OPEN ACCESS

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 Noorduyn,[1] Ise 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

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bjsports-2018-100065).

[1]Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam, The Netherlands
[2]Department of Orthopaedic Surgery, University Medical Centre Utrecht, Utrecht, The Netherlands
[3]Department of Health Sciences, VU University, Amsterdam, The Netherlands
[4]Department of Orthopaedic Surgery, Clinical Orthopaedic Research Centre – mN, Diakonessenhuis, Utrecht, The Netherlands
[5]Department of Orthopaedic Surgery, Mayo Clinic, Rochester, MN, United States

Correspondence to
Dr Victor A van de Graaf, Orthopaedic surgery, Joint Research, OLVG, Amsterdam, Netherlands; v.a.vandegraaf@olvg.nl

Accepted 31 May 2019
Published Online First 21 June 2019

Linked

► http://dx.doi.org/10.1136/bjsports-2019-101106

Check for updates

To cite: van de Graaf VA, van Dongen JM, Willigenburg NW, *et al*. *Br J Sports Med* 2020;**54**:538–546.

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 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 deimplementation 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.[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/m^2.

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 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 (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 Friction 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:

Presenteeism days = ((10 − performance level)/10) * 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%, a two-sided α of 0.05 and a non-inferiority margin of eight 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 five complete datasets (loss of efficiency <5%).[18] We

analysed each dataset separately as specified below. Pooled 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,[7] mechanical complaints (IKDC question six), the affected meniscus (medial, lateral or both), BMI (in three categories: <25, 25–30 or ≥30 kg/m^2), age, gender and education level (high vs low).[19] 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*.[5] 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,[20 21] 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).[5] Bosmans *et al* based this margin on two visits to a primary healthcare provider, one outpatient visit and 3 days of absenteeism,[5] 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.[5] 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 1.65 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 (€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 €2500/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

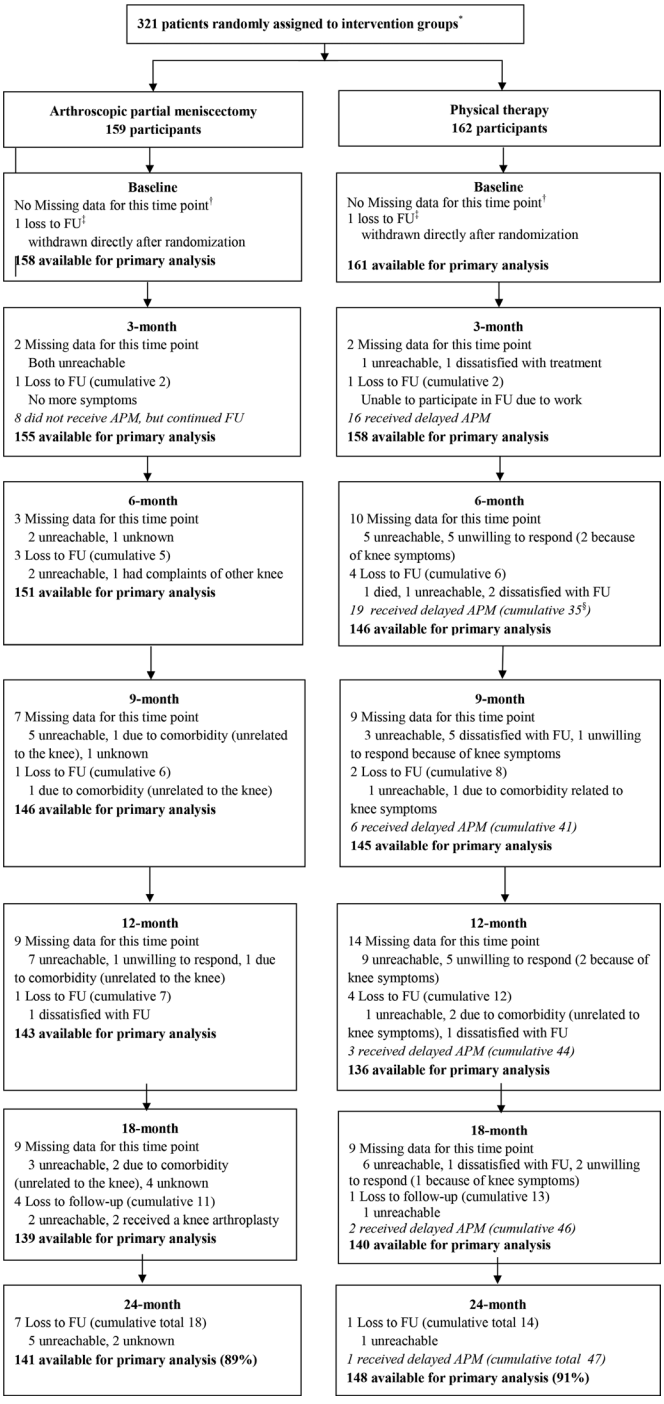

Figure 1 Flow of patients through the trial. *The number of patients screened for eligibility was not available. †Missing data refer to data that was missing at a specific time point, while patients remained available for the remaining follow-up moments. ‡Loss to follow-up refers to actual drop-out from the study, for example, patients who did not participate at any of the remaining time points (cumulative numbers are total number of drop-outs). §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. APM, arthroscopic partial meniscectomy; FU, follow-up.

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 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*	56 (35.4)	67 (41.6)
Imaging†		
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[42]	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‡	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.
*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.[7] Kellgren-Lawrence grade 4 was an exclusion criterion.
APM, arthroscopic partial meniscectomy; BMI, body mass index; EQ-5D-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

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

Analysis	Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE plane (%)				
	PT	APM		€	Points	€/point	NE*	SE†	SW‡	NW§	Non-inferiority region
Main analysis – imputed dataset	161	158	IKDC (range: 0–100)	−1803 (−3008 to −838)	−4.0 (−8.3 to 0.2)	449	0.0	2.5	97.5	0.0	97.3
	161	158	QALYs (range: 0–1)	−1803 (−3008 to −838)	−0.029 (−0.074 to 0.016)	61 584	0.0	9.8	90.1	0.0	89.0
SA1 – complete cases	89	81	IKDC (range: 0–100)	−976 (−1365 to −589)	−7.0 (−11.5 to 2.5)	139	0.0	0.1	99.9	0.0	66.9
	89	81	QALYs (range: 0–1)	−976 (−1365 to −589)	−0.024 (−0.077 to 0.030)	40 667	0.0	18.9	81.1	0.0	88.8
SA2 – human capital approach	161	158	IKDC (range: 0–100)	−1866 (−3129 to −871)	−4.0 (−8.3 to 0.2)	465	0.0	2.5	97.4	0.0	97.3
	161	158	QALYs (range: 0–1)	−1866 (−3129 to −871)	−0.029 (−0.074 to 0.016)	63 741	0.0	9.8	90.1	0.0	89.0
SA3 – healthcare perspective	161	158	IKDC (range: 0–100)	−1120 (−1767 to −707)	−4.0 (−8.3 to 0.2)	279	0.0	2.5	97.4	0.0	97.5
	161	158	QALYs (range: 0–1)	−1120 (−1767 to −707)	−0.029 (−0.074 to 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)	−3073 (−4280 to −2150)	−3.0 (−7.5 to 1.4)	1010	0.0	8.4	91.5	0.0	98.6
	97	150	QALYs (range: 0–1)	−3073 (−4280 to −2150)	0.021 (−0.024 to 0.065)	−148,866	0.0	81.6	18.4	0.0	99.9
Delayed APM	47	150	IKDC (range: 0–100)	525 (−1312 to 2272)	−6.0 (−13.1 to 1.1)	−88	1.8	2.0	25.4	70.8	43.4
	47	150	QALYs (range: 0–1)	525 (−1312 to 2272)	−0.108 (−0.185 to −0.032)	−4850	0.0	0.1	27.7	72.1	7.5

*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 less 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

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

*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.

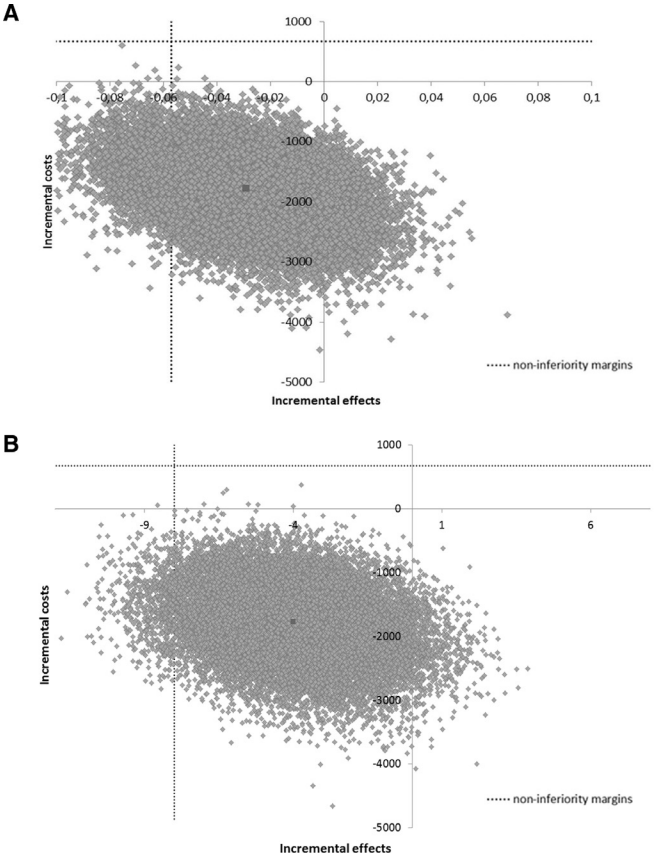

Figure 2 Cost-effectiveness planes, including non-inferiority margins, for quality-adjusted life-years (A) and the IKDC (B). IKDC, International Knee Documentation Committee; QALY, quality-adjusted life years.

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 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 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.

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 analysis is not performed, nor can non-inferiority be investigated.[22] Our data will further assist clinicians and healthcare decision makers in efficiently allocating already scarce healthcare resources[23] and will likely contribute to reducing healthcare costs.[24] Rongen and colleagues[4] 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 (ie, between €10 000 and €80 000 per QALY).[4] That study[4] has several limitations, as was illustrated previously.[25] 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.[22 23] 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.[5]

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[26] 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[6] due to differences in the applied analytical methods. These different methods include: (1) multiple imputation, which is recommended for economic evaluations,[27] 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[19]; (3) discounting for cost and effect data, which is recommended in economic evaluations[28]; 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[29 30]; this is considered the most valid method for estimating the clinical and financial implications of a healthcare intervention.[22 23] 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.[31] 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.[19] 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,[32] 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.[18] 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.[33]

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[34–40] 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,[41] 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 independent committee members in the adjustment of the sample size.

Collaborators Scholtes VAB, PhD; Mutsaerts ELAR, MD, PhD; Wolkenfelt J, MD; Krijnen MR, MD, PhD; van Deurzen DFP, MD; Moojen DJF, 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. Antonius Ziekenhuis, Utrecht, The Netherlands; Neeter C, PhD, Fysiken Physical Therapy, Amsterdam, The Netherlands; Kerkhoffs GMMJ, MD, PhD, professor of Orhopaedic Surgery, Department of Orthopaedic Surgery, Academic Medical Centre, Amsterdam, The Netherlands; Peters RW, MD, Department of Trauma Surgery, Academic Medical Centre, Amsterdam, The Netherlands, van den Brand ICJB, MD; de Vos-Jakobs S, MD; Spoor AB, MD; Gosens T, MD, PhD; Department of Orthopaedic Surgery, Elisabeth Tweesteden Ziekenhuis, Tilburg, The Netherlands; Rezaie W, MD, Department of Orthopaedic Surgery, Onze Lieve Vrouw Ziekenhuis, Aalst, Belgium; Hofstee DJ; Burger BJ, Department of Orthopaedic Surgery, Noordwest Ziekenhuisgroep, Alkmaar, The Netherlands; Haverkamp D, MD, PhD, Department of Orthopaedic Surgery, Xpert Orthopedie, Amsterdam, The Netherlands; Vervest AMJS, MD, PhD; van Rheenen TA, MD; Wijsbek AE, MD, Department of Orthopaedic Surgery, Tergooi Hospital, Hilversum, The Netherlands; van Arkel ERA, MD, PhD, Department of Orthopaedic Surgery, Haaglanden Medical Centre, The Hague, The Netherlands; Thomassen BJW, PhD, The Hague University of Applied Sciences, The Hague, The Netherlands; Sprague S, PhD, Division of Orthopaedic Surgery, Department of Surgery, Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; and Mol BWJ, MD, PhD, professor Obstetrics & Gynaecology, Department of Obstetrics and Gynaecology, School of Medicine, Monash University, Melbourne, Australia.

Contributors VvdG and RP were cochief investigators. VvdG, AdG, DS, MvT and RP were involved in the design of the study and its implementation. VvdG, NWW, AdG, DS, MvT and RP designed the prespecified statistical analysis plan. JCAN was the trial coordinator. IKB was responsible for monitoring, maintenance and blinding of the database. The members of the ESCAPE Research Group were responsible for study progress and data collection at site. JMvD, VvdG, NWW and MvT did the statistical analyses. VvdG and JMvD wrote the first draft of the manuscript; NWW, JCAN, IKB, AdG, DS, MvT and RP made revisions. All authors and collaborators read and approved the final manuscript.

Funding This study was funded by the Netherlands Organisation for Health Research and Development (in Dutch: ZonMw; grant number 837002009), Zilverenkruis Health Insurance (grant number Z436) and the foundation of medical research of the OLVG, Amsterdam (grant number 15u.025).

Competing interests All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: all authors had financial support from The Netherlands Organisation for Health Research and Development (in Dutch: ZonMw) for the submitted work; the Achmea Healthcare Foundation (in Dutch Stichting Achmea Gezonheidszorg fonds) and the foundation of medical research at the OLVG, Amsterdam, the Netherlands; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Patient consent for publication Not required.

Ethics approval The study was approved by the Medical Research Ethics Committees United (MEC-U; NL44188.100.13).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on request.

ORCID iD

Victor A van de Graaf http://orcid.org/0000-0001-5119-8887

REFERENCES

1. Järvinen TLN, Guyatt GH. Arthroscopic surgery for knee pain. *BMJ* 2016;354.
2. Siemieniuk RAC, Harris IA, Agoritsas T, *et al*. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: a clinical practice guideline. *BMJ* 1982;2017. j.
3. Järvinen TLN, Sihvonen R, Englund M. Arthroscopy for degenerative knee—a difficult habit to break? *Acta Orthopaedica* 2014;85:215–7.
4. Rongen JJ, Govers TM, Buma P, *et al*. 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* 2017.
5. Bosmans JE, de Bruijne MC, van Hout HPJ, *et al*. Practical guidelines for economic evaluations alongside equivalence trials. *Value in Health* 2008;11:251–8.
6. van de Graaf VA, Noorduyn JCA, Willigenburg NW, *et al*. Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears. *JAMA* 2018;320:1328–37.
7. Kellgren JH, Lawrence JS. Radiological assessment of Osteo-Arthrosis. *Annals of the Rheumatic Diseases* 1957;16:494–502.
8. van de Graaf VA, Scholtes VAB, Wolterbeek N, *et al*. 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:e014381.
9. van Arkel ERA, van Essen A, Koëter S, *et al*. Artroscopie van de knie: indicatie en behandeling. Dutch orthopedic association guideline 2011. website Dutch orthopedic association, 2012. Available: www orthopeden
10. Irrgang JJ, Anderson AF, Boland AL, *et al*. Development and validation of the International knee documentation Committee subjective knee form. *Am J Sports Med* 2001;29:600–13.
11. Crawford K, Briggs KK, Rodkey WG, *et al*. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007;23:839–44.
12. van de Graaf VA, Wolterbeek N, Scholtes VA, *et al*. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. *Am J Sports Med* 2014.
13. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
14. M. Versteegh M, M. Vermeulen K, M. A. A. Evers S, *et al*. Dutch tariff for the five-level version of EQ-5D. *Value in Health* 2016;19:343–52.

15. Hakkaart-van RL, van der Linden N, Bouwmans CA, *et al*. *Bijlage 1: Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg*. Zorginstituut Nederland, 2015.
16. Zorgautoriteit N, Onderhoud DBC, 2017. Available: https://www.opendisdata.nl
17. Bouwmans C, Krol M, Severens H, *et al*. The iMTA productivity cost questionnaire: a standardized instrument for measuring and Valuing health-related productivity losses. *Value Health* 2015;18:753–8.
18. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Statist. Med.* 2011;30:377–99.
19. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ.* 2004;13:461–75.
20. 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;14:1523–32.
21. Luo N, Johnson JA, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. *Medical Care* 2010;48:365–71.
22. van Dongen JM, van Wier MF, Tompa E, *et al*. Trial-based economic evaluations in occupational health: principles, methods, and recommendations. *J Occup Environ Med* 2014;56:563–72.
23. Shiell A, Donaldson C, Mitton C. Health economic evaluation. *J Epidemiol Community Health* 2002;56:85–8.
24. Stone JA, Salzler MJ, Parker DA, *et al*. Degenerative meniscus tears - assimilation of evidence and consensus statements across three continents: state of the art. *J Isakos* 2017;2:108–19.
25. Katz JN, Losina E. The cost-effectiveness of arthroscopic partial meniscectomy: comparing apples and oranges. *Osteoarthritis Cartilage* 2017.
26. Rongen JJ, Rovers MM, van Tienen TG, *et al*. 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 and Cartilage* 2017;25:23–9.
27. MacNeil Vroomen J, Eekhout I, Dijkgraaf MG, *et al*. Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? *Eur J Health Econ* 2016;17:939–50.
28. Husereau D, Drummond M, Petrou S, *et al*. Consolidated health economic evaluation reporting standards (cheers) statement. *BMJ* 2013;346:f1049.
29. Drummond MF, Sculpher MJ, Claxton K, *et al*. *Methods for the economic evaluation of health care programmes*. New York: Oxford University Press, 2005.
30. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ* 2011;342:d1548.
31. Hakkaart-van RL, Tan SS, Bouwmans CA. Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidzorg Diemen. *College voor Zorgverzekeringen* 2010.
32. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Statist. Med.* 1999;18:681–94.
33. Stewart WF, Ricci JA, Chee E. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA* 2003;290:2443–54.
34. Gauffin H, Tagesson S, Meunier A, *et al*. Knee arthroscopic surgery is beneficial to middle-aged patients with meniscal symptoms: a prospective, randomised, single-blinded study. *Osteoarthritis and Cartilage* 2014;22:1808–16.
35. Herrlin S, Hållander M, Wange P, *et al*. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007;15:393–401.
36. Herrlin SV, Wange PO, Lapidus G, *et al*. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? a five year follow-up. *Knee Surg Sports Traumatol Arthrosc* 2013;21:358–64.
37. Katz JN, Brophy RH, Chaisson CE, *et al*. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med* 2013;368:1675–84.
38. Sihvonen R, Paavola M, Malmivaara A, *et al*. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med* 2013;369:2515–24.
39. Yim J-H, Seon J-K, Song E-K, *et al*. A comparative study of meniscectomy and Nonoperative treatment for degenerative horizontal tears of the medial meniscus. *Am J Sports Med* 2013;41:1565–70.
40. Kise NJ, Risberg MA, Stensrud S, *et al*. 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.
41. Rongen JJ, van Tienen TG, Buma P, *et al*. Meniscus surgery is still widely performed in the treatment of degenerative meniscus tears in the Netherlands. *Knee Surg Sports Traumatol Arthrosc* 2017;41.
42. Brittberg M, Winalski CS. Evaluation of cartilage injuries and repair. *The Journal of Bone and Joint Surgery-American Volume* 2003;85:58–69.