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 1,2, Johanna M van Dongen,3 Nienke W Willigenburg,1 Julia C A Noordyn,1 Ike K Butter,1 Arthur de Gast,4 Daniel B F Saris,2,5 Maurits W van Tulder,3 Rudolf W Poolman,1 ESCAPE Research Group

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 nine 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 (−€1803; 95% CI −€3008 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 as 1.00 at a willingness to pay of €0/unit of effect (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 deimplementation of APM in patients with non-obstructive meniscal tears.

Trial registration numbers NCT018750719 and NTR3908.

INTRODUCTION

Each year, approximately 2 million arthroscopic knee surgeries are performed in the world, associated with $4 billion of direct medical costs.1 Even though a clinical important benefit of surgery over conservative treatment has not been demonstrated,2 the number of arthroscopic surgeries is decreasing slower than expected.3

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.4 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 (ie, 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 (ie, non-inferior).5

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 nine Dutch hospitals.6 We excluded patients with a locked knee, an anterior cruciate ligament rupture, severe osteoarthritis (Kellgren-Lawrence 4)7 and a body mass index (BMI) >35 kg/m2.8

© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.
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.6,8

Interventions

Physical therapy

After randomisation, we referred participants to one of the participating primary care PT clinics, and treatment 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 (online supplementary appendix A). Participating PT clinics were instructed about the protocol prior to the first study participant referral. Additionally, participants completed a home exercise programme (online supplementary appendix A). 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 for 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 (online supplementary appendix A). Participants were only referred for PT in case of swelling or signs of atrophy, as advised by the Dutch Orthopaedic Association Guidelines.9

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,10–12 with a range from 0 to 100, in which 100 indicated no limitations in daily or sporting activities. The EuroQol five-dimensional five-level questionnaire (EQ-5D-5L) was used to measure health-related quality of life (HR-QoL).13 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.14 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%.15

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,15 and for surgeries, we used the average costs from all hospitals in the Netherlands, derived from the Dutch Healthcare Authority.16 Other healthcare costs included costs related to the use of primary healthcare (eg, general practitioner), secondary healthcare (eg, 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.15

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.15

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.15

For estimating absenteeism and presenteeism costs, we used the Productivity Cost Questionnaire.17 We valued the patients’ number of sickness absence days in accordance with the Fric tion Cost Approach (FCA; friction period=12 weeks) using gender-specific price weights.15 For presenteeism costs, we asked participants to report the total number of days that they went to work while experiencing health complaints and to report 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:

\[
\text{Presenteeism days} = \left( (10 - \text{performance level}) / 10 \right) \times \text{number of days with health complaints}.
\]

Presenteeism days were valued using gender-specific price weights.15

For estimating unpaid productivity costs, we asked participants to report the total number of hours they were unable to perform unpaid tasks (eg, chores, volunteer work and educational activities), which were valued using a Dutch recommended shadow price.15

We converted all costs to Euros 2016 using consumer price indices and discounted costs occurring after 12 months at a rate of 4%.15

Other prespecified 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 six). 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 a SD of 18 points on the IKDC, a power of 90%, 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 five complete datasets (loss of efficiency <5%).18 We
analysed each dataset separately as specified below. Poolled estimates were calculated using Rubin’s rules.18

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, mechanical complaints (IKDC question six), the affected meniscus (medial, lateral or both), BMI (in three categories: <25, 25–30 or ≥30 kg/m²), age, gender and education level (high vs low).13 Subsequently, we calculated 95% CIs surrounding all cost differences using Bias Corrected and Accelerated (BCA) bootstrapping (5000 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 (5000 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 (V.14) with a level of significance of p<0.05. The unadjusted cost and effect differences and ICERs were calculated and presented in online supplementary appendix B.

The deviations from the original trial protocol can be found in online supplementary appendix C.

Sensitivity analyses (SAs)
We performed four 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 three groups: (1) participants assigned to APM who received APM, (2) participants assigned to PT who completed the PT protocol (eg, ≥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 et al.2 For this, we defined a non-inferiority margin of eight points for the IKDC, which is consistent with estimates of the smallest detectable change of this outcome.11 For QALYs, a non-inferiority margin of 0.057 was chosen,22 which is based on the assumption that a minimal clinically important difference in utility is sustained for 1 year (ie, 1 * 0.057). As universally accepted non-inferiority margins for societal costs are currently lacking, we used the margin suggested by Bosmans et al of €670 (ie, €500 converted to Euros 2016).2 Bosmans et al based this margin on two visits to a primary health-care provider, one outpatient visit and 3 days of absenteeism,1 which we deemed appropriate for the condition under study as well. We estimated the proportion of CE pairs within these margins (ie, 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.1 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.6 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 1.63 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 (€1964) (€1468; 95% CI €1347 to €1680). Mean total societal costs were also statistically significantly lower in the PT group (€3935) than in the APM group (€5991) (€1803; 95% CI €838 to €3008). 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 one point decrease on the IKDC in the PT group as compared with the APM group was associated with a societal cost saving of €449 (ie, 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 €2300/point improvement (online supplementary appendix D).

For QALYs, we found an ICER of 61,584, indicating that one QALY lost in the PT group as compared with the APM group was associated with a societal cost saving of €61 584 (ie, PT was less costly and less effective) (figure 2, table 2). The CEAC indicated that the probability of PT being cost-effective compared
221 patients randomly assigned to intervention groups

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

<table>
<thead>
<tr>
<th></th>
<th>APM group</th>
<th>PT group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>158</td>
<td>161</td>
</tr>
<tr>
<td>Age, years</td>
<td>57.6±6.5</td>
<td>57.3±6.8</td>
</tr>
<tr>
<td>Women</td>
<td>80 (50.6)</td>
<td>81 (50.3)</td>
</tr>
<tr>
<td>Right knee</td>
<td>88 (55.7)</td>
<td>81 (50.3)</td>
</tr>
<tr>
<td>Education level, beyond high school</td>
<td>67 (42.4)</td>
<td>86 (53.4)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.7±3.8</td>
<td>27.2±4.0</td>
</tr>
<tr>
<td>18.5≤BMI&lt;25</td>
<td>56 (35.4)</td>
<td>53 (32.9)</td>
</tr>
<tr>
<td>25≤BMI&lt;30</td>
<td>72 (45.6)</td>
<td>67 (41.6)</td>
</tr>
<tr>
<td>30≤BMI&lt;35</td>
<td>30 (19.0)</td>
<td>41 (25.5)</td>
</tr>
<tr>
<td>Mechanical complaints*</td>
<td>56 (35.4)</td>
<td>67 (41.6)</td>
</tr>
</tbody>
</table>

**Imaging**

<table>
<thead>
<tr>
<th><em>AF</em></th>
<th>N=158</th>
<th>N=161</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial</td>
<td>126 (79.7)</td>
<td>136 (84.5)</td>
</tr>
<tr>
<td>Lateral</td>
<td>30 (19.0)</td>
<td>25 (15.5)</td>
</tr>
<tr>
<td>Both</td>
<td>2 (1.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Type of tear on MRI42</td>
<td>N=151</td>
<td>N=152</td>
</tr>
<tr>
<td>Longitudinal vertical</td>
<td>5 (3.3)</td>
<td>5 (3.3)</td>
</tr>
<tr>
<td>Horizontal</td>
<td>80 (53.0)</td>
<td>69 (45.4)</td>
</tr>
<tr>
<td>Complex degenerative</td>
<td>47 (31.1)</td>
<td>58 (38.1)</td>
</tr>
<tr>
<td>Radial</td>
<td>13 (8.6)</td>
<td>10 (6.6)</td>
</tr>
<tr>
<td>Vertical flap</td>
<td>2 (1.3)</td>
<td>5 (3.3)</td>
</tr>
<tr>
<td>Unclassifiable</td>
<td>1 (0.7)</td>
<td>5 (3.3)</td>
</tr>
<tr>
<td>Horizontal flap</td>
<td>3 (2.0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>OA level‡</td>
<td>N=150</td>
<td>N=149</td>
</tr>
<tr>
<td>0: None OA</td>
<td>18 (12.0)</td>
<td>15 (10.1)</td>
</tr>
<tr>
<td>1: Doubtful</td>
<td>81 (54.0)</td>
<td>74 (49.7)</td>
</tr>
<tr>
<td>2: Minimal</td>
<td>45 (30.0)</td>
<td>55 (36.9)</td>
</tr>
<tr>
<td>3: Moderate</td>
<td>6 (4.0)</td>
<td>5 (3.3)</td>
</tr>
<tr>
<td>Knee function</td>
<td>N=158</td>
<td>N=161</td>
</tr>
<tr>
<td>IKDC score (0–100, worse to best)</td>
<td>44.8±16.6</td>
<td>46.5±14.6</td>
</tr>
<tr>
<td>EQ-SD-5L Index value</td>
<td>0.72±0.2</td>
<td>0.74±0.1</td>
</tr>
</tbody>
</table>

Data are n (%) or mean±SD.

* In contrast to locking of the knee joint, which was an exclusion criterion, mechanical complaints were allowed for inclusion.

†Although inclusion was based on clinical readings by different radiologists and orthopaedic surgeons, one radiologist read all radiographs post hoc and one radiologist read all MRIs post hoc. Some of the radiographs (6.3%) and MRIs (5.0%) were unavailable to the viewing radiologist.

‡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% jointspace narrowing) indicates severe osteoarthritis. Kellgren-Lawrence grade 0 was an exclusion criterion.

APM, arthroscopic partial meniscectomy; BMI, body mass index; EQ-SD-5L, EuroQol five-dimensional five-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.

**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 APM was 1.00, 0.99, and 0.40 at a willingness to pay of €0, €10 000 and €80 000/QALY, respectively (online supplementary appendix D).
Table 2  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

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Sample size</th>
<th>Outcome</th>
<th>$\Delta C$ (95% CI)</th>
<th>$\Delta E$ (95% CI)</th>
<th>ICER</th>
<th>Distribution CE plane (%)</th>
<th>Non-inferiority region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis – imputed dataset</td>
<td>161</td>
<td>IKDC (range: 0–100)</td>
<td>-1803 (−3008 to −838)</td>
<td>-4.0 (−8.3 to 0.2)</td>
<td>449</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>161</td>
<td>QALYs (range: 0–1)</td>
<td>-1803 (−3008 to −838)</td>
<td>-0.029 (−0.074 to 0.016)</td>
<td>61 584</td>
<td>0.0</td>
<td>9.8</td>
</tr>
<tr>
<td>SA1 – complete cases</td>
<td>89</td>
<td>IKDC (range: 0–100)</td>
<td>-976 (−1360 to −589)</td>
<td>-7.0 (−11.5 to 2.9)</td>
<td>139</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>QALYs (range: 0–1)</td>
<td>-976 (−1365 to −589)</td>
<td>-0.024 (−0.077 to 0.030)</td>
<td>40 667</td>
<td>0.0</td>
<td>18.9</td>
</tr>
<tr>
<td>SA2 – human capital approach</td>
<td>161</td>
<td>IKDC (range: 0–100)</td>
<td>-1866 (−3129 to −871)</td>
<td>-4.0 (−8.3 to 0.2)</td>
<td>465</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>161</td>
<td>QALYs (range: 0–1)</td>
<td>-1866 (−3129 to −871)</td>
<td>-0.029 (−0.074 to 0.016)</td>
<td>63 741</td>
<td>0.0</td>
<td>9.8</td>
</tr>
<tr>
<td>SA3 – healthcare perspective</td>
<td>161</td>
<td>IKDC (range: 0–100)</td>
<td>-1120 (−1767 to −707)</td>
<td>-4.0 (−8.3 to 0.2)</td>
<td>279</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>161</td>
<td>QALYs (range: 0–1)</td>
<td>-1120 (−1767 to −707)</td>
<td>-0.029 (−0.074 to 0.016)</td>
<td>38 269</td>
<td>0.0</td>
<td>9.8</td>
</tr>
<tr>
<td>SA4 – as-treated analysis PT (≥16 sessions)</td>
<td>97</td>
<td>IKDC (range: 0–100)</td>
<td>-3073 (−4280 to −2150)</td>
<td>-3.0 (−7.5 to 1.4)</td>
<td>1010</td>
<td>0.0</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>QALYs (range: 0–1)</td>
<td>-3073 (−4280 to −2150)</td>
<td>0.021 (−0.024 to 0.065)</td>
<td>-148,866</td>
<td>0.0</td>
<td>81.6</td>
</tr>
<tr>
<td>Delayed APM</td>
<td>47</td>
<td>IKDC (range: 0–100)</td>
<td>525 (−1312 to 2272)</td>
<td>-6.0 (−13.1 to 1.1)</td>
<td>-88</td>
<td>1.8</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>QALYs (range: 0–1)</td>
<td>525 (−1312 to 2272)</td>
<td>-0.108 (−0.185 to −0.032)</td>
<td>-4850</td>
<td>0.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

*Refers to the NE of the CE plane, indicating that PT is more effective and more costly than APM.
†Refers to the SE of the CE plane, indicating that PT is more effective and less costly than APM.
‡Refers to the SW of the CE plane, indicating that PT is less effective and more costly than APM.
§Refers to the NW of the CE plane, indicating that PT is less effective and more costly than APM.
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.
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

<table>
<thead>
<tr>
<th>Cost category</th>
<th>PT (n=161) mean (SEM)</th>
<th>APM (n=158) mean (SEM)</th>
<th>Cost difference crude, mean (95% CI)</th>
<th>Cost difference adjusted*, mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention costs</td>
<td>488 (10)</td>
<td>1964 (73)</td>
<td>−1476 (−1682 to −1370)</td>
<td>−1468 (−1680 to −1347)</td>
</tr>
<tr>
<td>Other healthcare costs</td>
<td>1527 (145)</td>
<td>1238 (205)</td>
<td>289 (−301 to 689)</td>
<td>347 (−276 to 726)</td>
</tr>
<tr>
<td>Primary care</td>
<td>407 (49)</td>
<td>734 (185)</td>
<td>−326 (−950 to −81)</td>
<td>−309 (−954 to −1347)</td>
</tr>
<tr>
<td>Secondary care</td>
<td>1114 (126)</td>
<td>499 (51)</td>
<td>615 (393 to 928)</td>
<td>655 (436 to 935)</td>
</tr>
<tr>
<td>Medication</td>
<td>6 (1)</td>
<td>5 (1)</td>
<td>1 (−2 to 4)</td>
<td>1 (−2 to 4)</td>
</tr>
<tr>
<td>Paid help costs</td>
<td>29 (12)</td>
<td>151 (60)</td>
<td>−122 (−333 to −42)</td>
<td>−134 (−358 to −49)</td>
</tr>
<tr>
<td>Informal care costs</td>
<td>290 (58)</td>
<td>573 (140)</td>
<td>−282 (−648 to −62)</td>
<td>−216 (−489 to −8)</td>
</tr>
<tr>
<td>Absenteeism costs</td>
<td>225 (48)</td>
<td>337 (51)</td>
<td>−112 (−238 to 12)</td>
<td>−83 (−200 to 35)</td>
</tr>
<tr>
<td>Presenteeism costs</td>
<td>424 (73)</td>
<td>328 (60)</td>
<td>96 (−77 to 265)</td>
<td>118 (−44 to 285)</td>
</tr>
<tr>
<td>Unpaid productivity costs</td>
<td>952 (169)</td>
<td>1402 (218)</td>
<td>−449 (−988 to −49)</td>
<td>−369 (−845 to −79)</td>
</tr>
<tr>
<td>Total</td>
<td>3935 (334)</td>
<td>5991 (504)</td>
<td>−2056 (−3343 to −1002)</td>
<td>−1803 (−3008 to −838)</td>
</tr>
</tbody>
</table>

*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.

APM, arthroscopic partial meniscectomy; PT, physical therapy; n, number of.

Figure 2  Cost-effectiveness planes, including non-inferiority margins, for quality-adjusted life-years (A) and the IKDC (B). IKDC, International Knee Documentation Committee; QALY, quality-adjusted life years.

DISCUSSION

In this first trial-based economic evaluation in patients with non-obstructive meniscal tears, the total societal costs of PT was 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 QALYs. When we excluded participants 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.

Comparison with other studies

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 evaluation is not performed, nor can non-inferiority be investigated. Our data will further assist clinicians and healthcare decision makers in efficiently allocating already scarce healthcare resources and will likely contribute to reducing healthcare costs. Rongen and colleagues 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 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).

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 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).
willingness to pay in the Netherlands (ie, between €10 000 and €80 000 per QALY).
That study has several limitations, as was illustrated previously. 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 are prospectively collected and few assumptions are made. 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 (€4407) based on their hospital records, whereas we determined these costs (€1935) 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.

During our 2-year follow-up, only five patients progressed to having a knee arthroplasty (three in the PT group and two 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 threefold increase in the risk for future knee arthroplasty after APM. 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 due to differences in the applied analytical methods. These different methods include: (1) multiple imputation, which is recommended for economic evaluations, versus full maximum likelihood estimation, which is often used in longitudinal data analyses; (2) correcting for the possible correlation between costs and effects (eg, by using SUR analyses), which is recommended in economic evaluations; (3) discounting for cost and effect data, which is recommended in economic evaluations; 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 of study**
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 practice while simultaneously reducing the risk of selection bias and results in the most reliable estimates of costs and effects; this is considered the most valid method for estimating the clinical and financial implications of a healthcare intervention.

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. 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. 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, which is considered 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. 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.

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 eight 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 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, 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 deimplementation of APM in patients with non-obstructive meniscal tears.

What are the findings?

► Physical therapy was cheaper but less effective than arthroscopic partial meniscectomy. The probabilities that physical activity is cost-effective and that physical activity is non-inferior to arthroscopic partial meniscectomy were relatively high for knee function and quality-adjusted life years.

How might it impact on clinical practice in the future?

► The results from this economic evaluation extend the results from the ESCAPE trial and other clinical trials that conclude that APM should not be treatment of first choice in patients with non-obstructive meniscal tears. Our results are important for policy makers (government) and decision makers (funders such as insurers and health maintenance organisations) who can influence whether there should be reimbursement for various treatments. Our data support further and faster deimplementation of arthroscopic partial meniscectomy from clinical guidelines.

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 clinical collaborators.

Collaborators Scholtes VAB, PhD; Mutsaerts ELAR, MD; PhD; Wolkenfelt J, MD; Krijnen MR, MD; PhD; van Deurzen DIF; MD; Moojen DJ; MD; PhD; Bloembergen CH, MD; Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam, The Netherlands; Snijders T, MD; Halma JJ, MD; Department of Orthopaedic Surgery, Clinical Orthopaedic Research Centre – mN, Diakonessenhuis, Utrecht, The Netherlands; Wolterbeek N, PhD, Department of Orthopaedic Surgery, St. Antoniusziekenhuis, Utrecht, The Netherlands; Neeter C, PhD, Fysiken Physical Therapy, Hilversum, The Netherlands; van Arkel ERA, MD, PhD, Department of Orthopaedic Surgery, Amsterdam, The Netherlands; Vervest AMJS, MD, PhD; van Rheenen J, MD, PhD, Department of Orthopaedic Surgery, Xpert Orthopedie, Amsterdam, The Netherlands; Vervest AMJS, MD, PhD; van Rheenen VA, MD; Wijnsbeek AE, MD, Department of Orthopaedic Surgery, Rotterdam, The Netherlands; van den Brand ICJB, MD; de Vos-Jakobs S, MD; Spoor AB, MD; Gossens T, MD; PhD; Department of Orthopaedic Surgery, Elisabeth Tweekesteden Ziekenhuis, Tilburg, The Netherlands; Rezaie W, MD, Department of Orthopaedic Surgery, Onze Lieve Vrouwe Ziekenhuis, Aalst, Belgium; Hofstee DJ; Burger BJ, Department of Orthopaedic Surgery, Noordwest Ziekenhuis, Alkmaar, The Netherlands; Haerkamp D, MD, PhD, Department of Orthopaedic Surgery, Xpert Orthopedie, Amsterdam, The Netherlands; Veerbeek AE, MD, PhD; van Rheenen VA, MD; Gossens T, MD, PhD; Department of Orthopaedic Surgery, Utrecht, The Netherlands; van den Brand ICJB, MD, Department of Orthopaedic Surgery, Rotterdam, The Netherlands; Irrgang JH, Lawrence JS. Radiological assessment of Osteo-Arthritis. Annals of the Rheumatic Diseases 1957;16:494–502.

REFERENCES
Original research


