Knee arthroscopy versus conservative management in patients with degenerative knee disease: a systematic review

Romina Brignardello-Petersen,1,2 Gordon H Guyatt,1 Rachelle Buchbinder,3,4 Rudolf W Poolman,5 Stefan Schandelmaier,1,5,6 Yaping Chang,1 Behnam Sadeghirad,1,7 Nathan Evaniew,8 Per O Vandvik9,10

ABSTRACT
Objective: To determine the effects and complications of arthroscopic surgery compared with conservative management strategies in patients with degenerative knee disease.
Design: Systematic review.
Main outcome measures: Pain, function, adverse events.
Data sources: MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials (CENTRAL), Google Scholar and Open Grey up to August 2016.
Eligibility criteria: For effects, randomised clinical trials (RCTs) comparing arthroscopic surgery with a conservative management strategy (including sham surgery) in patients with degenerative knee disease. For complications, RCTs and observational studies.
Review methods: Two reviewers independently extracted data and assessed risk of bias for patient-important outcomes. A parallel guideline committee (BMJ Rapid Recommendations) provided input on the design and interpretation of the systematic review, including selection of patient-important outcomes. We used the GRADE approach to rate the certainty (quality) of the evidence.
Results: We included 13 RCTs and 12 observational studies. With respect to pain, the review identified high-certainty evidence that knee arthroscopy results in a very small reduction in pain up to 3 months (mean difference =5.4 on a 100-point scale, 95% CI 2.0 to 8.8) and very small or no pain reduction up to 2 years (mean difference =3.1, 95% CI −0.2 to 6.4) when compared with conservative management. With respect to function, the review identified moderate-certainty evidence that knee arthroscopy results in a very small improvement in the short term (mean difference =4.9 on a 100-point scale, 95% CI 1.5 to 8.4) and very small or no improved function up to 2 years (mean difference =3.2, 95% CI −0.5 to 6.8). Alternative presentations of magnitude of effect, and associated sensitivity analyses, were consistent with the findings of the primary analysis. Low-quality evidence suggested a very low probability of serious complications after knee arthroscopy.

Strengths and limitations of this study
- This in an update of previously published systematic reviews on the topic.
- This review is linked to a BMJ Rapid Recommendations project. We conducted the review directed by a guideline panel that included patient representatives. This guideline panel provided detailed input with regards to the patients, interventions and outcomes and the interpretation of the results from this review.
- We included seven new studies, analysed data focusing on clinical interpretability and explicitly assessed the certainty in the estimates of effect.
- We performed meta-analyses using different measures of effect, and conducted subgroup and sensitivity analyses that strengthened our conclusions.

Conclusions: Over the long term, patients who undergo knee arthroscopy versus those who receive conservative management strategies do not have important benefits in pain or function.

Trial registration number: PROSPERO CRD42016046242.

INTRODUCTION
As a result of degenerative knee disease (osteoarthritis in the knee which can involve the joint lining and/or menisci), ~25% of people over 45 years, experience pain and other symptoms that may be severe and negatively impact quality of life (QoL).1–3 Total knee arthroplasty is the only definitive therapy available, but is reserved for patients with severe disease who fail conservative management.

In the USA, arthroscopic knee surgery in people with degenerative knee disease is the most common ambulatory orthopaedic
procedure, and the ninth most commonly performed ambulatory procedure overall. Such surgery results in transient increase in pain and the necessity for restriction in activities for a period of 2–12 weeks. Current guidelines recommend against arthroscopic lavage and/or debridement for patients with symptomatic knee osteoarthritis, but do not make specific recommendations for or against partial meniscectomy in those with degenerative meniscal tears (with or without other concomitant degenerative changes).7 8 Further, many orthopaedic surgeons suggest that patients with mechanical symptoms and meniscal tears—typically locking or catching of the knee—may benefit from arthroscopic partial meniscectomy.9 10

Our systematic review informs the second BMJ Rapid Recommendations, a new BMJ series of trustworthy clinical practice recommendations published in response to potentially practice-changing evidence.12 A trial that compared the outcomes of exercise therapy versus knee arthroscopic partial meniscectomy in 140 middle-aged patients with degenerative meniscal tears, published in July 2016 triggered this systematic review.13 Previous systematic reviews addressing the impact of arthroscopic knee surgery did not consider all patient-important outcomes; did not consider patient importance when addressing patient-reported outcomes such as pain, function and QoL; and did not include all currently available randomised clinical trials (RCTs).14 15

To determine the effects and complications in patients with symptomatic degenerative knee disease, we performed a systematic review and meta-analysis of arthroscopic surgery with debridement, and/or partial meniscectomy compared with conservative management strategies.

METHODS

Readers can access the protocol of this systematic review in International prospective registry of systematic reviews (PROSPERO) (CRD42016046242). According to the BMJ Rapid Recommendations process, a guideline panel provided critical oversight to the review and identified populations, subgroups and outcomes of interest. The panel included eight content experts and front-line clinicians (three orthopaedic surgeons, one rheumatologist, one epidemiologist, one general practitioner and two physiotherapists), four methodologists (three of them whom are also front-line clinicians and general internists) and three patients with lived experience of degenerative knee disease.

All patients received personal training and support to optimise contributions throughout the guideline development process. The patient panel members led the interpretation of the results based on what they expected the typical patient values and preferences to be, as well as the variation between patients. We also considered patients’ values and preferences by using the minimally important difference (MID) to interpret the results obtained in the meta-analyses. These MIDs were obtained from a systematic review of studies in which patients were directly asked about the magnitude of change they had experienced, and whether that change was trivial, small but important, or larger. Clinical experts who were part of the team of that systematic review judged the applicability of such studies to the target population and raised no concerns.

Eligibility criteria

For the effects of arthroscopic surgery, we included RCTs comparing arthroscopic surgery, including any or all of debridement and/or partial meniscectomy to any conservative management strategy (exercise therapy, injections, drugs, sham surgery) in patients with symptomatic degenerative knee disease (defined as persistent knee pain that affects the patient’s QoL and does not respond to conservative treatment), with or without osteoarthritis, of any age. We excluded studies that enrolled patients with acute trauma and those that enrolled fewer than 10 patients. For the complications of arthroscopic surgery, we also included observational studies (OS) (cohort studies, registry studies and case series) in patients with degenerative knee disease undergoing arthroscopic surgery, with or without a comparison group. We excluded studies published before the year 2000 when considering complications (but not effects).

Literature search

We performed an update of a previously published systematic review, including MEDLINE (PubMed), EMBASE (Ovid) and Cochrane Central Register of Controlled Trials (CENTRAL) (see online supplementary appendix 1) from 1 January 2014 to 16 August 2016. In addition, we constructed specific search strategies for these three databases for one outcome not studied in the previous review (nerve damage), with no date limits. We also searched for grey literature using the first 500 hits from Google Scholar and Open Grey. We did not limit any of the searches by language of publication.

Study selection and data abstraction

Teams of two reviewers, working independently, performed all study selection and data abstraction using standardised forms and reviewed the titles and abstract of all the references resulting from the searches. We retrieved and reviewed the full text of all references identified as potentially eligible by at least one reviewer. We also reviewed the full text of all references excluded at the full text screening stage in the prior review.17 We included all studies judged as eligible by the two reviewers. Reviewers resolved disagreements by discussion.

Reviewers abstracted characteristics of eligible studies including study design, number of patients enrolled, age and sex distribution, number of patients followed-up, whether partial meniscectomy was performed,
cointerventions, and outcomes, including pain, function, QoL and knee replacement. When authors reported results from more than one measure of pain or function, we decided a priori to use only the measure ranked highest in a hierarchy of patient-reported outcomes specific to the patients of interest. When studies had more than two arms, we only used the data from the arms relevant to this study. The review addressed these outcomes at 3 months or less, and at the longest follow-up reported.

The review addressed complication outcomes of mortality, venous thromboembolism (VTE), infection and nerve damage. Reviewers abstracted the absolute number of patients who experienced the outcomes over the follow-up period. When studies did not report VTE but reported pulmonary embolism and deep-vein thrombosis separately instead, we used these numbers to estimate the number of VTEs, considering the potential overlap due to patients experiencing both. We examined these outcomes over the 3 months following surgery.

**Summary measures and data synthesis**

We summarised continuous outcomes (pain, function and QoL) at the study level using the difference in change from baseline between groups. When baseline mean and SD per group at baseline and follow-up, but not change measures, were available, we assumed a within group correlation of 0.5 to estimate the SD of the change from baseline per study arm. If arm-level data were not reported, we abstracted the difference in change from baseline between the groups. When SDs at follow-up were not reported, we assumed the same SD as at baseline. When no SDs were available, we used the weighted average from all the other RCTs measuring the outcome with the same instrument. When studies reported medians and IQRs, we converted to means and SDs.

We performed meta-analyses, and present results for patient-reported continuous outcomes in two ways. First, we transformed all scores to the scale of an index instrument, the highest in the hierarchy and pooled results of all studies using the mean difference as the summary measure. This resulted in scores that could range from 0 to 100, in which higher scores signified better outcomes (less pain, better function, better QoL). Second, we used the MID of each of the instruments to determine the proportion of patients who achieved this difference. When no credible MID was found for a particular instrument, we used the MID of the index instrument. Data for time-to-knee replacement was not available, so we summarised the data for knee replacement using the proportion of patients who had the outcome per group and pooled those data using relative risk as the summary measure. These meta-analyses were performed using random effects models using the Hartung-Knapp-Sidik-Jonkman method. All analyses were performed using an intention-to-treat approach. When authors did not report data in a way that allowed incorporation in it the meta-analyses, we summarised the results narratively.

For complications, we used the number of patients having the event and the total number of patients undergoing knee arthroscopy, and pooled these data using a generalised linear mixed-effects model that allowed inclusion of studies with no events without a continuity correction.

We planned to perform four subgroup analyses for the outcomes pain and function: trials in which there was >50% of patients with radiographic osteoarthritis (defined as Kellgren-Lawrence grades 2–4) versus trials with ≤50% of patients with radiographic osteoarthritis; trials in which patients were blinded versus not blinded; trials in which meniscectomy was performed versus those in which it was not; and trials in which a control group received an active intervention (eg, exercise therapy, injections) versus control groups without such interventions (eg, waiting list, no treatment). We performed sensitivity analyses for calculating the difference in patients who achieve a change higher than the MID in two ways: (1) using the lowest and highest value of the MID of each instrument, based on the range of the MIDs that were deemed credible, and (2) calculating the standardised mean difference and then transforming the standardised mean difference into a risk difference (this method does not use an MID). All data analyses used the package meta in the software R, V.3.3.1. (R: a language and environment for statistical computing (program). Vienna, Austria: R Foundation for Statistical Computing, 2016).

**Certainty of the evidence assessments**

We assessed the certainty of the estimates of effect (quality of evidence) using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. We considered potential limitations in risk of bias, inconsistency, imprecision, indirectness and publication bias. We used a modification of the Cochrane Risk of bias tool to assess the risk of bias of the studies informing on the effects of arthroscopic surgery, and the relevant items of the Methodological Index for Non-Randomised Studies (MINORS) tool to assess the risk of bias of the studies informing on the complications of knee arthroscopy. All authors, in consultation with the parallel BMJ Rapid Recommendation guideline panel participated in, and
came to consensus regarding, certainty of estimates ratings.

The median of the change in score in the control arm from the studies that reported this information and did not use sham surgery as a control provided estimates of expected outcome in the control group (which is the equivalent of the baseline risk in dichotomous outcomes), which informed calculation of absolute estimates of effect. Summary of findings tables created using MAGICapp summarised key information for all patient-important outcomes.

RESULTS

Of 710 unique references screened in title and abstract, 149 articles underwent full text screening, of which 13 RCTs informing the effects of knee arthroscopy and 15 studies informing the complications of knee arthroscopy (12 OS and three RCTs) proved eligible (figure 1).

Effects

Study characteristics

The 13 eligible RCTs were published between 1993 and 2016, recruited a median of 119 patients, and enrolled patients with mean age from 48.9 to 62.8 years old and a sex distribution from 5% to 81.7% women. Two studies performed sham surgery in the control group, while most of the other studies used exercise therapy. Table 1 presents details of study characteristics.

**Table 2** presents the GRADE summary of findings for effects of knee arthroscopy compared with control. Patients who underwent arthroscopic surgery had a change in pain scores larger on average than patients who received control, in the short-term (5.4 points on a 100-point scale, 95% CI 2.0 to 8.8, n=10 studies, 1231 patients, see online supplementary appendix figure S1) and long-term (3.1 95% CI −0.2 to 6.4, n=8 studies, 1097 patients, see online supplementary appendix figure S2). The MID for this outcome measured with the index instrument (KOOS pain subscale) was 12 points. Using the MIDs specific to each instrument, (Devji T, et al. Submitted for publication 2016) 12.4% more patients receiving arthroscopy achieved an improvement in pain greater than the MID (n=11 studies, 1102 patients) in the short term.

Over the first 3 months of follow-up, the median average of improvement in pain was 15 points in patients who received conservative management versus 20 points in patients who underwent knee arthroscopy; over the long term, the median average improvement 19 points in patients who received conservative management versus 22 points in patients who underwent knee arthroscopy.

Patients who underwent arthroscopic surgery had an improvement in function score that was, on average, 4.9 points larger on a 100-point scale than patients who received control in the short term (95% CI 1.5 to 8.4, n=7 studies, 964 patients, see online supplementary appendix figure S3), and 3.2 points larger (95% CI −0.5
<table>
<thead>
<tr>
<th>Study</th>
<th>Number of patients enrolled</th>
<th>Comparator</th>
<th>Patients age (mean)</th>
<th>% females</th>
<th>ROA &gt;50%*</th>
<th>Pain measure†</th>
<th>Baseline mean intervention (SD)</th>
<th>Baseline mean control (SD)</th>
<th>Function measure†</th>
<th>Baseline mean (SD)</th>
<th>Baseline mean control (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang, 1993</td>
<td>34</td>
<td>Close needle joint lavage</td>
<td>62.8</td>
<td>71.6</td>
<td>Y</td>
<td>AIMS pain</td>
<td>65 (20)</td>
<td>61 (21)</td>
<td>AIMS physical function</td>
<td>23 (16)</td>
<td>17 (10)</td>
</tr>
<tr>
<td>Gauffin, 2014</td>
<td>150</td>
<td>Exercise therapy</td>
<td>54.5</td>
<td>27.3</td>
<td>N</td>
<td>KOOS pain</td>
<td>55 (18)</td>
<td>58 (18)</td>
<td>KOOS ADL</td>
<td>65 (18)</td>
<td>68 (22)</td>
</tr>
<tr>
<td>Herrlin, 2007</td>
<td>96</td>
<td>Exercise therapy</td>
<td>54</td>
<td>38.9</td>
<td>N</td>
<td>KOOS pain</td>
<td>56 (18)</td>
<td>63 (21)</td>
<td>KOOS ADL</td>
<td>68 (21)</td>
<td>73 (20)</td>
</tr>
<tr>
<td>Katz, 2013</td>
<td>351</td>
<td>Exercise therapy</td>
<td>58.4</td>
<td>56.7</td>
<td>Y</td>
<td>KOOS pain</td>
<td>54 (16)</td>
<td>53 (16)</td>
<td>WOMAC function</td>
<td>37 (18)</td>
<td>38 (18)</td>
</tr>
<tr>
<td>Kirkley, 2008</td>
<td>188</td>
<td>Exercise therapy</td>
<td>59.6</td>
<td>62.9</td>
<td>Y</td>
<td>WOMAC pain</td>
<td>52 (21)</td>
<td>43 (24)</td>
<td>WOMAC function</td>
<td>51 (21)</td>
<td>43 (23)</td>
</tr>
<tr>
<td>Kise, 2016</td>
<td>140</td>
<td>Exercise therapy</td>
<td>49.6</td>
<td>39</td>
<td>Y</td>
<td>KOOS pain</td>
<td>68 (15)</td>
<td>63 (21)</td>
<td>KOOS ADL</td>
<td>80 (16)</td>
<td>75 (22)</td>
</tr>
<tr>
<td>Moseley, 2002</td>
<td>119</td>
<td>Sham surgery</td>
<td>52.8</td>
<td>5</td>
<td>Y</td>
<td>SF-36 body pain</td>
<td>39 (19)</td>
<td>38 (18)</td>
<td>SF-36 physical function</td>
<td>42 (22)</td>
<td>47 (23)</td>
</tr>
<tr>
<td>Osteras, 2012</td>
<td>17</td>
<td>Exercise therapy</td>
<td>49.7</td>
<td>23.6</td>
<td>N</td>
<td>VAS</td>
<td>37 (10)</td>
<td>35 (17)</td>
<td>NM</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Saeed, 2015</td>
<td>120</td>
<td>Hyaluronic acid injection</td>
<td>NR</td>
<td>81.7</td>
<td>NR</td>
<td>Knee society score‡</td>
<td>NR</td>
<td>NR</td>
<td>Knee society score‡</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Sihvonen, 2013</td>
<td>146</td>
<td>Sham surgery</td>
<td>52</td>
<td>39</td>
<td>N</td>
<td>VAS</td>
<td>58 (20)</td>
<td>61 (20)</td>
<td>Lysholm knee score‡</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Stensrud, 2015</td>
<td>82</td>
<td>Exercise therapy</td>
<td>48.9</td>
<td>35.4</td>
<td>N</td>
<td>Ordinal scale</td>
<td>NR</td>
<td>NR</td>
<td>Ordinal scale</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Vermesan, 2013</td>
<td>114</td>
<td>Steroid injection</td>
<td>58.4</td>
<td>79.2</td>
<td>NR</td>
<td>Oxford knee score‡</td>
<td>NR</td>
<td>NR</td>
<td>Oxford knee score‡</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Yim, 2013</td>
<td>108</td>
<td>Exercise therapy</td>
<td>56.8</td>
<td>79.4</td>
<td>N</td>
<td>VAS</td>
<td>52 (18)</td>
<td>49 (15)</td>
<td>Lysholm score‡</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Based on Kellgren-Lawrence classification. Grades 2–4 were considered radiographic OA.
†All measures were converted to 0–100 scale. Higher scores mean less pain and better function.
‡Instrument combines pain and function together.
ADL, Function in Daily Living; AIMS, Arthritis Impact Measurement Scale; KOOS, Knee injury and Osteoarthritis Outcome Score; NA, not applicable; NM, not measured; NR, not reported; ROA, Radiographic osteoarthritis; SF-36, 36-Item Short-Form Survey; VAS, Visual Analogue Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index.
to 6.8, n=6 studies, 843 patients see online supplementary appendix figure S4) in the long term. The MID for this outcome measured with the index instrument (KOOS ADL subscale) was 8 points.69 The probability of achieving a change in function higher than the MID was 13.4% higher in patients receiving arthroscopy (n=6 studies, 835 patients) in the short term.

In the short term, patients who received conservative management achieved a median average improvement in function of 9 points, versus 14 points in patients who underwent knee arthroscopy; over the long term, the median average improvement was 10 points in patients who received conservative management versus 13 points in patients who underwent knee arthroscopy.

We were able to perform subgroup analyses according to blinding of patients and proportion of patients with radiographic osteoarthritis >50% for both of these outcomes. None of the analyses showed differences in results between groups (see online supplementary appendix figures S5–12). All RCTs performed partial meniscectomy as part of the intervention when needed, and all used active comparators. Therefore, we did not perform subgroup analyses for these variables.

Sensitivity analyses showed that for short-term pain and short-term function, results using the upper and lower limits of the MID estimate, and the approach using the standardised mean difference, in all cases yielded lower estimates of the numbers with important benefit from arthroscopy than did our primary analysis (see online supplementary appendix 2).

Changes in QoL scores were similar between patients undergoing knee arthroscopy and patients receiving control. In the short term, the difference in change from baseline scores was 6 points greater for knee arthroscopy (95% CI −1.5 to 13.5, n=1 study, 120 patients). In the long term, the difference in change from baseline was 2.1 points (95% CI −1.0 to 5.2, n=2 studies, 209 patients, see online supplementary appendix figure S13). The MID for the index instrument (EQ-5D) is 15 points.60 The median average of improvement in QoL was 8 points in patients who received conservative management versus 14 points in patients who underwent knee arthroscopy in the short term; and 10.3 points in patients who received conservative management versus 12.4 points in patients who underwent knee arthroscopy.

The risk of undergoing knee replacement up to 1 year after the intervention was 1.89 times higher in patients undergoing knee arthroscopy (95% CI 0.51 to 7, n=6 studies, 497 patients, see online supplementary appendix figure S14).

Complications

Study characteristics

The studies included in the complications systematic review reported data from a median of 20 770 patients. Average patient age ranged from 42 to 62 years, and the proportion of women from 39% to 64.6% (n=1 study, 12 426 patients). Table 3 presents detailed study characteristics.

Complications of knee arthroscopy

Table 4 provides a GRADE summary of findings for the complications of knee arthroscopy. Patients who underwent knee arthroscopy have an extremely small risk of death, that is, (<1 in 1000, 95% CI 0 to 1, n=7 studies, 454 086 patients, see online supplementary appendix figure S17); a risk of VTE of five in 1000 (95% CI 2 to 10, n=11 studies, 1 119 920 patients, see online supplementary appendix figure S18); a risk of infection of 2 in 1000 (95% CI 1 to 4, n=5 studies, 603 838 patients, see online supplementary appendix figure S19) and an extremely small risk of nerve damage (<1 in 1000, 95% CI 0 to 1, n=1 study, 12 426 patients).

Certainty of the evidence

The estimates of complications of knee arthroscopy had low certainty. All studies suffered from risk of bias concerns, mainly due to the retrospective nature of the data collection (using data that had not been collected for the purposes of the study) (see online supplementary appendix figure S20). The studies informing mortality, VTE and infection showed inconsistent results from a clinical and statistical perspective, which resulted in rating down the certainty for the pooled estimate. Finally, the only study informing nerve damage included patients with arthroscopy of the shoulder as well,57 and therefore warranted rating down this estimate for indirectness. There was no evidence of publication bias (see online supplementary appendix figure S21). Table 4 presents details regarding the assessments of the certainty of the complications of knee arthroscopy per outcome.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Timeframe</th>
<th>Study results and measurements</th>
<th>Absolute effect estimates</th>
<th>Certainty in effect estimates (Quality of evidence)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain (difference in change from baseline) 3 months</strong></td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS pain subscale, Scale: 0–100 high better, minimally important difference 12) Data from 1231 patients in 10 studies Follow-up 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Function (difference in change from baseline) 3 months</strong></td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS ADL subscale, Scale: 0–100, high better minimally important difference 8) Based on data from 964 patients in 7 studies Follow-up 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Quality of life (difference in change from baseline) 3 months</strong></td>
<td>Measured by: EQ-5D VAS Scale: 0–100, high better minimally important difference 15 Based on data from 120 patients in one study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pain and function up to 3 months</strong></td>
<td>Based on data from 316 patients in 3 studies Follow-up up to 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Timeframe</th>
<th>Study results and measurements</th>
<th>Conservative management</th>
<th>Arthroscopy</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain (difference in change from baseline) 3 months</strong></td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS pain subscale) Scale: 0–100 high better, minimally important difference 12) Data from 1231 patients in 10 studies Follow-up 3 months</td>
<td>15.0 Points (mean)</td>
<td>20.0 Points (mean)</td>
<td>High</td>
<td>On average, knee arthroscopy results in very small extra reduction in pain scores when compared with control</td>
</tr>
<tr>
<td><strong>Function (difference in change from baseline) 3 months</strong></td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS ADL subscale, Scale: 0–100, high better minimally important difference 8) Based on data from 964 patients in 7 studies Follow-up 3 months</td>
<td>9.0 Points (mean)</td>
<td>14.0 Points (mean)</td>
<td>Moderate</td>
<td>Knee arthroscopy may increase function change slightly more than control</td>
</tr>
<tr>
<td><strong>Quality of life (difference in change from baseline) 3 months</strong></td>
<td>Measured by: EQ-5D VAS Scale: 0–100, high better minimally important difference 15 Based on data from 120 patients in one study</td>
<td>8.0 Points (mean)</td>
<td>14.0 Points (mean)</td>
<td>Low</td>
<td>Knee arthroscopy may have, on average, little or no difference on QoL change, compared with control</td>
</tr>
<tr>
<td><strong>Pain and function up to 3 months</strong></td>
<td>Based on data from 316 patients in 3 studies Follow-up up to 3 months</td>
<td>Three studies evaluated the effects of knee arthroscopy in pain and function using measures that combined these two outcomes together or that could not be pooled. One study reported a difference in change from baseline in the Oxford knee score that favoured arthroscopy by 4.9 points (95% CI 3.61 to 6.20, 114 patients) over steroids injections. A second study reported no differences in the median in an overall self-assessment based on a 7-point ordinal</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Outcome Timeframe</th>
<th>Study results and measurements</th>
<th>Absolute effect estimates</th>
<th>Certainty in effect estimates (Quality of evidence)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long term</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain (difference in change from baseline) 1–2 years</td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS pain subscale minimally important difference 12) Scale: 0–100, high better Based on data from 1097 patients in 8 studies Follow-up 2 years</td>
<td>19.0 Points (mean) 22.0 Points (mean) Difference: mean difference 3.13 more (95% CI 0.17 fewer—6.43 more)</td>
<td>High</td>
<td>On average, knee arthroscopy results in no difference or a very small reduction, in pain</td>
</tr>
<tr>
<td>Function (difference in change from baseline) 1–2 years</td>
<td>Measured by: different instruments converted to scale of index instrument (KOOS ADL subscale minimally important difference 8) Scale: 0–100, high better Based on data from 843 patients in 6 studies Follow-up 2 years</td>
<td>10.0 Points (mean) 13.0 Points (mean) Difference: mean difference 3.16 more (95% CI 0.48 less—6.8 more)</td>
<td>Moderate</td>
<td>Owing to serious risk of bias and borderline imprecision On average, knee arthroscopy probably does results in no improvement or a very small improvement, in function</td>
</tr>
<tr>
<td>Quality of life (difference in change from baseline) 1–2 years</td>
<td>Measured by: EQ-5D VAS, 15D (converted to EQ-5D scale, MID 15) Scale: 0–100, high better Based on data from 269 patients in 2 studies Follow-up 2 years</td>
<td>10.3 Points (mean) 12.4 Points (mean) Difference: mean difference 2.12 more (95% CI 0.96 fewer—5.21 more)</td>
<td>High</td>
<td>On average, knee arthroscopy does not result in an important improvement in quality of life</td>
</tr>
<tr>
<td>Knee replacement 1–2 years</td>
<td>Relative risk: 1.89 (95% CI 0.51 to 7.0) Based on data from 497 patients in 2 studies Follow-up 1 year</td>
<td>12 Per 1000 23 Per 1000 Difference: 11 more per 1000 (95% CI 107 more—6 fewer)</td>
<td>Moderate</td>
<td>Owing to serious imprecision On average, knee arthroscopy does not result in an increase in the risk of knee replacement</td>
</tr>
<tr>
<td>Pain and function 1–2 years</td>
<td>Based on data from 114 patients in one study Follow-up 1 year</td>
<td>One study measured pain and function using a composite score. The study showed that patients who receive arthroscopy have a change in Oxford knee score 2.6 points higher than patients receiving steroids injections (95% CI 1.14 to 4.06)</td>
<td>Moderate</td>
<td>Owing to serious risk of bias Knee arthroscopy probably has little or no difference on pain and function</td>
</tr>
</tbody>
</table>

15D, the Health Related Quality of Life 15-Dimension questionnaire; ADL, Function in Daily Living; KOOS, Knee injury and Osteoarthritis Outcome Score; VAS, Visual Analogue Scale.
This systematic review provides high-quality evidence that patients with degenerative knee disease who undergo arthroscopy experience, on average, very small benefits in pain, function and QoL over periods of up to 3 months when compared with patients who receive a conservative management strategy (table 2). Results up to 2 years failed to show benefits in pain or function, and excluded any but very small benefits (table 2). The median of the average pain change in patients receiving conservative management was 15 points in the short term and 19 points in the long term (MID 12 points). Patients receiving arthroscopy had average change 5.4 points higher in the short term, and 3.1 points higher in the long term. These differences were not patient important. Thus, whether patients receive arthroscopy or not, the clinical trial experience suggests, on average, a small benefit in pain reduction over the short and long term.

### DISCUSSION

This systematic review provides high-quality evidence that patients with degenerative knee disease who undergo arthroscopy experience, on average, very small benefits in pain, function and QoL over periods of up to 3 months when compared with patients who receive a conservative management strategy (table 2). Results up to 2 years failed to show benefits in pain or function, and excluded any but very small benefits (table 2). The median of the average pain change in patients receiving conservative management was 15 points in the short term and 19 points in the long term (MID 12 points). Patients receiving arthroscopy had average change 5.4 points higher in the short term, and 3.1 points higher in the long term. These differences were not patient important. Thus, whether patients receive arthroscopy or not, the clinical trial experience suggests, on average, a small benefit in pain reduction over the short and long term.

### TABLE 3

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Number of patients</th>
<th>Age (mean)</th>
<th>% females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basques, 2015</td>
<td>Retrospective cohort (registry)</td>
<td>17 774</td>
<td>53</td>
<td>46.9</td>
</tr>
<tr>
<td>Bohensky, 2014</td>
<td>Retrospective cohort (registry)</td>
<td>139 031</td>
<td>NR</td>
<td>42.5</td>
</tr>
<tr>
<td>Cancienne, 2016</td>
<td>Prospective cohort</td>
<td>173 216</td>
<td>NR</td>
<td>64.6</td>
</tr>
<tr>
<td>Hame, 2012</td>
<td>Retrospective cohort (registry)</td>
<td>314 578</td>
<td>NR</td>
<td>62</td>
</tr>
<tr>
<td>Hetroni, 2011</td>
<td>Retrospective cohort (registry)</td>
<td>418 323</td>
<td>45.5</td>
<td>46.8</td>
</tr>
<tr>
<td>Hoppener, 2006</td>
<td>Retrospective cohort (registry)</td>
<td>335</td>
<td>42</td>
<td>43.3</td>
</tr>
<tr>
<td>Jameson, 2011</td>
<td>Retrospective cohort (registry)</td>
<td>261 446</td>
<td>46</td>
<td>40.7</td>
</tr>
<tr>
<td>Katz, 2013</td>
<td>RCT</td>
<td>174</td>
<td>59</td>
<td>55.9</td>
</tr>
<tr>
<td>Kise, 2016</td>
<td>RCT</td>
<td>70</td>
<td>48.9</td>
<td>39</td>
</tr>
<tr>
<td>Krych, 2015</td>
<td>Retrospective cohort (registry)</td>
<td>12 595</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Maletis, 2012</td>
<td>Retrospective cohort (registry)</td>
<td>20 770</td>
<td>44</td>
<td>42.8</td>
</tr>
<tr>
<td>Sihvonen, 2013</td>
<td>RCT</td>
<td>70</td>
<td>52</td>
<td>58</td>
</tr>
<tr>
<td>Wai, 2002</td>
<td>Retrospective cohort (registry)</td>
<td>14 391</td>
<td>62.4</td>
<td>49.9</td>
</tr>
<tr>
<td>Yacub, 2009</td>
<td>Retrospective cohort (registry)</td>
<td>12 426</td>
<td>NR</td>
<td>57.3</td>
</tr>
<tr>
<td>Yeranosian, 2013</td>
<td>Retrospective cohort (registry)</td>
<td>432 038</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

NR, not reported, RCT, randomised clinical trial.

### TABLE 4

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Timeframe</th>
<th>Study results and measurements</th>
<th>Absolute effect estimates</th>
<th>Certainty in effect estimates (Quality of evidence)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>3 months</td>
<td>Based on data from 454 086 patients in 7 studies, Follow-up 3 months</td>
<td>0 Per 1000 0 Per 1000 Difference: &lt;1 more per 1000 (95% CI 0 more—1 more)</td>
<td>Low Owing to serious risk of bias and serious inconsistency</td>
<td>Arthroscopy may have an extremely small risk of mortality</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>3 months</td>
<td>Based on data from 1 119 920 patients in 11 studies, Follow-up 3 months</td>
<td>0 Per 1000 5 per 1000 Difference: 5 more per 1000 (95% CI 2 more—10 more)</td>
<td>Low Owing to serious risk of bias, owing to serious inconsistency</td>
<td>Arthroscopy may have a small risk for venous thromboembolism</td>
</tr>
<tr>
<td>Infection</td>
<td>3 months</td>
<td>Based on data from 603 838 patients in 5 studies, Follow-up 3 months</td>
<td>0 Per 1000 2 per 1000 Difference: 2 more per 1000 (95% CI 1 more—4 more)</td>
<td>Low Owing to serious risk of bias, owing to serious inconsistency</td>
<td>Arthroscopy may have a very small risk for infection</td>
</tr>
<tr>
<td>Nerve damage</td>
<td>3 months</td>
<td>Based on data from 12 426 patients in one study, Follow-up 3 months</td>
<td>0 Per 1000 0 Per 1000 Difference: &lt;1 more per 1000 (95% CI 0 more—1 more)</td>
<td>Low Owing to serious risk of bias, owing to serious indirectness</td>
<td>Arthroscopy may have an extremely small risk of nerve damage</td>
</tr>
</tbody>
</table>
The results for function proved similar, with very small average differences in the short term, and no convincing evidence of benefit in the long term (table 2). Patients who received a conservative management strategy had a median average change of 9 points in the short term and 10 points in the long term. (MID 8 points). Risk of bias limitations leave this evidence less secure (moderate quality) than for pain.

Study results provide high-quality evidence that the benefits of arthroscopic surgery on QoL over the long term are minimal, if they exist at all (table 2). Low-quality evidence raises the possibility of a higher risk of knee replacement with arthroscopic surgery.

We found a low risk of serious adverse effects in patients undergoing knee arthroscopy. The risk of mortality and nerve damage may be close to 0, while the risk of VTE and infection may be five and two in 1000 patients, respectively. We have low certainty in this evidence, however, because the studies included were likely to be biased and showed results that were inconsistent.

Our systematic review has particular strengths. First, it provides the most comprehensive and trustworthy body of evidence up to date, including 10 studies not included in the most recent prior review. While the conclusions of our systematic review may not be qualitatively different from the conclusions of previous reviews addressing the same question, we believe that all the additions in terms of studies included and methods for summarising, presenting and appraising the evidence strengthen the conclusions derived from this body of evidence considerably. Second, this systematic review was developed in parallel with a BMJ Rapid Recommendation according to predefined standards, methods and processes. Extensive input from content experts and patients in the guideline panel throughout the process secured appropriate selection of outcomes and analyses as well as appropriate interpretation of the results from the systematic review. The rapid recommendations published together with our linked systematic review should provide clinicians and their patients with optimal guidance in practice and will also allow other guideline organisations to reuse or adapt content to their contexts, if needed. Third, by converting all the instruments to the scale of an index instrument we do overcome the potential limitations of using the standardised mean difference (namely, the analysis depending on a similar SD across studies, and the resulting measure of effect being difficult to interpret), and provide an estimate of the proportion of patients who would achieve a minimally important change per arm, and the difference between these proportions. This allows incorporating patients’ values and preferences explicitly when interpreting the results. A rigorous linked systematic review of studies addressing the issue informed our estimates of the minimally important change (Devji T, et al. Submitted for publication 2016) and our results were robust to accounting for the uncertainty in the MID, as well as to calculating the proportion who might benefit using an approach relying on the standardised mean difference. Fourth, we provide an explicit and transparent assessment of the certainty in the absolute estimates of effect, which considers limitations of the evidence with regards to risk of bias, inconsistency, imprecision, indirectness and publication bias.61

Our review is limited by suboptimal reporting in many of the original studies, requiring imputing SDs and, in a number of studies, estimating correlations between baseline and follow-up. It is possible that there is a subgroup of patients—for instance, those with locking symptoms—who do achieve substantial benefit from arthroscopic knee surgery. The available studies do not, however, provide evidence of any such subgroup. The burden of proof now rests with those who claim that such a subpopulation exists, with compelling RCT evidence required to substantiate the claim.

In summary, our results provide low-quality evidence that knee arthroscopy is a safe procedure with a low risk of complications and moderate to high-quality evidence that the procedure provides very small benefits in pain and function over conservative therapy in the short term. The evidence fails to support a persistence of these benefits over the long term. Patients and their healthcare providers must trade-off the marginal short-term benefits against the burden of the surgical procedure (pain, swelling, limited mobility, restriction of activities, over a period of 2–6 weeks).

Author affiliations
1Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2Evidence-Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Santiago, Chile
3Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
4Monash Department of Clinical Epidemiology, Cabrini Institute, Malvern, Victoria, Australia
5Department of Orthopedic Surgery and Traumatology, Joint Research, Onze Lieve Vrouwe Gasthuis (OLVG), Amsterdam, The Netherlands
6Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel, Switzerland
7HIV/STI Surveillance Research Center, and WHO Collaborating Center for HIV Surveillance, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Medical University Campus, Haff-Bagh Highway, Kerman, Iran
8Division of Orthopaedics, Department of Surgery, McMaster University, Hamilton, Ontario, Canada
9Department of Medicine, Innlandet Hospital Trust-Division Gjøvik, Gjøvik, Norway
10Faculty of Medicine, Institute of Health and Society, University of Oslo, Oslo, Norway

Acknowledgements We thank members of the rapid recommendations panel for critical feedback on outcome and subgroup selection, GRADE judgments and manuscript feedback, including Reed Siemieniuk (panel chair and internist), Ian A. Harris (orthopaedic surgeon), Martin Englund (epidemiologist), Casey Quinlan (patient representative), Hazel M Wilson (patient representative), Anne Lydiatt (patient representative), Lyubov Lybryn (patient liaison expert), Nina Rydland (physiotherapist), Stijn van de Welde (physiotherapist), Thomas Agoritsas (methodologist, internist) and Annette Kristiansen (methods editor and internist).
Contributors GHG and POV conceived the study idea. RB-P performed the literature search. SS, BS, YC, NE and RB-P performed screening, data abstraction and risk of bias assessments. RB-P performed the data analysis. RB-P, RB and GHG interpreted the data analysis. RB-P and GHG interpreted the data performed certainty of evidence assessments. RB-P wrote the first draft of the manuscript. GHG, POV, RB and RP critically revised the manuscript. All authors approved the final version of the manuscript. RB-P had full access to all of the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis. RB-P is guarantor.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. RB is funded by an Australian National Health and Medical Research Council (NHMRC) Senior Principal Research Fellowship.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf

Disclaimer All authors declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other financial relationships that could appear to have influenced the submitted work.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Extra data are available in the publication of the BMJ Rapid Recommendation in MAGICapp.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES


Linked articles in this BMJ Rapid Recommendations cluster

  Summary of the results from the Rapid Recommendation process

  Review of all available randomised trials that assessed the benefits of knee arthroscopy compared with non-operative care and observational studies that assessed risks

  Review addressing what level of individual change on a given scale is important to patients (minimally important difference). The study informed sensitivity analyses for the review on net benefit, informed discussions on patient values and preferences, and was key to interpreting the magnitude of effect sizes and the strength of the recommendation

- MAGICapp (www.magicapp.org)
  Expanded version of the results with multilayered recommendations, evidence summaries, and decision aids for use on all devices