Supplementary data

LITERATURE REVIEWS

Question 1

McMurray's test

In the systematic review of Smith et al. (2015) the pooled sensitivity for McMurray's test was 61% (95% CI 45–74) and the pooled specificity 84% (CI 69–92). The positive likelihood ratio (LR+) was 3.2 (CI 1.7–5.9) and the negative likelihood ratio (LR–) was 0.52 (CI 0.34–0.81). In other words, information from McMurray's test contributes only mildly to the probability that the patient in question has a meniscal injury.

Apley's test

In the systematic review of Smith et al. (2015), the Apley's test data was not pooled due to insufficient data. 2 studies included in the review (Karachalios et al. 2005, Rinonapoli et al. 2011) investigated the diagnostic accuracy of Apley's test with varying results. For instance, values of sensitivity were found ranging from 41% to 84%. Furthermore, Karachalos et al. (2005) used as the gold standard MRI and Rinonapoli et al. (2011) arthroscopy.

Thessaly's test

In the systematic review of Smith et al. (2015) the pooled sensitivity for Thessaly's test at 20° was 75% (CI 53–89) and the pooled specificity 87% (CI 65–96). The positive likelihood ratio (LR+) was 5.6 (CI 1.5–21.0) and the negative likelihood ratio (LR–) was 0.28 (CI 0.11–0.71). The data of the Thessaly test at 5° was not pooled due to insufficient data.

Goossens et al. (2015) reported a sensitivity of 64% (CI 60–68) for medial and/or lateral meniscus tears. The corresponding specificity was 53% (CI 43–63). The reported LR+ and LR– were 1.37 (CI 1.10–1.70) and 0.68 (CI 0.59–0.78).

Although the results of Goossens are not in line with the pooled estimates from Smith et al. (2015), we suggest that Thessaly's test contributes only mildly to the probability that a patient has a meniscal injury.

Joint line tenderness (JLT) test

In the systematic review of Smith et al. (2015) the pooled sensitivity for the JLT test at 20° was 83% (CI 73–90) and the pooled specificity 83% (CI 61–94). The LR+ was 4.0 (CI 2.1–7.5) and LR– was 0.23 (CI 0.12 to 0.44). In other words, information from the JLT test contributes only mildly to the probability that the patient in question has a meniscal injury.

Combination of physical tests

Goosens et al. (2015) also evaluated the combination of the Thessaly test followed by the McMurray test.

The combination of tests contributes only mildly to the probability of meniscal injury. Although the sensitivity of a combined negative test score is higher than 70%, this result contributes only mildly to the probability of having a meniscus injury. With other words, not much information is gained after performing the Thessaly test followed by the McMurray test.

Level of evidence

There are 4 levels of evidence: high, moderate, low, and very low. RCTs start with a high level of evidence.

McMurray's test: The level of evidence for McMurray's test is moderate, as 1 study used MRI as the reference standard, which consequently was the largest study in the meta-analysis (indirectness).

Apley's test: As the results of Apley's test could not be pooled because of insufficient data, the level of evidence could not be evaluated.

Thessaly's test (at 20°): The level evidence for Thessaly's test is moderate, as 1 study used MRI as the reference standard, which consequently was the largest study in the meta-analysis (indirectness).

Joint line tenderness test: The level evidence for the JLT test is moderate, as 1 study used MRI as the reference standard, which consequently was the largest study in the meta-analysis (indirectness).

Thessaly's test (at 20°) followed by McMurray's test: the level of evidence is high.

Question 2

1. MRI

Medial meniscal injury

In the systematic review of Phelan et al. (2016), the pooled sensitivity and specificity of MRI for detection of medical meniscal tears were 0.89 (CI 0.83–0.94) and 0.88 (CI 0.82–0.93) respectively. This means that 11% of patients with meniscal tears could be missed, and 12% of patients could have meniscal tears while the MRI diagnosis was normal. The pooled LR+ was 8.0 (CI 4.7–13.4) and the pooled LR– was 0.1 (CI 0.7–0.2).

In the systematic review of Smith et al. (2016), the pooled sensitivity and specificity of 3T MRI to diagnose medial meniscal injury were 0.94 (CI 0.91–0.96) and 0.79 (CI 0.75–0.83) respectively.

Lateral meniscal injury

In the systematic review of Phelan et al. (2016), the pooled sensitivity and specificity of MRI for detection of lateral meniscal tears were 0.78 (CI 0.66–0.87) and 0.95 (CI 0.91–0.97) respectively. The pooled LR+ was 14.5 (CI 8.7–24.3) and the pooled LR– was 0.2 (CI 0.2–0.4).

In the systematic review of Smith et al. (2016), the pooled sensitivity and specificity of 3T MRI to diagnose lateral meniscal injury were 0.81 (CI 0.75–0.85) and 0.87 (CI 0.84–0.89) respectively.

Knee cartilage lesions

In the systematic review of Zhang et al. (2013), the pooled sensitivity and specificity of MRI for detection of knee cartilage lesions were 0.75 (CI 0.62–0.84) and 0.94 (CI 0.89–0.97) respectively. The pooled LR+ was 13 (CI 6.5–24) and the pooled LR– was 0.27 (CI 0.17–0.42).

In the systematic review of Quatman et al. (2011), no meta-analysis was performed. The sensitivity of MRI for detection of articular cartilage abnormalities among included studies ranged from 0.29 to 0.96 and the specificity ranged from 0.50 to 1.00.

2. Ultrasonography (US)

Meniscal injury

The meta-analysis of Dai et al. (2015) showed a moderate pooled sensitivity of 0.88 (CI 0.84–0.91) and a high specificity of 0.90 (CI 0.86–0.93) of US in diagnosing meniscal injury. This means that 12% of patients with meniscal injury could be missed, and 10% of patients could have meniscal injury while the ultrasonography diagnosis is normal. The pooled LR+ was 7 (CI 4–12) and the LR– was 0.17 (CI 0.10–0.26). There was moderate to high heterogeneity of these values (73% for sensitivity and 61% for specificity). Therefore, a sensitivity analysis was performed by excluding each study. This analysis decreased the heterogeneity, but the results were similar to the overall results.

The systematic review of Phelan et al. (2016) included only 3 studies evaluating the diagnostic accuracy of ultrasound in the diagnosis of meniscal injury and only one study in the diagnosis of ACL injury. The results of these study were not shown, and no meta-analysis was performed. However, all three studies were included in the meta-analysis of Dai et al. (2015).

In the systematic review of Xia et al. (2016) the pooled sensitivity and specificity for diagnosing meniscal injury using US were 0.78 (CI 0.75–0.80) and 0.84 (CI 0.82–0.86) respectively. However, in this pooled analysis the data of 9 studies published before 2006 were also included. The diagnostic performance of US was specified for each included study for different meniscal injuries (lateral, medial, total).

Level of evidence

MRI and meniscal injury: The level of evidence for diagnosing meniscal injury was downgraded by 1 level because of limitations in the study design (risk of bias, due to patient selection (nonrandomized) and interpretation of MRI).

MRI and chondral lesions: The level of evidence for diagnosing chondral lesions was downgraded by 2 levels because of limitations in the study design (risk of bias, due to patient selection (nonrandomized) and interpretation of MRI) and inconsistency of results (wide variance of point estimates across studies).

Ultrasonography: The level of evidence for diagnosing meniscal injury was not downgraded.

Question 3

RCTs

Vives et al. (2003) compared the accuracy of nonenhanced MRI with that of intraarticular contrast-enhanced direct MRI arthrography and intravenous contrast-enhanced indirect MRI arthrography for detection 10 of recurrent meniscal tears. 41 patients previously treated for a meniscal tear were randomized into 3 groups: conventional MRI, indirect arthrography (intravenous contrast), and direct arthrography (intraarticular contrast). All patients underwent a second-look arthroscopy (i.e., the gold standard).

White et al. (2002) investigated the accuracy of conventional MRI, direct MRI arthrography, and indirect MRI arthrography in assessment of possible recurrent or residual meniscal tears. 364 patients were prospectively examined. However, only 94 patients underwent a second-look arthroscopic surgery (i.e., the gold standard). It was unclear why only 94 patients underwent a second-look arthroscopy and whether these patients were a representative (randomized) sample.

Observational studies

Kececi et al. (2015) evaluated the diagnostic value of direct MRI arthrography in detection of re-torn or unhealed menisci that were previously repaired. 24 symptomatic patients were included, all of whom underwent a second-look arthroscopy (i.e., the gold standard). Authors decided to include patients who received an arthroscopy for both diagnostic and therapeutic purposes.

Magee (2014) assessed the accuracy of conventional MRI and direct MR-arthrography in the diagnosis of meniscal retears as compared with arthroscopy. 100 patients were included. All patients underwent a second-look arthroscopy (i.e., the gold standard).

1. Accuracy

All included studies reported data on the accuracy of the arthrography. LR+ and LR– were calculated using the reported sensitivity and specificity.

The accuracy of a direct MR arthrography was high in all 4 studies. Results were consistent across all parameters of accuracy.

The accuracy of an indirect MR arthroscopy was reported by only 2 studies. 1 study (Vives et al. 2003) reported that the accuracy was high (LR+ > 10 and a LR– < 0.1). The accuracy in another study (White et al. 2002) was somewhat lower, however, and pointed in the same direction.

3 of the 4 included studies evaluated the diagnostic accuracy of the conventional MRI. All 3 studies suggested that the diagnostic accuracy is moderate.

Overall, direct MR arthrography seemed to be able to diagnose a recurrent meniscal tear in a patient with complaints after a therapeutic arthroscopy. There was no difference in the results from an RCT or a cohort study.

2. Costs

None of the included studies reported any data on costs.

3. Clinical outcome: meniscal retears

Vives et al. (2003) and Kececi et al. (2015) did not report the number of recurrent meniscal tears. Others (White et al. 2002) found 71 recurrent meniscal tears; however, it was unclear among how many patients. Magee (2014) reported that among 100 patients 94 had a meniscal retear. The results from the last 2 studies cannot be pooled or compared to draw a conclusion.

Level of evidence

Accuracy: The level of evidence was downgraded by 2 levels because of risk of bias (in 3 of 4 studies the results of the arthroscopy (i.e., the gold standard) were not interpreted blinded from the imaging results), imprecision (less than 300 patients included in total).

Costs: None of the included studies reported data on the costs.

Clinical outcome: It was not possible to assess the level of evidence as the results from 2 studies could not be pooled or compared.

Question 4

Patients with (acute) traumatic meniscus injury

No studies were included as none of the studies met the selection criteria.

Patients with a degenerative meniscal tear

Brignardello-Petersen et al. (2017) in a meta-analysis determined the effects and complications of arthroscopic surgery compared with nonoperative management strategies in patients with degenerative knee disease. 13 RCTs were included to inform on effects of knee arthroscopy and 15 studies (12 observational studies and 3 RCTs) provided data on the complications of knee arthroscopy.

1. Pain

Short-term benefits (< 3 months) were reported in 10 RCTs. The pooled difference in change from baseline was on average 5.4 (CI 1.9–8.8). Long-term benefits (1 to 2 years) were reported in 8 RCTs. The pooled difference in change from baseline was on average 3.1 (CI 0.2–6.4). The benefits of arthroscopy in pain scores, both short and long-term, were no different from that of nonoperative treatment.

***2. Functio*n**

Short-term data on function was available in 7 studies and long-term data in 6 studies. The mean score difference from baseline in function after 3 months was 4.9 (CI 1.5–8.4) in favor of arthroscopy and after 1 to 2 years 3.2 (CI 0.48–6.8).

3. Complications

In line with the recommendation, the working group chose to report the outcomes venous thromboembolism (VTE) and infections as potential complications.

The difference in proportion of patients with a VTE between arthroscopy versus nonoperative management was 5 per 1,000 patients (CI 2–10). Arthroscopy may have a small risk of VTE.

For infections, the difference between arthroscopy versus conservative management was 2 per 1,000 patients (CI 1–4). Arthroscopy may have a very small risk of infection.

Level of evidence

Pain: The level of evidence for the outcome pain (both short- and long-term) was not downgraded. Although risk of bias due to lack of blinding was a concern in most trials, trials with a low risk of bias reported similar results to those in which there were risk of bias concerns.

Function: The level of evidence for the outcome function (both short- and long-term) was downgraded by 1 level due to serious risk of bias and borderline imprecision.

Complications: The level of evidence for the outcomes VTE and infections were both downgraded by 2 levels due to serious risk of bias and serious inconsistency. There was no evidence of publication bias.

Question 5

1. Pain

Østerås et al. (2014) measured pain using a visual analogue scale (VAS), which ranged from 0 to 10 (none to most pain). Østerås et al. (2014) reported that pain was lower in patients who received physical therapy after arthroscopy compared with patients who did not receive a postoperative rehabilitation program at 12 months' follow-up. Compared with baseline, the mean difference between the intervention and control group was –1.0 (CI –1.3 to-–0.6) at 12 months' follow-up. Analyses were adjusted for baseline score.

2. Psychological problems

Østerås et al. (2014) measured symptoms of anxiety and depression via the Hospital Anxiety and Depression Scale (HADS), which ranged from 0 to 21 (least to worst). Østerås et al. (2014) reported that there were fewer psychological problems in patients who received physical therapy after arthroscopy compared with patients who did not receive a postoperative rehabilitation program at 12 months' follow-up. Compared with baseline, the mean difference between the intervention and control group was –0.7 (CI –1.1 to –0.3) at 12 months' follow-up. Analyses were adjusted for baseline score.

3. Function

Østerås et al. (2014) measured function via the Knee injury and Osteoarthritis Outcome Score 10 (KOOS), which ranged from 0 to 100 (worst to best function). However, the results suggested a decrease in KOOS score in both groups, but the results were interpreted as a beneficial effect for function. Because of this discrepancy, the results were not described.

Østerås et al. (2014) also measured function with a one-leg hop test. The pretest values were 85.6% (SD 7.8) in the group who received physiotherapy and 73.2% (SD 8.5). At 12 months' follow-up, the values were 96.7% (SD 5.1) and 81.4% (SD 8.3), respectively. The mean difference at 12 months' follow-up and adjusted for baseline values was 3.3 (CI 0.6–6.1), meaning that the group who received physiotherapy performed the test better than the group who did not receive physiotherapy.

4. Range of motion

Østerås et al. (2014) reported no data on the range of motion.

5. Muscle strength

Østerås et al. (2014) also measured strength as quadriceps muscle strength using a five-repetition maximum on a leg extension bench. Østerås et al. (2014) reported that strength as measured by the quadriceps muscle strength was better in patients who received physical therapy after arthroscopy compared with patients who did not receive a postoperative rehabilitation program at 12 months' follow-up. Compared with baseline, the mean difference between the intervention and control group was 4.4 (CI 3.2–5.6) at 12 months' follow-up. Analyses were adjusted for baseline score. The group who received physiotherapy were able to press more weight at 12 months follow-up than the group who did not receive physiotherapy.

Levels of evidence

The level of evidence for the outcome measures pain, psychological problems, and strength were downgraded by 2 levels due to a relatively small sample of patients (N = 75) and risk of bias (Østerås et al. 2014). Risk of bias was suspected because of unclear or lack of blinding regarding the treatment allocation for participants, care providers, and outcome assessors. In addition, there was significant dropout during the study and an intention-to-treat analysis was not performed.

Brignardello-Petersen R, Guyatt G H, Buchbinder R, et al. Knee arthroscopy versus conservative management in patients with degenerative knee disease: a systematic review. BMJ Open 2017; 7(5): e016114. doi: 10.1136/bmjopen-2017-016114.

Brouwers M C, Kho M E, Browman G P, Burgers J S, Cluzeau F, Feder G, Fervers B, Graham I D, Grimshaw J, Hanna S E, Littlejohns P, Makarski J, Zitzelsberger L. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010; 182(18): E839-E842.

Dai H, Huang Z G, Chen Z J, Liu J-x. Diagnostic accuracy of ultrasonography in assessing meniscal injury: meta-analysis of prospective studies. J Orthop Sci 2015; 20(4): 675-81.

Goossens P, Keijsers E, van Geenen R J, Zijta A, van den Broek M, Verhagen A P, Scholten-Peeters G G M. Validity of the Thessaly test in evaluating meniscal tears compared with arthroscopy: a diagnostic accuracy study. J Orthop Sports Phys Ther 2015; 45(1): 18-24, B1. doi: 10.2519/jospt.2015.5215.

Karachalios T, Hantes M, Zibis A H, Chambat P, Djian P, Hardy P, Potel J-F, Servien E, Seil R. Diagnostic accuracy of a new clinical test (the Thessaly test) for early detection of meniscal tears. J Bone Joint Surg Am 2005; 87: 955-62.

Kececi B, Bicer E K, Arkun R, Argin M, Taskiran E. The value of magnetic resonance arthrography in the evaluation of repaired menisci. Eur J Orthop Surg Traumatol 2015; 25(1): 173-9. doi: 10.1007/s00590-014-1450-z.

Magee T. Accuracy of 3-Tesla MR and MR arthrography in diagnosis of meniscal retear in the post-operative knee. Skeletal Radiol 2014; 43(8): 1057-64. doi: 10.1007/s00256-014-1895-5.

Østerås H, Østerås B, Torstensen T A. Is postoperative exercise therapy necessary in patients with degenerative meniscus? A randomized controlled trial with one year follow-up. Knee Surg Sports Traumatol Arthrosc 2014; 22(1): 200-6.

Phelan N, Rowland P, Galvin R, et al. A systematic review and meta-analysis of the diagnostic accuracy of MRI for suspected ACL and meniscal tears of the knee. Knee Surg Sports Traumatol Arthrosc 2016; 24(5): 1525-39.

Quatman C E, Hettrich C M, Schmitt L C, et al. The clinical utility and diagnostic performance of magnetic resonance imaging for identification of early and advanced knee osteoarthritis: a systematic review. Am J Sports Med 2011; 39(7): 1557-68.

Rinonapoli G, Carraro A, Delcogliano A. The clinical diagnosis of meniscal tear is not easy: reliability of two clinical meniscal tests and magnetic resonance imaging. Int J Immunopathol Pharmacol 2011; 24: 39-44.

Smith B E, Thacker D, Crewesmith A, et al. Special tests for assessing meniscal tears within the knee: a systematic review and meta-analysis. Evid Based Med 2015; 20(3): 88-97. doi: 10.1136/ebmed-2014-110160.

Smith C, McGarvey C, Harb Z, Back D, Houghton R, Davies A, Ajuied A. Diagnostic efficacy of 3-T MRI for knee injuries using arthroscopy as a reference standard: a meta-analysis. Am J Roentgenol 2016; 207(2): 369-77.

Vives M J, Homesley D, Ciccotti M G, Schweitzer M E. Evaluation of recurring meniscal tears with gadolinium-enhanced magnetic resonance imaging: a randomized, prospective study. Am J Sports Med 2003; 31(6): 868-73.

White L M, Schweitzer M E, Weishaupt D, Kramer J, Davis A, Marks P H. Diagnosis of recurrent meniscal tears: prospective evaluation of conventional MR imaging, indirect MR arthrography, and direct MR arthrography. Radiology 2002; 222(2): 421-9.

Xia X P, Chen H L, Zhou B. Ultrasonography for meniscal injuries in knee joint: a systematic review and meta-analysis. J Sports Med Phys Fitness 2016; 56(10): 1179-87.

Zhang M, Min Z, Rana N, Ji H, Zhang M. Accuracy of magnetic resonance imaging in grading knee chondral defects. Arthroscopy 2013; 29(2): 349-56.