from Australia. Of the participants, 163 (84%) were orthopaedic

surgeons, while 31 (16%) were residents in orthopaedic surgery. A total of 101 (52%) participants were experienced knee surgeons and 93 (48%) participants were residents, less experienced or had no knee expertise. An overview of the surgeons' characteristics is provided in Supplement 23.

Figure 2 presents the results of the predicted outcome per treatment group. Overall, 50.0% (95% CI, 39.6% to 60.4%) of all predictions were correct. This percentage was similar between experienced knee surgeons and the other surgeons, 50.4% (95% CI, 48.6% to 52.2%) vs 49.5% (95% CI, 48.0% to 51.1%), respectively (p=.58).

The percentage of correct predictions was 66.0% (95% CI, 57.0% to 75.0%) in the group of treatment responders vs 34.0% (95% CI, 21.3% to 46.6%) in the group of treatment non-responders (p<.001).

Table 1 presents an overview of the survey results for each patient profile. Overall, 21.6% of surgeons chose APM and 78.4% of surgeons chose exercise therapy as the preferred treatment. There was no difference in treatment preference between the level of experience, with 23.7% of experienced knee surgeons choosing APM as the preferred treatment compared with 19.5% of the other surgeons.

Fifty-one per cent of the surgeons reported evidence-based medicine to be more important than personal experience in their clinical decision making, and 77% considered themselves to be completely up to date with the literature for treatment of meniscal tears. The available evidence was convincing to 74% of the participants, and 76% felt confident in choosing between APM and exercise therapy. Seventy-seven per cent indicated that exercise therapy is a good option as initial treatment for non-obstructive meniscal tears, and 89% disagreed with APM being a good option as initial treatment. A complete overview of the results, as well as the distribution per expertise group, is presented in Supplement 24.

Patient characteristics that direct surgeons towards APM include bucket handle tears (94% of surgeons), knee locking (82%), failed non-operative treatment (82%), traumatic aetiology (76%) and age <45 years (74%), while characteristics that direct surgeons towards exercise therapy include moderate to severe osteoarthritis (96%), degenerative aetiology (92%), no obstructive complaints (88%), age >45 years (87%) and obesity (79%). Education level, gender and location of tear do not affect treatment choice. An overview of the results of all characteristics that were presented in the survey is shown in Supplement 25.

Discussion

The survey results indicate that orthopaedic surgeons are unable to identify whether a patient with a non-obstructive meniscal tear will benefit from APM or exercise therapy. The percentage of correct predictions was similar to prediction expected by chance alone, regardless of clinical expertise.

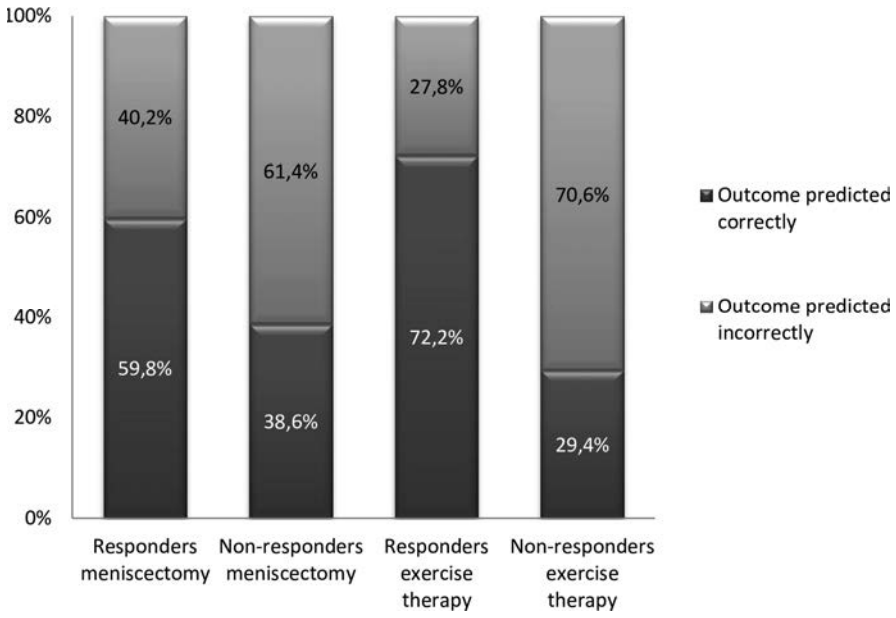


Figure 2. Predicted outcome per treatment group.
The figure above demonstrates the distribution between the correct and incorrect predictions for each of the categorised profiles (responders and non-responders to meniscectomy and exercise therapy). A correct prediction in the 'responders meniscectomy' and 'responders exercise therapy' corresponds to the options 'some improvement' and 'strong improvement', and a correct prediction in the 'non-responders meniscectomy' and 'non-responders exercise therapy' corresponds to the options 'strong deterioration', 'mild deterioration' and 'no relevant difference' from the 5-point Likert Scale.

Comparison with literature

The present survey is the first study that determines whether orthopaedic surgeons are able to predict treatment outcome in patients with meniscal tears based on patient profiles. In a recent systematic review evaluating clinicians' general expectations of any treatment, test or screening test, the authors reported that clinicians often have inaccurate expectations of treatment response.[175] With an underestimation of the harms and an overestimation of the benefits, the authors concluded that these inaccurate predictions are likely to result in suboptimal clinical management choices.[175] These results are comparable to the findings in this study, in which the surgeons' ability to predict the outcome was poorer in a group of treatment non-responders, that is, an overestimation of treatment response in this group.

Table 1. Results for responders/non-responders per treatment group

Treatment response	Patient profile[a]	Δ IKDC per case	Surgeon preferred treatment 'APM' (%)	Correctly identified as (non-)responder (%)
Responders to APM	18	70.1	20.1	62.9
	12	67.8	14.4	43.8
	15	64.4	33.0	69.1
	17	62.1	19.6	59.8
	6	58.6	21.1	63.4
Group average		64.6 (58.9–70.3)	21.6 (13.1–30.2)	59.8 (47.9–71.6)
Non-responders to APM	14	-6.9	39.7	22.7
	1	-8.05	7.2	43.8
	2	-9.2	25.3	26.3
	5	-11.5	52.6	23.7
	19	-21.8	1.0	76.3
Group average		-11.5 (-19.0, -4.0)	25.2 (-1.7–52.0)	38.6 (10.3–66.8)
Responders to exercise therapy	7	62.1	16.0	78.9
	10	56.3	47.4	49.5
	9	54.0	5.2	81.4
	13	51.7	5.2	76.8
	8	46.0	19.6	74.2
Group average		54.0 (46.7–61.4)	18.7 (-2.8–40.2)	72.2 (56.1–88.2)
Non-responders to exercise therapy	11	-8.1	7.2	21.6
	20	-8.1	7.2	18.6
	3	-12.6	43.8	38.1
	16	-12.6	37.6	44.8
	4	-24.1	9.8	23.7
Group average		-13.1 (-21.3, -4.9)	21.1 (-1.3–3.5)	29.4 (15.2–43.6)

Group average is expressed as percentage with CI.

[a]The patient profile numbers match the patient profile numbers in Supplement 21.

Abbreviations: APM, arthroscopic partial meniscectomy; CI, confidence interval; IKDC, International Knee Documentation Committee.

We have found only 1 study that determined whether orthopaedic surgeons are able to predict a treatment response.[176] In patients with sarcomas, orthopaedic oncologists were also incapable of accurately predicting the outcome of limb salvage surgery.[176]

Furthermore, it has previously been shown in a population with knee disorders that surgeons tend to be (over)optimistic with respect to treatment outcome.[177,178] This was supported by our findings, as two-thirds of the non-responders were expected to respond well.

Implications

In the present study, 89% of orthopaedic surgeon participants disagreed with APM being a good option for the initial treatment. However, APM was chosen as the preferred treatment in 22% of the cases. Interestingly, the percentage of respondents who recommended meniscectomy was highest (25.2%) for the descriptions of patients who did not benefit from surgery (non-responders to surgery).

This discrepancy—the greater propensity to recommend surgery for those patient descriptions that were associated with non-responders to treatment—and the poor ability of orthopaedic surgeons to predict who will respond well after surgery, suggests that surgeons should rely more on objective evidence from the literature when choosing treatment modalities in middle-aged patients with non-obstructive meniscal tears.

The participating surgeons were mainly focused on knee-specific characteristics that influenced treatment outcome. Among the most chosen variables that directed surgeons towards meniscectomy were obstructive complaints and traumatic aetiology. However, these convictions are not supported by the most recent literature. Obstructive complaints are associated with poor treatment response in general and meniscectomy in these patients has no added benefit over sham surgery.[58,179] There is no difference in improvement from meniscectomy between patients with a traumatic or a degenerative aetiology.[180] These misconceptions contribute to the large numbers of meniscectomies still performed.

Instead, considering the whole person in clinical decision making—by including characteristics such as education level, gender and activity level—may improve the surgeon's predictive ability. Psychological, mental health and socioeconomic variables are known to influence a person's health status and mobility in patients with other knee injuries.[181,182] Future research should focus on the effects of these variables and on finding other variables that influence treatment outcome in patients with meniscal tears.

Of the 22% of cases in which surgeons recommended meniscectomy as the preferred treatment, the mean expected change from meniscectomy was 4.3 points (on the 5-point Likert Scale), whereas the mean expected change from the non-preferred exercise therapy in these cases was 3.2 points (mean difference 1.1 points). This information provides insight into the criteria used by surgeons for deciding that surgery is indicated. The participating surgeons in this survey considered it sufficient to recommend meniscectomy if they expected no relevant change from exercise therapy (3.2 points on the 5-point Likert Scale) and a mild improvement from meniscectomy as compared with exercise therapy (1.1 points on the 5-point Likert Scale).

In a European survey, prior to the publication of the randomised controlled trials,[3,4,21,22,40,41,116,130] 75% of surgeons recommended APM as the first treatment in patients with knee osteoarthritis and meniscal tears.[57] In the present study, 22% of orthopaedic

surgeons chose APM as the preferred treatment in patients over 45 years old with a non-obstructive meniscal tear. These numbers demonstrate the willingness to change clinical practice from an initial surgical approach towards a conservative approach. However, the most recent data do not show a similar decrease in the number of APMs performed.[20,167] To further reduce the number of APMs, more effort is needed such as the implementation of administrative measures or (local) policy changes, which were earlier proven to be effective in reducing the number of knee arthroscopies in Norway and Australia.[183,184]

With 36% and 33% failures of initial exercise therapy (those who underwent meniscectomy during follow-up) reported in the literature after 1 year and 2 years, respectively,[3,130] especially this group of patients should be identified in the outpatient clinic. However, in the present study, the surgeons' ability to predict the outcome was poorest in the non-responders to treatment. According to the literature, there is currently insufficient evidence to allow prediction at an individual level in patients with meniscal tears. Future studies, such as prediction models and individual patient data meta-analyses, could help improving identification of treatment (non-)responders.

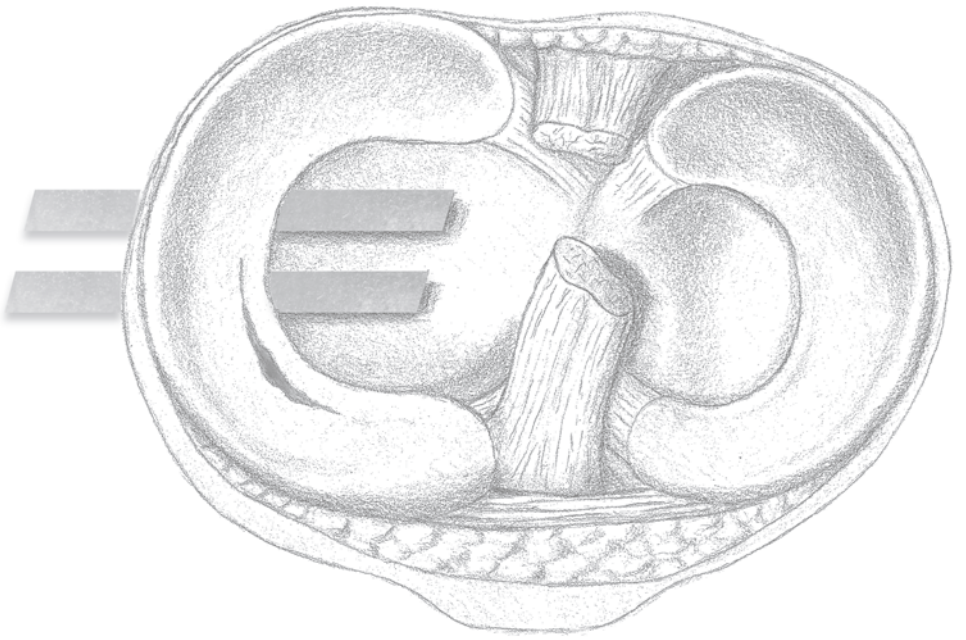
Strengths and limitations

To our knowledge, this is the first survey of orthopaedic surgeons to glean insight into their perception and expectations of treating patients with meniscal tears. By selecting the 5 patients with the best and worse outcomes per treatment group, we aimed for the most unarguable distinction between treatment responders and non-responders for the surgeons.

This study has several limitations. First, a digital survey was the only feasible way to obtain predictions of many surgeons for each of the 20 patient profiles. We made an effort to provide the most relevant information for clinical decision making, but acknowledge that this is still different from a real patient. Predictions in a real clinical setting might be more accurate. However, the majority of patients in a real clinical setting will have a more moderate treatment outcome (compared with the top-5 responders and bottom-5 non-responders included in this survey), which is likely harder to predict. Second, although we retrieved 194 responses, the response rate was low (17%) — this raises the potential for selection bias. Most surgeons indicated that they were all up to date with the literature and they were convinced by the evidence (Supplement 24). A higher response rate therefore might have led to a higher percentage of preferred APM and a worse prediction capability. Third, the poor prediction ability could also be due to unknown variables that determine the outcome after treatment of meniscal tears. The duration of symptoms, radiographic level of knee osteoarthritis and the amount of resected meniscus are known variables to be associated with the outcome following meniscectomy.[185] Only the level of osteoarthritis was provided to the surgeons. Knowledge of the other variables might have increased the accuracy of their prediction.

Conclusion

Surgeons' criteria used for deciding that surgery is indicated in a sample of patients with degenerative meniscal tears resulted in a prediction as accurate as a coin toss. This was true regardless of a surgeon's experience. This suggests that non-surgical management is appropriate as first-line therapy in these patients. We respectfully recommend that orthopaedic surgeons should rely more on the objective evidence from the literature when choosing treatment options.



9

Discussion

Key findings

Chapter 2. Pooling data of previous RCTs revealed that partial meniscectomy has only a short-term small benefit of unknown clinical importance over non-surgical treatment in patients with degenerative meniscal tears.

Chapter 3. The IKDC 'Subjective Knee Form' is a reliable and valid measurement instrument when evaluating treatment outcome in patients with a meniscal tear.

Chapter 4. We designed and initiated a multicentre randomised clinical trial and published the trial protocol to improve transparency for our healthcare and research community.

Chapter 5. In our randomised clinical trial, we found that supervised exercise therapy is non-inferior to partial meniscectomy for improving patient-reported knee function in patients with degenerative meniscal tears.

Chapter 6. Supervised exercise therapy is more cost-effective than arthroscopic partial meniscectomy in the treatment of degenerative meniscal tears.

Chapter 7. The IKDC 'Subjective Knee Form' is a responsive measurement instrument when evaluating treatment outcome in patients with a meniscal tear.

Chapter 8. Even experienced orthopaedic surgeons are unable to predict outcomes of partial meniscectomy and supervised exercise therapy.

Implications of this work

Change in practice

The work in this thesis will aid and direct surgeons in their clinical decision-making when treating patients with degenerative meniscal tears. It should also find its way to patients and into society in a wider sense to educate people on the limitations of partial meniscectomy and the benefit of initial conservative treatment. This will allow for a better discussion between patient and care provider.

We showed that supervised exercise therapy is both non-inferior to and more cost-effective than partial meniscectomy and should be considered as the primary choice in patients with degenerative meniscal tears (*Chapters 5 and 6*). Given the widespread use of partial meniscectomies, millions of ineffective procedures associated with high costs and potential harm for the patient can therefore be prevented.

With the ongoing rise of our healthcare costs, a shortage of resources arises. Therefore, we should aim to improve the quality of healthcare while simultaneously controlling the costs. Healthcare providers should select only the treatment with the optimal balance between treatment outcomes and costs according to the concept of value-based healthcare.

Osteoarthritis after meniscal surgery

We did not find a difference in the progression of osteoarthritis in the Escape trial over a 24-month follow-up (*Chapter 5*). However, we believe that this follow-up period is too short for any conclusions on this point.

Prediction of treatment outcome

With approximately 30% of patients who do not sufficiently respond to initial supervised exercise therapy (*Chapter 5*), surgeons are committed to select this subgroup in advance to avoid a long and ineffective treatment. How can we recognise these patients in the outpatient clinic? After careful consideration of a patient's history, physical examination and radiological findings, a surgeon estimates the expected outcome of different treatment options and discusses and selects the best option with the patient. We examined the ability of surgeons to predict the treatment response to partial meniscectomy and exercise therapy in patients with degenerative meniscal tears. Regrettably, we found that surgeons cannot predict treatment outcome in these patients based on their characteristics (*Chapter 8*).

Measuring treatment outcome

Due to the limited number of clinical trials available, one objective was to increase the body of evidence for healthcare efficiency (i.e. both effectivity and cost-

effectiveness). The previous trials used different measurement instruments for their primary outcome. Therefore, interpretation and comparison of study findings are difficult. Selection of a measurement instrument should be based on the quality of its measurement properties and only the best instrument is justified to use. That is why we compared the measurement properties of the Dutch-language versions of the IKDC, KOOS, and WOMAC. We found that the IKDC 'Subjective Knee Form' is a reliable, valid and responsive measurement instrument when evaluating treatment outcome in patients with a meniscal tear. This implies that the IKDC, rather than the KOOS or WOMAC, should be used to assess functional outcome in patients with meniscal tears.

In our trial protocol (*Chapter 4*), we defined the threshold for non-inferiority based on the smallest detectable difference (SDD) instead of the – unknown – minimal important change (MIC). Therefore, the findings both in our trial and economic evaluation were open to discussion. Besides the risk of being underpowered, our conclusions may have been different if the true MIC was smaller than the SDD. For example, if the true MIC was lower than 8, the results could have been inconclusive. For a better interpretation of the treatment effect and study findings, we calculated the minimal important change (MIC) in our study population. Data were available for 298 patients who completed baseline and 6-month follow-up questionnaires. We found an MIC of 10.9 points (*Chapter 7*). The fact that the MIC is higher than the smallest detectable change that we used in our power analysis (10.9 points vs 8 points) indicates that the sample size was sufficient. Furthermore, this strengthens the conclusions of non-inferiority of exercise therapy compared to partial meniscectomy.

Comparison with literature

Change in practice

Despite the growing number of studies indicating that surgery is not beneficial over an initial conservative approach, the decline in the number of meniscectomies is less than expected. One commonly heard reason for this is a surgeon's belief of being able to identify which patients will benefit from surgery.

The first RCT that compared the effectiveness of partial meniscectomy to supervised exercise therapy dates back to 2007.[19] Herrlin and colleagues found no difference in improvement of knee function or knee pain between both treatments. It was not until 6 years later that the next RCTs reported similar results.[3,4,21,22] The effect of these publications on the number of partial meniscectomies performed has been reported for many countries. Smaller than expected declines in the number of partial meniscectomies have been reported in Denmark, England, Finland, the Netherlands, Norway, Sweden, and the United States.[14,20,167,184,186,187] The rates of partial meniscectomies between 2012 and 2015 in Switzerland even remained unchanged.[188]

Osteoarthritis after meniscal surgery

Many of the partial meniscectomies are still performed in patients with knee osteoarthritis.[189] Especially in these knees, resecting part of the meniscus can accelerate the progression of osteoarthritis due to a change in leg alignment and an increase in joint load.[5,190] This is known to contribute to the increased risk of osteoarthritis after total meniscectomy,[15,191,192] but is also believed for partial meniscectomies.[5,7,193,194] However, there is debate whether or not these findings are clinically relevant.[50,192]

Several studies concluded that the rates of total knee replacement increased with 14% to 17% within 24 months after partial meniscectomy.[195,196] These data are not supported by the findings from the RCTs after 24months follow-up that reported knee replacement rates up to 3% after partial meniscectomy.[3,4,116,197] This could be explained by selection bias of participants in a trial, although this might not completely explain this difference. Only 1 RCT published the 5-year rates of knee replacement.[198] Katz and colleagues reported a 7.1% rate of knee replacements in 351 participants with no statistically significant difference between treatment groups according to the intention-to-treat analysis (9.2% in the meniscectomy group vs 5.1% in the exercise group, hazard ratio 2.0; 95% CI, 0.84–4.9).[198] However, in the as-treated analysis, the authors found a 5-times higher chance of receiving a total knee arthroplasty in surgically treated participants compared to those treated conservatively.[198] It therefore seems that partial meniscectomy does result in an increased need for knee replacement surgery, although participants who had a meniscectomy may have been more familiar with the process of surgery and therefore be more inclined to opt for TKR.[198] Also, in the study of Katz, participants who underwent TKR had a higher baseline pain score compared to the other participants. This may also have affected the higher TKR rates.

Prediction of treatment outcome

Certain characteristics directing surgeons towards a partial meniscectomy, such as mechanical complaints, appear to be a poor predictor of the outcome.[58,179] This may, in part, explain why surgeons have difficulty predicting the treatment response in these patients.

Not only surgeons have difficulty predicting the outcome for patients with degenerative meniscal tears. Also, a computer-based model, combining 18 characteristics, could not predict the change in knee function nor identify any subgroups that are more or less likely to benefit from meniscal surgery.[199]

Despite this, there are some suggestions that a complex meniscal tear, larger extrusion, medial cartilage injuries, a larger meniscal excision, and obesity are prognostic factors for worse knee function after partial meniscectomy.[67,200]

The foregoing does not provide useful information to help determine who will likely be a non-responder to conservative treatment. When looking more specifically at this

group of interest, 1 study found that these patients had a shorter duration of symptoms and higher baseline pain scores. The authors emphasised that rigorous supervised exercise therapy before partial meniscectomy did not compromise surgical outcome.[201] Conflicting evidence is reported for the location of the tear, gender, type of tear, and age.[67,202,203]

So for now, it remains difficult and unreliable to predict which patients will not sufficiently benefit from exercise therapy, although the duration of symptoms and the level of baseline scores might play a part.

Exercise therapy vs partial meniscectomy

The large increase in the number of meniscectomies between 2000 and 2010 was not supported by any scientific evidence. *Chapter 2* describes the results from our systematic literature search and meta-analysis of level-1 studies comparing surgery to non-surgical treatment in patients with degenerative meniscal tears, with knee function as the primary outcome. We included 6 RCTs with a total of 773 patients. Pooling the studies resulted in a small, statistically significant, benefit of partial meniscectomy for knee function and knee pain up to 6 months. However, the clinical relevance of these differences is unknown. We found no differences between the treatment groups for any of the outcomes at longer follow-up.

The body of evidence against the use of partial meniscectomy continues to grow. After the publication of our meta-analysis, several new RCTs published their results.[116,130,204,205] In 2019, Abram and colleagues published the results of a meta-analysis that included these new trials.[206] The authors compared partial meniscectomy to physical therapy and found small improvements from partial meniscectomy in knee pain and knee function at 6 months or longer, and a moderate improvement from partial meniscectomy in knee-specific quality of life up to 12 months after partial meniscectomy.[206] However, when the authors compared partial meniscectomy to sham or placebo surgery, the authors found no more differences between treatment groups for knee pain, knee function, or knee-specific quality of life.[206] Although these findings highlight the placebo effect of this surgical intervention, the additional benefits of meniscectomy over physical therapy may explain why orthopaedic surgeons are still finding it hard to convince patients not to have knee arthroscopy.

Strengths and limitations

Strengths

The work in this thesis aims to change the treatment of degenerative meniscal tears. The published evidence at the time of the start of this work was not only scarce but also only focused on clinical outcome measures. Therefore, we aimed to contribute to this lack of evidence by designing and conducting a methodologically robust RCT with

the first trial-based economic evaluation. By working with some of the leading experts from different backgrounds, we were able to put our ideas into practice and realise this project.

We published the protocol of this trial to improve transparency, which is considered one of the most important elements in research. The published study protocol provides a full overview of the methods used in our study. This will offer a thorough assessment of selective reporting and internal and external validity.[207] Publication of a trial protocol increases the benefits and decreases the risks for patients.

We used a non-inferiority trial design. We chose this design after Herrlin and colleagues found no difference between the meniscectomy and supervised exercise therapy group.[19] As supervised exercise therapy has potential advantages over surgery, including a non-invasive nature with fewer adverse events and lower costs, we believe that this design is appropriate.

With different outcome measurement instruments used in the performed and ongoing trials, we first determined which measurement instrument has the best properties in patients with degenerative meniscal tears without knee locking, according to the COSMIN criteria. Using the best measurement instrument will result in a better and more uniform interpretation of the outcome of treatment of meniscal injuries.

The trial-based design for our economic evaluation is considered the most valid approach for estimating the clinical and financial implications of a healthcare intervention.[140,141] By combining the clinical and financial data, we can determine whether partial meniscectomy or supervised exercise therapy provides better value for money.

Finally, we converted eminence-based into evidence-based practice. One frequently heard reason why orthopaedic surgeons continued to perform partial meniscectomies was their belief in being able to predict who would benefit more from this surgery. We were able to test this belief and found that surgeons' predictions of the outcome of treatment in patients with a degenerative meniscal tear were as accurate as tossing a coin.

Limitations

Study design

We conducted the trial and economic evaluation in several teaching hospitals in the Netherlands among patients between 45 and 70 years of age with non-obstructive meniscal tears. The survey was distributed among Dutch and Australian orthopaedic surgeons. Conducting the exact same trial in another patient population or another country with a different healthcare system may change the outcome. This limits the generalisability of our results. The findings in this thesis should therefore be interpreted in this context.

Our trial (*Chapter 4*) is vulnerable to performance bias due to the unblinded study design. We expect this to result in an overestimation of the true surgery effect due to

the higher treatment expectation in this group. We therefore expect that the true effect of supervised exercise therapy may even be higher. This is supported by the results of a blinded placebo-controlled trial, in which the authors found a larger (statistically not significant) improvement from the placebo treatment compared to the surgically treated group.[4]

The combination of meniscectomy and supervised exercise therapy may be more effective than partial meniscectomy alone. However, we chose to follow the recommendations from our national guideline and meniscectomy was only followed by supervised exercise therapy on indication. This may have influenced our findings and makes it harder to compare our results to the studies that did combine meniscectomy with supervised exercise therapy. Nevertheless, when we compared the trials in which meniscectomy was followed by supervised exercise therapy [4,19,205] to the trials with meniscectomy alone [21,116,197], we found similar results. Therefore, we believe that this approach with meniscectomy alone does not jeopardise the results.

Although a trial-based approach is considered the most valid method for estimating the clinical and financial implications of a healthcare intervention, economic evaluation studies often require large sample sizes.[140,141] Since these numbers are very hard to obtain in surgical trials, these trials – including ours (*Chapter 6*) – risk being underpowered.

In the online survey (*Chapter 8*) surgeons were presented patient profiles derived from the Escape trial (*Chapter 4*). This digital representation may not realistically represent the patient in the outpatient clinic. Furthermore, although we aimed to provide all relevant information, we cannot exclude that some information found relevant by surgeons may have been missing. We acknowledge that this may have influenced the predicting capacity of surgeons. However, by using only the extreme cases – those who responded best and worst to both treatments – we oversimplified the clinical reality, in which the vast majority of patients will be closer to the average.

Data collection and analysis

Because we did not keep a screening log in the trial (*Chapters 5 and 6*), we are unaware of the inclusion rate. With a recruitment phase of 27 months, this threatens the external validity of our findings. Low inclusion rates can originate both from patients and physicians, which can lead to selection bias, including self-selection.

The use of self-reported data in our trial (*Chapters 5 and 6*), clinimetrical studies (*Chapters 3 and 7*) and survey (*Chapter 8*), may have been susceptible to socially desirable answers and recall bias.

Bias may also have occurred during the data analyses. Although we planned to adjust for pre-specified potential confounders in the study protocol, we did not pre-specify that we would also test for effect modification of these confounders. Therefore, these results were described as exploratory since such unplanned analyses can lead to coincidental associations.

Study results

Due to the prospective follow-up (*Chapters 5, 6, and 7*), changes in participant responses may also be attributed to repeated measurements, known as the testing effect. Although we believe that the intervals were sufficient not to remember the answers from the previous measurement, this may have influenced the internal validity.

The treatment effect associated with arthroscopic surgery of the knee can, at least in part, be attributed to a placebo effect.[4] Placebo effects can be modified and substantially enhanced by a variety of factors that alter beliefs and expectations.[208] We do not know to what extent a placebo effect affected the results from our trial due to the absence of a sham-surgery group.

Although we anticipated a 25% cross-over rate in the sample size calculation of our clinical trial, 29% of the participants received delayed meniscectomy due to the persistence of symptoms during the follow-up of the trial. Seventy-five cent of these crossed over during the first 6 months of the follow-up. Postponing the option for delayed surgery with at least 6 months after starting the exercise protocol may help to reduce the number of surgeries further.

Implications for future research

The work is far from done and future research should aim to:

Further reduce the number of unnecessary meniscectomies

The evidence against the use of partial meniscectomy is more and more convincing. However, the decline in the number of partial meniscectomies is smaller than expected. Therefore, the implementation of these findings into practice needs more attention. The barriers for changing practice should be explored with all stakeholders, including surgeons, patients, general practitioners, physical therapists, policymakers and insurance companies.

The costs of treatment should never be the most important motivation for a physician to steer towards or withhold treatment. That said, financial reimbursement is known to influence professional decision-making.[209] Unfortunately, the reimbursement of our healthcare does not distinguish between evaluated (evidence-based) and unevaluated (eminence-based) healthcare.

Local clinician-led policies have already been proven effective in reducing the number of inappropriate surgeries in Australia.[183] Rigorously changing the nation-wide reimbursement to a system in which only properly evaluated (and effective) healthcare is reimbursed, would therefore be the ultimate goal.

Determine the long-term effects on the progression of osteoarthritis

The long-term effects of partial meniscectomy compared to supervised exercise therapy on the progression of osteoarthritis remain unclear. Longer follow-up from the published RCTs will provide more and valuable insight in how these numbers compare to the conservatively treated group. Currently, the 5-year follow-up of the Escape trial is ongoing.

Gain more insight into the subgroup of non-responders to conservative treatment

With up to 30% of non-responders to PT, further evidence is required to determine the characteristics of this subgroup. Unfortunately, the identification of subgroups that do (or do not) benefit from partial meniscectomy has failed thus far. Individual trials may have been too small to perform valid and reliable subgroup analyses. Identifying these subgroups is key in further reducing the number of ineffective partial meniscectomies. Based on the characteristics of a given group, conservative treatment could be 'personalized', targeted at their characteristics.[210]

Currently, an individual patient data meta-analysis using the original individual participant data of published RCTs is being performed.[211] This study aims to identify any subgroups by increasing power.

The identification of these subgroups may help to adjust our treatment protocols. Alternatives include adjustment of the exercise protocol and correction of knee malalignment with a brace or surgically to decompress the meniscus.

Better inform our patients of the true outcome of both surgery and conservative treatment

As patients tend to have over-optimistic expectations from partial meniscectomy,212 patients should be better informed about the outcome. Patients should be better informed about the advantages and disadvantages of all treatment options. For this, a decision aid is currently being developed, which will better prepare patients for their outpatient visit.

What was already known about this topic

In patients with degenerative meniscal tears:

- Partial meniscectomy is not superior over non-surgical treatment.
- When the torn meniscus is (partly) surgically removed, the joint contact area decreases and as a consequence the joint peak load increases.
- Partial meniscectomies continue to be performed in large numbers.

What this thesis adds

In patients with degenerative meniscal tears:

- Pooling data of previous RCTs revealed that partial meniscectomy has only a short-term small benefit of unknown clinical importance over non-surgical treatment.
- The IKDC 'Subjective Knee Form' is a reliable, valid and responsive measurement instrument when evaluating treatment outcome.
- Supervised exercise therapy is non-inferior for patient-reported knee function and more cost-effective than partial meniscectomy.
- Surgeons are unable in their ability to predict outcomes of partial meniscectomy and supervised exercise therapy.

Conclusions

1. The International Knee Documentation Committee Subjective knee form is a reliable, valid and responsive measurement instrument for evaluating knee function in the treatment of degenerative meniscal tears.

2. Supervised exercise therapy is non-inferior to and more cost-effective than arthroscopic partial meniscectomy in the treatment of degenerative meniscal tears.

3. Surgeons cannot reliably select the patients who are expected to benefit from partial meniscectomy and should therefore rely more on the evidence from the literature when considering treatment options.

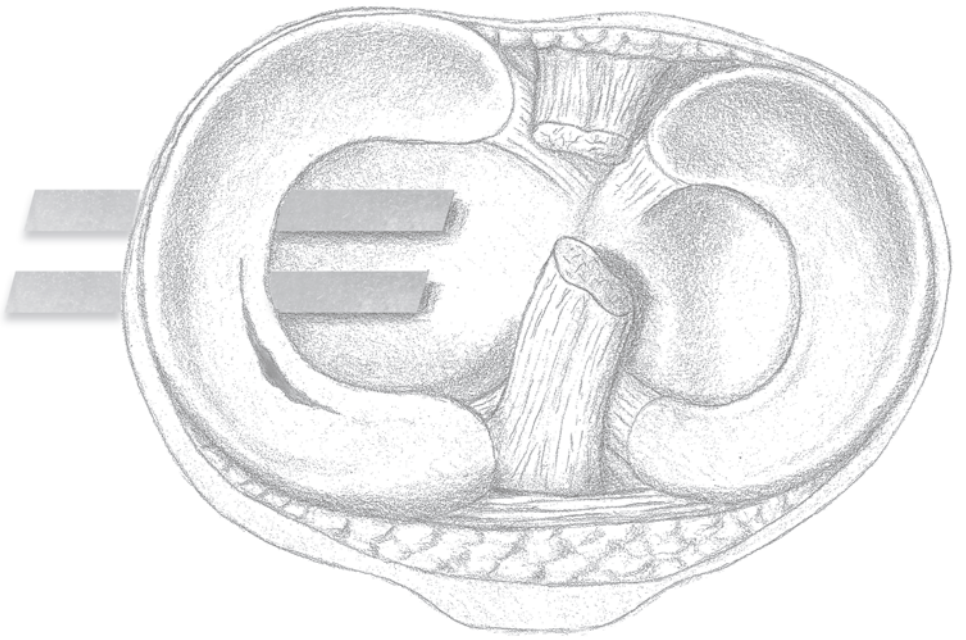
4. There is considerable evidence that supervised exercise therapy should be proposed as treatment of first choice in patients with (non-obstructive) degenerative meniscal tears.

5. Unnecessary surgeries will continue to be performed, as long as 1) orthopaedic surgeons are unconvinced by the existing evidence, 2) the guidelines do not provide a more uniform consensus, and 3) our healthcare system does not distinguish between evidence-based and eminence-based reimbursement of care.

6. Given the current widespread use of partial meniscectomy, future research should focus on the subgroup of non-responders to conservative treatment to further reduce the numbers of unnecessary surgeries.

Therefore, we conclude that:

Supervised exercise therapy, compared to arthroscopic partial meniscectomy, provides better value for money in the initial treatment of degenerative meniscal tears.



10

Appendices

Appendix 1 Literature search, May 2016

Search in MEDLINE (927)

(menisc* AND (injury OR injuries OR tear OR tears OR lesion* OR laesion* OR rupture)) AND ((menisc* AND (surgery OR surgical OR repair)) OR meniscect*) AND (conservative OR (physical therap*) OR (physio therap*) OR physiotherap* OR "SHAM" OR ((delayed OR timing OR "time factors"[Mesh]) AND (surgery OR surgical)))

Search in Embase (903)

menisc*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]
injury.mp.
injuries.mp.
tear.mp.
tears.mp.
lesion*.mp.
laesion*.mp.
rupture.mp.
2 or 3 or 4 or 5 or 6 or 7 or 8
menisc*.mp.
surgery.mp.
surgical.mp.
repair.mp.
11 or 12 or 13
10 and 14
meniscect*.mp.
15 and 16
conservative.mp.
(physical therapy OR physical therapies).mp.
(physio therapy OR physio therapies).mp.
physiotherapy.mp.
SHAM.mp.
delayed.mp.
timing.mp.
time-factors.mp.
surgery.mp.
surgical.mp.
23 or 24 or 25
26 or 27
28 and 29

18 or 19 or 20 or 21 or 22 or 30
15 or 16
1 and 9
31 and 32 and 33

Search in Cochrane (115)

(menisc* AND (injury OR injuries OR tear OR tears OR lesion* OR laesion* OR rupture)) AND ((menisc* AND (surgery OR surgical OR repair)) OR meniscect*) AND ((conservative OR (physical therap*) OR (physio therap*) OR physiotherap* OR "SHAM" OR ((delayed OR timing OR time) AND (surgery OR surgical)))

Search in the NHS Centre for Reviews and Dissemination (CRD) (30)

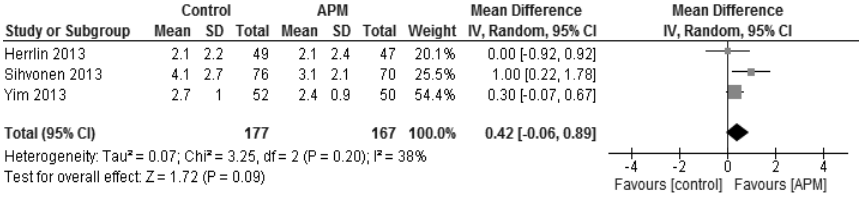
(meniscus OR menisci) AND (injury OR injuries OR tear OR tears OR lesion OR lesions OR laesion OR laesions OR rupture) AND (surgery OR surgical OR repair OR meniscectomy)

Search in PEDro (22)

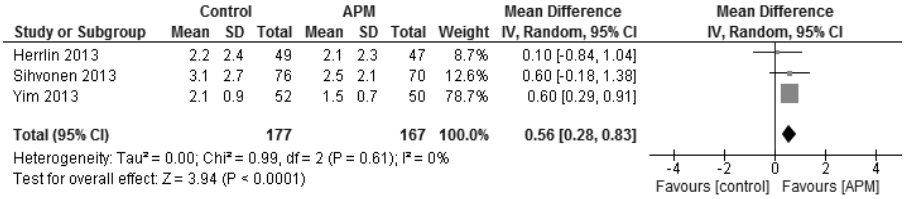
(meniscus)

The search strings for the different sources are presented.

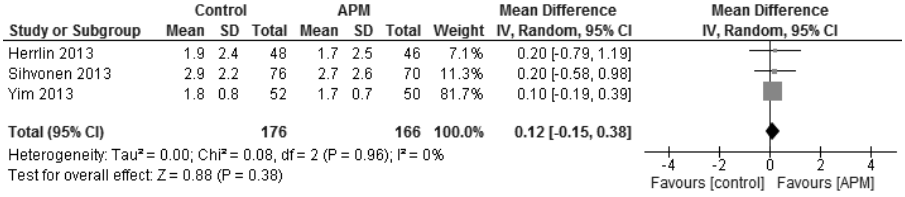
Appendix 2 Forest plots of secondary outcome pain



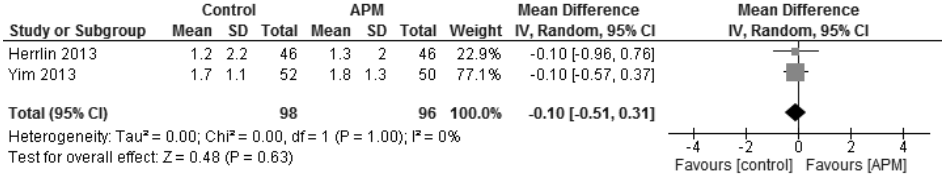
Appendix 2.1 Forest plot of comparison VAS and NRS during activity



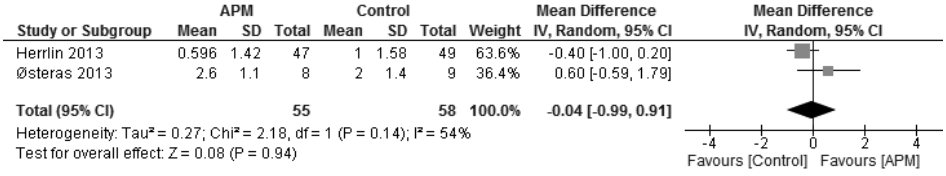
Appendix 2.2 Forest plot of comparison VAS and NRS during activity – 6 months



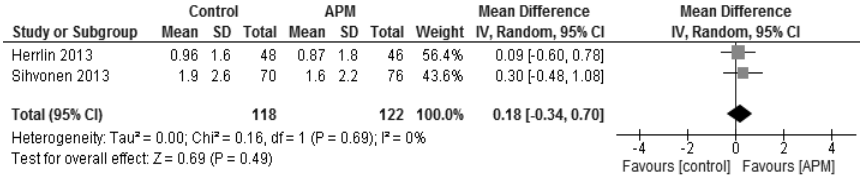
Appendix 2.3 Forest plot of comparison VAS and NRS during activity



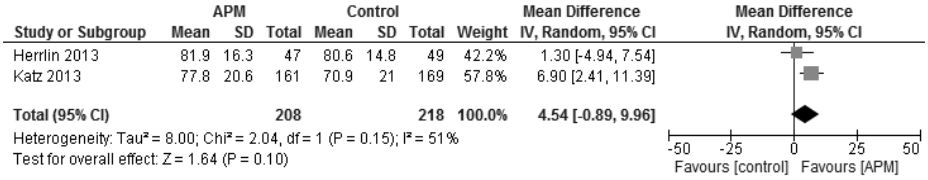
Appendix 2.4 Forest plot of comparison VAS and NRS during activity – 24 months



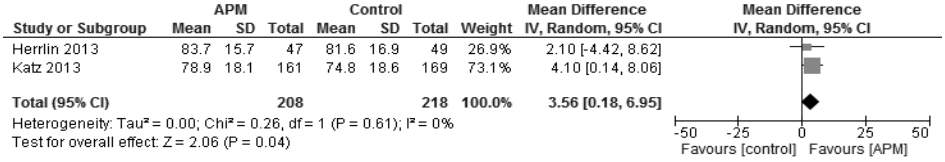
Appendix 2.5 Forest plot of comparison VAS – Rest 2-3 months



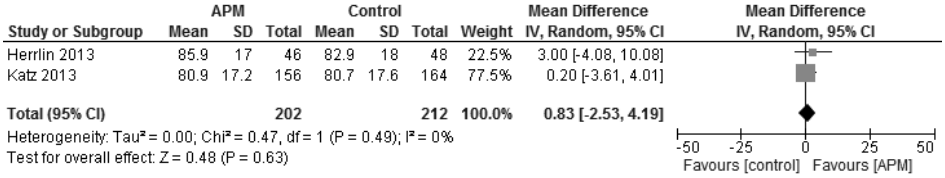
Appendix 2.6 Forest plot of comparison VAS and NRS in rest 12 months



Appendix 2.7 Forest plot of comparison KOOS Pain dimension 2-3 months



Appendix 2.8 Forest plot of comparison KOOS pain dimension – 6 months

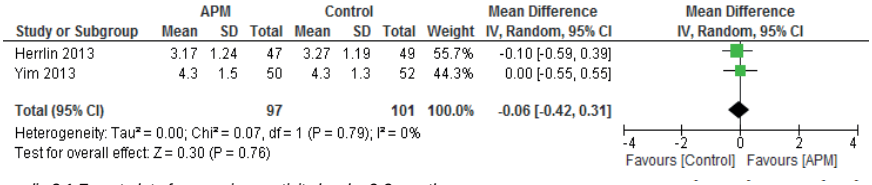


Appendix 2.9 Forest plot of comparison KOOS Pain dimension – 12 months

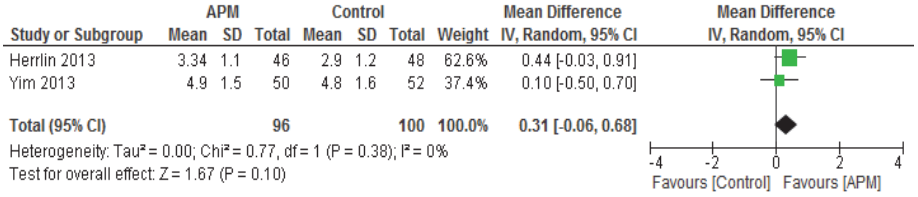
Statistically significant favourable results for surgery were found at 6 months (pooling VAS and NRS; and pooling KOOS pain data).

Abbreviations: APM, Arthroscopic Partial Meniscectomy; CI, Confidence Interval; df, Degrees of Freedom; I^2, level of heterogeneity (50-75% substantial, >75% considerable); KOOS, Knee injury and Osteoarthritis Outcome Score; NRS, Numeric Rating Scale; VAS, Visual Analogue Scale; p, probability level; SD, Standard Deviation.

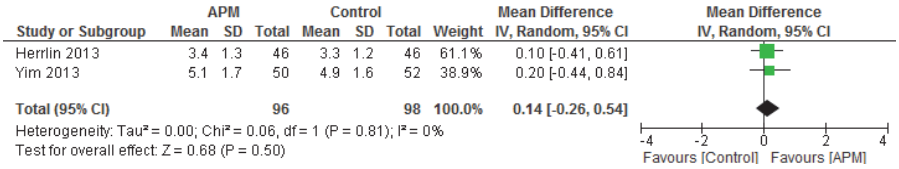
Appendix 3 Forest plots of secondary outcome Activity level



Appendix 3.1 Forest plot of comparison activity level – 2-3 months



Appendix 3.2 Forest plot of comparison activity level – 12 months



Appendix 3.3 Forest plot of comparison activity level – 24 months

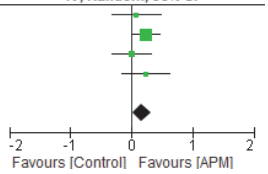
No statistically significant differences were found for change in activity level.

Abbreviations: APM, Arthroscopic Partial Meniscectomy; CI, Confidence Interval; df, Degrees of Freedom; I^2, level of heterogeneity (50-75% substantial, >75% considerable); p, probability level; SD, Standard Deviation.

Appendix 4 Forest plots of pooled primary outcomes

	APM			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Herrlin 2013	88.4	16.1	47	87.1	17.3	49	14.3%	0.08 [-0.32, 0.48]	
Katz 2013	85.3	17.5	161	81	17.9	169	48.8%	0.24 [0.03, 0.46]	
Sihvonen 2013	82.8	15.8	70	82.7	14.7	76	21.7%	0.01 [-0.32, 0.33]	
Yim 2013	84.1	7.7	50	82.3	7.8	52	15.1%	0.23 [-0.16, 0.62]	
Total (95% CI)			**328**			**346**	**100.0%**	**0.17 [0.01, 0.32]**	

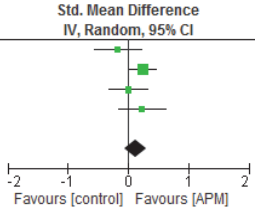
Heterogeneity: Tau² = 0.00; Chi² = 1.70, df = 3 (P = 0.64); I² = 0%
Test for overall effect: Z = 2.14 (P = 0.03)



Appendix 4.1 Comparison physical function (WOMAC, KOOS, LKSS) – 6 months

	APM			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Herrlin 2013	80.2	16.8	47	82.9	14.7	49	17.1%	-0.17 [-0.57, 0.23]	
Katz 2013	85.3	17.5	161	81	17.9	169	41.1%	0.24 [0.03, 0.46]	
Sihvonen 2013	82.8	15.8	70	82.7	14.7	76	23.9%	0.01 [-0.32, 0.33]	
Yim 2013	84.1	7.7	50	82.3	7.8	52	17.9%	0.23 [-0.16, 0.62]	
Total (95% CI)			**328**			**346**	**100.0%**	**0.11 [-0.07, 0.30]**	

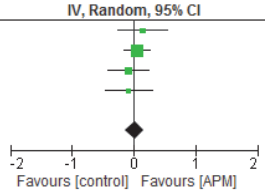
Heterogeneity: Tau² = 0.01; Chi² = 3.99, df = 3 (P = 0.26); I² = 25%
Test for overall effect: Z = 1.22 (P = 0.22)



Appendix 4.2 Comparison physical function (WOMAC, LKSS) – 6 months

	APM			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Herrlin 2013	89.2	16.5	46	86.7	18.1	48	14.2%	0.14 [-0.26, 0.55]	
Katz 2013	86.3	15.9	156	85.5	16.3	164	48.4%	0.05 [-0.17, 0.27]	
Sihvonen 2013	82.2	14.3	70	83.4	12.5	76	22.0%	-0.09 [-0.41, 0.24]	
Yim 2013	83.5	7.8	50	84.1	6.5	52	15.4%	-0.08 [-0.47, 0.31]	
Total (95% CI)			**322**			**340**	**100.0%**	**0.01 [-0.14, 0.16]**	

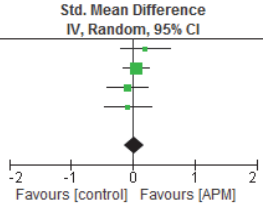
Heterogeneity: Tau² = 0.00; Chi² = 1.12, df = 3 (P = 0.77); I² = 0%
Test for overall effect: Z = 0.15 (P = 0.88)



Appendix 4.3 Comparison physical function (WOMAC, KOOS, LKSS) – 12 months

	APM			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Herrlin 2013	84.9	17	46	81.6	15.9	48	14.1%	0.20 [-0.21, 0.60]	
Katz 2013	86.3	15.9	156	85.5	16.3	164	48.4%	0.05 [-0.17, 0.27]	
Sihvonen 2013	82.2	14.3	70	83.4	12.5	76	22.0%	-0.09 [-0.41, 0.24]	
Yim 2013	83.5	7.8	50	84.1	6.5	52	15.4%	-0.08 [-0.47, 0.31]	
Total (95% CI)			**322**			**340**	**100.0%**	**0.02 [-0.13, 0.17]**	

Heterogeneity: Tau² = 0.00; Chi² = 1.52, df = 3 (P = 0.68); I² = 0%
Test for overall effect: Z = 0.25 (P = 0.80)



Appendix 4.4 Forest plot of comparison physical function – 12 months

Statistically significant favourable results for surgery were found at 6 months (pooling KOOS, LKSS, and WOMAC data).

Abbreviations: ADL, Activities of Daily Living; APM, Arthroscopic Partial Meniscectomy; CI, Confidence Interval; df; Degrees of Freedom; I^2, level of heterogeneity (50-75% substantial, >75% considerable); KOOS, Knee injury and Osteoarthritis Outcome Score; LKSS, Lysholm Knee Scoring Scale; p, probability level; SD, Standard Deviation; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Appendix 5 Exercise therapy protocol

The exercise program for both groups performed during 8 weeks

Time (week)	Exercises	repetitions or time
0-8	stationary bicycling for warming up and cooling down or cardiovascular training	gradual increase 7-15 min or longer
0-8	pully, strap around healthy ankle, stay and keep balance on injured side, move healthy leg forward, backward and sideward by standing in all 4 directions	3x12
0-4	calf raises on a leg press	3x12
0-8	standing hip extension in a "multi-hip" trainings device	3x12
0-4	balance on wobble board on both feet	
0-8	stair walking, walking, running, jumping according the patients ICF challenging with throwing a ball	10 min
5-Aug	calf raises standing on one leg	3x12
1-Aug	leg press, place the shinbone horizontal and the knee starting at 110˚, unilateral	3x12
5-Aug	lunges (according the needs of the patient) with < 90˚ knee flexion	3x12
5-Aug	balance on wobble board on one foot challenging with throwing a ball	3 min
5-Aug	cross-trainer as cardiovascular and cooling down training	10 min or more

footnote: By all exercises is it important to keep the patients individual needs and limitations focused by using the ICF. The uninjured side is also less trained as usual and therefore both sides should be trained. Besides training of the lower extremity, "core stability" training is of importance for good posture positioning and moving. The active rehabilitation program is designed around cardiovascular (circulation), coordination and balance, and closed chain strength exercises. Shearing forces in the knee are less using closed chain exercises compared to open chained exercises. The closed chain exercises activate both agonists and antagonists around the knee joint resulting in a direct rotatory movement and prevent in shearing forces seen by open chained exercises. (Heijne 2004, 2006 studied the role of open and closed exercises in the rehabilitation after a reconstruction of the anterior cruciate ligament and advised to be careful with open chained exercises in the early start of rehabilitation).

Home exercise program

All participants were instructed with the following exercises twice a week during a minimum period of 2 months:

- One leg standing during 60 seconds;
- A step-down exercise comprising 3, 9, 10 repetitions.

Appendix 6 Deviations from the original trial protocol

The following updates were recorded in the study protocol (and updated in the registries) during the conduct of this trial:

1. Participating centres

The trial has not started in the VU Medical Centre and in the Sint Antonius hospital. In both hospitals the staff of the orthopaedic surgery department could not agree on participating in this trial.

a. Several centres have been included during the conduct of this trial:
b. Noord-West Ziekenhuis groep, Alkmaar (November 2013)
c. Jan van Goyen Medical Center, Amsterdam (July 2013)
d. Elisabeth Tweesteden Ziekenhuis, Tilburg (June 2014)
e. Slotervaart Ziekenhuis, Amsterdam (August 2014)
f. Tergooi Ziekenhuis, Hilversum (September 2014)
g. Medisch Centrum Haaglanden, the Hague (April 2015)

2. Follow-up outcomes

The outcome physical examination at 24 months follow-up was removed from the study protocol in 2014 since it was believed that this outcome would have no added value.

3. Interim analysis sample size by independent committee

In August 2015, we performed an interim analysis to recalculate our sample size. Initially the sample size was based on a power of 90%, an alpha of .05, a standard deviation (SD) of 18 points and a non-inferiority threshold of 8 points on the IKDC 'Subjective Knee Form'. We calculated that with 20% loss to follow-up after 24 months and 25% delayed APM in PT group, 201 patients would be needed per group in this equivalence type RCT. This meant a total of 402 patients.

However, the SD was based on the reported standard deviation by Crawford and colleagues, who found an SD of 20 points on the International Knee Documentation Committee (IKDC) in a group of postoperative patients. Although we expected the SD to be smaller in our group after longer (24 months) post-enrolment, we used the SD of 20 for our sample size calculation to prevent the risk of being underpowered.

After 100 inclusions at 12 months post-enrolment, we performed an interim analysis to recalculate our SD and to prevent unnecessary inclusions. We found an SD of 17.5 and recalculated our sample size with an SD of 18 points. We found that we would need 160 patients per group, 320 patients in total.

This recalculation was done and approved by an independent committee (August 2015).

4. Updated statistical plan

In 2015 we updated the original statistical plan. We added prof. J.W.R. Twisk, leading expert in the field of mixed model analysis, to our research group, and replaced the originally intended

General Estimation Equation analyses by Mixed Modelling:

Data analysis

Effectiveness analysis

To investigate the clinical effectiveness of both treatment groups, we will use linear mixed-model analysis for continuous outcomes. Logistic generalized estimation equation analysis will be used for dichotomous outcomes. This method takes into account the dependency of observations within a patient, and the fact that not all patients may be assessed at each time point (missing data). All analyses will be carried out on an intention-to-treat and per/protocol basis, as well as cross-over analysis.

In the primary linear mixed model, the outcome variable studied (e.g., physical function on the IKDC) will be analysed as a dependent variable. To investigate the effect at the different time points, we will analyse the model, according to a 4-level structure (treatment group, centre, patient and time, in which time will be treated as a categorical variable to assess the treatment effects at the different time points). Time will be included as a dummy variable (reference is baseline T0), and 4 interaction terms will be analysed (T2Xgroup; T3Xgroup; T5Xgroup; T7Xgroup). To investigate the overall effect of both treatments (irrespective of time), we will also analyse the model according to a 3-level structure (treatment group, centre, patient). The baseline outcome will be included as a covariate in all models.

Besides analysing the basic model (e.g., analysis of main effects for treatment group and time and a time-by-treatment interaction), we will also control for possible confounders, by adding them as covariates (e.g., body mass index, gender, profession, ASA classification, the affected meniscus, the type of tear and the status of OA according to Kellgren and Lawrence Grading Scale for Osteoarthritis). Covariates are defined as resulting in more than 10% change in the parameter estimate of time-by-treatment interaction.

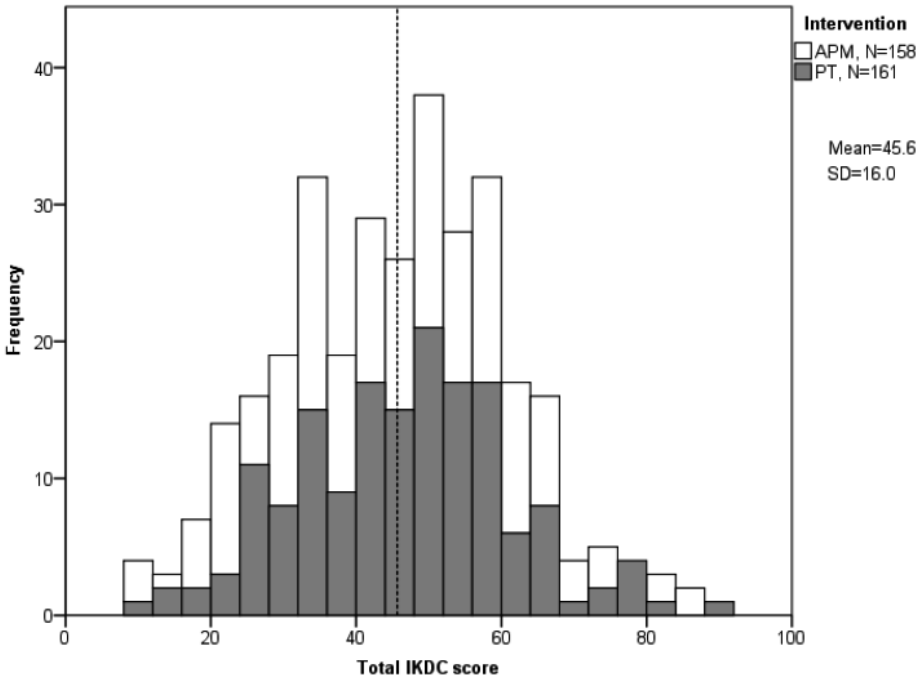
In the secondary linear mixed models, the outcome variables studied (e.g., general health on the RAND-36, quality of life on the EQ-5D-5L, level of activity on the TAS, knee pain on the question 10 of IKDC, the correlation between a patient's expectation and satisfaction, productivity losses on the Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness (TiC-P), muscle strength, range of motion and squatting) will be analysed in a similar way.

The estimated main effects for treatment at different assessment points under these different models are reported as in differences in means with 95% CIs for continuous outcomes, and ORs with 95% CIs for dichotomous outcomes.

At the time points 3 months (T2), 6 months (T3), 12 months (T5) and 2 years (T7), we will describe the incidence of revisions (intervention group) or treatment failures (=delayed APM, control group) using descriptives.

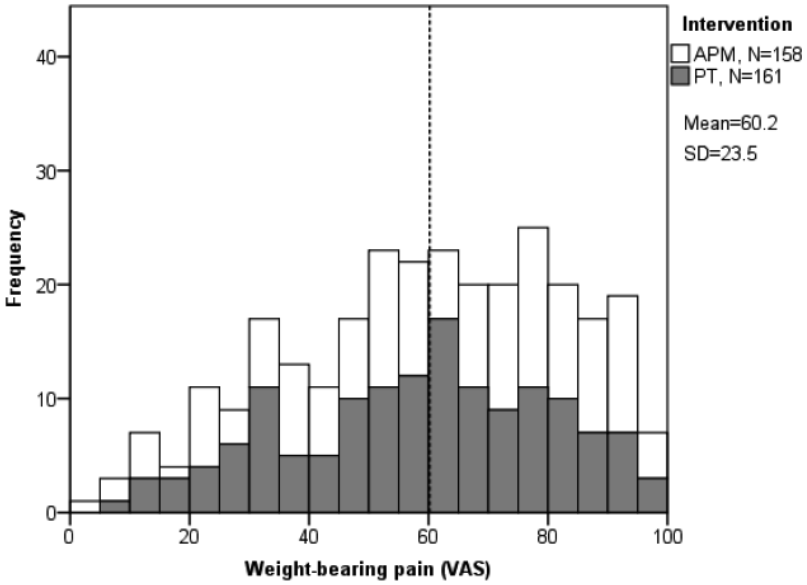
After 2 years (T7), we will compare the incidence of development or progression of OA between groups using a $\chi 2$ test (or Fisher's exact as appropriate).

Appendix 7 Distribution of IKDC score for knee function



In white: APM group, in grey: PT group. Scores range 0–100, with higher scores indicating better knee function. The vertical dashed line represents the mean baseline IKDC score for all participants. Abbreviations: APM, arthroscopic partial meniscectomy; IKDC, International Knee Documentation Committee; N, number; PT, physical therapy; SD, standard deviation.

Appendix 8 Distribution of VAS score for knee pain during weight-bearing



In white: APM group, in grey: PT group. Scores range 0–100, with higher scores indicating more pain. The vertical dashed line represents the mean baseline VAS score for all participants. Abbreviations: APM, arthroscopic partial meniscectomy; N, number; PT, physical therapy; SD, standard deviation; VAS, Visual Analogue Scale.

Appendix 9 Adjusted intervention effects for the IKDC score of knee function

Outcome	IKDC Score					
	APM	**N**	**PT**	**N**	**Delayed APM**	**N**
Primary outcome						
Intention-to-treat						
Baseline	44.8	158	46.5	161	-	-
Overall[c]	66.5	590	64.2	588	-	-
3 months	59.9	155	60.0	158	-	-
6 months	64.7	151	63.2	146	-	-
12 months	70.7	143	66.4	136	-	-
24 months	71.5	141	67.7	148	-	-
Secondary outcome						
As-treated[d]						
Baseline	43.9	150	48.6	97	40.8	47
Overall[c]	66.2	561	66.6	372	57.6	163
3 months	59.3	147	64.3	97	49.8	47
6 months	64.5	144	65.3	93	55.7	40
12 months	70.5	137	67.5	90	63.8	35
24 months	71.1	133	69.2	92	63.0	41

[a] The between-group difference for different follow-up moments and as overall effect with additional correction for the following potential confounders: sex, age in years, BMI (3 categories: ´normal´: 18.5<BMI<25; 'overweight': 25≤BMI<30; 'obesity': 30≤BMI<35 kg/m2), education level (high vs low),[24] osteoarthritis level (Kellgren-Lawrence classification),[14] location of tear (medial, lateral, both), mechanical complaints (derived from the IKDC), and baseline pain during weight-bearing. Positive values signify that patients did better with (delayed) APM.

[b] P-values for non-inferiority, indicating the probability of a >8 points difference in IKDC scores between groups in favour of (delayed) APM.

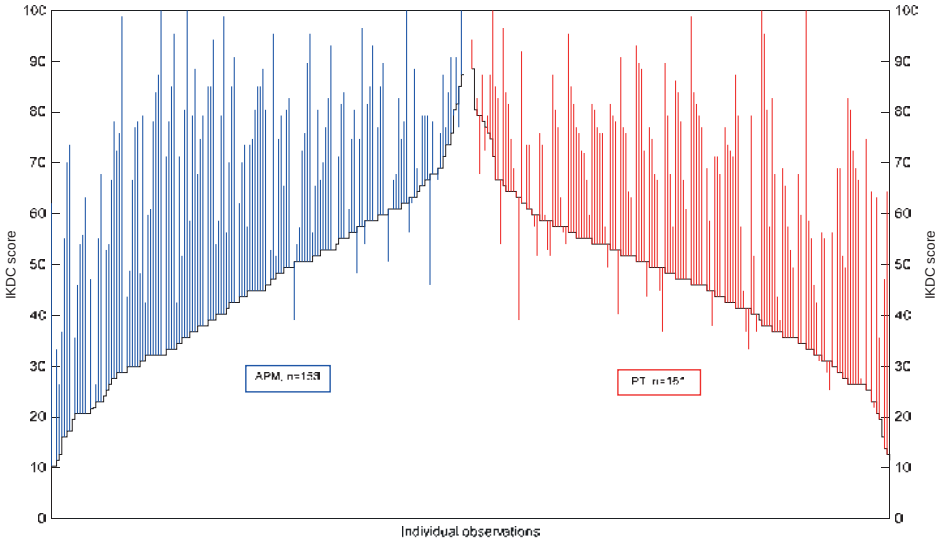
[c] Overall estimate over 24 months refers to the overall IKDC score within each group and between groups including all time points. The overall effect is based on a model including intervention group, baseline IKDC score and potential confounders as specified in [b] above.

APM vs PT		Delayed APM vs PT	
Between-group difference[a] (97.5% CI)	**p-value for non-inferiority[b]**	**Between-group difference[a] (97.5% CI)**	**p-value for non-inferiority[b]**
-	-	-	-
4.4 (-∞ – 7.5)	.01	-	-
1.1 (-∞ – 5.0)	<.001	-	-
4.2 (-∞ – 8.1)	.03	-	-
7.1 (-∞ – 11.1)	.33	-	-
5.3 (-∞ – 9.3)	.09	-	-
-	-	-	-
2.9 (-∞ – 6.5)	.003	-6.5 (-∞ – 1.5)	.005
-2.1 (-∞ – 2.4)	<.001	-10.9 (-∞, -4.8)	<.001
2.8 (-∞ – 7.4)	.01	-7.1 (-∞, -0.75)	<.001
6.3 (-∞ – 10.9)	.24	-1.9 (-∞ – 4.8)	.002
4.6 (-∞ – 9.2)	.07	-3.4 (-∞ – 3.0)	<.001

[d] In the as-treated model, participants were analysed who adhered to their randomised treatment in 3 groups, 1) participants allocated to the APM group who received APM, 2) participants allocated to the PT group who completed the PT protocol without having APM during the follow-up period (e.g. ≥16 PT sessions; <16 sessions was considered a protocol violation), and 3) participants allocated to the PT group who had APM during follow-up (delayed APM group). Patients who were randomised for APM but did not have surgery and patients who were randomised to PT, but did not complete the PT protocol and did not have delayed APM were discarded from the as-treated analysis.

Abbreviations: APM, arthroscopic partial meniscectomy; IKDC, International Knee Documentation Committee (range 0 (most limitations) to 100 (no limitations)); N, number of observations; PT, physical therapy; 97.5% CI, 97.5% confidence interval, i.e. a 1-sided alpha of 2.5%.

Appendix 10 Parallel line plot of individual change scores in IKDC for knee function from baseline to 24 months



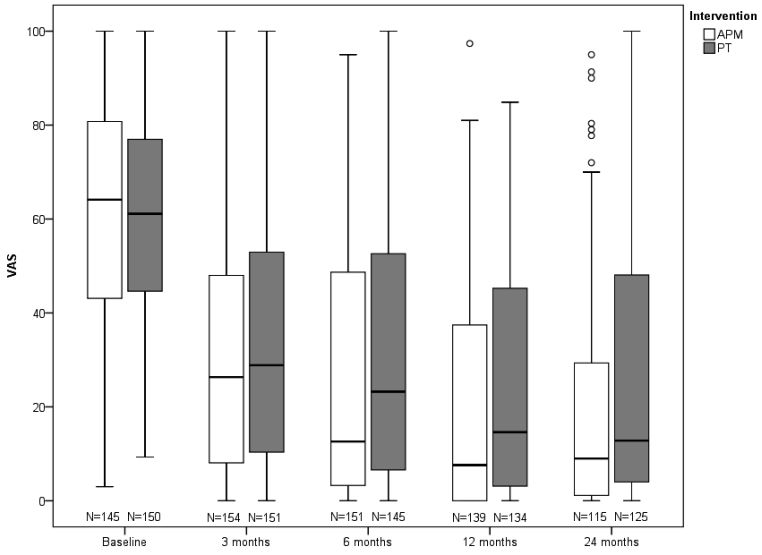
Parallel line plot of individual observations. The vertical lines represent the difference in individual IKDC score (range 0-100) from baseline (which is represented by the black line) to 24 months. In blue the individual APM data are shown, in red the individual PT data are shown. Positive scores indicate improvement of knee function, whereas negative scores indicate deterioration of knee function. Abbreviations: APM, arthroscopic partial meniscectomy; IKDC, International Knee Documentation Committee; PT, physical therapy.

Appendix 11 Median IKDC points by treatment group

Measurement	Arthroscopic Partial Meniscectomy	Physical Therapy
Baseline	47.1 (32.2–57.5)	47.1 (36.6–56.7)
3 months	62.1 (48.7–75.4)	60.9 (49.4–72.4)
6 months	70.1 (58.0–83.3)	65.5 (51.0–80.0)
12 months	74.7 (62.9–88.8)	65.5 (50.1–80.9)
24 months	74.7 (63.9–85.5)	69.0 (58.0–79.9)

Values are median (IQR) IKDC Points.
Abbreviations: IKDC, International Knee Documentation Committee; IQR, interquartile range.

Appendix 12 Distribution of VAS score for knee pain during weight-bearing



This figure represents the results of the secondary outcome knee pain during weight-bearing on the VAS (range 0 (no pain) to 100 (worst pain imaginable) for intention-to-treat analysis. The data represents actual patient data at each time. In each comparison, the box indicates the range between the 25th and 75th percentile with the median indicated as a horizontal line within the box. The whiskers extend to the upper and lower adjacent values, the most extreme values that are within 1.5 * IQR beyond the 25th and 75th percentiles. Circles indicate points beyond these values. Median data for APM (IQR) are: Baseline 64.1(41.8–80.3), 3 months 26.4 (10.3–50.4), 6 months 12.6 (2.4–48.4), 12 months 7.6 (0–37.9), 24 months 9.0 (5.3–33.8) mm. Median data for PT (IQR) are: Baseline 61.1 (43.1–75.6), 3 months 28.9 (11.8–55.1), 6 months 23.3 (7.9–38.9), 12 months 14.6 (3.1–45.7), 24 months 12.8 (3.4–47.6) mm. Abbreviations: APM, arthroscopic partial meniscectomy; IQR, Interquartile range; PT, physical therapy; mm, millimetre; VAS, Visual Analogue Scale.

Appendix 13 Adjusted intervention effects for the VAS for weight-bearing pain

	VAS weight-bearing					
Outcome	**APM**	**N**	**PT**	**N**	**Delayed APM**	**N**
Exploratory outcome						
Intention-to-treat						
Baseline	61.1	145	59.3	150	-	-
Overall[c]	24.5	559	28.8	555	-	-
3 months	30.4	154	33.4	151	-	-
6 months	25.4	151	31.0	145	-	-
12 months	21.0	139	24.4	134	-	-
24 months	19.6	115	25.5	125	-	-
As-treated[d]						
Baseline	62.5	137	56.0	91	66.4	43
Overall[c]	24.9	531	24.7	349	39.8	154
3 months	30.8	146	27.5	91	48.0	46
6 months	26.0	144	28.0	92	40.8	40
12 months	21.0	133	21.0	89	31.4	34
24 months	20.5	108	21.6	77	36.0	34

[a] The between-group difference for different follow-up moments and as overall effect with additional correction for the following potential confounders: sex, age in years, BMI (3 categories: ´normal´: 18.5<BMI<25; 'overweight': 25≤BMI<30; 'obesity': 30≤BMI<35 kg/m2), education level (high vs low),[24] osteoarthritis level (Kellgren-Lawrence classification),[14] location of tear (medial, lateral, both), mechanical complaints (derived from the IKDC), and baseline pain during weight-bearing. Negative values signify that patients did better with APM.

[b] P-values with respect to zero (superiority testing).

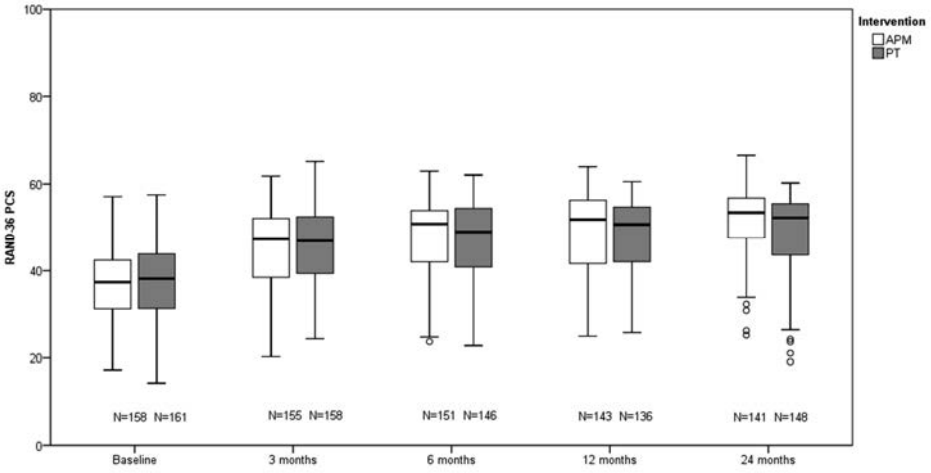
[c] Overall estimate over 24 months refers to the overall IKDC score within each group and between groups including all time points. The overall effect is based on a model including intervention group, baseline IKDC score and potential confounders as specified in [b] above.

	APM vs PT		Delayed APM vs PT	
	Between-group difference[a] (95% CI)	p-value[b]	Between-group difference[a] (95% CI)	p-value[b]
	-	-	-	-
	-6.7 (-11.3, -2.2)	.004	-	-
	-3.3 (-9.3–2.7)	.28	-	-
	-9.1 (-15.2, -3.0)	.003	-	-
	-7.0 (-13.3, -0.67)	.03	-	-
	-8.3 (-14.9, -1.7)	.01	-	-
	-	-	-	-
	-3.8 (-9.0–1.5)	.16	12.6 (5.4–19.8)	.001
	0.58 (-6.4–7.6)	.87	16.3 (6.8–25.8)	.001
	-6.5 (-13.6–0.57)	.07	11.6 (1.6–21.6)	.02
	-4.5 (-11.7–2.7)	.22	7.9 (-2.7–18.5)	.14
	-5.4 (-13.0–2.3)	.17	11.9 (1.5–22.3)	.03

[d] In the as-treated model, participants were analysed who adhered to their randomised treatment in 3 groups, 1) participants allocated to the APM group who received APM, 2) participants allocated to the PT group who completed the PT protocol without having APM during the follow-up period (e.g. ≥16 PT sessions; <16 sessions was considered a protocol violation), and 3) participants allocated to the PT group who had APM during follow-up (delayed APM group). Patients who were randomised for APM but did not have surgery and patients who were randomised to PT, but did not complete the PT protocol and did not have delayed APM were discarded from the as-treated analysis.

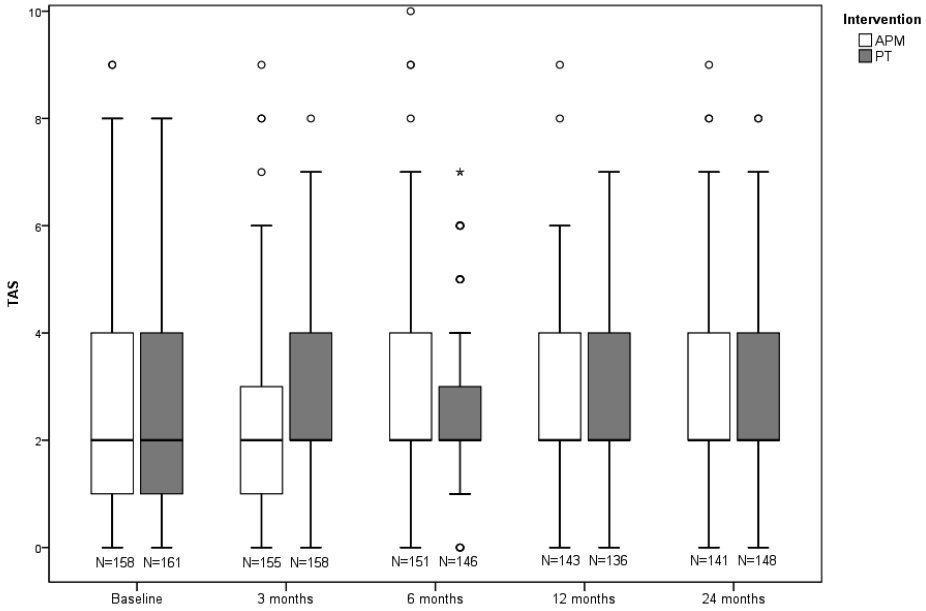
Abbreviations: APM, arthroscopic partial meniscectomy; N, number of observations; PT, physical therapy; VAS, visual analogue scale (range 0 (no pain) to 100 (worst pain imaginable)); 95% CI, 95% confidence interval, i.e. a 2-sided alpha of 5%.

Appendix 14 Distribution of RAND-36 PCS score for general health



This figure represents the results of the exploratory outcome general health on the RAND-36 PCS (range 0–100, worse to best) for intention-to-treat analysis. The data represents actual patient data at each time. In each comparison, the box indicates the range between the 25th and 75th percentile with the median indicated as a horizontal line within the box. The whiskers extend to the upper and lower adjacent values, the most extreme values that are within 1.5 * IQR beyond the 25th and 75th percentiles. Circles indicate points beyond these values. Median data for APM (IQR) are: Baseline 37.3 (31.6–43.0), 3 months 47.4 (40.5–54.3), 6 months 50.8 (44.8–56.7), 12 months 51.8 (44.1–59.3), 24 months 53.4 (48.8–58.0) points. Median data for PT (IQR) are: Baseline 38.2 (31.8–44.6), 3 months 47.0 (40.4–53.5), 6 months 48.9 (42.1–55.7), 12 months 50.7 (44.3–57.0), 24 months 52.2 (46.–58.2) points. Abbreviations: APM, arthroscopic partial meniscectomy; IQR, Interquartile range; PT, physical therapy; RAND-36 PCS, RAND-36 Physical Component Score.

Appendix 15 Distribution of TAS score for activity level



This figure represents the results of the secondary outcome activity level on the Tegner Activity Scale (range 0–10, worse to best) for intention-to-treat analysis. The data represents actual patient data at each time. In each comparison, the box indicates the range between the 25th and 75th percentile with the median indicated as a horizontal line within the box. The whiskers extend to the upper and lower adjacent values, the most extreme values that are within 1.5 * IQR beyond the 25th and 75th percentiles. Circles indicate points beyond these values. Median data for APM (IQR) are: Baseline 2.0 (1.0–4.0), 3 months 2.0 (1.0–3.0), 6 months 2.0 (2.0–4.0), 12 months 2.0 (2.0–4.0), 24 months 2.0 (2.0–4.0) points. Median data for PT (IQR) are: Baseline 2.0 (1.0–4.0), 3 months 2.0 (2.0–4.0), 6 months 2.0 (2.0–3.0), 12 months 2.0 (2.0–4.0), 24 months 2.0 (2.0–4.0) points. Abbreviations: APM, arthroscopic partial meniscectomy; IQR, Interquartile range; PT, physical therapy; TAS, Tegner Activity Level.

Appendix 16 Adverse events

Event[a]	APM (N=159)	PT (N=162)[b]
Serious adverse events		
Cardiovascular		
Acute myocardial infarction	0	1
Sudden death	0	1
Venous Thromboembolism	0	0
Neurological[c]	1	1
Alcoholic pancreatitis	0	1
Lymph node malignancy	1	0
Rectal polyp	1	0
Knee surgery		
Arthroscopy	3[d]	1[e]
Total knee arthroplasty	2	3
Other	1	0
Total	9	8
Non-serious adverse events		
Musculoskeletal		
Reactive arthritis	1	0
Knee pain resulting in extra consultation	6	2
Pain in back, hip or foot	2	0
Surgical site infection	0	0
Other	0	2
Total	9	4

[a] Adverse events were detected by patient reporting.

[b] Delayed APM in the PT group was not included as adverse event in this table.

[c] Neurological adverse events included an intracranial malignancy and radiation therapy for unknown disease.

[d] 1 participant received meniscus surgery on the other knee.

[e] 1 participant from the delayed APM group had a reoperation.

Abbreviations: APM, arthroscopic partial meniscectomy; N, number; PT, physical therapy.

Appendix 17 Results of interaction effects between intervention group and potential effect modifiers

Potential effect modifiers	Between-group difference[a]	Interaction with group	p-value for interaction	95% CI
BMI obesity (reference)[b]	10.7	-	-	4.7–16.8
BMI overweight[b]		-9.6	.01	-17.0, -2.2
BMI normal[b]		-9.4	.02	-17.1, -1.6
Location of tear lateral (reference)	9.1	-	-	1.7–16.6
Location of tear (medial)		-6.4	.12	-14.5–1.7
Location of tear (both)		-7.0	.39	-22.7–8.8
Education level low (reference)[121]	1.9	-	-	-2.1–5.9
Education level high[121]		4.2	.15	-1.5–9.9
OA level low (reference)[83]	3.7	-	-	0.00–7.5
OA level moderate to severe[83]		1.01	.74	-5.1–7.10
Mechanical complaints no (reference)	3.9	-	-	0.2–7.5
Mechanical complaints yes		-0.73	.81	-6.6–5.2
Sex male (reference)	4.3	-	-	0.3–8.4
Sex female		-1.5	.60	-7.2–4.2
Age[c]		0.14	.53	-0.29–0.57
Baseline VAS weight-bearing[c]		0.14	.03	0.01–0.27
Baseline IKDC score[c]		-0.12	.21	-3.1–0.07

The findings indicate that the effect of APM with respect to PT was significantly larger in participants with obesity (participants with obesity scored on average 10.7 points higher (p=.001) on the IKDC score after APM compared to PT)), and significantly larger in participants with higher baseline pain (for each millimetre increase in baseline pain, the effect of APM was 0.14 IKDC points larger (p-value for interaction .03)).

[a] Difference in IKDC score of patients in the APM group compared to the PT group. Positive values signify that patients did better with APM.

[b] BMI (3 categories: ´normal´: 18.5<BMI<25; 'overweight': 25≤BMI<30; 'obesity': 30≤BMI<35 kg/m2).

[c] continuous outcome.

Abbreviations: BMI, body mass index; 95% CI, 95% confidence interval; IKDC, International Knee Documentation Committee (range 0 (most limitations) to 100 (no limitations)); OA, osteoarthritis according to Kellgren Lawrence score (range 0 (no osteophytes or joint-space narrowing) to 4 (>50% joint-space narrowing)); VAS, visual analogue scale (range 0 (no pain) to 100 (worst pain imaginable)).

Appendix 18 Unadjusted differences in pooled mean costs and effects

Analysis	Sample size PT	Sample size APM	Outcome	ΔC (95% CI) €
Main analysis - Imputed dataset	161	158	IKDC (Range: 0–100)	-2056 (-3326, -1019)
	161	158	QALYs (Range: 0–1)	-2056 (-3326, -1019)

Differences in pooled mean costs and effects (95% CI), incremental cost-effectiveness ratios, distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes, and percentage of bootstrapped cost-effectiveness pairs located in the non-inferiority region of the cost-effectiveness planes for the unadjusted costs and effects.

[a] Refers to the northeast quadrant of the CE plane, indicating that PT is more effective and more costly than APM.

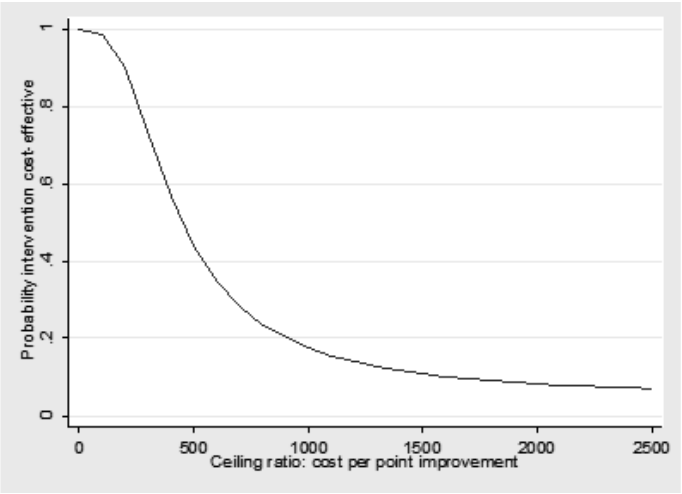
[b] Refers to the southeast quadrant of the CE plane, indicating that PT is more effective and less costly than APM.

ΔE (95% CI) Points	ICER €/point	Distribution CE plane (%)				
		NE[a]	SE[b]	SW[c]	NW[d]	Non-inferiority region
-3.8 (-8.0–0.5)	544	0.0	3.4	96.6	0.0	96.4
-0.024 (-0.071–0.023)	85 953	0.0	16.2	83.8	0.0	91.7

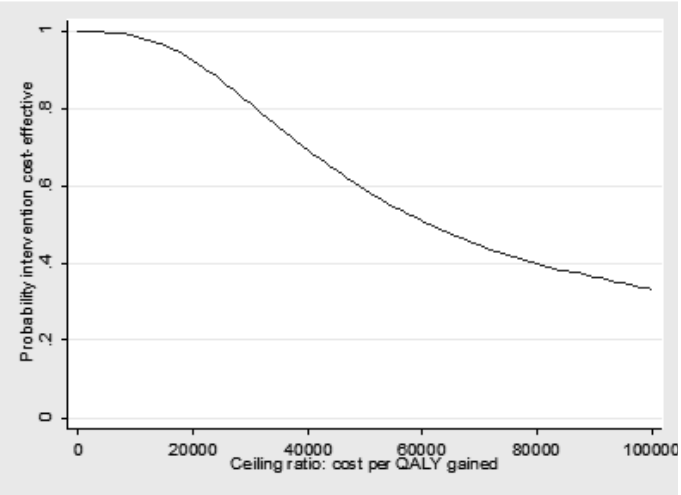
[c] Refers to the southwest quadrant of the CE plane, indicating that PT is less effective and less costly than APM.
[d] Refers to the northwest quadrant of the CE plane, indicating that PT is less effective and more costly than APM.
Abbreviations: APM, Arthroscopic Partial Meniscectomy; C, Costs; CE plane, Cost-Effectiveness plane; E, Effects; ICER, Incremental Cost-effectiveness Ratio; IKDC, International Knee Documentation Committee; PT, Physical Therapy; QALYs, Quality Adjusted Life Years; SA, Sensitivity Analysis.

Appendix 19 Cost-effectiveness acceptability curves

a)



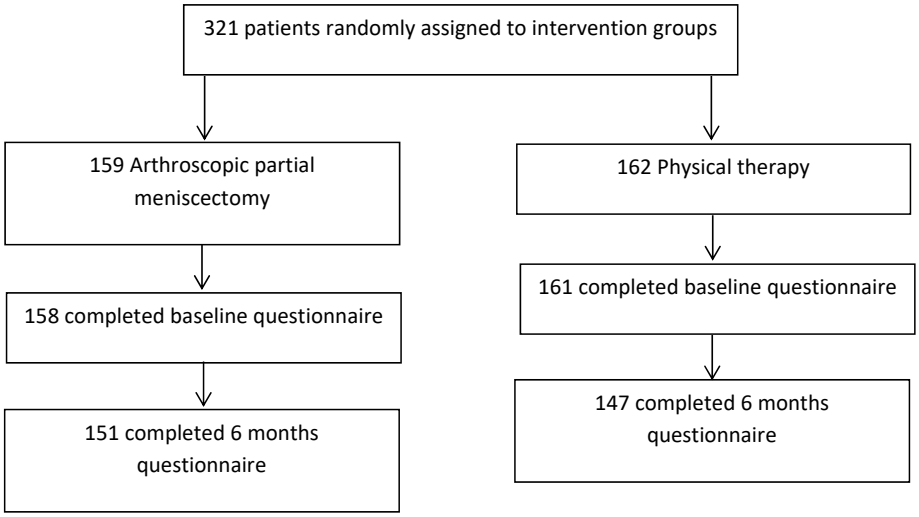
b)



Cost-effectiveness acceptability curves indicating the probability of PT being cost-effective in comparison with APM for different values of willingness to pay for the IKDC (a) and QALYs (b).

Abbreviations: APM, Arthroscopic Partial Meniscectomy; IKDC; International Knee Documentation Committee; PT, Physical Therapy; QALYs, Quality Adjusted Life Years.

Appendix 20 Patient inclusion and follow-up



Appendix 21 Patient profiles

Profile 1

Unemployed

50 y

BMI 23

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	1
VAS weight bearing (0-100)	7
IKDC score (0-100)	85

Physical exam	
Flexion	140°
Extension	+5°
Joint effusion	-
Joint line tenderness	-
McMurray	pain medial
Thessaly	-

Physical therapy in last 3 months	
No	

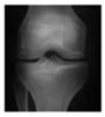
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 2

Professional cook

54 y

BMI 23

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	56
VAS weight bearing (0-100)	85
IKDC score (0-100)	57

Physical exam	
Flexion	120°
Extension	+10°
Joint effusion	+
Joint line tenderness	medial
McMurray	-
Thessaly	pain medial in 20°

Physical therapy in last 3 months	
No	

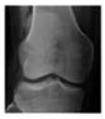
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 0



Profile 3

Functional manager

45 y

BMI 24

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	52
VAS weight bearing (0-100)	60
IKDC score (0-100)	67

Physical exam	
Flexion	140°
Extension	0°
Joint effusion	-
Joint line tenderness	medial and lateral
McMurray	pain lateral
Thessaly	pain lateral in 5°

Physical therapy in last 3 months	
No	

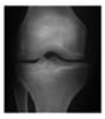
MRI

Medial meniscus	*Lateral meniscus*
Intact	Complex tear

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 4

Administrative assistant

50 y

BMI 23

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	1
VAS weight bearing (0-100)	7
IKDC score (0-100)	63

Physical exam	
Flexion	140°
Extension	0°
Joint effusion	-
Joint line tenderness	pain lateral
Mc Murray	-
Thessaly	pain lateral in 5°

Physical therapy in last 3 months	
No	

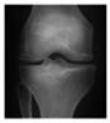
MRI

Medial meniscus	*Lateral meniscus*
Intact	Complex tear

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 5

Fuel operator

50 y

BMI 23.8

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	84
VAS weight bearing (0-100)	100
IKDC score (0-100)	51

Physical exam	
Flexion	135°
Extension	0°
Joint effusion	-
Joint line tenderness	-
Mc Murray	-
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
Yes, 11 session	

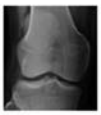
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 0



Profile 6

Hotel manager

60 y

BMI 31

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	45
VAS weight bearing (0-100)	77
IKDC score (0-100)	40

Physical exam	
Flexion	130°
Extension	0°
Joint effusion	+
Joint line tenderness	medial
Mc Murray	pain medial
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
No	

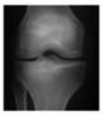
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 7

Taxi driver

62 y

BMI 26

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	66
VAS weight bearing (0-100)	51
IKDC score (0-100)	38

Physical exam	
Flexion	120°
Extension	0°
Joint effusion	-
Joint line tenderness	medial
Mc Murray	pain medial
Thessaly	-

Physical therapy in last 3 months	
No	

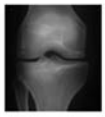
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 8

Sculptor

50 y

BMI 23

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	69
VAS weight bearing (0-100)	91
IKDC score (0-100)	37

Physical exam	
Flexion	130°
Extension	0°
Joint effusion	+
Joint line tenderness	medial
McMurray	pain medial
Thessaly	pain medial in 20°

Physical therapy in last 3 months	
No	

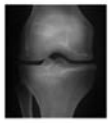
MRI

Medial meniscus	*Lateral meniscus*
Complex tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 9

Upholsterer

65 y

BMI 25

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	4
VAS weight bearing (0-100)	93
IKDC score (0-100)	26

Physical exam	
Flexion	110°
Extension	0°
Joint effusion	-
Joint line tenderness	-
McMurray	-
Thessaly	pain lateral in 5°

Physical therapy in last 3 months	
Yes, 2 sessions	

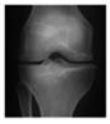
MRI

Medial meniscus	*Lateral meniscus*
Intact	Radial tear

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 10

Teacher primary school

59 y

BMI 27

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	14
VAS weight bearing (0-100)	68
IKDC score (0-100)	26

Physical exam	
Flexion	130°
Extension	0°
Joint effusion	+
Joint line tenderness	medial
McMurray	-
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
Yes, 20 sessions	

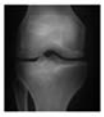
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 11

Composer / musician

62 y

BMI 33

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	19
VAS weight bearing (0-100)	61
IKDC score (0-100)	41

Physical exam	
Flexion	130°
Extension	0°
Joint effusion	-
Joint line tenderness	medial
McMurray	pain medial
Thessaly	pain medial in 5°
Thessaly	pain lateral in 20°

Physical therapy in last 3 months	
No	

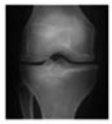
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 12

Unemployed

69 y

BMI 23

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	48
VAS weight bearing (0-100)	11
IKDC score (0-100)	32

Physical exam	
Flexion	140°
Extension	+5°
Joint effusion	-
Joint line tenderness	lateral
McMurray	pain lateral
Thessaly	pain lateral in 5°

Physical therapy in last 3 months	
Yes, 4 sessions	

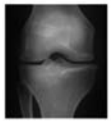
MRI

Medial meniscus	*Lateral meniscus*
Intact	Complex tear

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 13

Unemployed

62 y

BMI 32

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	67
VAS weight bearing (0-100)	90
IKDC score (0-100)	13

Physical exam	
Flexion	130°
Extension	+5°
Joint effusion	+
Joint line tenderness	medial
McMurray	pain medial
Thessaly	-

Physical therapy in last 3 months	
Yes, 2 sessions	

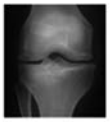
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 14

Pharmacy assistant

46 y

BMI 24

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	61
VAS weight bearing (0-100)	57
IKDC score (0-100)	63

Physical exam	
Flexion	140°
Extension	0°
Joint effusion	-
Joint line tenderness	-
McMurray	pain lateral
Thessaly	pain lateral in 5°

Physical therapy in last 3 months	
No	

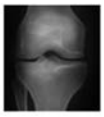
MRI

Medial meniscus	*Lateral meniscus*
Intact	Horizontal

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 15

Civil servant

49 y

BMI 22

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	25
VAS weight bearing (0-100)	81
IKDC score (0-100)	36

Physical exam	
Flexion	145°
Extension	0°
Joint effusion	+
Joint line tenderness	medial
McMurray	pain medial
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
Yes, 3 sessions	

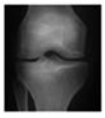
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Horizontal tear

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 16

Saleswoman

51 y

BMI 35

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	12
VAS weight bearing (0-100)	64
IKDC score (0-100)	49

Physical exam	
Flexion	135°
Extension	0°
Joint effusion	+
Joint line tenderness	medial
McMurray	pain medial
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
Yes, 9 sessions	

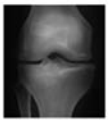
MRI

Medial meniscus	*Lateral meniscus*
Complex tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 2



Profile 17

Unemployed

66 y

BMI 26

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	67
VAS weight bearing (0-100)	90
IKDC score (0-100)	33

Physical exam	
Flexion	130°
Extension	-5°
Joint effusion	+
Joint line tenderness	medial
McMurray	pain medial
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
No	

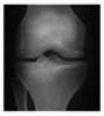
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 18

Varnish processor

51 y

BMI 28

Consultation

Complaints	
Pseudo-locking	+
VAS rest (0-100)	6
VAS weight bearing (0-100)	8
IKDC score (0-100)	29

Physical exam	
Flexion	140°
Extension	5°
Joint effusion	-
Joint line tenderness	medial and lateral
McMurray	click lateral
Thessaly	pain medial in 5°

Physical therapy in last 3 months	
No	

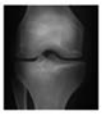
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Profile 19

Agent / salesman

63 y

BMI 25

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	10
VAS weight bearing (0-100)	11
IKDC score (0-100)	68

Physical exam	
Flexion	135°
Extension	0°
Joint effusion	+
Joint line tenderness	lateral
McMurray	pain lateral
Thessaly	-

Physical therapy in last 3 months	
No	

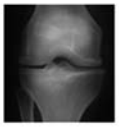
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Horizontal tear

X-ray

Kellgren-Lawrence (OA)

Grade 3



Profile 20

Unemployed

51 y

BMI 28

Consultation

Complaints	
Pseudo-locking	-
VAS rest (0-100)	42
VAS weight bearing (0-100)	42
IKDC score (0-100)	60

Physical exam	
Flexion	120°
Extension	0
Joint effusion	-
Joint line tenderness	medial
McMurray	-
Thessaly	pain medial in 20°

Physical therapy in last 3 months	
No	

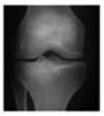
MRI

Medial meniscus	*Lateral meniscus*
Horizontal tear	Intact

X-ray

Kellgren-Lawrence (OA)

Grade 1



Appendix 22 Survey content

Biographics

Years of experience as orthopaedic surgeon:

- Less than 5 years
- Between 5 and 15 years
- More than 15 years

Field of expertise (more options possible):

- Knee surgery
- Hip surgery
- Shoulder/elbow surgery
- Back surgery
- Foot/ankle surgery
- Trauma surgery
- Arthroscopic surgery
- Sports injuries
- Infectiology
- Other: ……..

In clinical decision making, what relative weighting do you give to evidence from your personal experience and of those around you, compared to evidence from clinical research?

Experience based Published clinical research

The evidence

The following statements concern treatment (meniscectomy versus physical therapy) in patients between 45 and 70 years old with a non-obstructive meniscal tear

I am completely up to date with the literature on this topic

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

I think the evidence on this topic is very strong

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

I feel very confident in choosing between both treatment options

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

In my opinion, *meniscectomy* is a good option for the initial treatment of patients between 45 and 70 years old with a non-obstructive meniscal tear

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

In my opinion, *physical therapy* is a good option for the initial treatment of patients between 45 and 70 years old with a non-obstructive meniscal tear

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

Patient profiles 1-20 (example profile 1)

In the next section you will be presented 20 patient profiles. On the basis of every profile we kindly ask you to answer 3 questions.

Explanation of terms and abbreviations used

Complaints

- Pseudo-locking
  - › A "catching" sensation which inhibits moving but quickly disappears. No true (irretrievable) knee locking.

- VAS pain (0-100)
  - › Visual analogue scale to express pain. A score of 0 corresponds to no pain at all and a score of 100 reflects the worst possible pain. This score is divided into a score at rest and during weight-bearing.
- IKDC (0-100)
  - › International Knee Documentation Committee (IKDC) 'Subjective Knee Form'. The maximum of 100 points reflects no restrictions in daily and sports activities and the absence of symptoms. A change of more than 8.8 points in IKDC score is considered to exceed the smallest detectable change. The mentioned score in the patient profile corresponds to the baseline score (before treatment).

Physical exam

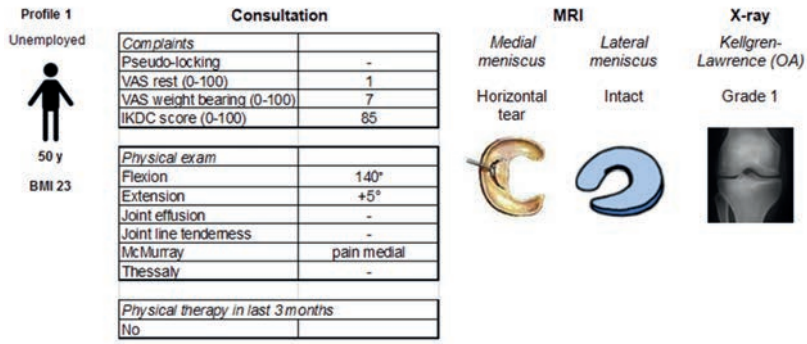
- Flexion of the knee; in degrees.
- Extension of the knee; in degrees. +5° means hyperextension, and -5° means a limitation of extension.
- Joint effusion; the presence of joint effusion. + means yes, - means no.
- Joint line tenderness; the presence of medial or lateral joint line tenderness during compression.
- McMurray; expressed as the presence of pain, a click, or both.
- Thessaly test; expressed for 5° or 20° knee flexion. If the Thessaly test is positive (e.g. pain) in 5°, it is assumed that it is positive in 20°. If 20° is presented, the test in 5° was negative.

MRI

All MRIs were viewed by 1 radiologist and tears were classified according to the ISAKOS classification.

X-ray (OA)

Osteoarthritis on Kellgren-Lawrence scale. In this study, patients with grade 4 were excluded. The miniature picture is a standardised image from the classification and does not represent the authentic X-ray of the described patient.



Would you prefer meniscectomy or physical therapy as treatment in this particular patient?

- Arthroscopic partial meniscectomy (APM)
- Physical therapy (PT)

What would you think that will be the effect of your treatment of choice on knee function after two years?

- Strong deterioration (at least 20 points on IKDC)
- Mild deterioration (10-20 points on IKDC)
- No relevant difference (-10 to +10 points on IKDC)
- Some improvement (10-20 points on IKDC)
- Strong improvement (at least 20 points on IKDC)

What will the outcome be if the other treatment would be applied?

- Strong deterioration (at least 20 points on IKDC)
- Mild deterioration (10-20 points on IKDC)
- No relevant difference (-10 to +10 points on IKDC)
- Some improvement (10-20 points on IKDC)
- Strong improvement (at least 20 points on IKDC)

Patient characteristics

On this page we would like to ask you which of the following patient characteristics affect your treatment choice and if so, in which direction?

- Younger patients (approximately <45 years)
  - › APM
  - › PT
  - › Unaffected

The same lay-out was used for all following characteristics

- Older patients (approximately >45 years)
- Normal BMI (18,5–25 kg/m2)
- Obesity (BMI>25 kg/m2)
- Absence of obstructive/locking complaints
- Presence of obstructive/locking complaints
- Medial tear
- Lateral tear
- Longitudinal-vertical (ISAKOS)
- Horizontal tear (ISAKOS)
- Radial tear (ISAKOS)
- Vertical flap tear (ISAKOS)

- Complex tear (ISAKOS)
- Bucket handle tear
- No-mild osteoarthritis (Kellgren Lawrence 0–2)
- Moderate-severe osteoarthritis (Kellgren Lawrence 3–4)
- Lower education level
- Higher education level
- Poor baseline physical function (IKDC approximately <30)
- Good baseline physical function (IKDC approximately >50)
- Low activity level (Tegner 1–3)
- High activity levels (Tegner >3)
- Lower levels of pain (VAS <7)
- Higher levels of pain (VAS >7)
- Male gender
- Female gender
- A patient's wish for practicing sports
- Traumatic aetiology
- Degenerative aetiology
- Failed conservative treatment

Appendix 23 Characteristics of surgeons

Characteristic	Number	Percentage
Years of experience as orthopaedic surgeon		
Less than 5 years	35	18.0
Between 5 and 15 years	57	29.4
More than 15 years	71	36.6
Resident orthopaedic surgery	30	15.5
Estimated number of performed knee arthroscopies during career:		
Less than 10	5	2.6
Less than 50	21	10.8
Between 50 and 150	33	17.0
More than 150	134	69.1
Fields of expertise (more options possible):		
Knee surgery	146	75.3
Hip surgery	107	55.2
Shoulder/elbow surgery	49	25.3
Back surgery	18	9.3
Foot/ankle surgery	35	18.0
Trauma surgery	79	40.7
Arthroscopic surgery	75	38.7
Sports injuries	69	35.6
Infectiology	11	5.7
Paediatric surgery	5	2.6
Hand/wrist surgery	5	2.6

Appendix 24 Opinion towards the literature

In clinical decision making, what relative weighting do you give to evidence from your personal experience and of those around you compared to evidence from clinical research?	Chosen by all surgeons (%)	Chosen by experienced[a] knee surgeons (%)	Chosen by other orthopaedic surgeons (%)
Experience based	0.5	0	1.1
	12.4	12.9	11.8
	38.1	42.6	33.3
	44.3	37.6	51.6
olished clinical research	4.6	6.9	2.2
I'm completely up to date with literature on this topic			
Strongly disagree	0.5	0	1.1
Disagree	5.2	0	10.8
Neither agree nor disagree	17	13.9	20.4
Agree	64.4	71.3	57
Strongly agree	12.9	14.9	10.8
The evidence on this topic is convincing			
Strongly disagree	0.5	0	1.1
Disagree	8.2	11.9	4.3
Neither agree nor disagree	17	14.9	19.4
Agree	63.9	63.4	64.5
Strongly agree	10.3	9.9	10.8
I feel very confident in choosing between both treatment options			
Strongly disagree	1	0	2.2
Disagree	6.7	5	8.6
Neither agree nor disagree	16.5	10.9	22.6
Agree	65.5	71.3	59.1
Strongly agree	10.3	12.9	7.5
In my opinion. meniscectomy is a good option for the initial treatment of patients between 45 and 70 years old with a non-obstructive meniscal tear			
Strongly disagree	39.2	38.6	39.8
Disagree	49.5	47.5	51.6
Neither agree nor disagree	8.2	9.9	6.5
Agree	3.1	4	2.2
Strongly agree	0	0	0
In my opinion. physical therapy is a good option for the initial treatment of patients between 45 and 70 years old with a non-obstructive meniscal tear			
Strongly disagree	1.5	2	1.1
Disagree	7.2	10.9	3.2
Neither agree nor disagree	13.9	16.8	10.8
Agree	48.5	48.5	48.4
Strongly agree	28.9	21.8	36.6

[a] Experienced knee surgeons are orthopaedic surgeons with at least 5 years of experience in knee surgery.

Appendix 25 Results of the treatment choice affecting patient characteristics

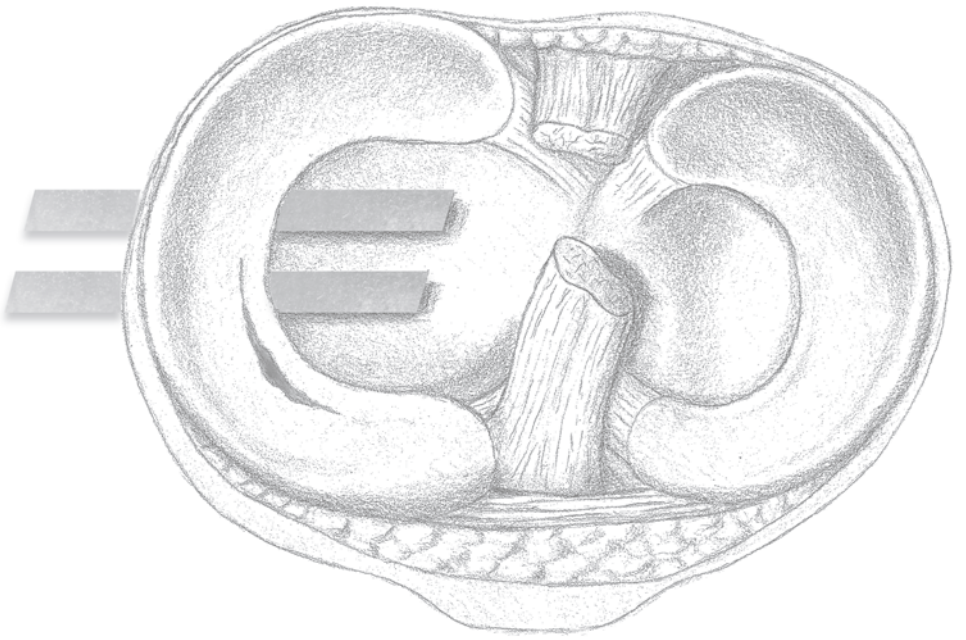
Patient characteristics	Proportion of surgeons directed towards APM (%)	Proportion of surgeons directed towards PT (%)	Proportion of surgeons for who characteristic did not influence treatment choice (%)
Higher education level	6	6	88
Lower education level	2	11	88
Male gender	10	5	85
Female gender	2	14	84
Medial tear	21	7	72
Lateral tear	11	26	63
A patient's wish for practicing sports	33	8	59
Normal BMI (18.5–25 kg/m2)	39	5	56
Higher levels of pain (VAS >7)	29	15	56
Radial tear (ISAKOS)	26	20	54
Longitudinal-vertical (ISAKOS)	36	10	54
Good baseline physical function (IKDC approximately >50)	30	19	52
Horizontal tear (ISAKOS)	7	42	51
Horizontal flap tear (ISAKOS)	38	16	46
High activity levels (Tegner >3)	49	6	45
Lower levels of pain (VAS <7)	5	50	45
Vertical flap tear (ISAKOS)	51	7	43
Complex tear (ISAKOS)	26	31	43
Low activity level (Tegner 1–3)	1	63	36
Poor baseline physical function (IKDC approximately <30)	10	56	34
No-mild osteoarthritis (KL 0–2)	61	8	31
Younger patients (approximately <45 years)	74	1	25
Traumatic aetiology	76	3	22
Obesity (BMI>25 kg/m2)	2	79	20
Presence of obstructive/locking complaints	82	3	15
Failed conservative treatment	82	3	15
Older patients (approximately >45 years)	1	87	13
Absence of obstructive/locking complaints	1	88	11

Appendix 25. (continued)

Patient characteristics	Proportion of surgeons directed towards APM (%)	Proportion of surgeons directed towards PT (%)	Proportion of surgeons for who characteristic did not influence treatment choice (%)
Degenerative aetiology	1	92	7
Bucket handle tear	94	0	6
Moderate-severe osteoarthritis (KL 3–4)	1	96	3

The percentages were generated based on information from the section '*Patient characteristics*' from the survey (Supplement 2).

Abbreviations: BMI, Body Mass Index; IKDC, International Knee Documentation Committee; ISAKOS, International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine; KL, Kellgren Lawrence; VAS, Visual Analogue Scale.



11

References

Dutch summary /
Nederlandse samenvatting

Acknowledgements /
Dankwoord

Curriculum vitae

References

1. Fox AJ, Bedi A, Rodeo SA. The basic science of human knee menisci: structure, composition, and function. Sports Health. 2012;4(4):340-51.
2. Herwig J, Egner E, Buddecke E. Chemical changes of human knee joint menisci in various stages of degeneration. Ann Rheum Dis. 1984;43(4):635-40.
3. Katz JN, Brophy RH, Chaisson CE, de Chaves L, Cole BJ, Dahm DL, Donnell-Fink LA, Guermazi A, Haas AK, Jones MH, Levy BA, Mandl LA, Martin SD, Marx RG, Miniaci A, Matava MJ, Palmisano J, Reinke EK, Richardson BE, Rome BN, Safran-Norton CE, Skoniecki DJ, Solomon DH, Smith MV, Spindler KP, Stuart MJ, Wright J, Wright RW, Losina E. Surgery versus physical therapy for a meniscal tear and osteoarthritis. N Engl J Med. 2013;368(18):1675-84.
4. Sihvonen R, Paavola M, Malmivaara A, Itala A, Joukainen A, Nurmi H, Kalske J, Järvinen TL; Finnish Degenerative Meniscal Lesion Study Group. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. N Engl J Med. 2013;369(26):2515-24.
5. Thorlund JB, Holsgaard-Larsen A, Creaby MW, Jorgensen GM, Nissen N, Englund M, Lohmander LS. Changes in knee joint load indices from before to 12 months after arthroscopic partial meniscectomy: a prospective cohort study. Osteoarthritis Cartilage. 2016;24(7):1153-9.
6. McDermott ID, Amis AA. The consequences of meniscectomy. J Bone Joint Surg Br. 2006;88(12):1549-56.
7. Hall M, Wrigley TV, Metcalf BR, Cicuttini FM, Wang Y, Hinman RS, Dempsey AR, Mills PM, Lloyd DG, Bennell KL. Do Moments and Strength Predict Cartilage Changes after Partial Meniscectomy? Med Sci Sports Exerc. 2015;47(8):1549-56.
8. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med. 2002;347(2):81-8.
9. Kim S, Bosque J, Meehan JP, Jamali A, Marder R. 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(11):994-1000.
10. Lazic S, Boughton O, Hing C, Bernard J. Arthroscopic washout of the knee: a procedure in decline. Knee. 2014;21(2):631-4.
11. Järvinen TL, Sihvonen R, Englund M. Arthroscopy for degenerative knee--a difficult habit to break? Acta Orthop. 2014;85(3):215-7.
12. Katz JN, Brownlee SA, Jones MH. The role of arthroscopy in the management of knee osteoarthritis. Best Pract Res Clin Rheumatol. 2014;28(1):143-56.
13. Crossley KM, Kemp JL, Ratzlaff C, Roos EM. Time to stop meniscectomies for degenerative tears: Practice must catch up with evidence [Blog]: Br J Sports Med. Published online: June 15, 2014.
14. Thorlund JB, Hare KB, Lohmander LS. Large increase in arthroscopic meniscus surgery in the middle-aged and older population in Denmark from 2000 to 2011. Acta Orthop. 2014;85(3):287-92.
15. Englund M, Roemer FW, Hayashi D, Crema MD, Guermazi A. Meniscus pathology, osteoarthritis and the treatment controversy. Nat Rev Rheumatol. 2012;8(7):412-9.
16. Cullen KA, Hall MJ, Golosinskiy A. Ambulatory surgery in the United States, 2006. Natl Health Stat Report. 2009(11):1-25.

17. Englund M, Guermazi A, Gale D, Hunter DJ, Aliabadi P, Clancy M, Felson DT. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. N Engl J Med. 2008;359(11):1108-15.

18. Hunter DJ, Zhang YQ, Niu JB, Tu X, Amin S, Clancy M, Guermazi A, Grigorian M, Gale D, Felson DT. The association of meniscal pathologic changes with cartilage loss in symptomatic knee osteoarthritis. Arthritis Rheum. 2006;54(3):795-801.

19. Herrlin S, Hallander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. Knee Surg Sports Traumatol Arthrosc. 2007;15(4):393-401.

20. Rongen JJ, van Tienen TG, Buma P, Hannink G. Meniscus surgery is still widely performed in the treatment of degenerative meniscus tears in The Netherlands. Knee Surg Sports Traumatol Arthrosc. 2018;26(4):1123-9.

21. Yim JH, Seon JK, Song EK, Choi JI, Kim MC, Lee KB, Seo HY. A comparative study of meniscectomy and nonoperative treatment for degenerative horizontal tears of the medial meniscus. Am J Sports Med. 2013;41(7):1565-70.

22. Herrlin SV, Wange PO, Lapidus G, Hallander M, Werner S, Weidenhielm L. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. Knee Surg Sports Traumatol Arthrosc. 2013;21(2):358-64.

23. Elattrache N, Lattermann C, Hannon M, Cole B. New England journal of medicine article evaluating the usefulness of meniscectomy is flawed. Arthroscopy. 2014;30(5):542-3.

24. Prasad V, Vandross A, Toomey C, Cheung M, Rho J, Quinn S, Chacko SJ, Borkar D, Gall V, Selvaraj S, Ho N, Cifu A. A decade of reversal: an analysis of 146 contradicted medical practices. Mayo Clin Proc. 2013;88(8):790-8.

25. Smith QW, Street RL Jr., Volk RJ, Fordis M. Differing levels of clinical evidence: exploring communication challenges in shared decision making. Introduction. Med Care Res Rev. 2013;70(1 Suppl):3S-13S.

26. Palmer S, Byford S, Raftery J. Economics notes: types of economic evaluation. BMJ. 1999;318(7194):1349.

27. Parkinson B. Pharmaceutical Policy in Australia. Sydney: Centre for Health Economic Research and Evaluation; University of Technology; 2013.

28. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. New York: Oxford University Press; 2005 2005.

29. Higuchi H, Kimura M, Shirakura K, Terauchi M, Takagishi K. Factors affecting long-term results after arthroscopic partial meniscectomy. Clin Orthop Relat Res. 2000(377):161-8.

30. Mesiha M, Zurakowski D, Soriano J, Nielson JH, Zarins B, Murray MM. Pathologic characteristics of the torn human meniscus. Am J Sports Med. 2007;35(1):103-12.

31. van Arkel ERA, van Essen A, Koëter S. Nederlandse Orthopaedische Vereniging. Richtlijn Artroscopie van de knie: indicatie en behandeling. Dutch Orthopedic Association guideline 2010. Available from www.orthopeden.org.

32. Khan M, Evaniew N, Bedi A, Ayeni OR, Bhandari M. Arthroscopic surgery for degenerative tears of the meniscus: a systematic review and meta-analysis. CMAJ. 2014;186(14):1057-64.

33. Thorlund JB, Juhl CB, Roos EM, Lohmander LS. Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. BMJ. 2015;350:h2747.

34. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-

analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.

35. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.
36. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
37. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer T, Varonen H, Vist GE, Williams JW Jr., Zaza S. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490.
38. Takeshima N, Sozu T, Tajika A, Ogawa Y, Hayasaka Y, Furukawa TA. Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? BMC Med Res Methodol. 2014;14:30.
39. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.
40. Stensrud S, Risberg MA, Roos EM. Effect of exercise therapy compared with arthroscopic surgery on knee muscle strength and functional performance in middle-aged patients with degenerative meniscus tears: a 3-mo follow-up of a randomized controlled trial. Am J Phys Med Rehabil. 2015;94(6):460-73.
41. Østeras H, Østeras B, Torstensen TA. Medical exercise therapy, and not arthroscopic surgery, resulted in decreased depression and anxiety in patients with degenerative meniscus injury. J Bodyw Mov Ther. 2012;16(4):456-63.
42. Ebert JR, Smith A, Wood DJ, Ackland TR. A comparison of the responsiveness of 4 commonly used patient-reported outcome instruments at 5 years after matrix-induced autologous chondrocyte implantation. Am J Sports Med. 2013;41(12):2791-9.
43. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol. 2000;27(11):2635-41.
44. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003;1:64.
45. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, Bombardier C, Felson D, Hochberg M, van der Heijde D, Dougados M. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis. 2005;64(1):29-33.
46. Andersson-Molina H, Karlsson H, Rockborn P. Arthroscopic partial and total meniscectomy: A long-term follow-up study with matched controls. Arthroscopy. 2002;18(2):183-9.
47. Englund M. Meniscal tear--a feature of osteoarthritis. Acta Orthop Scand Suppl. 2004;75(312):1-45, backcover.
48. Meredith DS, Losina E, Mahomed NN, Wright J, Katz JN. Factors predicting functional and radiographic outcomes after arthroscopic partial meniscectomy: a review of the literature. Arthroscopy. 2005;21(2):211-23.

49. Mills PM, Wang Y, Cicuttini FM, Stoffel K, Stachowiak GW, Podsiadlo P, Lloyd DG. Tibio-femoral cartilage defects 3-5 years following arthroscopic partial medial meniscectomy. Osteoarthritis Cartilage. 2008;16(12):1526-31.

50. Petty CA, Lubowitz JH. Does arthroscopic partial meniscectomy result in knee osteoarthritis? A systematic review with a minimum of 8 years' follow-up. Arthroscopy. 2011;27(3):419-24.

51. Roos H, Lauren M, Adalberth T, Roos EM, Jonsson K, Lohmander LS. Knee osteoarthritis after meniscectomy: prevalence of radiographic changes after twenty-one years, compared with matched controls. Arthritis Rheum. 1998;41(4):687-93.

52. Salata MJ, Gibbs AE, Sekiya JK. A systematic review of clinical outcomes in patients undergoing meniscectomy. Am J Sports Med. 2010;38(9):1907-16.

53. Williams RJ 3rd, Warner KK, Petrigliano FA, Potter HG, Hatch J, Cordasco FA. MRI evaluation of isolated arthroscopic partial meniscectomy patients at a minimum five-year follow-up. HSS J. 2007;3(1):35-43.

54. Chatain F, Robinson AH, Adeleine P, Chambat P, Neyret P. The natural history of the knee following arthroscopic medial meniscectomy. Knee Surg Sports Traumatol Arthrosc. 2001;9(1):15-8.

55. Crevoisier X, Munzinger U, Drobny T. Arthroscopic partial meniscectomy in patients over 70 years of age. Arthroscopy. 2001;17(7):732-6.

56. Roos EM, Ostenberg A, Roos H, Ekdahl C, Lohmander LS. 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(4):316-24.

57. Mayr HO, Rueschenschmidt M, Seil R, Dejour D, Bernstein A, Suedkamp N, Stoehr A. Indications for and results of arthroscopy in the arthritic knee: a European survey. Int Orthop. 2013;37(7):1263-71.

58. Sihvonen R, Englund M, Turkiewicz A, Järvinen TL; Finnish Degenerative Meniscal Lesion Study Group. Mechanical Symptoms and Arthroscopic Partial Meniscectomy in Patients With Degenerative Meniscus Tear: A Secondary Analysis of a Randomized Trial. Ann Intern Med. 2016;164(7):449-55.

59. Losina E, Dervan EE, Paltiel AD, Dong Y, Wright RJ, Spindler KP, Mandl LA, Jones MH, Marx RG, Safran-Norton CE, Katz JN. Defining the Value of Future Research to Identify the Preferred Treatment of Meniscal Tear in the Presence of Knee Osteoarthritis. PLoS One. 2015;10(6):e0130256.

60. Mather RC 3rd, Garrett WE, Cole BJ, Hussey K, Bolognesi MP, Lassiter T, Orlando LA. Cost-effectiveness analysis of the diagnosis of meniscus tears. Am J Sports Med. 2015;43(1):128-37.

61. van de Graaf VA. Early Surgery Versus Conservative Therapy for Meniscal Injuries in Older Patients (ESCAPE). Available from clinicaltrialsgov/show/NCT01850719. Registered May 9, 2013.

62. Katz JN, Chaisson CE, Cole B, Guermazi A, Hunter DJ, Jones M, Levy BA, Mandl LA, Martin S, Marx RG, Safran-Norton C, Roemer FW, Skoniecki D, Solomon DH, Spindler KP, Wright J, Wright RW, Losina E. The MeTeOR trial (Meniscal Tear in Osteoarthritis Research): rationale and design features. Contemp Clin Trials. 2012;33(6):1189-96.

63. Hare KB, Lohmander LS, Christensen R, Roos EM. Arthroscopic partial meniscectomy in middle-aged patients with mild or no knee osteoarthritis: a protocol for a double-blind, randomized sham-controlled multi-centre trial. BMC Musculoskelet Disord. 2013;14:71.

64. Järvinen TL. Development of knee osteoarthritis after arthroscopic partial resection of degenerative meniscus tear. Available from clinicaltrialsgov/show/NCT01052233. Registered January 19, 2010.

65. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45.

66. Biedert RM. Treatment of intrasubstance meniscal lesions: a randomized prospective study of four different methods. Knee Surg Sports Traumatol Arthrosc. 2000;8(2):104-8.

67. Erdil M, Bilsel K, Sungur M, Dikmen G, Tuncer N, Polat G, Elmadag NM, Tuncay I, Asik M. Does obesity negatively affect the functional results of arthroscopic partial meniscectomy? A retrospective cohort study. Arthroscopy. 2013;29(2):232-7.

68. Kijowski R, Woods MA, McGuine TA, Wilson JJ, Graf BK, De Smet AA. Arthroscopic partial meniscectomy: MR imaging for prediction of outcome in middle-aged and elderly patients. Radiology. 2011;259(1):203-12.

69. Kim JR, Kim BG, Kim JW, Lee JH, Kim JH. Traumatic and non-traumatic isolated horizontal meniscal tears of the knee in patients less than 40 years of age. Eur J Orthop Surg Traumatol. 2013;23(5):589-93.

70. Rathleff CR, Cavallius C, Jensen HP, Simonsen OH, Rasmussen S, Kaalund S, Østgaard SE. Successful conservative treatment of patients with MRI-verified meniscal lesions. Knee Surg Sports Traumatol Arthrosc. 2015;23(1):178-83.

71. Sihvonen R, Paavola M, Malmivaara A, Järvinen TL. 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(3).

72. van Meer BL, Meuffels DE, Vissers MM, Bierma-Zeinstra SM, Verhaar JA, Terwee CB, Reijman M. Knee injury and Osteoarthritis Outcome Score or International Knee Documentation Committee Subjective Knee Form: which questionnaire is most useful to monitor patients with an anterior cruciate ligament rupture in the short-term? Arthroscopy. 2013;29(4):701-15.

73. Bekkers JE, de Windt TS, Raijmakers NJ, Dhert WJ, Saris DB. Validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions. Osteoarthritis Cartilage. 2009;17(11):1434-9.

74. Engelhart L, Nelson L, Lewis S, Mordin M, Demuro-Mercon C, Uddin S, McLeod L, Cole B, Farr J. Validation of the Knee Injury and Osteoarthritis Outcome Score subscales for patients with articular cartilage lesions of the knee. Am J Sports Med. 2012;40(10):2264-72.

75. Greco NJ, Anderson AF, Mann BJ, Cole BJ, Farr J, Nissen CW, Irrgang JJ. Responsiveness of the International Knee Documentation Committee Subjective Knee Form in comparison to the Western Ontario and McMaster Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal articular cartilage defects. Am J Sports Med. 2010;38(5):891-902.

76. Briggs KK, Kocher MS, Rodkey WG, Steadman JR. 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(4):698-705.

77. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, Richmond JC, Shelborne KD. Development and validation of the International Knee Documentation Committee subjective knee form. Am J Sports Med. 2001;29(5):600-13.

78. Haverkamp D, Sierevelt IN, Breugem SJ, Lohuis K, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med. 2006;34(10):1680-4.

79. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther. 1998;28(2):88-96.

80. de Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. Health Qual Life Outcomes. 2008;6:16.

81. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15(12):1833-40.

82. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.

83. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494-502.

84. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury. 2011;42(3):236-40.

85. Tanner SM, Dainty KN, Marx RG, Kirkley A. Knee-specific quality-of-life instruments: which ones measure symptoms and disabilities most important to patients? Am J Sports Med. 2007;35(9):1450-8.

86. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol. 2006;59(10):1033-9.

87. World Health Organization. 54th World Health Assembly Geneva; International Classification of Functioning, Disability and Health: ICF. World health report. Geneva, Switzerland: WHO; 2001.

88. Roos EM, Roos HP, Lohmander LS. WOMAC Osteoarthritis Index--additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and MacMaster Universities. Osteoarthritis Cartilage. 1999;7(2):216-21.

89. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version. Scand J Med Sci Sports. 1998;8(6):439-48.

90. Englund M, Roos EM, Roos HP, Lohmander LS. Patient-relevant outcomes fourteen years after meniscectomy: influence of type of meniscal tear and size of resection. Rheumatology (Oxford). 2001;40(6):631-9.

91. Ericsson YB, Roos EM, Dahlberg L. Muscle strength, functional performance, and self-reported outcomes four years after arthroscopic partial meniscectomy in middle-aged patients. Arthritis Rheum. 2006;55(6):946-52.

92. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. Arthroscopy. 2007;23(8):839-44.

93. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes. 2003;1:17.

94. Ryser L, Wright BD, Aeschlimann A, Mariacher-Gehler S, Stucki G. A new look at the Western Ontario and McMaster Universities Osteoarthritis Index using Rasch analysis. Arthritis Care Res. 1999;12(5):331-5.

95. Sun Y, Sturmer T, Gunther KP, Brenner H. Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee--a review of the literature. Clin Rheumatol. 1997;16(2):185-98.

96. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development and testing of a reduced WOMAC function scale. J Bone Joint Surg Br. 2003;85(5):706-11.

97. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am. 2007;89(4):780-5.

98. van de Graaf VA, Wolterbeek N, Mutsaerts EL, Scholtes VA, Saris DB, de GA, Poolman RW. Arthroscopic Partial Meniscectomy or Conservative Treatment for Nonobstructive Meniscal Tears: A Systematic Review and Meta-analysis of Randomized Controlled Trials. Arthroscopy. 2016;32(9):1855-65.

99. Englund M, Guermazi A, Lohmander LS. The meniscus in knee osteoarthritis. Rheum Dis Clin North Am. 2009;35(3):579-90.

100. van de Graaf VA, Wolterbeek N, Scholtes VA, Mutsaerts EL, Poolman RW. Reliability and Validity of the IKDC, KOOS, and WOMAC for Patients With Meniscal Injuries. Am J Sports Med. 2014;42(6):1408-16.

101. VanderZee KI, Sanderman R, Heyink JW, de HH. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. Int J Behav Med. 1996;3(2):104-22.

102. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.

103. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res. 1985(198):43-9.

104. Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther. 1999;22(3):144-8.

105. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.

106. Lamers LM, Stalmeier PF, McDonnell J, Krabbe PF, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. Ned Tijdschr Geneeskd. 2005;149(28):1574-8.

107. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med. 1999;18(6):681-94.

108. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998;18(2 Suppl):S68-80.

109. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. Health Econ. 2004;13(5):405-15.

110. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemens B, Tiemens M, Zwieten L, van Rossum L, van Vré E, Tiemans B. Manual Trimbos/iMTA questionnaire for costs associated with psychiatric illness (TiC-P). Institute for Medical Technology Assessment, Erasmus Universiteit Rotterdam. 2002.

111. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. Pharmacoeconomics. 1996;10(5):460-6.

112. Tan SS, Bouwmans CA, Rutten FF, Hakkaart-van Roijen L. Update of the Dutch Manual for Costing in Economic Evaluations. Int J Technol Assess Health Care. 2012;28(2):152-8.

113. Steiner CA, Karaca Z, Moore BJ, Imshaug MC, Pickens G. Surgeries in hospital-based ambulatory surgery and hospital inpatient settings, 2014. Available from www.hcup-us.ahrq.gov/reports/statbriefs/sb223-Ambulatory-Inpatient-Surgeries-2014.pdf.

114. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev. 2015;1:CD004376.

115. Siemieniuk RAC, Harris IA, Agoritsas T, Poolman RW, Brignardello-Petersen R, Van de Velde S, Buchbinder R, Englund M, Lytvyn L, Quinlan C, Helsingen L, Knutsen G, Olsen NR, Macdonald H, Hailey L, Wilson HM, Lydiatt A, Kristiansen A. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: a clinical practice guideline. BMJ. 2017;357:j1982.

116. Kise NJ, Risberg MA, Stensrud S, Ranstam J, Engebretsen L, Roos EM. Exercise therapy versus arthroscopic partial meniscectomy for degenerative meniscal tear in middle-aged patients: randomised controlled trial with two year follow-up. BMJ. 2016;354:i3740.

117. Stahel PF, Wang P, Hutfless S, McCarty E, Mehler PS, Osgood GM, Makary MA. Surgeon Practice Patterns of Arthroscopic Partial Meniscectomy for Degenerative Disease in the United States: A Measure of Low-Value Care. JAMA Surg. 2018;153(5):494-6.

118. Anderson AF, Irrgang JJ, Kocher MS, Mann BJ, Harrast JJ. The International Knee Documentation Committee Subjective Knee Evaluation Form. Normative Data. Am J Sports Med. 2006;34(1):8.

119. Huskisson EC. Measurement of pain. Lancet. 1974;2(7889):1127-31.

120. Brittberg M, Winalski CS. Evaluation of cartilage injuries and repair. J Bone Joint Surg Am. 2003;85-A Suppl 2:58-69.

121. United Nations Educational Scientific and Cultural Organization (UNESCO) Institue for statistics. ISCED mappings; 2016. Available from www.uis.unesco.org/en/isced-mappings.

122. Stone JA, Salzler MJ, Parker DA, Becker R, Harner CD. Degenerative meniscus tears - assimilation of evidence and consensus statements across three continents: state of the art. Journal of ISAKOS. 2017;2(2):108-19.

123. Beaufils P, Becker R, Kopf S, Englund M, Verdonk R, Ollivier M, Seil R. Surgical management of degenerative meniscus lesions: the 2016 ESSKA meniscus consensus. Knee Surg Sports Traumatol Arthrosc. 2017;25(2):335-46.

124. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Neyret P, Richmond JC, Shelbourne KD. Responsiveness of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med. 2006;34(10):1567-73.

125. Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. J Shoulder Elbow Surg. 2009;18(6):927-32.

126. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum. 2001;45(4):384-91.

127. Järvinen TL, Guyatt GH. Arthroscopic surgery for knee pain. BMJ. 2016;354:i3934.

128. Rongen JJ, Govers TM, Buma P, Rovers MM, Hannink G. Arthroscopic meniscectomy for degenerative meniscal tears reduces knee pain but is not cost-effective in a routine health

care setting: a multi-center longitudinal observational study using data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2018;26(2):184-194.

129. Bosmans JE, de Bruijne MC, van Hout HP, Hermens ML, Ader HJ, van Tulder MW. Practical guidelines for economic evaluations alongside equivalence trials. Value Health. 2008;11(2):251-8.
130. Van de Graaf VA, Noorduyn JCA, Willigenburg NW, Butter IK, de Gast A, Mol BW, Saris DBF, Twisk JWR, Poolman RW; Escape Research Group. Effect of early surgery vs physical therapy on knee function among patients with non-obstructive meniscal tears: the ESCAPE randomized clinical trial. JAMA. 2018;320(13):1328-37.
131. van de Graaf VA, Scholtes VA, Wolterbeek N, Noorduyn JC, Neeter C, van Tulder MW, Saris DB, de Gast A, Poolman RW; ESCAPE Research Group. Cost-effectiveness of Early Surgery versus Conservative Treatment with Optional Delayed Meniscectomy for Patients over 45 years with non-obstructive meniscal tears (ESCAPE study): protocol of a randomised controlled trial. BMJ Open. 2016;6(12):e014381.
132. Versteegh M, Vermeulen M, Evers AA, de Wit GA, Prenger R, Stolk A. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52.
133. Hakkaart-van Roijen L, van der Linden N, Bouwmans CA, Kanters T, Tan SS. Bijlage 1: Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. Zorginstituut Nederland. 2015.
134. Nederlandse Zorgautoriteit. DBC Onderhoud; 2017. Available from www.opendisdata.nl.
135. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. Value Health. 2015;18(6):753-8.
136. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.
137. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. Health Econ. 2004;13(5):461-75.
138. Luo N, Johnson J, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. Med Care. 2010;48(4):365-71.
139. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res. 2005;14(6):1523-32.
140. van Dongen JM, van Wier MF, Tompa E, Bongers PM, van der Beek AJ, van Tulder MW, Bosmans JE. Trial-based economic evaluations in occupational health: principles, methods, and recommendations. J Occup Environ Med. 2014;56(6):563-72.
141. Shiell A, Donaldson C, Mitton C, Currie G. Health economic evaluation. J Epidemiol Community Health. 2002;56(2):85-8.
142. Katz JN, Losina E. The cost-effectiveness of arthroscopic partial meniscectomy: comparing apples and oranges. Osteoarthritis Cartilage. 2018;26(2):152-153.
143. Rongen JJ, Rovers MM, van Tienen TG, Buma P, Hannink G. Increased risk for knee replacement surgery after arthroscopic surgery for degenerative meniscal tears: a multi-center longitudinal observational study using data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2017;25(1):23-9.

144. MacNeil VJ, Eekhout I, Dijkgraaf MG, van HH, de Rooij SE, Heymans MW, Bosmans JE. Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? Eur J Health Econ. 2016;17(8):939-50.

145. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ. 2013;346:f1049.

146. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. BMJ. 2011;342:d1548.

147. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. New York Oxford University Press, editor2005.

148. Hakkaart-van Roijen L, Tan SS, Bouwmans CA. Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidzorg Diemen:: College voor Zorgverzekeringen; 2010 2010.

149. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. JAMA. 2003;290(18):2443-54.

150. Gauffin H, Tagesson S, Meunier A, Magnusson H, Kvist J. Knee arthroscopic surgery is beneficial to middle-aged patients with meniscal symptoms: a prospective, randomised, single-blinded study. Osteoarthritis Cartilage. 2014;22(11):1808-16.

151. Abram SGF, Middleton R, Beard DJ, Price AJ, Hopewell S. Patient-reported outcome measures for patients with meniscal tears: a systematic review of measurement properties and evaluation with the COSMIN checklist. BMJ Open. 2017;7(10):e017247.

152. Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther. 2016;20(2):105-13.

153. Harris JD, Brand JC, Cote MP, Faucett SC, Dhawan A. Research Pearls: The Significance of Statistics and Perils of Pooling. Part 1: Clinical Versus Statistical Significance. Arthroscopy. 2017;33(6):1102-12.

154. de Vet HC, Terluin B, Knol DL, Roorda LD, Mokkink LB, Ostelo RW, Hendriks EJ, Bouter LM, Terwee CB. Three ways to quantify uncertainty in individually applied "minimally important change" values. J Clin Epidemiol. 2010;63(1):37-45.

155. Devji T, Guyatt GH, Lytvyn L, Brignardello-Petersen R, Foroutan F, Sadeghirad B, Buchbinder R, Poolman RW, Harris IA, Carrasco-Labra A, Siemieniuk RAC, Vandvik PO. Application of minimal important differences in degenerative knee disease outcomes: a systematic review and case study to inform BMJ Rapid Recommendations. BMJ Open. 2017;7(5):e015587.

156. de Vet HCW, Terwee CB, Mokkink L, Knol D. Measurement in Medicine: A Practical Guide.: Cambridge University Press; 2011 2011.

157. van der Kloot WA, Oostendorp RA, van der Meij J, van den Heuvel J. [The Dutch version of the McGill pain questionnaire: a reliable pain questionnaire]. Ned Tijdschr Geneeskd. 1995;139(13):669-73.

158. Oak SR, Strnad GJ, Bena J, Farrow LD, Parker RD, Jones MH, Spindler KP. Responsiveness Comparison of the EQ-5D, PROMIS Global Health, and VR-12 Questionnaires in Knee Arthroscopy. Orthop J Sports Med. 2016;4(12):2325967116674714.

159. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407-15.

160. de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, Boers M, Bouter LM. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. Qual Life Res. 2007;16(1):131-42.

161. Terwee CB, Roorda LD, Knol DL, de Boer MR, de Vet HC. Linking measurement error to minimal important change of patient-reported outcomes. J Clin Epidemiol. 2009;62(10):1062-7.

162. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-57.

163. Tigerstrand Grevnerts H, Terwee CB, Kvist J. The measurement properties of the IKDC-subjective knee form. Knee Surg Sports Traumatol Arthrosc. 2015;23(12):3698-706.

164. van der Linde JA, van Kampen DA, van Beers LWAH, van Deurzen DFP, Saris DBF, Terwee CB. The Responsiveness and Minimal Important Change of the Western Ontario Shoulder Instability Index and Oxford Shoulder Instability Score. J Orthop Sports Phys Ther. 2017;47(6):402-10.

165. van der Wal RJP, Heemskerk BTJ, van Arkel ERA, Mokkink LB, Thomassen BJW. Translation and Validation of the Dutch Western Ontario Meniscal Evaluation Tool. J Knee Surg. 2017;30(4):314-22.

166. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res. 2003;12(4):349-62.

167. Abram SGF, Judge A, Beard DJ, Wilson HA, Price AJ. Temporal trends and regional variation in the rate of arthroscopic knee surgery in England: analysis of over 1.7 million procedures between 1997 and 2017. Has practice changed in response to new evidence? Br J Sports Med. 2019;53(24):1533-1538.

168. Frellick M. Skip Arthroscopy for Degenerative Knee Disease, Report Says. 2018. Available from www.medscape.com/viewarticle/879922.

169. Hohmann E, Glatt V, Tetsworth K, Cote M. Arthroscopic Partial Meniscectomy Versus Physical Therapy for Degenerative Meniscus Lesions: How Robust Is the Current Evidence? A Critical Systematic Review and Qualitative Synthesis. Arthroscopy. 2018;34(9):2699-708.

170. Johnson DC. Why I disagree with 2017 knee arthroscopy recommendations. MedStar Washington Hospital center, 2017. Available from www.blog.medstarwashington.org/2017/07/26/arthroscopic-knee-surgery-guidelines/.

171. Kartus JT. Editorial Commentary: Fresh Evidence That There Is Still No Reliable Evidence on How to Best Treat Knee Degenerative Meniscus Tears. Arthroscopy. 2018;34(9):2711-2.

172. Leopold SS. Editorial: Appropriate Use? Guidelines on Arthroscopic Surgery for Degenerative Meniscus Tears Need Updating. Clin Orthop Relat Res. 2017;475(5):1283-6.

173. Castor EDC. Castor Electronic Data Capture. 2019. Available from www.castoredc.com.

174. Van de Graaf VA, Van Dongen JM, Willigenburg NW, Noorduyn JC, Butter IK, De Gast A, Saris DB, Van Tulder MW, Poolman RW; Escape Research Group. How do the costs of physical therapy and arthroscopic partial meniscectomy compare? A trial-based economic evaluation of two treatments in patients with meniscal tears alongside the ESCAPE study. Br J Sports Med. 2019;Jun 21. pii: bjsports-2018-100065 [Epub ahead of print].

175. Hoffmann TC, Del Mar C. Clinicians' Expectations of the Benefits and Harms of Treatments, Screening, and Tests: A Systematic Review. JAMA Intern Med. 2017;177(3):407-19.

176. Kolk S, Cox K, Weerdesteyn V, Hannink G, Bramer J, Dijkstra S, Jutte P, Ploegmakers J, van de Sande M, Schreuder H, Verdonschot N, van der Geest I. Can orthopedic oncologists predict functional outcome in patients with sarcoma after limb salvage surgery in the lower limb? A nationwide study. Sarcoma. 2014;2014:436598.

177. Hoher J, Bach T, Munster A, Bouillon B, Tiling T. Does the mode of data collection change results in a subjective knee score? Self-administration versus interview. Am J Sports Med. 1997;25(5):642-7.

178. Roos EM. Outcome after anterior cruciate ligament reconstruction--a comparison of patients' and surgeons' assessments. Scand J Med Sci Sports. 2001;11(5):287-91.

179. Thorlund JB, Juhl CB, Ingelsrud LH, Skou ST. Risk factors, diagnosis and non-surgical treatment for meniscal tears: evidence and recommendations: a statement paper commissioned by the Danish Society of Sports Physical Therapy (DSSF). Br J Sports Med. 2018;52(9):557-65.

180. Thorlund JB, Englund M, Christensen R, Nissen N, Pihl K, Jorgensen U, Schjerning J, Lohmander LS. Patient reported outcomes in patients undergoing arthroscopic partial meniscectomy for traumatic or degenerative meniscal tears: comparative prospective cohort study. BMJ. 2017;356:j356.

181. Dunn WR, Wolf BR, Harrell FE Jr., Reinke EK, Huston LJ, MOON Knee Group, Spindler KP. Baseline predictors of health-related quality of life after anterior cruciate ligament reconstruction: a longitudinal analysis of a multicenter cohort at two and six years. J Bone Joint Surg Am. 2015;97(7):551-7.

182. Sainio P, Martelin T, Koskinen S, Heliovaara M. Educational differences in mobility: the contribution of physical workload, obesity, smoking and chronic conditions. J Epidemiol Community Health. 2007;61(5):401-8.

183. Chen HY, Harris IA, Sutherland K, Levesque JF. A controlled before-after study to evaluate the effect of a clinician led policy to reduce knee arthroscopy in NSW. BMC Musculoskelet Disord. 2018;19(1):148.

184. Holtedahl R, Brox JI, Aune AK, Nguyen D, Risberg MA, Tjomsland O. Changes in the rate of publicly financed knee arthroscopies: an analysis of data from the Norwegian patient registry from 2012 to 2016. BMJ Open. 2018;8(6):e021199.

185. Eijgenraam SM, Reijman M, Bierma-Zeinstra SMA, van Yperen DT, Meuffels DE. Can we predict the clinical outcome of arthroscopic partial meniscectomy? A systematic review. Br J Sports Med. 2018;52(8):514-21.

186. Essilfie A, Kang HP, Mayer EN, Trasolini NA, Alluri RK, Weber AE. Are Orthopaedic Surgeons Performing Fewer Arthroscopic Partial Meniscectomies in Patients Greater Than 50 Years Old? A National Database Study. Arthroscopy. 2019;35(4):1152-9 e1.

187. Mattila VM, Sihvonen R, Paloneva J, Fellander-Tsai L. Changes in rates of arthroscopy due to degenerative knee disease and traumatic meniscal tears in Finland and Sweden. Acta Orthop. 2016;87(1):5-11.

188. Muheim LLS, Senn O, Fruh M, Reich O, Rosemann T, Neuner-Jehle SM. Inappropriate use of arthroscopic meniscal surgery in degenerative knee disease. Acta Orthop. 2017;88(5):550-5.

189. Bergkvist D, Dahlberg LE, Neuman P, Englund M. Knee arthroscopies: who gets them, what does the radiologist report, and what does the surgeon find? An evaluation from southern Sweden. Acta Orthop. 2016;87(1):12-6.

190. Yoon KH, Lee SH, Bae DK, Park SY, Oh H. Does varus alignment increase after medial meniscectomy? Knee Surg Sports Traumatol Arthrosc. 2013;21(9):2131-6.

191. Pengas IP, Assiotis A, Nash W, Hatcher J, Banks J, McNicholas MJ. Total meniscectomy in adolescents: a 40-year follow-up. J Bone Joint Surg Br. 2012;94(12):1649-54.

192. Englund M, Lohmander LS. Risk factors for symptomatic knee osteoarthritis fifteen to twenty-two years after meniscectomy. Arthritis Rheum. 2004;50(9):2811-9.

193. Papalia R, Del Buono A, Osti L, Denaro V, Maffulli N. Meniscectomy as a risk factor for knee osteoarthritis: a systematic review. Br Med Bull. 2011;99:89-106.

194. Collins LE, Losina E, Guernazi A, Katz JN. The risk of osteoarthritis progression after arthroscopic partial meniscectomy (APM); data from an RCT of APM vs. physical therapy. OARSI 2017: P. S59.

195. Abram SGF, Judge A, Beard DJ, Price AJ. Rates of knee arthroplasty within one-year of undergoing arthroscopic partial meniscectomy in England: temporal trends, regional and age-group variation in conversion rates. Osteoarthritis Cartilage. 2019;27(10):1420-9.

196. Rodriguez-Merchan EC, Garcia-Ramos JA, Padilla-Eguiluz NG, Gomez-Barrena E. Arthroscopic Partial Meniscectomy for Painful Degenerative Meniscal Tears in the Presence of Knee Osteoarthritis in Patients Older than 50 Years of Age: Predictors of an Early (1 to 5 Years) Total Knee Replacement. Arch Bone Jt Surg. 2018;6(3):203-11.

197. van de Graaf VA, Noorduyn JCA, Willigenburg NW, Butter IK, de Gast A, Mol BW, Saris DBF, Twisk JWR, Poolman RW; Escape Research Group. Effect of Early Surgery vs Physical Therapy on Knee Function Among Patients With Nonobstructive Meniscal Tears: The ESCAPE Randomized Clinical Trial. JAMA. 2018;320(13):1328-37.

198. Katz JN, Shrestha S, Losina E, Jones MH, Marx RG, Mandl LA, Levy BA, MacFarlane LA, Spindler KP, Silva GS, MeTe ORI, Collins JE. Five-year outcome of operative and non-operative management of meniscal tear in persons greater than 45 years old. Arthritis Rheumatol. 2019;Aug 20. [Epub ahead of print].

199. Pihl K, Ensor J, Peat G, Englund M, Lohmander S, Jorgensen U, Nissen N, Fristed JV, Thorlund JB. Wild goose chase - no predictable patient subgroups benefit from meniscal surgery: patient-reported outcomes of 641 patients 1 year after surgery. Br J Sports Med. 2020;54(1):13-22.

200. Kise NJ, Aga C, Engebretsen L, Roos EM, Tariq R, Risberg MA. Complex Tears, Extrusion, and Larger Excision Are Prognostic Factors for Worse Outcomes 1 and 2 Years After Arthroscopic Partial Meniscectomy for Degenerative Meniscal Tears: A Secondary Explorative Study of the Surgically Treated Group From the Odense-Oslo Meniscectomy Versus Exercise (OMEX) Trial. Am J Sports Med. 2019;47(10):2402-11.

201. Katz JN, Wright J, Spindler KP, Mandl LA, Safran-Norton CE, Reinke EK, Levy BA, Wright RW, Jones MH, Martin SD, Marx RG, Losina E. Predictors and Outcomes of Crossover to Surgery from Physical Therapy for Meniscal Tear and Osteoarthritis: A Randomized Trial Comparing Physical Therapy and Surgery. J Bone Joint Surg Am. 2016;98(22):1890-6.

202. Gowd AK, Lalehzarian SP, Liu JN, Agarwalla A, Christian DR, Forsythe B, Cole BJ, Verma NN. Factors Associated With Clinically Significant Patient-Reported Outcomes After Primary Arthroscopic Partial Meniscectomy. Arthroscopy. 2019;35(5):1567-75 e3.

203. Haviv B, Bronak S, Kosashvili Y, Thein R. Which patients are less likely to improve during the first year after arthroscopic partial meniscectomy? A multivariate analysis of 201 patients with prospective follow-up. Knee Surg Sports Traumatol Arthrosc. 2016;24(5):1427-31.

204. Roos EM, Hare KB, Nielsen SM, Christensen R, Lohmander LS. Better outcome from arthroscopic partial meniscectomy than skin incisions only? A sham-controlled randomised trial in patients aged 35-55 years with knee pain and an MRI-verified meniscal tear. BMJ Open. 2018;8(2):e019461.

205. Gauffin H, Sonesson S, Meunier A, Magnusson H, Kvist J. Knee Arthroscopic Surgery in Middle-Aged Patients With Meniscal Symptoms: A 3-Year Follow-up of a Prospective, Randomized Study. Am J Sports Med. 2017;45(9):2077-84.

206. Abram SGF, Hopewell S, Monk AP, Bayliss LE, Beard DJ, Price AJ. Arthroscopic partial meniscectomy for meniscal tears of the knee: a systematic review and meta-analysis. Br J Sports Med. 2019;Feb 22. pii: bjsports-2018-100223 [Epub ahead of print].

207. Chan AW, Hrobjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. Trials. 2018;19(1):116.

208. Tracey I. Getting the pain you expect: mechanisms of placebo, nocebo and reappraisal effects in humans. Nat Med. 2010;16(11):1277-83.

209. Lingard EA, Berven S, Katz JN, Kinemax Outcomes Group. Management and care of patients undergoing total knee arthroplasty: variations across different health care settings. Arthritis Care Res. 2000;13(3):129-36.

210. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. BMJ. 2018;363:k4245.

211. Pihl K, Roos EM, Nissen N, JoRgensen U, Schjerning J, Thorlund JB. Over-optimistic patient expectations of recovery and leisure activities after arthroscopic meniscus surgery. Acta Orthop. 2016;87(6):615-21.

Dutch summary / Nederlandse samenvatting

De behandeling van degeneratieve meniscusscheuren moet ernstig worden heroverwogen. Vanaf de beginjaren 2000 is de partiële meniscectomie de meest uitgevoerde operatie binnen de orthopedie. De eerste gerandomiseerde trial die de partiële meniscectomie vergeleek met een conservatieve behandeling, vond geen verschil tussen beide groepen in patiënt-gerapporteerde kniefunctie.[19] Deze studie vormde de basis voor de totstandkoming van dit proefschrift.

Het doel van dit proefschrift is om te bepalen of partiële meniscectomie of fysiotherapie doelmatiger is in de behandeling van degeneratieve meniscusscheuren.

Hoofdstuk 2 beschrijft de resultaten van een systematische literatuurstudie en een meta-analyse van level-1 studies (RCTs). Deze studies vergeleken de arthroscopische partiële meniscectomie met een conservatieve of placebo behandeling bij patiënten met een degeneratieve meniscusscheur, met kniefunctie als primaire uitkomst. Zes RCTs met in totaal 773 patiënten werden geïncludeerd. Het poolen van deze studies resulteerde in een kleine, statistisch significant betere uitkomst na partiële meniscectomie voor kniefunctie en knie pijn tot 6 maanden. Na 6 maanden werd geen verschil tussen de groepen gevonden.

Door het beperkte aantal studies, zoals beschreven in hoofdstuk 2, was een van de doelen van dit proefschrift om meer bewijslast te verkrijgen op het gebied van effectiviteit en doelmatigheid. Hiervoor is de Escape trial geïnitieerd. De in hoofdstuk 2 beschreven trials gebruiken verschillende meetinstrumenten als primaire uitkomstmaat. Het gebruik van verschillende meetinstrumenten maakt de interpretatie en het vergelijken van de onderzoeksresultaten complex. Volgens de COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) criteria dient de selectie van een meetinstrument gebaseerd te worden op de kwaliteit van diens meeteigenschappen. Er wordt aangeraden alleen het instrument met de beste eigenschappen te gebruiken. Hierdoor is bij de opzet van de Escape studie eerst een klinimetrische studie verricht. *Hoofdstuk 3* beschrijft de resultaten van dit onderzoek, waarin de meeteigenschappen "betrouwbaarheid" en "validiteit" van de Nederlandstalige versies van de International Knee Documentation Committee (IKDC), de Knee Injury and Osteoarthritis Outcome Score (KOOS), en de Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) vragenlijsten zijn vergeleken in een groep van 75 patiënten met een meniscusscheur. De IKDC bleek het meest betrouwbaar en valide van deze 3 vragenlijsten. Deze resultaten impliceren dat de IKDC het meest geschikt is voor de evaluatie van patiënten met een meniscusscheur.

Hoofstuk 4 beschrijft het onderzoeksprotocol van de Escape trial, een multicenter gerandomiseerd doelmatigheidsonderzoek naar de behandeling van degeneratieve meniscusscheuren. Een gepubliceerd onderzoeksprotocol biedt de mogelijkheid om een volledig overzicht te geven van de gebruikte methoden in een studie. Dit draagt bij

aan objectieve beoordeling van selectief rapporteren en interne en externe validatie, hetgeen essentieel is voor transparantie in onderzoek.[207]

In *hoofdstuk 5* worden de resultaten beschreven van de klinische uitkomsten van de Escape trial. Het doel van dit onderzoek was om te bepalen of fysiotherapie niet-inferieur is (d.w.z. niet minder effectief) ten opzichte van partiële meniscectomie in het verbeteren van zelf-gerapporteerde kniefunctie bij patiënten met een meniscusscheur. Tussen november 2013 en november 2015 werden in 9 ziekenhuizen 321 patiënten geïncludeerd en willekeurig ingedeeld voor fysiotherapie of partiële meniscectomie. Op basis van de resultaten van hoofdstuk 3 werd gekozen voor de IKDC als primaire uitkomstmaat, waarbij gekozen werd voor een non-inferiority drempel van 8 punten op de IKDC. Over een periode van 24 maanden verbeterde beide groepen significant in knie functioneren, met een verschil van 3.6 punten tussen beide groepen (97.5% CI, -∞ tot 6.5; p-waarde voor non-inferiority = .001). 29% van de patiënten die geloot werden voor fysiotherapie, herstelde onvoldoende en werden gedurende de looptijd van de studie alsnog geopereerd. Op basis van deze resultaten concluderen we dat fysiotherapie niet-inferieur is aan partiële meniscectomie voor het verbeteren van knie functioneren gedurende 2 jaar follow-up, en daarmee een goed alternatief is voor partiële meniscectomie bij patiënten met een degeneratieve meniscusscheur.

Hoofdstuk 6 beschrijft de resultaten van de economische analyse van de Escape studie. In dezelfde 321 patiënten als beschreven in *hoofdstuk 5*, bleek fysiotherapie veel goedkoper (€1803; 95% CI, −€3008, −€838) maar ook iets minder effectief (4.0 punten op de IKDC; 95% CI, -8.3–0.2 / 0.029 QALYs; 95% CI, -0.074–0.016) dan partiële meniscectomie na 24 maanden. Daarmee is de waarschijnlijkheid dat fysiotherapie kosteneffectief is hoog, tegen een acceptabele betalingsbereidheid (willingness to pay) voor zowel knie functioneren (IKDC) en voor kwaliteit aangepaste levensjaren (QALYs). Deze bevindingen komen overeen met de resultaten uit *hoofdstuk 5* en impliceren dat een verdere afname van het aantal partiële meniscectomieën gerechtvaardigd is.

Om de in *hoofdstuk 5 en 6* gevonden behandeleffecten te evalueren en de bevindingen gedegen te kunnen interpreteren, hebben we in *hoofdstuk 7* de meeteigenschappen "responsiviteit" en "de minimale belangrijke verandering" (minimal important change, MIC) berekend vanuit de Escape trial. Data van 298 patiënten waren hiervoor bruikbaar. Uit de analyses bleek dat de IKDC goed in staat is om het verschil in kniefunctie te meten bij patiënten in de behandeling van een degeneratieve meniscusscheur. De MIC bleek 10.9 punten op de IKDC. Deze was nog niet beschreven in deze patiëntgroep. Daardoor was de non-inferiority drempel in de Escape studie gebaseerd op het kleinst detecteerbare verschil (smallest detectable change, SDC), namelijk 8 punten op de IKDC. Het feit dat de minimale belangrijke verandering (MIC = 10.9 punten) hoger is dan het kleinst detecteerbare verschil gebruikt in onze power berekening (SDC = 8 punten), versterkt de conclusies van hoofdstukken 5 en 6.

Ondanks de groeiende bewijslast tegen de partiële meniscectomie, daalt het aantal meniscectomieën minder hard dan verwacht. Een frequent gehoord argument

hiervoor is de overtuiging van orthopedisch chirurgen dat zij denken vooraf te kunnen voorspellen welke patiënten meer gebaat zijn bij een operatie. In *hoofdstuk 8* hebben we dit onderzocht door een survey af te nemen onder Nederlandse en Australische orthopedisch chirurgen. Het doel van deze survey was om te onderzoeken of orthopedisch chirurgen in staat zijn om de uitkomst van conservatieve en operatieve behandeling van degeneratieve meniscusscheuren te voorspellen. Deelnemende chirurgen werden gevraagd om van 20 patiëntprofielen een behandelvoorkeur te geven en vervolgens een voorspelling te doen omtrent de uitkomst van beide behandelingen. Orthopedisch chirurgen bleken slechts in 50% van de gevallen in staat te zijn de uitkomst juist te voorspellen, wat overeenkomt met de kans wanneer je een munt opgooit. Dit bleek onafhankelijk van klinische ervaring. Deze resultaten weerspreken daarmee het veelgehoorde argument dat orthopedisch chirurgen in staat zijn om de uitkomst te voorspellen van de behandeling van degeneratieve meniscusscheuren.

Conclusies

1. De IKDC vragenlijst is een betrouwbare, valide en responsieve vragenlijst in de evaluatie van de behandeling van meniscusscheuren.
2. Fysiotherapie is niet-inferieur aan, en kosten-effectiever dan, partiële meniscectomie in de behandeling van patiënten met een degeneratieve meniscusscheur.
3. Orthopedisch chirurgen blijken slecht in staat om de behandeluitkomst te voorspellen van patiënten met een degeneratieve meniscusscheur en zouden hun behandelkeuze meer moeten baseren op de resultaten uit de literatuur dan op hun eigen ervaring.
4. Fysiotherapie is de behandeling van eerste keus bij patiënten met een niet-obstructieve degeneratieve meniscusscheur.
5. Onnodige operaties zullen blijven worden uitgevoerd, zolang 1) orthopedisch chirurgen niet overtuigd zijn van de bevindingen uit de literatuur, 2) de richtlijnen geen uniforme consensus bereiken, en 3) ons gezondheidszorgsysteem geen onderscheid maakt tussen het financieren van doelmatige en niet-doelmatige zorg.
6. Toekomstig onderzoek moet zich richten op de subgroep van patiënten die onvoldoende verbeteren na een behandeling met fysiotherapie, om zo het aantal (onnodige) operaties verder te kunnen reduceren.

Daarom luidt de conclusie van dit proefschrift:

Fysiotherapie, in vergelijking met een arthroscopische partiële meniscectomie, leidt tot doelmatigere zorg in de behandeling van degeneratieve meniscusscheuren.

Acknowledgements / Dankwoord

Dit proefschrift, of beter gezegd: dit levenswerk, had nooit gerealiseerd kunnen worden zonder de hulp van velen. Mijn dank is groot aan iedereen die heeft bijgedragen aan de totstandkoming van dit proefschrift. In het bijzonder wil ik alle deelnemers aan de Escape studie bedanken. Hoe mooi een idee ook kan zijn, zonder uw bijdrage zou dit alles slechts bij een idee gebleven zijn. Dank voor het keer op keer invullen van de (eindeloze) vragenlijsten. Daarnaast wil ik ook een woord van dank richten tot de deelnemende ziekenhuizen; het AUMC (locatie AMC), het Diakonessenhuis Utrecht, Noordwest ziekenhuisgroep (locatie Alkmaar), het Elisabeth Tweesteden ziekenhuis (locatie Tilburg), het Tergooi (locatie Hilversum), het medisch centrum Jan van Goyen, OLVG Amsterdam (locatie Oost én West), en Xpert Orthopedie (het Slotervaart ziekenhuis); zonder uw inzet bij de inclusies, follow-up, en dataverzameling was dit proefschrift nooit tot stand gekomen.

Daarnaast wil ik nog een aantal mensen in het bijzonder bedanken:

Geachte professor Saris, beste Daniël, de eerste opzet voor dit project begon al 8 jaar geleden, en vanaf het begin heb je me gesteund en de vrijheid gegeven om aan dit lange proces mijn eigen invulling te geven. Hiervoor ben ik je zeer dankbaar. Dank voor al je hulp, feedback, overleg momenten, en bovenal dat je mijn promotor wilde zijn. Zonder jouw hulp was dit niet gelukt.

Beste Rudolf, dit bijzondere project is al in 2011 tot stand gekomen toen ik als oudste co onderzoek wilde doen, en bij wie kon ik beter aankloppen dan bij de kersverse opleider en hoofd van JointResearch. Je legde direct mijn onderzoek bij de uittredend opleider stil en je opdracht aan mij was: "Ga de literatuur over de meniscus maar eens op een rij zetten". En kijk waar dit toe geleid heeft. Veel dank voor al je hulp, inspiratie, humor, onze bila's, en dat je me keer op keer weer on-track kreeg als ik weer op een zijspoor raakte. Ik ben zeer vereerd dat je dit traject als mede-promotor afsluit.

Beste Arthur, ook jij bent vanaf het eerste uur bij dit project betrokken. Dank voor al je hulp en bevlogenheid. Ik waardeer je persoonlijke touch en heb genoten van onze fietstochten en onze memorabele rit over de A2.

Beste Nienke, als laatste ben jij toegetreden tot mijn promotiecommissie, maar zonder jouw hulp had dit boekwerk er heel anders uitgezien. Dan voor je kritische en vooral gedetailleerde revisies, keer op keer op keer, en dank dat je me steeds weer wist te motiveren als ik het eigenlijk wel weer goed vond (lees: er helemaal klaar mee was).

Beste BW (Mol), onlosmakelijk ben je verbonden geraakt aan dit project, en keer op keer dook jij steeds weer op. Wat een belangrijke en onvergetelijke bijdrage heb jij geleverd aan dit proefschrift. Vanaf de ZonMw aanvraag tot aan de mooiste publicatie van allemaal, jouw bijdrage is hier essentieel in geweest en ik durf te stellen (jammer

dat dit proefschrift geen stellingen kent), dat dit zonder jouw hulp niet gelukt was. Nog steeds ben je betrokken bij onze meest recente projecten en ik hoop dat dit voorlopig zo zal blijven. Dank voor al je hulp!

Hooggeleerde leden van de beoordelingscommissie, prof. dr. L.P.H. Leenen, prof. dr. F.J.G. Backx, prof. dr. C. Veenhof, prof. dr. R.L. Diercks, veel dank voor de moeite die u genomen hebt dit manuscript te beoordelen en uiteraard het goed te keuren. Ik kijk met veel plezier en interesse uit naar 13 maart 2020. Dear prof. dr. A.J. Krych thank you for being part of the assessment committee. Such a great honour and I'm looking forward to March 13, 2020.

Julia, jij nam mijn plaats in op het moment dat ik de kliniek weer in ging. Vanaf het moment dat jij mijn werkzaamheden overnam, wist ik dat het goed zou komen. Je hebt al die tijd de volle 100% gegeven en bent cruciaal geweest voor de voltooiing van de Escape studie. Ontzettend knap hoe je dit de afgelopen jaren naast je functie als fysiotherapeut hebt gedaan en geweldig dat je de uit dit proefschrift voortkomende vraagstukken hebt omgezet tot een eigen promotie traject. Dank je voor alles!

Vanessa en Nienke (Wolterbeek), jullie waren vooral betrokken in de beginfase van dit project, maar niet minder belangrijk! Jullie hebben me wegwijs gemaakt in de onderzoekswereld en aan de voet gestaan van mijn eerste publicaties. Dank daarvoor!

Escape team – Ise, Coen, Tess, Olivier, Max, en Micheline – stuk voor stuk hebben jullie bijgedragen aan dit boek. Jullie zijn de drijvende krachten achter de Escape! Dank voor jullie inzet en het bereikte resultaat!

Principal Investigators – Ewoud van Arkel, Igor van den Brand, Daniël Haverkamp, Dirk-Jan Hofstee, Gino Kerkhoffs, Eduard Mutsaerts, Rolf Peters, Anton Vervest, en Julius Wolkenfelt – dank voor jullie begeleiding, hulp, inclusies, en natuurlijk deelname aan de Escape studie. Het was geen gemakkelijk project, maar dankzij jullie hulp is het een succes geworden!

Escape Research Group – vele van jullie zijn hiervoor al genoemd, maar nog niet genoemd zijn: Matthijs Krijnen, Derek van Deurzen, Dirk Jan Moojen, Thom Snijders, Jelle Halma, Camille Neeter, Suzanne de Vos, Andy Spoor, Taco Gosens, Wahid Rezaie, Bart Burger, Thijs van Rheenen, Anne Wijsbek, Bregje Thomassen, Sheila Sprague, en Maurits van Tulder – dank voor jullie bijdrage en hulp bij de voltooiing van de Escape studie.

(oud-)Collega OLVG'ers (staf, AIOS, ANIOS, PA'ers, en poli-assistenten): dank voor jullie hulp en bijdrage aan de totstandkoming van dit proefschrift. De naam 'Escape' kunnen jullie waarschijnlijk niet meer horen, maar ik ben trots dat ik dit project vanuit het OLVG heb mogen uitvoeren en dat ik sinds 2011 aan het OLVG verbonden heb mogen zijn.

VVV –Erik, Frank, Michael, René, Roland, Sonny en Thijmen – vrienden van het eerste uur. Sommigen van jullie ken ik al sinds de peuterspeelzaal en de vriendschap die wij met elkaar hebben opgebouwd is uniek. Wellicht dat dit boekwerk enig inzicht en begrip

brengt in waarom ik er met regelmaat niet bij ben, maar weet dat ik jullie vriendschap heel hoog heb zitten en bijzonder waardeer.

Plak en Nøl, LB9'ers. De vriendschap die we (met onze peren) hebben is me enorm dierbaar, en ik hoop op nog veel kerstmarkten, wijnborrels (een vervolgcursus?), en game-of-thrones marathons met elkaar.

Paragoni – Bram, Chris, Emiel, Jan, Jochem, Joris, Jos, Klaas en Pieter – onze vriendschap begon in 2004 in het mooie Utrecht. En ook al ben ik niet vaak meer in Utrecht te vinden (oké, behalve iedere dag voor werk dan), onze vriendschap is nog altijd zeer bijzonder en hecht! Met elkaar volwassen geworden, al vraag ik me bij Eus nog steeds wel eens af hoe hij het als vader redt. Ik ga zorgen dat ik na 13 maegt meer tijd voor jullie maak en ik hoop dat ik met dit proefschrift eindelijk eens kans maak op de titel 'oprichter van het jaar'! Natuurlijk wil ik ook een woord richten tot jou, Boov. Betrokken en altijd op de hoogte. Jij wist oprecht vaak nog beter hoe de zaken ervoor stonden dan ik dat zelf wist. Hoe je dat deed is me nog steeds een raadsel. Grote steun en toeverlaat, ik mis je vriend!

Beachboys – professor Schuurs en grote Kaj – wat een eer dat jullie mijn paranimfen willen zijn. Sinds 2015 is er een hechte vriendschap tussen ons ontstaan en nadat Kaj en ik al jouw bruidspersoneel mochten zijn, vind ik het heel bijzonder en ben ik mega trots dat jullie straks naast mij op het podium staan. Het voelt toch ook een beetje als trouwen in mijn eentje. Gappies, ik hoop nog op heel veel meer (surf)trips met elkaar; Kaj, wij gaan ons Montreal-Hollywood-Costa Rica-Panama AO-avontuur dit jaar nog eens dunnetjes overdoen, en Schuurs, wij gaan door tot we onze chicks eruit skiën!

Thijs en Geer, Bart en Marjo, en Joop en Ing, al ruim 12 jaar mijn 2e familie en vanaf het eerste moment voel ik me bij jullie thuis! Dank voor alle warmte, kerstavonden, de kookkunsten van Joop, en natuurlijk de onvergetelijke reis naar Zuid-Afrika!

Marg en Sander; Lisan, Marten en Mies; en pap en mam: lieve ouders, zussen, schoonbroers, en mijn kleine nichtje. Marg en Lisan, mijn lieve zussen, ik ben ontzettend trots op jullie en ik ben heel blij met jullie mannen in onze familie! Jullie hadden geen betere kunnen uitkiezen. Pap en mam, het was vast niet altijd makkelijk met ons 3, maar kijk hoe jullie ons toch afgeleverd hebben. Ik weet hoe trots jullie zijn op ons, maar ik hoop dat jullie ook weten hoe trots ik op jullie ben. Mam, jij bent er altijd voor ons en je liefde voor ons komt voor alles, en pap, jij bent mijn grote voorbeeld!

Lou, mijn allerliefste en wie had dat ooit gedacht: mijn cover designer! Al ruim 12 jaar samen, en wat ben ik een gelukkig mens met jou. Mijn blije ei, we hebben dezelfde humor, dezelfde hobby's, en hoe freaky ook: we denken te vaak hetzelfde. Hopelijk gaan we samen nog veel meer avonturen beleven, te beginnen met Australië. Dank je voor je talloze revisies van mijn brakke Engels (of Nederlands); je hebt misschien wel net zoveel tijd in dit proefschrift zitten als ik. Je haalt het beste in mij naar boven, maakt me compleet en gelukkig. Ik hou van je!

Curriculum vitae

Victor Adrianus van de Graaf (1986) attended high school at Norbertus College in Roosendaal, from which he graduated in 2004. After initially wanting to join the Air Force, his lack of advanced chemistry got him into medical school at the Vrije Universiteit Amsterdam. During his time at university he lived in Utrecht, which is where he met the love of his life; Malou. Victor completed his final year of medicine (2011) in the orthopaedic surgery department at the OLVG in Amsterdam (supervisor: prof. dr. R.W. Poolman). This is where he laid the foundation for his thesis, which began when Rudolf Poolman asked him to look into the available research on the degenerative meniscus.

In 2012, he moved to Amsterdam with Malou and started as a non-training resident in orthopaedic surgery at the former Sint Lucas Andreas Ziekenhuis (SLAZ, supervisor; drs. J. Wolkenfelt), now OLVG West. After he successfully applied for a healthcare efficiency grant at ZonMw, he started as a PhD Candidate under supervision of prof. dr. D.B.F. Saris, prof. dr. R.W. Poolman, and dr. A. de Gast in the summer of 2013. After 1.5 year of fulltime research, he continued as a non-training resident in general surgery at the former SLAZ in October 2014. After he was admitted to the orthopaedic surgery residency, he stayed for another 1.5 year at the former SLAZ to complete the first part of his residency (supervisor: dr. B.C. Vrouenraets). He then continued his residency at the St Antonius Ziekenhuis (supervisor then: dr. M.R. Veen) and the University Medical Centre Utrecht (supervisor: dr. J.J. Verlaan). In May 2020, he will head back to the St Antonius Ziekenhuis (supervisor now: dr. M. van Dijk) to finish his residency in November 2021. After his residency, Victor intends to move to Australia (together with Malou) to work as a fellow and enjoy the Ozzie way of life.