PAUL F. BEATTIE, PT, PhD, FAPTA[1] • **DAVID G. GREATHOUSE,** PT, PhD, FAPTA[2,3]

Edith Holmes Met the Challenges to Lead *JOSPT* to the Next Level

Every once in a while, we savor the opportunity to marvel at the exhilarating feeling of overcoming challenges. This impact is what our profession lives for—it is the enduring core of physical therapy—the moment a patient takes those first independent steps, regains confidence, and returns to the field. But it is not just at the patient-care level where huge challenges must be overcome.

Nineteen years ago, the *Journal of Orthopaedic and Sports Physical Therapy* (*JOSPT*) was struggling as a low-visibility endeavor managed by a large publishing corporation. Forward-looking representatives of the Orthopaedic and Sports Academies (then known as "sections") decided to shift management, publishing *JOSPT* in-house as a nonprofit 501.C3 venture. The challenge was to maintain a fiscally sound journal in the rapidly changing environment of medical publishing without benefiting from the endless resources of a large corporate publisher, ie, how to develop and maintain the "economy of small."

Faced with the challenges of minimal infrastructure, small revenue streams, and low impact factor, Ms Edith Holmes joined the *JOSPT* family as the journal's Executive Director/Publisher. Her always welcoming demeanor and tireless, unselfish, and visionary work have guided *JOSPT* to become one of the world's leading musculoskeletal rehabilitation journals. After 19 years at the helm, we are privileged to honor Ms Holmes upon her retirement as Executive Director/Publisher.

Ms Holmes set the standard by integrating and communicating effectively with a diverse group of stakeholders. These stakeholders often had divergent opinions about the process of publishing, editorial review, management, and the overall product provided by the journal. In uniting this group to create and continuously deliver usable material of high value to *JOSPT* readers, Ms Holmes blended content excellence with publishing excellence. Her management strategy was to provide stable governance and exceptional customer service, to nurture and mentor a highly skilled publishing team, and to communicate continuously with the Board of Directors and the editorial team.

Now in its 43rd year, *JOSPT* is one of the go-to journals for clinicians and researchers in musculoskeletal rehabilitation (including orthopaedic and sports physical therapy). Its impact factor is continuously among the highest in the rehabilitation field. More than 80 000 unique users visit the *JOSPT* website every month. The journal receives more than 1000 submissions each year, and with 36 international partners, the publications in *JOSPT* are disseminated to a global audience. In addition to publishing high-quality manuscripts, *JOSPT* also provides numerous ancillary services, including opportunities to Read for Credit, and clinical information specifically developed for patients.

In 1979, *JOSPT*'s founding editors James Gould and George Davies challenged orthopaedics and sports physical therapy to create a peer-reviewed research journal that would be an industry leader. Under Edith Holmes' leadership, that challenge has been met.

On behalf of the *JOSPT* Board of Directors, publishing staff, and editorial team, we thank our dear friend and colleague, Ms Edith Holmes, for her superb work, inspiration, and commitment to *JOSPT*. She has given us all a moment to savor the exhilarating impact of overcoming challenges.

"Be great in act, as you have been in thought."
~ William Shakespeare

SYNOPSIS: After 19 years at the helm, we are privileged to honor Ms. Edith Holmes upon her retirement as Executive Director/Publisher of *JOSPT*. Ms. Holmes set the standard by integrating and communicating effectively with a diverse group of stakeholders. Ms. Holmes is honored for her superb work, inspiration, and commitment to *JOSPT*. *J Orthop Sports Phys Ther 2022;52(4):169. doi:10.2519/jospt.2022.11258*

KEY WORDS: *challenges, JOSPT, leadership, management*

DATA SHARING: There are no data in this paper.

AUTHOR CONTRIBUTIONS: Both authors contributed equally to writing and revising the work.

PATIENT AND PUBLIC INVOLVEMENT: No patients or public partners were involved in the work.

[1]Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, SC. [2]Clinical Electrophysiology Services, Texas Physical Therapy Specialists, Live Oak, TX. [3]Army-Baylor University Doctoral Program in Physical Therapy, San Antonio, TX. Dr Beattie is a Distinguished Clinical Professor Emeritus and is a Wilderness Emergency Medical Technician. He also served as a member of the *JOSPT* Board of Directors from 2007 to 2014 and from 2015 to 2022. Dr Greathouse is the Director of Clinical Electrophysiology Services and is a board-certified specialist in clinical electrophysiologic physical therapy. He also served as president of the *JOSPT* Board of Directors from 2005 to 2013. Dr Beattie has no competing interests or conflict of interest. Dr Greathouse has no financial disclosure or conflict of interest in the preparation and submission of this manuscript.

DANIEL L. BELAVY, PhD[1,2] • SCOTT D. TAGLIAFERRI, MSc ClinExPhys[1] • PAUL BUNTINE, MBBS FACEM MClinRes[3,4]
TOBIAS SAUERESSIG, DipI-Vw[5] • CLAIRE SAMANNA, MSc ClinExPhys[1] • THOMAS MCGUCKIAN, PhD[6]
CLINT T. MILLER, PhD[7] • PATRICK J. OWEN, PhD[1]

Reducing Low-Value Imaging for Low Back Pain: Systematic Review With Meta-analysis

OBJECTIVE: To examine the effectiveness of implementing interventions to improve guideline-recommended imaging referrals in low back pain.

DESIGN: Systematic review with meta-analysis.

LITERATURE SEARCH: We searched MEDLINE, EMBASE, the Cumulative Index to Nursing and Allied Health Literature, Web of Science Core Collection, and the Cochrane Central Register of Controlled Trials from inception to June 14, 2021, as well as Google Scholar and reference lists of relevant systematic reviews published in the last 10 years. We conducted forward and backward citation tracking.

STUDY SELECTION CRITERIA: Randomized controlled or clinical trials in adults with low back pain to improve imaging referrals.

DATA SYNTHESIS: Bias was assessed using the Cochrane Risk of Bias 2 tool. Data were synthesized using narrative synthesis and random-effects meta-analysis (Hartung-Knapp-Sidik-Jonkman method). We assessed the certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation approach.

RESULTS: Of the 2719 identified records, 8 trials were included, with 6 studies eligible for meta-analysis (participants: N = 170 460). All trials incorporated clinician education; 4 included audit and/or feedback components. Comparators were no-intervention control and passive dissemination of guidelines. Five trials were rated as low risk of bias, and 2 trials were rated as having some concerns. There was low-certainty evidence that implementing interventions to improve guideline-recommended imaging referrals had no effect (odds ratio [95% confidence interval]: 0.87 [0.72, 1.05]; I^2 = 0%; studies: n = 6). The main finding was robust to sensitivity analyses.

CONCLUSION: We found low-certainty evidence that interventions to reduce imaging referrals or use in low back pain had no effect. Education interventions are unlikely to be effective. Organizational- and policy-level interventions are more likely to be effective. *J Orthop Sports Phys Ther 2022;52(4):175-191. Epub 05 Feb 2022. doi:10.2519/jospt.2022.10731*

KEY WORDS: *back pain, chronic pain, controlled before-after studies, CT scan, diagnostic imaging, implementation science, interrupted time series analysis, low back pain, magnetic resonance, meta-analysis, radicular pain, radiography, randomized controlled trial, sciatica, systematic review, x-ray*

Adhering to clinical guidelines reduces health care costs and may improve outcomes for people with low back pain (LBP). Unnecessary imaging[20] can increase the cost of health care, impair patient outcomes, and expose patients to unnecessary radiation.[6] Studies have examined the impact of the elements of guideline-adherent care on costs[17,35] and patient outcomes[5,7]; a reduction in costs can be expected.

Evidence-based clinical guidelines consistently recommend against routine imaging for nonspecific LBP.[45] The majority (58%) of primary care guidelines support imaging when serious pathology is suspected; 42% endorse imaging if results are likely to change or direct treatment decisions.[45] Primary care guidelines from Canada and Finland also recommend imaging if LBP persists beyond 4 to 6 weeks.[45] Recommendations for imaging radicular LBP are less established. Imaging is recommended in cases of severe neurological compromise and when pain associated with radicular symptoms or spinal stenosis without severe neurological compromise persists following 1 month of conservative treatment.[11] Collectively, guidelines for nonspecific and radicular LBP tend to recommend that clinicians avoid imaging.

Adherence to clinical guidelines in primary care and the emergency department is variable. One systematic review[39] of radiation-emitting imaging (ie, x-rays, CT scans) for LBP estimated that 57% of x-rays and 46% of CT scans in primary care and the emergency department were

[1]School of Exercise and Nutrition Sciences, Institute for Physical Activity and Nutrition, Deakin University, Geelong, Australia. [2]Division of Physiotherapy, Department of Applied Health Sciences, Hochschule für Gesundheit (University of Applied Sciences), Bochum, Germany. [3]Emergency Department, Box Hill Hospital, Eastern Health, Melbourne, Australia. [4]Eastern Health Clinical School, Monash University, Melbourne, Australia. [5]Physio Meets Science, Leimen, Germany. [6]School of Behavioural and Health Sciences, Faculty of Health Sciences, Australian Catholic University, Melbourne, Australia. [7]School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia. This study was registered with PROSPERO (International Prospective Register of Systematic Reviews; registration number CRD42020215141). This study was supported by internal institutional funding only. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Daniel L. Belavy, Division of Physiotherapy, Department of Applied Health Sciences, Hochschule für Gesundheit (University of Applied Sciences), Gesundheitscampus 6-8, 44801 Bochum, Germany. E-mail: daniel.belavy@hs-gesundheit.de, belavy@gmail.com

inappropriately ordered (with low-to-very low certainty evidence). A systematic review[16] of studies published since 2010 showed that 1 in 4 people with LBP were referred for imaging in primary care and that 1 in 3 people with LBP were referred for imaging in emergency departments (with moderate-to-high certainty evidence). Another recent systematic review[31] of imaging in patients referred for LBP estimated that 32% (with moderate certainty evidence) of imaging was inappropriate. Overall, the data suggest strong potential for cost savings and reducing harm if clinical practice adheres to guidelines regarding imaging.

We aimed to build on a previous systematic review[32] (published in 2015; search was conducted in June 2014) to review approaches to implementing guidelines for imaging or imaging referrals in primary care and emergency department settings. We focused on randomized controlled or clinical trials (RCTs), as these represent the highest level of evidence. As a secondary goal, to further inform future work, we also collated information from prospective, nonrandomized interventional studies relevant to the review area.

METHODS

This review was completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[38] The review was registered prospectively in the PROSPERO (International Prospective Register of Systematic Reviews) database (registration number CRD42020215141).

Information Sources and Search Strategy

Five databases (MEDLINE [no limits], EMBASE [excluding MEDLINE], the Cumulative Index to Nursing and Allied Health Literature [excluding MEDLINE], Web of Science Core Collection [excluding MEDLINE], and the Cochrane Central Register of Controlled Trials [excluding MEDLINE, EMBASE, and trial registrations]) were electronically searched for research published from database inception to June 22, 2020. We updated the search on June 14, 2021. The search terms and strategy can be found in **SUPPLEMENTAL TABLE 1**. To locate additional relevant records, we also searched Google Scholar and included the reference lists of relevant systematic reviews (identified via the *Cochrane Database of Systematic Reviews* and Google Scholar) published in the last 10 years. The reference lists of included studies were checked for potentially relevant articles. In addition, forward citation tracking of included studies was performed by adding articles that cited the included studies in Web of Science to screening. Furthermore, reference lists (double-screened) of studies excluded solely on the basis of study design (eg, interrupted time series analyses, controlled before-after studies) were also assessed (backward citation tracking only) for potentially relevant articles.

Study Selection

All results of the search were screened to exclude duplicates. Independent screening of the titles and abstracts of the remaining studies considering predetermined eligibility criteria was completed by 2 independent reviewers (C.S. and T.M.). The full-text reports of articles that seemed eligible after this first screening were screened again. Any disagreements were adjudicated by S.D.T. and discussed with the project team as necessary.

Eligibility Criteria

Inclusion criteria followed the participants, interventions, comparators, outcomes and study design (PICOS) framework.[38] The participants (P) were adults (r18 years of age) with LBP. LBP was defined as back pain with or without leg pain where there were no specific spinal pathologies (ie, vertebral fracture, malignancy, spinal infection, axial spondyloarthritis, cauda equina syndrome).[4] Spondylolisthesis, spondylosis, disc herniation, disc degeneration, scoliosis, deformity (eg, hemivertebrae), and radicular syndromes (eg, radicular pain [leg pain or sciatica], radiculopathy, spinal stenosis) were included.[4] Failed back surgery syndrome was included, as this is not a specific disease.[12] If a study classed the population as otherwise unspecified "back pain," our predetermined criteria included this population. No limitation was placed on interventions (I), and these were classified according to procedures[42] used by the Cochrane Effective Practice and Organisation of Care (EPOC) group into professional, financial, organizational, patient-oriented, structural, or regulatory interventions. Comparators (C) that were considered were, per Cochrane EPOC procedures,[42] no-intervention control group, standard practice control group, and/or untargeted activity. Imaging referrals or use was included as outcomes (O). In regard to the study design (S), only full-text reports of analytical studies published in English were considered for inclusion. Studies published in a peer-reviewed journal (ie, gray literature excluded) with a parallel arm (individually designed or cluster-designed) RCT design were eligible. Studies that were prospective interventional trials but not randomized (eg, controlled before-after or interrupted time series designs) and excluded solely on the basis of not being an RCT were not included in the main analysis but were kept for extraction to address a secondary goal. No restrictions were placed on the date of publication.

Data Collection and Data Items

Trial data extraction was completed by 2 independent assessors (C.S. and T.M.). Extracted information included relevant publication information (ie, author, title, year, journal), study design, study funding, conflicts of interest, number of participants, participant characteristics (eg, age and sex), intervention details (eg, duration, type, frequency), cluster details for cluster-randomized trials (eg, number of clusters), and outcome measures. We also extracted data on participants' pain intensity and disability as well as any adverse events from included trials, where available. Extracted data

were the number or percentage of imaging use or referrals and the total number of participants or appointment sessions. Data from all available time points were extracted. When data were presented in figures only, rather than numerical data within text, data were extracted by generating a screenshot, loading this in ImageJ (Version 1.48v; https://imagej.nih.gov/ij/) to measure the length (in pixels) of the axes to calibrate and then the length (in pixels) of the data points of interest.[57]

In all instances where the data required for meta-analysis were not available, the authors were contacted a minimum of 3 times over a 4-week period to request the information. Similarity between extracted data from the 2 independent assessors was evaluated through custom spreadsheets set up in Microsoft Excel. Any discrepancies were discussed by the assessors, with disagreements adjudicated by S.D.T. Interventions were classified per their components as clinician education and/or workshops ("ed"); passive dissemination ("diss"); audit and/or feedback approaches ("au.fe"); reminders ("re"); administrative, electronic medical records system, policy, or organizational changes ("org"); and/or patient education ("pat.ed"). For prospective, nonrandomized interventional designs, 1 of the 2 assessors (C.S. or T.M.) extracted the data, the lead author (D.L.B.) cross-checked this information, and S.D.T. adjudicated in case of conflicts.

Risk of Bias in Individual Studies

The Cochrane Collaboration Risk of Bias 2 tool[51] was used to examine potential bias from the randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and the selection of the reported results for individually randomized and cluster-randomized trials. Each domain was assessed as risk of bias and labeled as "low risk," "some concerns," or "high risk." For each source of bias, trials were then classified as low risk, some concerns, or high risk, as per the overall algorithm. The 2 assessors evaluated this independently, and any disagreements for the risk of bias were adjudicated by S.D.T. Nonrandomized interventional studies were not assessed for risk of bias.

Synthesis of Results

The evidence synthesis for this review was conducted in accordance with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) guidelines.[23]

Statistical Analysis

For the statistical analysis of included randomized studies, we created 2 categories of comparators: (1) multifaceted intervention and (2) no-intervention control group and standard practice control group according to the EPOC guidelines.[42] Our primary outcome measure was imaging referrals or use. As effect size, we used the odds ratio (OR), as it has favorable statistical properties compared to the risk ratio.[15] For an easier interpretation of the OR, we transformed it into a risk difference, with corresponding 95% confidence intervals (CIs). For this transformation, we used an assumed comparator risk based on a systematic review and meta-analysis[16] of 24.8% (95% CI, 19.3% to 31.1%).

Cluster RCTs were handled by calculating a design effect to correct for clustering of the trials. The design effect is approximately $1 + (M - 1) \times ICC$, where M is the average cluster size and ICC is the intracluster (or intraclass) correlation coefficient.[26] The sample size and the number of events were divided by the design effect to adjust sample sizes and the number of events for the corresponding trial[26] and to avoid overestimating the precision of the estimate. We used either published ICCs[36] or the most conservative value from the available ICCs to inform a choice for other studies without available ICCs. We also performed a sensitivity analysis with a range of different ICCs to check the robustness of the results.[26]

A random-effects meta-analysis was used for dichotomous outcomes with a Paule-Mandel estimator for the between-study variance T^2; an SSW (or the sum of squares within groups) estimator for the overall effect, with weights that depended only on the studies' effective sample sizes; and a 95% CI for the overall effect based on the Hartung-Knapp-Sidik-Jonkman method. We used this method because it outperforms the standard random-effects method and other methods.[1] Measures of heterogeneity were Cochran's Q and the resulting chi-square statistic and I^2. To assess the amount of heterogeneity, 95% prediction intervals were used if there were at least 10 trials in the meta-analysis.[8] Publication bias was assessed via funnel plots, Egger's test, and trim-and-fill methods if at least 10 trials were included in the meta-analysis.[41] We performed a sensitivity analysis via outlier identification and influence analysis.[56] All calculations and graphics were performed with software R[46] and extension packages Meta[3] and dmetar.[25]

RESULTS

A summary of the systematic review process is shown in **SUPPLEMENTAL FIGURE 1**. There were 2719 records (after the removal of 1859 duplicates) included in the initial title and abstract screening. Following the title and abstract screening, there were 101 studies included in the full-text screening. The examination of full texts resulted in 93 studies being excluded (**SUPPLEMENTAL TABLE 2**) and 8 trials[10,13,14,17,18,22,35,49] being included (**TABLES 1** and **2**). Of these, 6 trials[10,13,14,18,22,49] were eligible for meta-analysis (participants: N = 170 460), but 2 trials[17,35] could not be included in quantitative synthesis, as we were unable to extract or acquire data required for quantitative synthesis (ie, the total number or percentage of imaging referrals or use relative to the total number of patients or appointments).

Study Characteristics

Population The sample sizes included in the intervention phases of the trials

TABLE 1 Characteristics of Studies Included in the Review

Author	Year	Study Design	N INT	N CON	Type LBP	N Clusters INT	N Clusters CON	Cluster Type	INT Type[a]	INT Period	INT	CON	Funding	Study Conclusions
Cherkin et al[10,b]	2018	Cluster RCT	2106 (943 baseline, 1163 study period)	2534 (1061 baseline, 1473 study period)	Mixed	3	3	Primary care clinics	Professional (ed)	5 months	Six training sessions on the use of the STarT Back tool and on the diagnosis and management of pain	No intervention	None or public/ nonprofit	The intervention was not effective in reducing the rates of imaging for LBP.
Coombs et al[13]	2021	Cluster RCT	1392	3233	Mixed	4[c]	4[c]	Hospital emergency departments	Professional (ed|au.fe)	20 months	Multifaceted: 1. Education seminars 2. Education materials 3. Provision of non-opioid pain management 4. Fast-track referral 5. Audit and/or feedback	Usual care	None or public/ nonprofit	A multifaceted intervention for guideline implementation did not improve imaging rates.
Dey et al[14]	2004	Cluster RCT	1049	1138	Acute	12	11	Health centers	Professional (ed|diss)	8 months	Outreach visits with interactive discussion between the guideline team and general practitioners on guidelines plus a poster including guideline recommendations	No intervention	Not reported	The intervention did not result in a reduction in imaging referral.
Eccles et al[17]	2001	Cluster RCT	Unclear	-	Unclear	Audit/feedback = 60, reminder = 61, reminder and audit/ feedback = 62	61	Practices	Professional (ed|au.fe|re)	6 months	Audit and/or feedback, educational reminder messages, audit and/or feedback and educational reminder messages	Postal dissemination of guidelines only	None or public/ nonprofit	Educational reminders appear (significance not reported) to be an effective intervention to reduce imaging referral rates for LBP, but audit and/or feedback does not.
Fenton et al[18]	2016	Cluster RCT	Not reported	-	Unclear	30	31	Medical residents	Professional (ed|au.fe)	1 year	Standardized patient instructors with education of patientcentered care, feedback, and focus groups	Standard patient encounter without feedback and guideline dissemination via e-mail	None or public/ nonprofit	The SP-based intervention to improve patient centeredness did not improve imaging rates.

French et al[22]	2013	Cluster RCT	77 716 (visit to practice with any clinical condition)	75 226 (visit to practice with any clinical condition)	Acute	45	47	Practices	Professional (ed)	8 months	Interactive workshops including lectures, small-group discussions, and activities related to acute low back pain management	Postal dissemination of guidelines	None or public/ non-profit	There was no difference in imaging referral rates between intervention and control groups.
Kerry et al[35]	2000	Cluster RCT	Not reported for low back pain	-	Unclear	33	36	Practices	Professional (diss|au.fe)	6 months	Distribution of clinical guidelines and individual feedback on referral rates	No intervention	None or public/ non-profit	The combination of guideline dissemination and feedback on referrals appeared to be an effective strategy for reducing lumbar spine imaging.
Schectman[49,d]	2003	Cluster RCT	2232 (1014 baseline, 1218 study year)	1834 (1006 baseline, 828 study year)	Acute	Unclear	Unclear	Practices	Professional (ed|diss|au.fe)	1 year	90-minute clinical guideline, educational session, educational vignettes, audit report on care during baseline year, and a copy of the guidelines	No intervention, patient education materials (pamphlet and videotape)	None or public/ non-profit	Imaging rates appeared (significance not reported; nonsignificant in our calculations; see FIGURE 1) to reduce change from 31% to 19% for x-ray and from 7.6% to 5.6% for CT/MRI in the "Physician Education and Feedback" group as well as change from 21% to 18% for x-ray and from 5.6% to 7.1% for CT/MRI in the "Physician Control" group.

Abbreviations: CON, control; CT/MRI, CT scan/magnetic resonance imaging; INT, intervention; LBP, low back pain; RCT, randomized controlled or clinical trial; SP, standardized patient.
[a]*The specific classification of each intervention is described as follows: "ed," clinician education/workshops; "diss," passive dissemination; "au.fe," audit and feedback to clinicians; "re," reminders to clinicians; "org," organizational change, administrative changes, electronic medical records system changes, and/or government policy change; "pat.ed," patient education.*
[b]*Cherkin et al, 2018: Imaging outcomes include all participants who visited clinics (not just those who provided patient-reported outcomes).*
[c]*Coombs et al, 2021: Step-wedged cluster RCT with 4 clusters total, with control and intervention periods.*
[d]*Schectman et al, 2003: The authors were contacted to confirm the numbers in the analysis. The numbers differ from prior reviews but were confirmed by the authors. For clusters randomized, a total of 14 practices were utilized; however, it was unclear in terms of the total number of clusters randomized to each arm.*

TABLE 2 Outcomes of Studies Included in the Review

Author	Year	Outcome	Baseline INT	Follow-up INT	Baseline CON	Follow-up CON	*P* Value	n INT for Analysis[a]	N INT for Analysis[a]	n CON for Analysis[a]	N CON for Analysis[a]	ICC[b]
Cherkin et al[10,c]	2018	Proportion and 95% confidence interval for imaging use	0.16 (0.08, 0.25)	0.22 (0.11, 0.34)	0.14 (0.07, 0.22)	0.18 (0.09, 0.27)	.578	256	1163	265	1473	-
Coombs et al[13]	2021	Number and percentage of use	-	327 (23.5)	-	774 (23.9)	.290	327	1392	774	3233	0.016
Dey et al[14]	2004	Total number and % of referrals for x-ray Baseline values only reported as % of 365 individuals	20.40%	158 (15.1)	21.10%	156 (13.7)	.62	158	1049	156	1138	0.018
Eccles et al[17]	2001	Mean and standard deviation imaging requests per 1000 patients	Audit/feedback: 7.24 (4.8) Reminder: 7.31 (5.2) Audit/feedback and reminder: 8.30 (5.1)	Audit/feedback: 5.97 (4.2) Reminder: 5.14 (3.7) Audit/feedback and reminder: 5.23 (3.7)	7.53 (4.1)	6.80 (4.3)	Not reported	-	-	-	-	-
Fenton et al[18,d]	2016	Incidence of spinal CT/MRI rates per 100 visits: Raw number of xray/CT/MRI utilization against total visits	2.3: 841/7796	2.5: 813/8829	2.2: 875/7742	2.2: 1092/10 582	.78	813	8829	1092	10 582	-
French et al[22]	2013	Number of referrals and rate per 1000 patients (both CT and x-ray referrals)		1117 (14.4)		1264 (16.8)	.244	1117	77 716	1264	75 226	X-ray: 0.004 (0003, 0.006) CT: 0.003 (0.002, 0.005)
Kerry et al[35]	2000	% change in referrals		15% reduction		5% increase	Not reported Difference 95% CI between groups = 20% (3 to 37)	-	-	-	-	-
Schectman et al[49,e]	2003	Total utilization % of patients by episode of care; % not consistent with guidelines	LS x-ray total: 31 LS CT or MRI total: 7.6 LS x-ray not consistent: 14.5 LS CT or MRI not consistent: 5.7	LS x-ray total: 19 LS CT or MRI total: 5.6 LS x-ray not consistent: 8.1 LS CT or MRI not consistent: 3.5	LS x-ray total: 21 LS CT or MRI total: 5.6 LS x-ray not consistent: 8.2 LS CT or MRI not consistent: 3.5	LS x-ray total: 18 LS CT or MRI total: 7.1 LS x-ray not consistent: 8.6 LS CT or MRI not consistent: 5.4	Not reported for imaging alone	X-ray: 231 CT/MRI: 68	1218	X-ray: 149 CT/MRI: 59	828	-

Abbreviations: CI, confidence interval; CT/MRI, CT scan/magnetic resonance imaging; ICC, intracluster correlation coefficient; LS, lumbar spine.
[a]The n INT/CON number refers to the raw number of imaging use or referrals, whereas the N INT/CON number refers to the total number of participants. The numbers used in analysis reported in the table are those from the primary analysis prior to ICC adjustment.
[b]ICC was used as per the Cochrane guidelines to calculate the effective sample size from cluster-randomized trials.
[c]Cherkin et al, 2018: Imaging outcomes include all participants who visited clinics (not just those who provided patient-reported outcomes). Calculations for the main data prior to ICC adjustment were as follows: Intervention: 1163 × 0.22 = 255.86 and Control: 1473 × 0.18 = 265.14. To get the raw numbers for imaging use, the proportion of use per arm was multiplied by the total number of participants in each arm.
[d]Fenton et al, 2016: Authors provided raw data for utilization of x-ray/CT/MRI for back pain against the total number of visits for any clinical condition.
[e]Schectman et al, 2003: The authors were contacted to confirm the numbers in the analysis. The numbers presented here differ from those presented in prior reviews but were confirmed by Professor Schectman for the current review; thus, the values used here are correct. Calculations for the main data prior to ICC adjustment were as follows: X-ray Intervention: 1218 × 0.19 = 231.42, X-ray Control: 828 × 0.18 = 149.04, CT/MRI Intervention: 1218 × 0.056 = 68.208, and CT/MRI Control: 828 × 0.071 = 58.788. These calculations were done to get the raw numbers for imaging use, by multiplying the percentage of use per arm by the total number of participants in each arm.

ranged from 2046 to 152 942. The number of patients included in 2 trials[17,35] was unclear. Cluster sizes ranged from 4 to 244. The number of clusters in 1 trial[49] was unclear. Attempts to contact authors to obtain this information were not successful. Three trials[14,22,49] examined patients with acute LBP; 2 trials[10,13] examined patients of mixed pain duration, and the majority of patients had less than 3 months of pain. In 3 trials,[17,18,35] the duration of pain was unclear.

Intervention All trials[10,13,14,17,18,22,49] incorporated education and/or workshop component interventions for clinicians, with 5 trials[13,17,18,35,49] incorporating audit and/or feedback components and 1 trial[17] including additional reminders for clinicians. One trial[49] also included educational materials for patients. Another trial[13] implemented fast-track referral systems and non-opioid pain management.

Comparator Five trials[10,13,14,35(p200),49] used a no-intervention control for clinicians, and 3 trials[17,18,22] used passive dissemination of guidelines via posts or e-mails. One trial[49] implemented patient education materials in a control group.

Outcome Seven trials[10,13,14,17,18,22,49] reported data as some form of proportion or percentage of imaging referrals or use. One trial[35] presented data as percentage change in imaging referrals. Three trials[17,35,49] reported an impact of the intervention on imaging rates but did not report statistical significance or lack thereof. The remaining trials reported no impact of the intervention on imaging.

All included studies were cluster RCTs. Seven trials[10,13,17,18,22,35,49] either received no funding or were funded by a public/not-for-profit organization. One trial[14] did not report funding sources.

Risk of Bias and GRADE Assessment

Six trials[10,13,14,18,22,35] were rated as low risk of bias, and 2 trials[17,49] were rated as having some concerns (**TABLE 3**). There was low risk of bias for randomization and timing of randomization in 75% and 88% of the trials, respectively; low risk of bias for deviations from interventions in 88% of the trials; low risk of bias for missing outcome data in 88% of the trials; low risk of bias for measurement of outcome in all trials; and, finally, low risk of bias for the selection of results in 88% of the trials. The certainty of evidence (using the GRADE framework) was rated for meta-analytic outcomes of imaging referrals or use as low. The main reasons for downgrading the evidence were study quality (1 level) and imprecision (1 level). Publication bias could not be assessed because there were fewer than 10 trials.[41]

Quantitative Analysis

Six trials[10,13,14,18,22,49] were included in the meta-analysis. For the primary analysis, we used a conservative ICC value of 0.02 for all trials that had no estimate, and for the other trials, we used the published estimate. For our primary outcome measure (imaging referrals or use), we estimated a nonsignificant OR of 0.87 (95% CI, 0.72 to 10.05; I^2 = 0%; 95% CI, 0% to 54.7%; n = 6 trials; GRADE: low certainty) in favor of the intervention group (**FIGURE 1**). The transformation of the OR into a risk difference with a baseline risk of 24.8% gives a nonsignificant number fewer than 1000 = 25 with a 95% CI of 56 to −10. For every 1000 patients with LBP, 248 are typically referred for imaging. With the intervention, 223 patients (25 patients fewer) for every 1000 would still be referred for imaging. This difference was not statistically significant.

We performed several sensitivity analyses (**TABLE 4**). First, we checked if there were potential outliers or influential trials and what impact removing these studies would have on the overall summary effect size. We identified 1 influential trial,[22] but removing this trial had no substantial effect. With the second sensitivity analysis, we removed 2 trials[18,22] because they reported imaging numbers and total visits for any clinical population (not back pain specific), whereas the other trials used number or percentage of imaging referrals or use for LBP. Excluding these 2 trials had a trivial impact on the effect size in favor of the control group (OR, 1.09;

TABLE 3 Risk of Bias for Cluster-Randomized Trials Using the Cochrane Risk of Bias 2 Tool

Author	Year	Domain 1a Randomization	Domain 1b Timing of Randomization	Domain 2 Deviations From Interventions	Domain 3 Missing Outcome Data	Domain 4 Measurement of Outcome	Domain 5 Selection of Results	Overall
Cherkin et al[10]	2018	Low	Low	Low	Low	Low	Low	Low
Coombs et al[13]	2021	Low	Low	Low	Low	Low	Low	Low
Dey et al[14]	2004	Low	Low	Low	Low	Low	Low	Low
Eccles et al[17]	2001	Some concerns	Low	Low	Low	Low	Low	Some concerns
Fenton et al[18]	2016	Low	Low	Low	Low	Low	Low	Low
French et al[22]	2013	Low	Low	Low	Low	Low	Low	Low
Kerry et al[35]	2000	Low	Low	Low	Low	Low	Low	Low
Schectman et al[49]	2003	Some concerns	Some concerns	Some concerns	Some concerns	Low	Some concerns	Some concerns

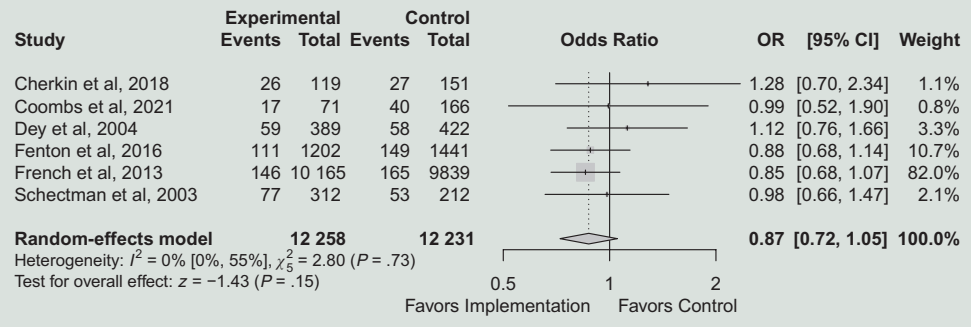

FIGURE 1. Forest plot for the meta-analysis investigating the effectiveness of implementation interventions vs control group (standard practice control group according to the Cochrane Effective Practice and Organisation of Care guidelines) for the primary outcome measure of imaging referrals or use. See also **TABLE 1** for more details on the included studies. Per the Cochrane guidelines,[26] the data (number of events/number of participants) of cluster-randomized trials are transformed via the design effect (1 + [M − 1] × ICC) prior to meta-analysis (see the Statistical Analysis section for more details). The raw outcome data from each study are in **TABLE 2**. Abbreviations: CI, confidence interval; OR, odds ratio.

95% CI, 0.85 to 1.38; I^2 = 0%; 95% CI, 0% to 26.9%; n = 4 trials; GRADE: low certainty) and did not change the overall conclusion. Two further sensitivity analyses checked if different ICC values would change the results, and they did not (**TABLE 4**). We conclude that the results of our main analysis are robust.

Secondary Goal: Nonrandomized Interventional Studies

Nonrandomized interventional studies that were excluded from the main review solely on the basis of study design are presented in **TABLE 5**. Three studies[21,52,58] were controlled before-after designs, and 15 studies[2,9,19,24,29,30,34,37,40,43,44,47,53-55] were interrupted time series designs. Six studies[9,24,47,52-54] contained components of clinician education and/or workshops; 9 studies,[19,21,29,34,40,47,52,53,58] passive dissemination; 1 study,[54] audit and/or feedback approaches; 10 studies,[2,19,24,30,34,37,43,44,54,55] administrative, electronic medical records system, policy, or organizational changes; and 2 studies,[43,52] patient education. The median percentage change in imaging associated with each type of intervention is presented in **SUPPLEMENTAL TABLE 3**. Interventions that included a component of administrative, electronic medical records system, policy, or organizational changes showed a median reduction in imaging of 6.6%, whereas interventions that did not include an aspect of this kind of process change showed a median reduction in imaging use of 2.5%.

Protocol Deviations From PROSPERO Registration

We initially planned to extract the mean and standard deviation of imaging use and pooled data through meta-analysis using standardized mean differences. However, at data extraction, we noted that the extraction of imaging use against the total number of participants/clinical consultations was more relevant for data analysis. Subsequently, we analyzed the

TABLE 4 Sensitivity Analyses

Outcome	Type of Sensitivity Analysis	Excluded Influential Studies	Meta-analytic Result of Main Analysis (OR [95% CI] I^2 [95% CI] Number of Studies)	Result of Sensitivity Analysis (OR [95% CI] I^2 [95% CI] Number of Studies)	Likely Impact on Meta-analytic Result
Imaging referrals or usage	Outlier and influential study analysis	French et al, 2013[22]	0.87 [0.72, 1.05] I^2 = 0.0% [0.0%, 54.7%] N = 6	0.96 [0.80, 1.15] I^2 = 0.0% [0.0%, 55.4%] N = 5	No substantial impact
Imaging referrals or usage	French et al, 2013,[22] and Fenton et al, 2016,[18] studies looked at imaging numbers against total visits for any clinical population, not just low back pain.	French et al, 2013,[22] and Fenton et al, 2016[18]	0.87 [0.72, 1.05] I^2 = 0.0% [0.0%, 54.7%] N = 6	1.09 [0.85, 1.38] I^2 = 0.0% [0.0%, 26.9%] N = 4	Trivial impact in favor of the control group. This does not change the nonstatistically significant result.
Imaging referrals or usage	ICC = 0.01 for studies not reporting an ICC	None	0.87 [0.72, 1.05] I^2 = 0.0% [0.0%, 54.7%] N = 6	0.88 [0.74, 1.04] I^2 = 0.0% [0.0%, 69.1%] N = 6	No substantial impact
Imaging referrals or usage	ICC = 0.004 for studies not reporting an ICC	None	0.87 [0.72, 1.05] I^2 = 0.0% [0.0%, 54.7%] N = 6	0.88 [0.74, 1.05] I^2 = 18.7% [0.0%, 63.4%] N = 6	No substantial impact

Abbreviations: CI, confidence interval; ICC, intracluster correlation coefficient; OR, odds ratio.

data using ORs. The numerical summary of nonrandomized interventional studies (**SUPPLEMENTAL TABLE 3**) was not pre-planned and added during the peer review process. One new cluster RCT,[50] identified via hand searching, was included during second-round revisions without updating the database search. There were no other protocol deviations.

DISCUSSION

This review examined the effects of interventions on optimizing imaging referrals in patients with LBP. All included RCTs examined clinician-level interventions, with some also providing resources to patients. Overall, we found low-certainty evidence that the interventions had no impact on imaging referrals or use in primary care or emergency departments. Nonetheless, the included studies provided some indications for more successful efforts in the future.

Clinician education and/or workshops was a key component of the interventions; a number of interventions included a passive dissemination component (ie, providing clinicians with information on guidelines). Education and passive dissemination are, in isolation, likely ineffective approaches. Some studies also incorporated audit and/or clinician feedback components.[17,18,28,33,35,48,49] None of the trials implemented organizational- and/or medical record–type interventions (eg, changes to medical record procedures to facilitate guideline-consistent management approaches).

To inform discussion of potentially effective interventions, we also collated and extracted nonrandomized interventional studies. Studies including organizational changes (eg, changed clinical practice structure and patient flow, enforcing stricter imaging ordering requirements, embedding changes in electronic medical records systems to guide clinicians) or that were subject to changes in government funding models appeared to show the largest and most consistent reductions in imaging rates (median reduction in imaging of 6.6% in these studies and of 2.5% in the remaining studies). While clinician and patient engagement and education are an important part of any intervention, we suggest that organizational, policy, and funding model changes are more likely to drive reductions in unnecessary imaging.

Effective interventions (eg, organizational changes) require urgent implementation, and we have provided suggestions on the basis of the current evidence. On a practical level, cluster RCTs are likely the most feasible high-quality study design to implement, as opposed to trials where patients are randomized to different interventions. Controlled before-after and interrupted time series designs are often easier to perform in an organizational setting, but randomized designs are required

TABLE 5 Nonrandomized, Prospective Trials on Imaging Rates (Excluded From Review Solely Due to Study Design)

Author	Year	Study Design	INT	INT Type[a]	CON	Outcome	Baseline INT	Follow-up INT	INT % Change	Baseline CON	Follow-up CON	CON % Change	INT-CON % Difference	*P* Value
Fine et al[19]	2017	ITS	Ontario (Canada) government policy change in 2012: changes to wording of billing codes, distribution of guidelines, development of online clinical decision aid tool	diss|org	Pre-implementation period	No. of tests per year	644 286	548 106	−14.9%	484 633	511 776	5.6%	6.9%	-
Kennedy et al[34]	2014	ITS	Ontario (Canada) government policy change in 2012: changes to wording of billing codes, distribution of guidelines, development of online clinical decision aid tool	diss|org	Pre-implementation period	No. of referrals per month	246 (20.1)	232.7 (38.3)	−5.4%	-	-	-	-	>.05
Klein et al[37]	2000	ITS	Specialist spine clinic with a physiatrist, nursing staff, and an exercise physiologist; linked with other clinics; updated practice guidelines and patient education materials. Practice guidelines used in clinical practice.	org	Pre-implementation period	% reduction in use	-	CT: 81% reduction MRI: 28% reduction	CT: 81% reduction MRI: 28% reduction	-	-	-	-	-
Matowe et al[40]	2002	ITS	Postal guideline dissemination	diss	Pre-implementation period	Change in referral use	147.7	Mean after: 124.5 Change in referral trend 13 months: −0.4	−15.7%	-	-	-	-	.001

Min et al[43]	2017	ITS	Electronic clinical decision support. A point-of-care checklist of accepted red flags was embedded in electronic medical record systems for lumbar imaging. Patient education summary with information on when imaging is appropriate.	org|pat.ed	Pre-implementation period	Proportion of patients referred for imaging	22%	17%	−5%	-	-	-	-	.0002
Rao et al[47]	2002	ITS	Education: seminars conducted by orthopedic surgeons and rehabilitation physicians. Distribution of guidelines and relevant literature.	ed|diss	Pre-implementation period	Total number	141/279, 51%	137/261, 52%	−2.8%	-	-	-	-	-
Suman et al[52]	2018	CBA	Multicomponent medical education, interdisciplinary communication and collaboration, and patient-physician communication. Additional online and offline education materials.	ed|diss|pat.ed	Usual care	Total number	200 (14%)	138 (11%)	−31%	145 (12%)	137 (11%)	−5.5%	0.7%	-
Werner et al[58]	2008	CBA	Media campaign including guideline dissemination, information about the project, waiting room posters and handouts. Materials also sent directly to health care providers.	diss	Control county (usual care)	N per 100 000	3387	3325	−1.8%	3250	5087	56.5%	41.9%	-
Hong et al[29]	2017	ITS	"Choosing Wisely" mass media campaign	diss	Nil	%	Low value: 25.15% High value: 37.22%	Low value: 22.07% High value: 36.14%	Low value: −3.1% High value: −1.1%	-	-	-	-	<.05

Table continues on page 186.

TABLE 5 Nonrandomized, Prospective Trials on Imaging Rates (Excluded From Review Solely Due to Study Design) (continued)

Author	Year	Study Design	INT	INT Type[a]	CON	Outcome	Baseline INT	Follow-up INT	INT % Change	Baseline CON	Follow-up CON	CON % Change	INT-CON % Difference	*P* Value
Freeborn et al[21]	1997	CBA	Passive dissemination and feedback on imaging requests	diss	Delayed dissemination	Test per 1000 visits of patients, split into internal medicine physicians and family practice physicians; into x-rays, CT scans, and MRI scans; and into guidelines only vs guidelines plus feedback.	Internal medicine physicians: X-ray: pre-guidelines, 8.52 (5.63); pre-feedback, 8.07 (6.15) CT scans: pre-guidelines, 2.76 (2.75); pre-feedback, 4.15 (3.15) MRI scans: pre-guidelines, 0.21 (0.76); pre-feedback, 0.35 (0.60) Family practice physicians: X-ray: pre-guidelines, 11.16 (5.32); pre-feedback, 10.51 (5.47) CT scans: pre-guidelines, 4.72 (3.51); pre-feedback, 4.30 (4.38) MRI scans: pre-guidelines, 0.25 (0.39); pre-feedback, 0.36 (0.61)	Internal medicine physicians: X-ray: post-guidelines, 8.93 (5.53); post-feedback, 8.66 (4.76) CT scans: post-guidelines, 3.66 (3.24); post-feedback, 2.86 (2.08) MRI scans: post-guidelines, 0.19 (0.50); post-feedback, 0.35 (0.73) Family practice physicians: X-ray: post-guidelines, 12.32 (6.53), post-feedback 11.15 (6.94). CT scans: post guideline 4.26 (4.75), post-feedback 3.36 (5.10) MRI scans: post-guidelines, 0.16 (0.32); post-feedback, 0.15 (0.23)	Internal medicine physicians: X-ray: post-guidelines, 4.8%; post-feedback, 7.3% CT scans: post-guidelines, 32.6%; post-feedback, −31.1% MRI scans: post-guidelines, −9.5%; post-feedback, 0% Family practice physicians: X-ray: post-guidelines, 10.4%; post-feedback, 6.1% CT scans: post-guidelines, −9.7%; post-feedback, −21.9% MRI scans: post-guidelines, −36%; post-feedback, −58.3%	Internal medicine physicians: X-ray: pre-guidelines, 9.06 (4.77); pre-feedback, 7.94 (4.13) CT scans: pre-guidelines, 2.32 (2.24); pre-feedback, 3.27 (2.52) MRI scans: pre-guideline 0.21 (0.67), pre-feedback 0.31 (0.68). Family practice physicians: X-ray: pre-guidelines, 8.57 (7.84); pre-feedback, 7.08 (4.87) CT scans: pre-guidelines, 3.38 (2.30); pre-feedback, 3.57 (2.06) MRI scans: pre-guidelines, 0.51 (0.99); pre-feedback, 0.17 (0.45)	Internal medicine physicians: X-ray: post-guidelines, 9.82 (5.49); post-feedback, 7.66 (4.70) CT scans: post-guidelines, 3.55 (2.8); post-feedback, 2.58 (2.25) MRI scans: post-guideline 0.33 (0.70), post-feedback 0.16 (0.45). Family Medicine Physicians: Xray: post-guideline 8.03(5.89), post-feedback 8.05 (5.15). CT scans: post-guidelines, 2.79 (2.82); post-feedback, 1.72 (1.54) MRI scans: post-guidelines, 0.42 (1.04); post-feedback, 0.40 (0.67)	Internal medicine physicians: X-ray: post-guidelines, 8.4%; post-feedback, −3.5% CT scans: post-guidelines, 53%; post-feedback, −21.1% MRI scans: post-guidelines, 57.1%; post-feedback, −48.4% Family practice physicians: Xray: post-guideline 6.3%, post-feedback 13.7%. CT scans: post guideline 17.5%, post-feedback 51.8% MRI scans: post-guideline 17.6%, post-feedback 135%	Internal medicine physicians: X-ray: post-guidelines, 9.5%; post-feedback, 12.3% CT scans: post-guidelines, 3.1%; post-feedback, 10.3% MRI scans: post-guidelines, 53.8%; post-feedback, 74.5% Family practice physicians: X-ray: post-guidelines, 42.2%; post-feedback, 32.3% CT scans: post-guidelines, 41.7%; post-feedback, 64.6% MRI scans: post-guidelines, 89.7%; post-feedback, 90.9%	>.05

Tahvonen et al[53]	2017	ITS	Passive dissemination via e-mail 3 times during study period. Educational lectures were offered. Article in local newspaper. Single radiology department.	ed|diss	Pre-implementation period	N per 100 000 inhabitants	603	After 2011 guidelines: 318, after 1-year follow-up: 285	−47.3%	-	-	-	-	<.001
Tracey et al[54]	1994	ITS	Chart audit, feedback to staff, education and updated referral requirements for radiology (referral accepted only if form correctly filled)	ed|au.fe|org	Pre-implementation period	n (%) in 2 months before guidelines and 2 months after	62/128 (48.4%)	50/184 (27.2%)	−21.2%	-	-	-	-	.0002
Moskowitz et al[44]	2000	ITS	Clinical guidelines restricting imaging referrals among nonradiologists	org	Pre-implementation period	N per 1000 patients	1995: 10.4	1997: 9.6	−7.7%	-	-	-	-	-
Baker et al[2]	1987	ITS	Special requisition test ordering forms restricting referral reasons to radiology	org	Pre-implementation period	n	Control year:1443	759	−47.4%	-	-	-	-	-
Ip et al[30]	2014	ITS	Integrated guidelines in electronic imaging requesting system	org	Pre-implementation period	N lumbar spine MRI scans per 100 office visits	5.3 per 100 (443/8437) office visits	3.7 per 100 (477/13 008) office visits	−30.2%	-	-	-	-	<.0001
Chandra et al[9]	2019	ITS	Education seminar for the Canadian version of the "Choosing Wisely" campaign	ed	Pre-implementation period	n, %, 95% CI	94 (12; 9.9-14.5)	108 (16.2;13.6-19.2)	14.9%	-	-	-	-	.023

Table continues on page 188.

TABLE 5 Nonrandomized, Prospective Trials on Imaging Rates (Excluded From Review Solely Due to Study Design) (continued)

Author	Year	Study Design	INT	INT Type[a]	CON	Outcome	Baseline INT	Follow-up INT	INT % Change	Baseline CON	Follow-up CON	CON % Change	INT-CON % Difference	*P* Value
Haig et al[24]	2019	ITS	Standardized clinical forms for patient encounter for assessment and management. Emergency department (single-site study) staff education implemented regularly by study leader. Standardized physical therapy protocol for patients with noncomplicated but significant acute back pain. Physiatry spine clinic referral for more complicated cases; this clinic consulted with patients within 48 h.	ed|org	Pre-implementation period	% prescribed before, during and after the intervention	Pre-intervention: MRI: 12% CT: 9% X-ray: 28%	During intervention: MRI: 9% CT: 13% X-ray: 29% 6-month follow-up: MRI: 9% CT: 16% X-ray: 34%	During intervention: MRI: −3% CT: 4% X-ray: 1% 6-month follow-up: MRI: −3% CT: 7% X-ray: 6%	-	-	-	-	MRI: .89 CT: .32 X-ray: .72
Venkatesh et al[55]	2021	ITS	E-QUAL Avoidable Imaging Initiative	org	Pre-implementation period	Imaging utilization rates (%)	Pre-intervention: X-ray: 36% CT: 20.1% MRI: 0.8%	During intervention: X-ray: 33.2% CT: 17.7% MRI: 0.7%	During intervention: X-ray: −2.8% CT: −2.4% MRI: −0.1%	-	-	-	-	X-ray: .095 CT: .09 MRI: .777

Abbreviations: CBA, controlled before-after; CT, computed tomography; ITS, interrupted time series; MRI, magnetic resonance imaging.
[a]The specific classification of each intervention is described as follows: "ed," clinician education/workshops; "diss," passive dissemination; "au.fe," audit and feedback to clinicians; "re," reminders to clinicians; "org," organizational change, administrative changes, electronic medical records system changes, and/or government policy change; "pat.ed," patient education.

as high-quality evidence for guiding practice in the future and informing guidelines. We recommend cluster RCTs of interventions that include a key component of organizational (eg, patient flow, changes to imaging referral requirements, and/or changes to electronic medical records systems) change.

Quality of reporting of imaging rates was varied. As recommended in the Cochrane guidelines,[27] future studies should report the total number of tests or referrals relative to the total number of appointments for the condition studied to allow for appropriate meta-analysis. Future cluster RCTs should always[27] report the ICC for their study. The ICC is a measure of how similar patients within the clusters are.[36] The ICC is required in meta-analysis to adjust the sample sizes for cluster size in RCTs. This was not done in a prior related review,[52] which risks overestimating[26] the precision of the included trials and the pooled main effect. Future RCTs should consider reporting how their interventions align with the EPOC guidelines.[42] One recent, potentially relevant cluster RCT[50] was excluded, as we judged that the intervention did not fall under EPOC-related patient interventions and that the therapist aspect did not clearly fall under EPOC criteria (see **SUPPLEMENTAL TABLE 2** for more details).

Limitations

Exploring the heterogeneity in the main data estimates (we implemented sensitivity analyses to examine the potential role of outliers, outcome type, and assumptions on ICC values) is a strength of our systematic review. We used more efficient[1] (ie, better coverage probability and less bias estimating the between-study variance τ) meta-analytic methods than in a prior review.[26] The main limitation is the limited pool of RCTs. Furthermore, the interventions were not all the same. Subgroup analyses on types of interventions or components of interventions were not possible due to the low number of studies.

CONCLUSION

We found low-certainty evidence that interventions to reduce imaging referrals or use in LBP had no effect. Education interventions are unlikely to be effective, and based on additional review of prospective, nonrandomized interventional studies, organizational- and policy-level interventions are more likely to be effective. ◉

KEY POINTS

FINDINGS: Education interventions are likely ineffective for reducing imaging in back pain.

IMPLICATIONS: Guideline implementation approaches from (cluster) RCTs are ineffective. Based on additional review of prospective, nonrandomized interventional studies, organizational- and policy-level interventions may be more effective.

CAUTION: There is a limited pool of RCTs.

STUDY DETAILS

AUTHOR CONTRIBUTIONS: Drs Belavy and Owen contributed to the conceptualization and provided resources. Scott D. Tagliaferri contributed to data curation and provided software assistance. Tobias Saueressig was involved in formal analysis and visualization. Dr Belavy did funding acquisition. Scott D. Tagliaferri, Claire Samanna, and Dr McGuckian were in charge of investigation. Drs Belavy, Owen, and Miller and Scott D. Tagliaferri, Paul Buntine, and Tobias Saueressig contributed to the methodology. Scott D. Tagliaferri and Drs Owen and Belavy were in charge of project administration as well as supervision. Dr Belavy, Scott D. Tagliaferri, and Tobias Saueressig drafted the manuscript. All authors reviewed and edited the draft and approved the final version of the manuscript.

DATA SHARING: Data underlying this study are available within the manuscript and in tabulated form in Supplemental Data. The statistical code for R is also available in Supplemental Data.

PATIENT AND PUBLIC INVOLVEMENT: There was no patient or public involvement in this research.

REFERENCES

1. Bakbergenuly I, Hoaglin DC, Kulinskaya E. Methods for estimating between-study variance and overall effect in meta-analysis of odds ratios. *Res Synth Methods*. 2020;11:426-442. https://doi.org/10.1002/jrsm.1404
2. Baker SR, Rabin A, Lantos G, Gallagher EJ. The effect of restricting the indications for lumbosacral spine radiography in patients with acute back symptoms. *Am J Roentgenol*. 1987;149:535-538. https://doi.org/10.2214/ajr.149.3.535
3. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health*. 2019;22:153-160. https://doi.org/10.1136/ebmental-2019-300117
4. Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. *Med J Aust*. 2017;206:268-273. https://doi.org/10.5694/mja16.00828
5. Becker A, Leonhardt C, Kochen MM, et al. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. *Spine*. 2008;33:473-480. https://doi.org/10.1097/BRS.0b013e3181657e0d
6. Berrington de González A, Darby S. Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *Lancet Lond Engl*. 2004;363:345-351. https://doi.org/10.1016/S0140-6736(04)15433-0
7. Bishop PB, Quon JA, Fisher CG, Dvorak MFS. The Chiropractic Hospital-based Interventions Research Outcomes (CHIRO) study: a randomized controlled trial on the effectiveness of clinical practice guidelines in the medical and chiropractic management of patients with acute mechanical low back pain. *Spine J*. 2010;10:1055-1064. https://doi.org/10.1016/j.spinee.2010.08.019
8. Borenstein M. *Common Mistakes in Meta-analysis: And How to Avoid Them*. Englewood, NJ: Biostat Inc; 2019.
9. Chandra K, Atkinson PR, Chatur H, Fraser J, Adams CL. To choose or not to choose: evaluating the effect of a Choosing Wisely knowledge translation initiative for imaging in low back pain by emergency physicians. *Cureus*. 2019;11:e4002. https://doi.org/10.7759/cureus.4002
10. Cherkin D, Balderson B, Wellman R, et al. Effect of low back pain risk-stratification strategy on patient outcomes and care processes: the MATCH randomized trial in primary care. *J Gen Intern Med*. 2018;33:1324-1336. https://doi.org/10.1007/s11606-018-4468-9

11. Chou R. Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Ann Intern Med*. 2011;154:181-189. https://doi.org/10.7326/0003-4819-154-3-201102010-00008
12. Clancy C, Quinn A, Wilson F. The aetiologies of failed back surgery syndrome: a systematic review. *J Back Musculoskelet Rehabil*. 2017;30:395-402. https://doi.org/10.3233/BMR-150318
13. Coombs DM, Machado GC, Richards B, et al. Effectiveness of a multifaceted intervention to improve emergency department care of low back pain: a stepped-wedge, cluster-randomised trial. *BMJ Qual Saf*. 2021;30:825-835. https://doi.org/10.1136/bmjqs-2020-012337
14. Dey P, Simpson CWR, Collins SI, et al. Implementation of RCGP guidelines for acute low back pain: a cluster randomised controlled trial. *Br J Gen Pract*. 2004;54:33-37.
15. Doi SA, Furuya-Kanamori L, Xu C, Lin L, Chivese T, Thalib L. Questionable utility of the relative risk in clinical research: a call for change to practice. *J Clin Epidemiol*. 2020. https://doi.org/10/S0895435620311719
16. Downie A, Hancock M, Jenkins H, et al. How common is imaging for low back pain in primary and emergency care? Systematic review and meta-analysis of over 4 million imaging requests across 21 years. *Br J Sports Med*. 2020;54:642-651. https://doi.org/10.1136/bjsports-2018-100087
17. Eccles M, Steen N, Grimshaw J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet*. 2001;357:1406-1409. https://doi.org/10.1016/S0140-6736(00)04564-5
18. Fenton JJ, Kravitz RL, Jerant A, et al. Promoting patient-centered counseling to reduce use of low-value diagnostic tests: a randomized clinical trial. *JAMA Intern Med*. 2016;176:191-197. https://doi.org/10.1001/jamainternmed.2015.6840
19. Fine B, Schultz SE, White L, Henry D. Impact of restricting diagnostic imaging reimbursement for uncomplicated low back pain in Ontario: a population-based interrupted time series analysis. *CMAJ Open*. 2017;5:E760-E767. https://doi.org/10.9778/cmajo.20160151
20. Flynn TW, Smith B, Chou R. Appropriate use of diagnostic imaging in low back pain: a reminder that unnecessary imaging may do as much harm as good. *J Orthop Sports Phys Ther*. 2011;41:838-846. https://doi.org/10.2519/jospt.2011.3618
21. Freeborn DK, Shye D, Mullooly JP, Eraker S, Romeo J. Primary care physicians' use of lumbar spine imaging tests: effects of guidelines and practice pattern feedback. *J Gen Intern Med*. 1997;12:619-625. https://doi.org/10.1046/j.1525-1497.1997.07122.x
22. French SD, McKenzie JE, O'Connor DA, et al. Evaluation of a theory-informed implementation intervention for the management of acute low back pain in general medical practice: the IMPLEMENT cluster randomised trial. *PLoS One*. 2013;8:e65471. https://doi.org/10.1371/journal.pone.0065471
23. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026
24. Haig A, Uren B, Loar S, et al. The impact of a complex consulting process with physiatry on emergency department management of back pain. *J Int Soc Phys Rehabil Med*. 2019;2:77-87. https://doi.org/10.4103/jisprm.jisprm_1_19
25. Harrer M, Cuijpers P, Furukawa T, Ebert DD. *Citation and Authors. metar: Companion R Package for the Guide "Doing Meta-Analysis in R." R package Version 0.0.9000*. 2019. https://dmetar.protectlab.org/authors.html
26. Higgins JPT, Thomas J, Chandler J. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, England: John Wiley & Sons Inc; 2019.
27. Higgins JPT, Thomas J, Chandler J, et al. Methods of analysis for cluster-randomized trials. In: Higgins JPT, Thomas J, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, England: John Wiley & Sons Inc; 2020.
28. Hoeijenbos M, Bekkering T, Lamers L, Hendriks E, van Tulder M, Koopmanschap M. Cost-effectiveness of an active implementation strategy for the Dutch physiotherapy guideline for low back pain. *Health Policy*. 2005;75:85-98. https://doi.org/10.1016/j.healthpol.2005.02.008
29. Hong AS, Ross-Degnan D, Zhang F, Wharam JF. Small decline in low-value back imaging associated with the 'Choosing Wisely' campaign, 2012–14. *Health Aff (Millwood)*. 2017;36:671-679. https://doi.org/10.1377/hlthaff.2016.1263
30. Ip IK, Gershanik EF, Schneider LI, et al. Impact of IT-enabled intervention on MRI use for back pain. *Am J Med*. 2014;127:512-518.e1. https://doi.org/10.1016/j.amjmed.2014.01.024
31. Jenkins HJ, Downie AS, Maher CG, Moloney NA, Magnussen JS, Hancock MJ. Imaging for low back pain: is clinical use consistent with guidelines? A systematic review and meta-analysis. *Spine J*. 2018;18:2266-2277. https://doi.org/10.1016/j.spinee.2018.05.004
32. Jenkins HJ, Hancock MJ, French SD, Maher CG, Engel RM, Magnussen JS. Effectiveness of interventions designed to reduce the use of imaging for low-back pain: a systematic review. *Can Med Assoc J*. 2015;187:401-408. https://doi.org/10.1503/cmaj.141183
33. Jensen CE, Riis A, Petersen KD, Jensen MB, Pedersen KM. Economic evaluation of an implementation strategy for the management of low back pain in general practice. *PAIN*. 2017;158:891-899. https://doi.org/10.1097/j.pain.0000000000000851
34. Kennedy SA, Fung W, Malik A, Farrokhyar F, Midia M. Effect of governmental intervention on appropriateness of lumbar MRI referrals: a Canadian experience. *J Am Coll Radiol*. 2014;11:802-807. https://doi.org/10.1016/j.jacr.2013.12.022
35. Kerry S, Oakeshott P, Dundas D, Williams J. Influence of postal distribution of The Royal College of Radiologists' guidelines, together with feedback on radiological referral rates, on X-ray referrals from general practice: a randomized controlled trial. *Fam Pract*. 2000;17:46-52. https://doi.org/10.1093/fampra/17.1.46
36. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med*. 2004;2:204-208. https://doi.org/10.1370/afm.141
37. Klein BJ, Radecki RT, Foris MP, Feil EI, Hickey ME. Bridging the gap between science and practice in managing low back pain: a comprehensive spine care system in a health maintenance organization setting. *Spine*. 2000;25:738-740. https://doi.org/10.1097/00007632-200003150-00015
38. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62:e1-e34. https://doi.org/10.1016/j.jclinepi.2009.06.006
39. Logan GS, Pike A, Copsey B, Parfrey P, Etchegary H, Hall A. What do we really know about the appropriateness of radiation emitting imaging for low back pain in primary and emergency care? A systematic review and meta-analysis of medical record reviews. *PLoS One*. 2019;14:e0225414. https://doi.org/10.1371/journal.pone.0225414
40. Matowe L, Ramsay CR, Grimshaw JM, Gilbert FJ, Macleod M-J, Needham G. Effects of mailed dissemination of the Royal College of Radiologists' guidelines on general practitioner referrals for radiography: a time series analysis. *Clin Radiol*. 2002;57:575-578. https://doi.org/10.1053/crad.2001.0894
41. Mavridis D, Salanti G. How to assess publication bias: funnel plot, trim-and-fill method and selection models. *Evid Based Ment Health*. 2014;17:30. https://doi.org/10.1136/eb-2013-101699
42. McAuley L, Ramsey C. *Cochrane Effective Practice and Organisation of Care Review Group; Data Collection Checklist*. Institute of Population Health, University of Ottawa. Available at: https://methods.cochrane.org/sites/methods.cochrane.org.bias/files/public/uploads/EPOC%20Data%20Collection%20Checklist.pdf. Accessed December 11, 2020.
43. Min A, Chan VWY, Aristizabal R, et al. Clinical decision support decreases volume of imaging for low back pain in an urban emergency department. *J Am Coll Radiol*. 2017;14:889-899. https://doi.org/10.1016/j.jacr.2017.03.005
44. Moskowitz H, Sunshine J, Grossman D, Adams L, Gelinas L. The effect of imaging guidelines

on the number and quality of outpatient radiographic examinations. *Am J Roentgenol*. 2000;175:9-15. https://doi.org/10.2214/ajr.175.1.1750009
45. Oliveira CB, Maher CG, Pinto RZ, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine J*. 2018;27:2791-2803. https://doi.org/10.1007/s00586-018-5673-2
46. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available at: https://www.R-project.org/
47. Rao J, Kroenke K, Mihaliak K, Eckert G, Weinberger M. Can guidelines impact the ordering of magnetic resonance imaging studies by primary care providers for low back pain? *Am J Manag Care*. 2002;8:27-35.
48. Riis A, Jensen CE, Bro F, et al. A multifaceted implementation strategy versus passive implementation of low back pain guidelines in general practice: a cluster randomised controlled trial. *Implement Sci*. 2016;11:143. https://doi.org/10.1186/s13012-016-0509-0
49. Schectman JM, Schroth WS, Verme D, Voss JD. Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *J Gen Intern Med*. 2003;18:773-780. https://doi.org/10.1046/j.1525-1497.2003.10205.x
50. Simula AS, Jenkins HJ, Hancock MJ, Malmivaara A, Booth N, Karppinen J. Patient education booklet to support evidence-based low back pain care in primary care—a cluster randomized controlled trial. *BMC Fam Pract*. 2021;22:178. https://doi.org/10.1186/s12875-021-01529-2
51. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. https://doi.org/10.1136/bmj.l4898
52. Suman A, Schaafsma FG, van de Ven PM, et al. Effectiveness of a multifaceted implementation strategy compared to usual care on low back pain guideline adherence among general practitioners. *BMC Health Serv Res*. 2018;18:358. https://doi.org/10.1186/s12913-018-3166-y
53. Tahvonen P, Oikarinen H, Niinimäki J, Liukkonen E, Mattila S, Tervonen O. Justification and active guideline implementation for spine radiography referrals in primary care. *Acta Radiol Stockh Swed 1987*. 2017;58:586-592. https://doi.org/10.1177/0284185116661879
54. Tracey NG, Martin JB, McKinstry CS, Mathew BM. Guidelines for lumbar spine radiography in acute low back pain: effect of implementation in an accident and emergency department. *Ulster Med J*. 1994;63:12-17.
55. Venkatesh AK, Scofi JE, Rothenberg C, et al. Choosing wisely in emergency medicine: early results and insights from the ACEP emergency quality network (E-QUAL). *Am J Emerg Med*. 2021;39:102-108. https://doi.org/10.1016/j.ajem.2020.01.029
56. Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods*. 2010;1:112-125. https://doi.org/10.1002/jrsm.11
57. Vucic K, Jelicic Kadic A, Puljak L. Survey of Cochrane protocols found methods for data extraction from figures not mentioned or unclear. *J Clin Epidemiol*. 2015;68:1161-1164. https://doi.org/10.1016/j.jclinepi.2014.11.016
58. Werner EL, Ihlebaek C, Laerum E, Wormgoor MEA, Indahl A. Low back pain media campaign: no effect on sickness behaviour. *Patient Educ Couns*. 2008;71:198-203. https://doi.org/10.1016/j.pec.2007.12.009

JONATHAN ORVIK GILADI, MSc[1] • ANDERS HOLSGAARD-LARSEN, PhD[2,3]
CLAUS VARNUM, MD, PhD[4,5] • JONAS BLOCH THORLUND, PhD[1,6]

Sports Participation and Performance 5 Years After Arthroscopic Partial Meniscectomy: A Retrospective Cohort Study of 288 Patients

OBJECTIVE: To investigate return to sport (RTS) approximately 5 years after arthroscopic partial meniscectomy (APM).

DESIGN: Retrospective cohort study.

METHODS: Knee Arthroscopy Cohort Southern Denmark patients were asked about RTS and reasons for non-RTS approximately 5 years (range, 4-6 years) after APM using online questionnaires. Patients engaged in their sport at the pre-injury level at follow-up were classified as "RTS" (or "returned to sport") and as being engaged in their sport with (1) full participation and performance, (2) reduced performance, or (3) both reduced participation and performance. Self-reported knee function was assessed using the Knee injury and Osteoarthritis Outcome Score (KOOS).

RESULTS: We included 288 patients (mean ± SD age, 49 ± 12 years; 44% women). Of these, 172 patients (60%; 95% confidence interval, 54%-65%) were classified as returned to sport, but only 42% (72/172) reported full participation and performance. Persistent problems with the operated knee were reported by 60% of the patients as the main reason for reduced participation or performance and by 70% of the patients as the main reason for not returning to pre-injury levels of their sport. Patients who had returned to sport, on average, improved by 10.1 points (95% confidence interval, 5.7-14.4) more in $KOOS_4$ scores from baseline to 5 years than non-RTS patients.

CONCLUSION: At approximately 5 years after APM, 6 in every 10 patients had returned to their sport at pre-injury levels, but only 1 in every 4 returned with full participation and performance, mainly due to persistent knee problems. Greater improvements in KOOS scores were observed in patients who were classified as returned to sport.
J Orthop Sports Phys Ther 2022;52(4):224-232. doi:10.2519/jospt.2022.10785

KEY WORDS: *arthroscopy, KOOS, meniscus, patient-reported outcome, return to sport*

Arthroscopic partial meniscectomy (APM) is one of the most frequently performed orthopedic procedures.[10,11] For younger athletes and recreationally active middle-aged individuals with meniscus injury alike, returning to sport is an important outcome. Few studies have investigated return to sport (RTS) after APM. Between 61% and 100% of young athletic patients[4,18,19] and 77% of middle-aged patients[1] return to sport.

RTS rates following hip arthroscopy are substantially lower than previously published[13] when one uses a clear and strict definition of RTS: the proportion of athletes who returned to their pre-injury sport at pre-injury levels. Reporting other important elements, such as the extent of sports participation and level of performance, changes in the type of sport, and cessation of sports due to hip and groin pain among those not engaged in their pre-injury sport at the pre-injury level, provided important insight into the different dimensions of RTS not previously reported. This methodology for reporting RTS is supported by international consensus.[3]

We aimed to investigate the proportion of patients returning to pre-injury levels of their sport and the extent of their sports participation and performance, approximately 5 years after APM. In addition, we investigated the differences in trajectories of patient-reported outcomes from before surgery to approximately 5 years after surgery between those who returned to pre-injury levels of their sport and those who did not.

[1]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark. [2]Department of Clinical Research, University of Southern Denmark, Odense, Denmark. [3]Department of Orthopaedics and Traumatology, Odense University Hospital, Odense, Denmark. [4]Department of Orthopaedic Surgery, Lillebaelt Hospital – Vejle, University Hospital of Southern Denmark, Vejle, Denmark. [5]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark. [6]Research Unit of General Practice, Department of Public Health, University of Southern Denmark, Odense, Denmark. This study was registered with ClinicalTrials.gov (identifier number NCT01871272). This study was supported by an individual postdoctoral grant awarded to Jonas Bloch Thorlund from the Danish Council for Independent Research/Medical Sciences and funds from the Region of Southern Denmark. The funders of the study had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, and approval of the manuscript. Jonas Bloch Thorlund holds a research grant from Pfizer outside the submitted work. The other authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Jonas Bloch Thorlund, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, 5230 Odense M, Denmark. E-mail: jthorlund@health.sdu.dk

METHODS

The STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) recommendations were followed to report this study.[30]

Patients

We included patients from the Knee Arthroscopy Cohort Southern Denmark (KACS) who participated in the approximately 5-year follow-up (range, 4-6 years). KACS is an observational cohort following patients undergoing arthroscopic meniscus surgery.[27] The recruitment procedure has been previously reported in detail.[27] In short, patients were consecutively recruited from 4 public hospitals in Denmark between February 1, 2013, and January 31, 2015. The KACS inclusion criteria were as follows: aged 18 years or older, undergoing arthroscopic surgery for a meniscus tear, able to read and understand Danish, and having an e-mail address. The KACS exclusion criteria were as follows: previous or planned anterior cruciate ligament (ACL) or posterior cruciate ligament reconstruction in either knee, any fracture in the lower extremities within 6 months prior to recruitment, or being unable to read and comprehend the online questionnaire. All study patients provided written informed consent, but the Regional Scientific Ethics Committee of Southern Denmark waived the need for ethical approval after reviewing the outline of KACS.

Data in the KACS cohort were self-reported via online questionnaires. Baseline information was collected approximately 2 weeks before surgery, and information from the consecutive follow-up assessments was collected at 12 weeks, 52 weeks, and approximately 5 years after surgery. Information about knee pathology was recorded by the operating surgeon during arthroscopy. A modified version of the International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine classification of meniscal tears[2] was used to classify meniscal pathology (tear type, tear location, tissue quality, etc).

We only included patients with complete data at baseline and at the approximately 5-year follow-up who underwent APM. Furthermore, we only included patients who were engaged in sports (all kinds of physical activities, whether organized in a club or performed individually) prior to their injury and who intended to return to their pre-injury levels of sport after surgery. This was determined by the following sequence of questions in the approximately 5-year follow-up questionnaire:

1. Patients were asked, "Did you participate in sports before your knee problems occurred, and you had your knee arthroscopy? (sports is all kinds of physical activities whether organized in a club or performed individually)" (yes/no). Patients answering "no" to this question were excluded.
2. The remaining patients were asked to "indicate the primary sport you participated in before your knee problems occurred, and you had arthroscopic surgery (primary sport is the sport you prioritized the most)" (free-text response). We categorized the types of sports according to Hefti et al,[12] modified to European sports types[8,12,17]: category 1 sports: sports with frequent pivoting, jumping, and cutting movements (eg, football and team handball); category 2 sports: sports where lateral movements are involved but with less frequent pivoting, jumping, and cutting movements than category 1 sports (eg, racket sports and volleyball); category 3 sports: straight-ahead activities where pivoting, jumping, and cutting movements are not performed (eg, running and cycling); and category 4 sports: sedentary individuals (ie, not applicable to patients included in this study).
3. Next, patients were asked, "What was the highest level of sport you participated in before your knee problems occurred, and you had arthroscopy?" with response options being "elite" (professional or highest competitive level in Denmark), "competitive" (engaged in organized matches or competitions), or "recreational" (no organized matches or competitive events) levels. Further questions in the approximately 5-year follow-up questionnaire were conditioned to be specific to the individual patient's stated pre-injury level of sports, which is indicated by "elite," "competitive," or "recreational."
4. Lastly, patients were asked, "Before you had surgery, did you intend to continue with your sport at minimum [elite/competitive/recreational] level after surgery?" (yes/no). Patients who answered "no" were excluded.

Outcomes

The main outcome was the proportion of patients who returned to sport and the extent of the patient's sports participation and performance at the approximately 5-year follow-up. To classify patients as "RTS" or "not RTS" (also hereinafter referred to as "returned to sport" and "not returned to sport," respectively), we asked, "Within the last 3 months, to what extent have you been able to engage in your sport at a minimum of [elite/competitive/recreational] level?" with response options being "full participation and performance"; "full participation, but reduced performance"; "reduced participation and reduced performance"; or "other." Patients answering "full participation and performance"; "full participation, but reduced performance"; or "reduced participation and reduced performance" were classified as RTS. Patients answering "other" were classified as not RTS. Patients reporting full participation and performance were additionally asked to "indicate to what extent you experience knee problems when you participate in your sport," with response

options ranging from "always" to "never" on a 5-point Likert scale. Patients reporting reduced participation, reduced performance, or both were asked, "Is your reduced participation/performance a result of problems with your knee?" with response options being "yes" or "no."

Patients who were classified as not RTS were asked, "Have you attempted to perform your sport at pre-injury level [elite/ competitive/recreational] at any time since surgery?" with response options being "Yes, but stopped because of problems with my knee"; "Yes, but problems with my knee were not the reason why I stopped"; "No, not been able to because of problems with my knee"; or "No, but not because of problems with my knee." Lastly, they were asked to "indicate what fits to your current level of participation in your sport," with

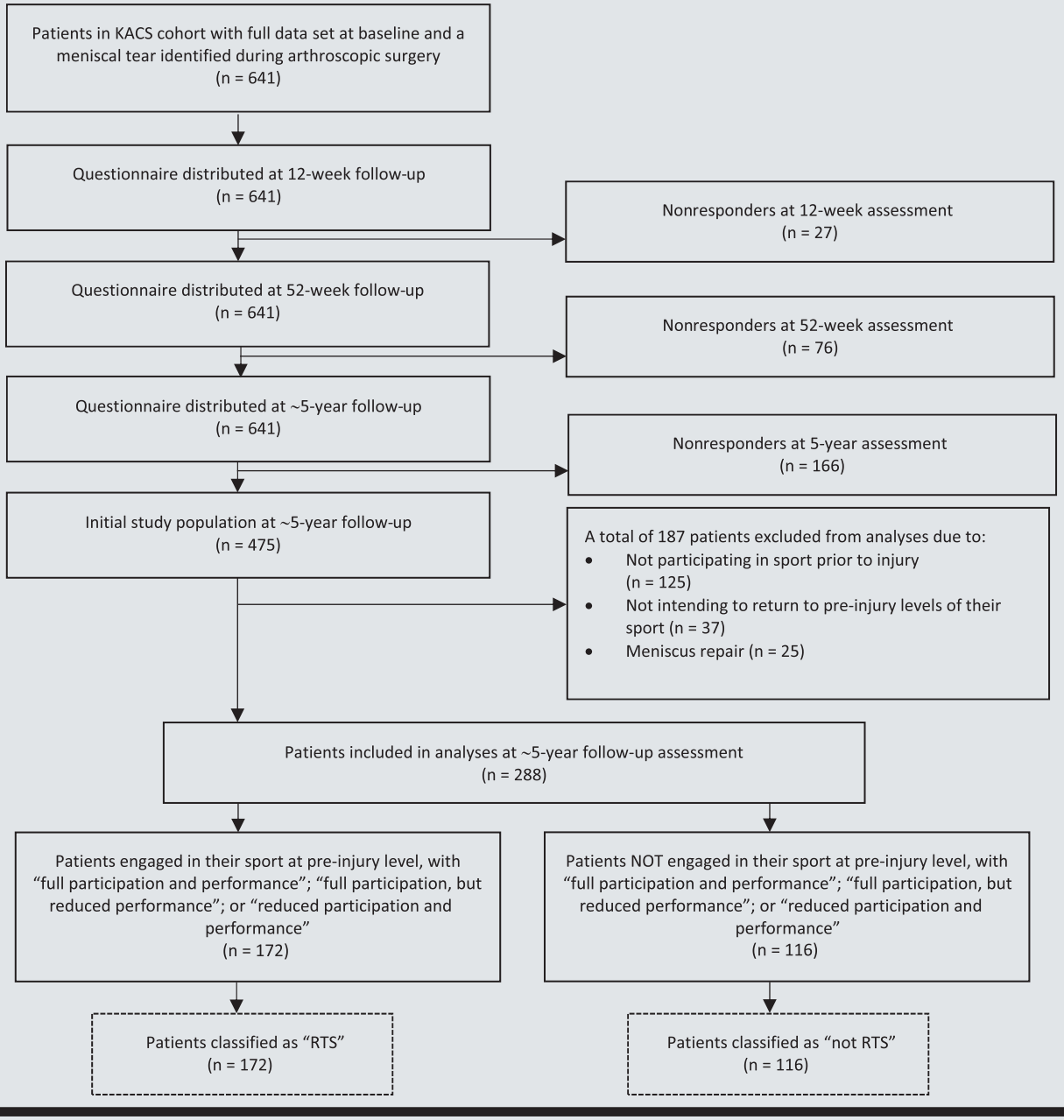

FIGURE 1. Flowchart of patient inclusion in the KACS cohort and classification of patients as "returned to sport" and "not returned to sport." Abbreviations: KACS, Knee Arthroscopy Cohort Southern Denmark; RTS, returned to sport.

response options being "I have engaged in it within the last 3 months, but at a lower activity level than [elite/competitive/recreational] because of problems with my knee"; "I have engaged in another sport within the last 3 months because of problems with my knee"; "I have not engaged in any sport within the last 3 months because of problems with my knee"; or "My change in activity level (including no participation in sport) and/or sport is not because of problems with my knee."

We assessed self-reported knee symptoms and function using the Knee injury and Osteoarthritis Outcome Score (KOOS).[24] The KOOS consists of 42 items in 5 separate subscales: Pain, Symptoms, Activities of Daily Living, Sport and Recreation, and knee-related Quality of Life. The subscales are separately scored from 0 to 100, where a score of 0 indicates extreme knee problems and a score of 100 indicates no knee problems.[22] In this study, $KOOS_4$ was used in addition to the 5 separate KOOS subscale scores to investigate between-group differences in outcomes. The $KOOS_4$ is the average of the 4 subscales, namely, Pain, Symptoms, Sport and Recreation, and knee-related Quality of Life, excluding the Activities of Daily Living subscale known to have ceiling effects in this population.[7,29] The $KOOS_4$ is responsive in trials assessing the effect of knee surgery to provide a single score for statistical purposes and to simplify interpretation.[9,16,26]

Statistical Analyses

Descriptive statistics were presented as means and standard deviations or numbers and percentages, as appropriate. Differences in baseline characteristics and RTS outcomes were assessed by an unpaired *t* test, a chi-square test, and the Fisher exact test, as appropriate. RTS, participation, and performance were reported as percentages with 95% confidence intervals (CIs). Between-group differences in the trajectories of KOOS scores were analyzed using a mixed linear model (restricted maximum likelihood estimation), with patient as a random effect and group (RTS vs not RTS), time (baseline, 12 weeks, 52 weeks, and approximately 5 years), and the Group × Time interaction as fixed effects, including age, sex, and body mass index as covariates. All statistical analyses were

TABLE 1 Patient Characteristics and Surgical Data (n = 288)[a]

Variable	Returned to Sport (RTS) Yes (n = 172)	No (n = 116)	*P* Value
Age, mean (SD), y	50.9 (12.1)	48.3 (12.2)	.067
Female, n (%)	74 (45)	52 (43)	.762
BMI, mean (SD), kg/m^2	26.8 (4.2)	26.2 (3.6)	.206
Symptom onset, n (%)			.502
Slowly evolved	53 (31)	31 (27)	
Semi-traumatic	76 (44)	49 (42)	
Traumatic	43 (25)	36 (31)	
Duration of symptoms, n (%)			.500
0-3 months	34 (20)	21 (18)	
4-6 months	47 (27)	34 (30)	
7-12 months	38 (22)	34 (29)	
13-24 months	31 (18)	14 (12)	
>24 months	22 (13)	13 (11)	
Meniscus tear location, n (%)			.549
Medial meniscus	130 (76)	82 (71)	
Lateral meniscus	27 (16)	24 (21)	
Both	15 (8)	10 (8)	
Meniscus tear type,[b] n (%)			.527
Complex	41 (28)	39 (39)	
Vertical flap	45 (31)	27 (27)	
Longitudinal-vertical	27 (18)	15 (15)	
Horizontal	12 (8)	7 (7)	
Horizontal flap	11 (8)	8 (8)	
Radial	10 (7)	4 (4)	
Root tear	0 (0)	1 (1)	
Meniscus tissue quality, n (%)			.974
Degenerative	56 (33)	40 (34)	
Non-degenerative	110 (64)	72 (62)	
Undetermined	6 (3)	4 (4)	
ACL status, n (%)			.071
Intact	156 (91)	97 (84)	
Ruptured	16 (9)	19 (16)	
Level of sport pre-injury, n (%)			.001
Competitive	36 (21)	45 (39)	
Recreational	136 (79)	71 (61)	
Sports category, n (%)			.116
Category 1	30 (18)	25 (22)	
Category 2	33 (19)	12 (10)	
Category 3	109 (63)	79 (68)	

Table continues on page 228.

TABLE 1 Patient Characteristics and Surgical Data (n = 288)[a] (continued)

Variable	Returned to Sport (RTS) Yes (n = 172)	No (n = 116)	*P* Value
KOOS score, mean (SD)			
$KOOS_4$	48.3 (14.6)	46.4 (15.4)	.289
Pain	58.3 (18.1)	56.9 (18.9)	.520
Symptoms	61.8 (17.7)	61.1 (19.8)	.754
ADL	67.0 (19.0)	66.0 (19.9)	.672
Sport/Rec	29.7 (22.3)	25.8 (21.7)	.144
QOL	43.5 (14.4)	41.9 (13.7)	.341

Abbreviations: ACL status, anterior cruciate ligament status determined by the operating surgeon during arthroscopy: ruptured defined as partial (n = 19) or fully ruptured (n = 16); ADL, Activities of Daily Living subscale; BMI, body mass index (kg/m²); Competitive, organized matches or competitions (n = 34) and professionals or athletes at the highest competitive level in Denmark (n = 2); KOOS, Knee injury and Osteoarthritis Outcome Score; QOL, knee-related Quality of Life subscale; Recreational, no organized matches or competitions; SD, standard deviation; Sport/Rec, Sport and Recreation subscale; Sports category, defined according to Hefti et al,[12] modified to European sports[8,17]: category 1: football (n = 40) and team handball (n = 15); category 2: racket sports (badminton, tennis, squash; n = 21), dancing (dancing, Zumba, aerobics; n = 14), martial arts (n = 4), and volleyball (n = 3); and category 3: running (n = 74), strength and conditioning (n = 36), cycling (n = 35), equestrian (n = 9), golf (n = 8), boating (rowing, kayak, sailing; n = 5), swimming (n = 4), triathlon (n = 2), and other sports (n = 15).
[a]*Percentages are rounded to the nearest whole number.*
[b]*Meniscus tear type, n = 247.*

performed using Stata/IC (Version 16.1; StataCorp LLC, College Station, TX).

RESULTS

Of the 641 KACS patients, 475 (74%) were assessed at the approximately 5-year follow-up. Furthermore, 187 patients were excluded due to having meniscus repair (n = 25), not participating in sports prior to injury (n = 125), or not intending to return to sport at pre-injury levels (n = 37), resulting in a sample of 288 patients (**FIGURE 1**). The majority of patients were men and slightly overweight (**TABLE 1**). A larger proportion of patients classified as RTS were engaged in recreational levels of sport compared to those classified as not RTS but were similar in all other in baseline characteristics.

At the approximately 5-year follow-up, 172 of the 288 patients (60%; 95% CI, 54%-65%) had returned to their pre-injury sports at pre-injury levels (**TABLE 2**). Of these 172 patients, 72 (42%; 95% CI, 35%-49%) were engaged in their pre-injury sport with full participation and performance, corresponding to 25% (95% CI, 20%-30%) of the total study population (**TABLE 2**). Of the 72 patients with full participation and performance, only 33 (46%; 95% CI, 35%-58%) reported never experiencing problems with their surgically treated knee when participating in their sport. These 33 patients represented 11% (95% CI, 7.9%-14.8%) of the total study population. Of the remaining 100 patients reporting either reduced performance or both reduced performance and participation, 60 (60%; 95% CI, 50%-69%) reported problems with the surgically treated knee as the main reason.

Problems with the surgically treated knee were also commonly reported by 116 patients as the reason for not returning to their sport at pre-injury levels (**FIGURE 3A**). Approximately half of these 116 patients had attempted to engage in their pre-injury sport at pre-injury levels since surgery, of which a majority were forced to stop due to knee problems (**FIGURE 3B**).

A larger proportion of patients engaged in recreational levels of sport returned to sport compared to patients who were engaged in competitive levels of sport ($P = .001$) (**TABLE 2**). A larger proportion of the recreational group reported full participation and performance ($P = .012$) (**TABLE 2**). In the older age group and in the intact ACL group, larger proportions reported full participation and performance (**TABLE 2**). There were no significant differences in RTS or the extent of participation and performance across the sex, symptom onset, meniscus tear location, and sports category groups (**TABLE 2**).

Patients in the RTS group had larger improvement in $KOOS_4$ scores over time (Group × Time interaction, $P < .001$) (**FIGURE 2**). The RTS group improved, on average, by 10.1 points (95% CI, 5.7-14.4) more than patients in the not RTS group from baseline to the approximately 5-year follow-up (**SUPPLEMENTAL TABLE 1**). Similar differences between the 2 groups were observed in all 5 KOOS subscale score trajectories and change scores (**SUPPLEMENTAL TABLES 1 and 2** and **FIGURE 1A-E**).

DISCUSSION

We investigated sports participation and performance in 288 patients approximately 5 years following APM. Sixty percent of patients had returned to their sport at the pre-injury level at follow-up. Previous studies have reported substantially higher proportions and rates of RTS.[1,4,18,19] Differences between the studied populations, our use of a strict definition of RTS, and the timing at which RTS was investigated might, in part, explain these discrepancies. Engaging in a different type and level of sport at follow-up was commonly reported among patients in our cohort, suggesting that not including the type and level of sport likely affects estimations of RTS.

Level of Participation and Performance

We investigated additional elements of RTS compared with previous studies and found that among patients who returned to their pre-injury level of sport,

TABLE 2 Overview of Return to Sport and the Extent of Participation and Performance at the Approximately 5-Year Follow-up[a]

	Returned to Sport (RTS)			Extent of Participation and Performance				
Variable	**Yes**	**No**	***P* Value**	**Full/Full**	**Full/Red.**	**Red./Red.**	**No RTS**	***P* Value**
All patients (n = 288), n (%)	172 (60)	116 (40)		72 (25)	41 (14)	59 (21)	116 (40)	
Sex, n (%)			.762					.331
Female (n = 126)	74 (59)	52 (41)		27 (21)	16 (13)	31 (25)	52 (41)	
Male (n = 162)	98 (60)	64 (40)		45 (28)	25 (15)	28 (17)	64 (40)	
Age group, n (%)			.276					.004
Younger (n = 58)	31 (53)	27 (47)		7 (12)	15 (26)	9 (15)	27 (47)	
Older (n = 230)	141 (61)	89 (39)		65 (28)	26 (11)	50 (25)	89 (39)	
Symptom onset, n (%)			.502					.544
Slowly evolved (n = 84)	53 (63)	31 (37)		25 (30)	10 (12)	18 (21)	31 (37)	
Semi-traumatic (n = 125)	76 (61)	49 (39)		28 (23)	18 (14)	30 (24)	49 (39)	
Traumatic (n = 79)	43 (54)	36 (46)		19 (24)	13 (16)	11 (14)	36 (46)	
Meniscus tear location, n (%)			.549					.816
Medial (n = 212)	130 (61)	82 (39)		57 (27)	28 (13)	45 (21)	82 (39)	
Lateral (n = 51)	27 (53)	24 (47)		10 (19)	8 (16)	9 (18)	24 (47)	
Both (n = 25)	15 (60)	10 (40)		5 (20)	5 (20)	5 (20)	10 (40)	
ACL status, n (%)			.071					.017
Intact (n = 253)	156 (62)	97 (38)		70 (28)	33 (13)	53 (21)	97 (38)	
Ruptured (n = 35)	16 (46)	19 (54)		2 (6)	8 (23)	6 (17)	19 (54)	
Level of sport pre-injury, n (%)			.001					.012
Competitive (n = 81)	36 (44)	45 (56)		15 (19)	9 (11)	12 (15)	45 (56)	
Recreational (n = 207)	136 (66)	71 (34)		57 (28)	32 (15)	47 (23)	71 (34)	
Sports category, n (%)			.116					.232
Category 1 (n = 55)	30 (55)	25 (45)		14 (26)	6 (11)	10 (18)	25 (45)	
Category 2 (n = 45)	33 (73)	12 (27)		11 (24)	15 (33)	7 (16)	12 (27)	
Category 3 (n = 188)	109 (58)	79 (42)		47 (25)	20 (11)	42 (22)	79 (42)	

Abbreviations: ACL status, anterior cruciate ligament status determined by the operating surgeon during arthroscopy: ruptured defined as partial (n = 19) or fully ruptured (n = 16); Competitive, organized matches or competitions (n = 34) and professionals or athletes at the highest competitive level in Denmark (n = 2); Full/Full, full participation and performance; Full/Red., full participation, but reduced performance; Red./Red., reduced participation and performance; Recreational, no organized matches or competitions; Sports category, defined according to Hefti et al,[12] modified to European sports[8,17]: category 1: football (n = 40) and team handball (n = 15); category 2: racket sports (badminton, tennis, squash; n = 21), dancing (dancing, Zumba, aerobics; n = 14), martial arts (n = 4), and volleyball (n = 3); and category 3: running (n = 74), strength and conditioning (n = 36), cycling (n = 35), equestrian (n = 9), golf (n = 8), boating (rowing, kayak, sailing; n = 5), swimming (n = 4), triathlon (n = 2), and other sports (n = 15); Younger, R40 years old; Older, >40 years old.

[a]*Percentages are rounded to the nearest whole number.*

the extent of their participation and performance varied. Less than half of the patients classified as RTS were engaged in their sport with full participation and performance; the majority reported either reduced participation or both reduced participation and performance. Our findings highlight the importance of investigating these additional elements to detail the full spectrum of RTS as an outcome. Furthermore, this is of clinical relevance when considering what constitutes a successful RTS for an individual patient. An individual fully participating in their sport at the pre-injury level with impaired performance or fully participating in a different sport may be inappropriately determined as successfully returned to sport. For others, reaching a level of participation in a different sport can represent a successful RTS.

Follow-up Time and Assessment of RTS

Investigating RTS approximately 5 years after surgery is longer than typically reported.[4,19,31] One study reporting RTS after APM in a middle-aged patient population reported that 77% returned to sport. However, this study had a short follow-up (6 months), which might not fully capture the ability to maintain sports participation and performance. This is relevant considering that some patients might return to sport early but are forced to stop at a later time due to knee problems and that other patients might not be able to fully return to sport until years after their surgery. As reflected by

the evolving $KOOS_4$ scores over the study period, some patients might continue to improve well beyond the initial time period after surgery (**FIGURE 2**). The $KOOS_4$ score trajectories differed between groups in favor of patients who returned to sport. The between-group difference seemingly increased over time and was similar for all KOOS subscale scores (**SUPPLEMENTAL FIGURE 1A-E**).

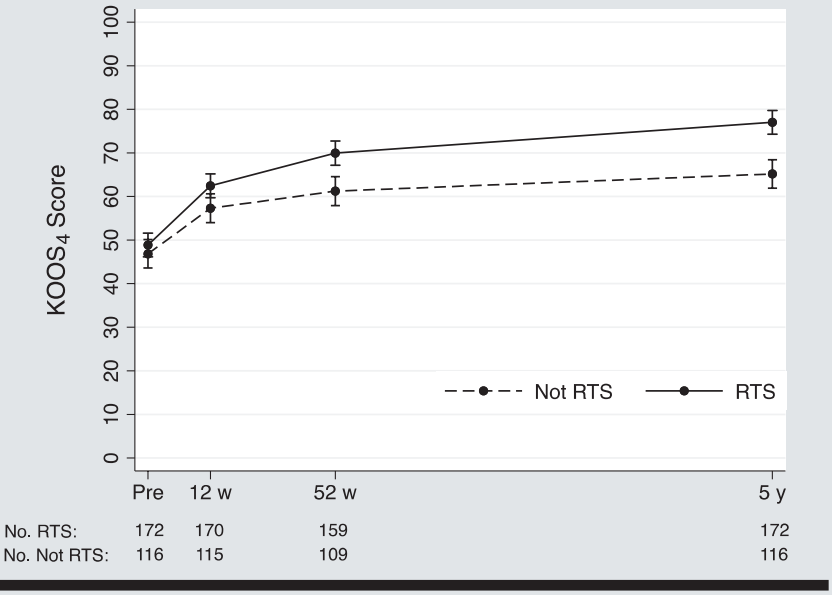

FIGURE 2. Adjusted $KOOS_4$ score trajectories from baseline to the approximately 5-year follow-up comparing patients who returned to sport and those who did not return to sport. Error bars displaying 95% confidence intervals. Model adjusted for age, sex, and body mass index (kg/m^2). The trajectories were significantly different between groups (Time × Group interaction, $P < .001$). Abbreviations: KOOS, Knee injury and Osteoarthritis Outcome Score; RTS, returned to sport.

Persistent Knee Problems Were Common

Among patients classified as RTS and reporting reduced participation or reduced participation and performance, 60% reported that this was due to problems with their operated knee. Among patients classified as not RTS, a substantial proportion (70%) either changed their level or type of sport or ceased sports completely due to problems with their operated knee (**FIGURE 3A**). These findings are supported by other studies reporting persistent knee pain after meniscus surgery requiring a decrease in sports engagement[5,6,15] and long-term symptoms and functional limitations.[23] The potential failure of complete symptom alleviation following surgery, along with the potential degenerative and detrimental effects of the surgery itself, is important to consider when surgery is proposed with the intent of keeping the middle-aged patient active and engaged in their sport for longer.[14,16,21,25] Our findings can inform shared clinical decision-making processes and underpin

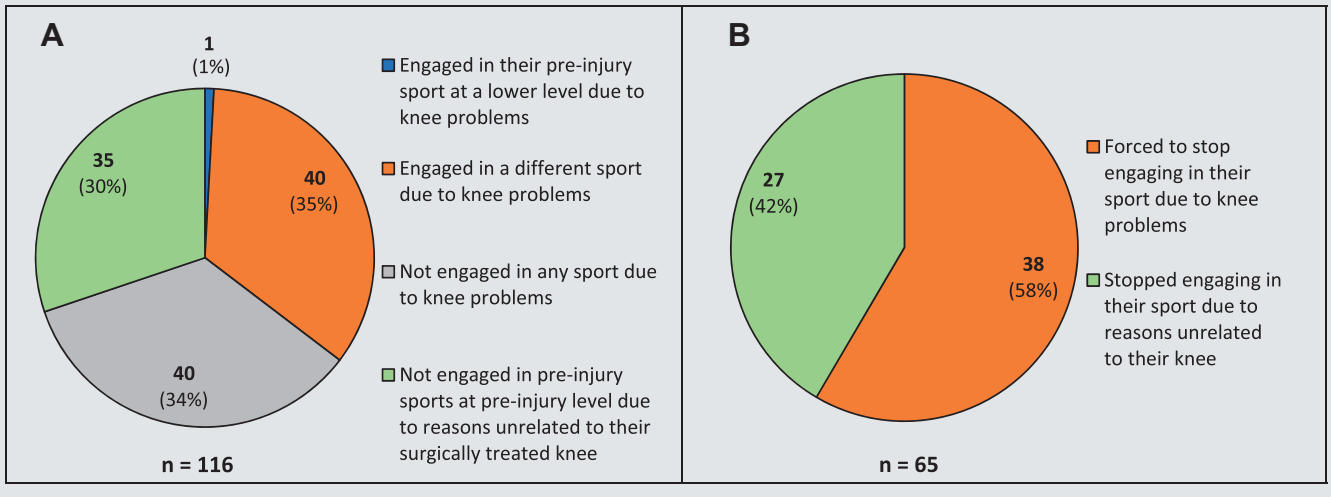

FIGURE 3. Overview of patients classified as "not returned to sport" (not RTS) (n = 116). (A) Of the 116 patients classified as not RTS, 1 (1%; 95% confidence interval [CI], 0.1%-6%) patient was engaged in pre-injury sport but at a lower level due to knee problems; 40 (35%; 95% CI, 26%-44%) patients were engaged in a different sport due to knee problems, whereas 40 (34%; 95% CI, 26%-44%) patients did not engage in any sport due to knee problems. Lastly, 35 (30%; 95% CI, 22%-39%) patients were not engaged in pre-injury sports at the pre-injury level due to reasons unrelated to their surgically treated knee. (B) Of the 116 patients classified as not RTS, 65 (56%; 95% CI, 47%-64%) had attempted to engage in their pre-injury sport at pre-injury levels at any time since surgery. Of these 65 patients, 38 (58%; 95% CI, 46%-70%) were forced to stop due to knee problems, and the remaining 27 (42%; 95% CI, 30%-54%) stopped engaging in their sport due to reasons unrelated to their knee.

realistic expectations following surgery, which often deviate from actual postoperative outcomes.[20]

Sports Demands May Impact RTS

A larger proportion of patients engaged in recreational sports returned to sport compared to patients engaged in competitive sports. This suggests that the level at which the sport is performed might affect the likelihood of RTS after APM. However, there was no difference in RTS between sports categories in our study (**TABLE 2**), and our results support a previous study of middle-aged patients.[1] Our results do not support the work of Aune et al,[4] who found a lower likelihood of RTS for players in positions with more frequent sprinting and cutting movements. Certain sport types might, to a larger degree, depend upon the function of the menisci.[22] However, all patients in the study by Aune et al[4] had lateral meniscus injuries and 51% had ACL reconstruction—both factors that may contribute to the relatively low RTS rates compared to other previous studies.[1,18,19]

In our study, patients with either a lateral meniscus injury or a ruptured ACL seemingly had lower RTS rates, but the differences did not reach statistical significance (**TABLE 2**). A larger proportion of older patients returned to sport with full participation and performance (**TABLE 2**), and although not statistically significant, a larger proportion of the older patients returned to sport compared to younger patients (61% vs 53%). However, the considerably larger proportion of older patients engaging in sport at a recreational level could explain this difference (**SUPPLEMENTAL TABLE 3**).

Limitations

Due to the retrospective design and the addition of questions regarding RTS to the 5-year follow-up questionnaire, there is a risk of recall bias. For instance, we asked about reasons for not being engaged in sport at the pre-injury level at the 5-year follow-up, but participation may have ceased much earlier. Furthermore, we cannot determine at which time point individuals returned to sport. Since study eligibility was dependent upon a response to the 5-year questionnaire, we could not determine a true follow-up rate. Because no outcome data were available for patients who did not respond, we cannot ascertain whether attrition was random or if there were systematic differences in RTS outcomes between them and the patients included in our study. Whether patients included in our study had better or worse outcomes, and hence the effect of a potential attrition bias on our results, is therefore unknown.

Our data were patient reported, and we cannot be certain if self-reported sports participation and performance corresponds to objective measures. Patients in the KACS cohort had a similar age and sex distribution compared with the population undergoing arthroscopic meniscus surgery in Denmark[28] at the time of the study. Indications for meniscal surgery may have changed over time; however, the current age mix of patients undergoing APM in Denmark is similar to that of the study period, although the number of patients undergoing APM has declined. Furthermore, the proportion of patients engaged in sports at the elite level was low (n = 2), limiting generalizability of the results to this population.

CONCLUSION

Among patients undergoing APM, 60% returned to pre-injury levels of their sport, the majority with either reduced participation or reduced performance. Persistent problems with the operated knee were commonly reported as the reason for the reduced participation or performance and for not being engaged in their sport at pre-injury levels at follow-up. Patients who returned to pre-injury levels of their sport improved more in self-reported knee function over the study period than those who did not. ◉

KEY POINTS

FINDINGS: Four in 10 patients did not return to their sport at pre-injury levels, and the majority of those who did reported reduced participation or performance. Persistent knee problems were reported by 6 in 10 patients as the reason for the reduced participation or performance and by 7 in 10 patients for not returning to pre-injury levels of their sport. Greater improvements in KOOS scores were observed in patients classified as returned to sport.

IMPLICATIONS: Studies investigating RTS after arthroscopic meniscus surgery should use clear definitions of RTS, where the type and level of sport are included, and investigate the extent of sports participation and performance. Our findings, which indicate that not all patients can expect to return to sport and that APM is not guaranteed to alleviate knee problems in all patients, can inform shared clinical decision-making processes and underpin realistic expectations following surgery.

CAUTION: The results may not generalize to adolescent or young adult patient populations and to those engaged in sports at very high levels.

STUDY DETAILS

AUTHOR CONTRIBUTIONS: Jonathan Orvik Giladi and Dr Thorlund conceived the idea of the study. Jonathan Orvik Giladi performed the statistical analysis. All authors took part in the interpretation of the results. Jonathan Orvik Giladi and Dr Thorlund drafted the first version of the manuscript. Drs Holsgaard-Larsen and Varnum provided critical feedback to the manuscript. All authors read and approved the final version of the manuscript.

DATA SHARING: There are no data available.

PATIENT AND PUBLIC INVOLVEMENT: There was no patient or public involvement.

ACKNOWLEDGMENTS: *We would like to acknowledge the efforts of all participating patients and orthopedic surgeons, nurses, and secretaries at the Department of Orthopaedics and Traumatology, Odense University Hospital,*

Odense and Svendborg, Denmark, and the Department of Orthopaedic Surgery, Lillebaelt Hospital, Vejle and Kolding, Denmark, for their assistance with patient recruitment and data collection. We would also like to acknowledge Kenneth Pihl, PT, PhD, for assisting with the 5-year follow-up of the KACS cohort and the statistical analysis.

REFERENCES

1. Agarwalla A, Gowd AK, Liu JN, et al. Predictive factors and duration to return to sport after isolated meniscectomy. *Orthop J Sport Med*. 2019;7:1-8. https://doi.org/10.1177/2325967119837940
2. Anderson AF, Irrgang JJ, Dunn W, et al. Interobserver reliability of the International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine (ISAKOS) classification of meniscal tears. *Am J Sports Med*. 2011;39:926-932. https://doi.org/10.1177/0363546511400533
3. Ardern CL, Glasgow P, Schneiders A, et al. 2016 consensus statement on return to sport from the First World Congress in Sports Physical Therapy, Bern. *Br J Sports Med*. 2016;50:853-864. https://doi.org/10.1136/bjsports-2016-096278
4. Aune KT, Andrews JR, Dugas JR, Cain EL. Return to play after partial lateral meniscectomy in National Football League athletes. *Am J Sports Med*. 2014;42:1865-1872. https://doi.org/10.1177/0363546514535069
5. Bonneux I, Vandekerckhove B. Arthroscopic partial lateral meniscectomy long-term results in athletes. *Acta Orthop Belg*. 2002;68:356-361.
6. Brophy RH, Gill CS, Lyman S, Barnes RP, Rodeo SA, Warren RF. Effect of anterior cruciate ligament reconstruction and meniscectomy on length of career in National Football League athletes. *Am J Sports Med*. 2009;37:2102-2107. https://doi.org/10.1177/0363546509349035
7. Collins NJ, Prinsen CAC, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. *Osteoarthr Cartil*. 2016;24:1317-1329. https://doi.org/10.1016/j.joca.2016.03.010
8. Eitzen I, Moksnes H, Snyder-Mackler L, Risberg MA. A progressive 5-week exercise therapy program leads to significant improvement in knee function early after anterior cruciate ligament injury. *J Orthop Sport Phys Ther*. 2010;40:705-721. https://doi.org/10.2519/jospt.2010.3345
9. Frobell RB, Roos EM, Roos HP, Ranstam J, Lohmander LS. A randomized trial of treatment for acute anterior cruciate ligament tears. *N Engl J Med*. 2010;363:331-342. https://doi.org/10.1056/NEJMoa0907797
10. Hall MJ, Schwartzman A, Zhang J, Liu X. Ambulatory surgery data from hospitals and ambulatory surgery centers: United States, 2010. *Natl Health Stat Report*. 2017;102:1-15.
11. Hamilton DF, Howie CR. Knee arthroscopy: influence of systems for delivering healthcare on procedure rates. *BMJ*. 2015;351:h4720. https://doi.org/10.1136/bmj.h4720
12. Hefti E, Müller W, Jakob RP, Stäubli H-U. Evaluation of knee ligament injuries with the IKDC form. *Knee Surg Sports Traumatol Arthrosc*. 1993;1:226-234. https://doi.org/10.1007/BF01560215
13. Ishøi L, Thorborg K, Kraemer O, Hölmich P. Return to sport and performance after hip arthroscopy for femoroacetabular impingement in 18- to 30-year-old athletes: a cross-sectional cohort study of 189 athletes. *Am J Sports Med*. 2018;46:2578-2587. https://doi.org/10.1177/0363546518789070
14. Katz JN, Shrestha S, Losina E, et al. Five-year outcome of operative and nonoperative management of meniscal tear in persons older than forty-five years. *Arthritis Rheumatol*. 2020;72:273-281. https://doi.org/10.1002/art.41082
15. Kim SG, Nagao M, Kamata K, Maeda K, Nozawa M. Return to sport after arthroscopic meniscectomy on stable knees. *BMC Sports Sci Med Rehabil*. 2013;5:2-5. https://doi.org/10.1186/2052-1847-5-23
16. Kise NJ, Risberg MA, Stensrud S, Ranstam J, Engebretsen L, Roos EM. Exercise therapy versus arthroscopic partial meniscectomy for degenerative meniscal tear in middle aged patients: randomised controlled trial with two year follow-up. *Br J Sports Med*. 2016;50:1473-1480. https://doi.org/10.1136/bjsports-2016-i3740rep
17. Moksnes H, Snyder-Mackler L, Risberg MA. Individuals with an anterior cruciate ligament-deficient knee classified as noncopers may be candidates for nonsurgical rehabilitation. *J Orthop Sport Phys Ther*. 2008;38:586-595. https://doi.org/10.2519/jospt.2008.2750
18. Nawabi DH, Cro S, Hamid IP, Williams A. Return to play after lateral meniscectomy compared with medial meniscectomy in elite professional soccer players. *Am J Sports Med*. 2014;42:2193-2198. https://doi.org/10.1177/0363546514540271
19. Osti L, Liu SH, Raskin A, Merlo F, Bocchi L. Partial lateral meniscectomy in athletes. *Arthrosc J Arthrosc Relat Surg*. 1994;10:424-430. https://doi.org/10.1016/S0749-8063(05)80194-X
20. Pihl K, Roos EM, Nissen N, JøRgensen U, Schjerning J, Thorlund JB. Over-optimistic patient expectations of recovery and leisure activities after arthroscopic meniscus surgery: a prospective cohort study of 478 patients. *Acta Orthop*. 2016;87:615-621. https://doi.org/10.1080/17453674.2016.1228411
21. Rongen JJ, Rovers MM, van Tienen TG, Buma P, Hannink G. Increased risk for knee replacement surgery after arthroscopic surgery for degenerative meniscal tears: a multi-center longitudinal observational study using data from the osteoarthritis initiative. *Osteoarthr Cartil*. 2017;25:23-29. https://doi.org/10.1016/j.joca.2016.09.013
22. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes*. 2003;1:1-8. https://doi.org/10.1186/1477-7525-1-64
23. Roos EM, Östenberg A, Roos H, Ekdahl C, Lohmander LS. Long-term outcome of meniscectomy: symptoms, function, and performance tests in patients with or without radiographic osteoarthritis compared to matched controls. *Osteoarthr Cartil*. 2001;9:316-324. https://doi.org/10.1053/joca.2000.0391
24. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sport Phys Ther*. 1998;28:88-96. https://doi.org/10.2519/jospt.1998.28.2.88
25. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic partial meniscectomy for a degenerative meniscus tear: a 5 year follow-up of the placebo-surgery controlled FIDELITY (Finnish Degenerative Meniscus Lesion Study) trial. *Br J Sports Med*. 2020;54:1332-1339. https://doi.org/10.1136/bjsports-2020-102813
26. Skou ST, Roos EM, Laursen MB, et al. A randomized, controlled trial of total knee replacement. *N Engl J Med*. 2015;373:1597-1606. https://doi.org/10.1056/NEJMoa1505467
27. Thorlund JB, Christensen R, Nissen N, et al. Knee Arthroscopy Cohort Southern Denmark (KACS): protocol for a prospective cohort study. *BMJ Open*. 2013;3:1-8. https://doi.org/10.1136/bmjopen-2013-003399
28. Thorlund JB, Hare KB, Lohmander LS. Large increase in arthroscopic meniscus surgery in the middle-aged and older population in Denmark from 2000 to 2011. *Acta Orthop*. 2014;85:287-292. https://doi.org/10.3109/17453674.2014.919558
29. van de Graaf VA, Wolterbeek N, Scholtes VAB, Mutsaerts ELAR, Poolman RW. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. *Am J Sports Med*. 2014;42:1408-1416. https://doi.org/10.1177/0363546514524698
30. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg*. 2014;12:1500-1524. https://doi.org/10.1016/j.ijsu.2014.07.014
31. Yeh PC, Starkey C, Lombardo S, Vitti G, Kharrazi FD. Epidemiology of isolated meniscal injury and its effect on performance in athletes from the National Basketball Association. *Am J Sports Med*. 2012;40:589-594. https://doi.org/10.1177/0363546511428601

JONAS BLOCH THORLUND, PhD[1,2] • MILENA SIMIC, PT, PhD[3] • KENNETH PIHL, PT, PhD[1] • DORTHE BANG BERTHELSEN, PT, MSc[4,5]
RICHARD DAY, MD[6,7] • BART KOES, PhD[1,8] • CARSTEN BOGH JUHL, PT, PhD[1,9]

Similar Effects of Exercise Therapy, Nonsteroidal Anti-inflammatory Drugs, and Opioids for Knee Osteoarthritis Pain: A Systematic Review with Network Meta-analysis

OBJECTIVE: To compare the effectiveness of opioids, nonsteroidal anti-inflammatory drugs (NSAIDs), and exercise therapy for knee osteoarthritis pain.

DESIGN: Systematic review with network meta-analysis.

LITERATURE SEARCH: We searched the databases MEDLINE, EMBASE, and Cochrane Central Register of Controlled Trials from inception to April 15, 2021. Web of Science was used for citation tracking.

STUDY SELECTION CRITERIA: Randomized controlled trials comparing exercise therapy, NSAIDs, and opioids in any combination for knee osteoarthritis pain.

DATA SYNTHESIS: Network meta-analysis comparing exercise therapy, NSAIDs, opioids, and placebo/control for knee osteoarthritis pain. Additional trials from previous reviews were included to create the external placebo/control anchor.

RESULTS: We included 13 trials (1398 patients) with direct comparisons, supplemented with data from 101 additional trials. The treatment effect of NSAIDs for knee osteoarthritis pain was similar to that of opioids (standardized mean difference [SMD], 0.02; 95% confidence interval [CI], −0.14 to 0.18; Grading of Recommendations, Assessment, Development and Evaluations [GRADE]: low certainty). Exercise therapy had a larger effect than NSAIDs (SMD, 0.54; 95% CI, 0.19 to 0.89; GRADE: very low certainty). No estimate could be made for exercise vs opioids due to the lack of studies. Exercise therapy ranked as the "best" intervention in the network meta-analysis, followed by NSAIDs, opioids, and placebo/control intervention (GRADE: low certainty).

CONCLUSION: Exercise therapy ranked as the best treatment for knee osteoarthritis pain, followed by NSAIDs and opioids. The difference between treatments was small and likely not clinically relevant, and the overall confidence in the ranking was low. The results highlight the limited evidence for comparative effectiveness between exercise therapy, NSAIDs, and opioids for knee osteoarthritis pain.
J Orthop Sports Phys Ther 2022;52(4):207-216. doi:10.2519/jospt.2022.10490

KEY WORDS: *analgesics, knee, osteoarthritis, pain, physiotherapy*

Opioids are frequently used to treat chronic musculoskeletal pain conditions.[18] In Sweden, 1 in 4 patients with knee or hip osteoarthritis (OA) have an opioid dispensed within a 12-month period, and a substantial proportion of patients have opioids prescribed within the first year of OA diagnosis.[43,44] The appropriateness for treating musculoskeletal pain conditions, such as OA, with opioids is the subject of strident debate due to risk of adverse events (AEs) and addiction.[8,16,42] Guidelines generally do not recommend opioids for knee OA pain unless other treatment options are exhausted, ineffective, or contraindicated.[28,30] The most recent guidelines from the Osteoarthritis Research Society International made a strong recommendation against any use of opioids for knee OA.[3]

Two other common treatments for knee OA pain are exercise therapy and nonsteroidal anti-inflammatory drugs (NSAIDs). The treatment effect estimated in systematic reviews and meta-analyses of opioids compared to placebo for knee OA pain does not seem to be larger than the treatment effect of NSAIDs compared to placebo or the treatment effect of exercise therapy compared to control interventions.[10,15,45] However, comparing effects across interventions obtained from randomized trials in pairwise meta-analysis is

[1]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark. [2]Research Unit of General Practice, Department of Public Health, University of Southern Denmark, Odense, Denmark. [3]Discipline of Physiotherapy, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia. [4]Department of Rehabilitation, Municipality of Guldborgsund, Nykøbing Falster, Denmark. [5]Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark. [6]Department of Clinical Pharmacology & Toxicology, St Vincent's Hospital Sydney, Sydney, Australia. [7]Faculty of Medicine, St Vincent's Clinical School, University of New South Wales, Sydney, Australia. [8]Department of General Practice, Erasmus University Medical Center, Rotterdam, the Netherlands. [9]Department of Occupational Therapy and Physiotherapy, Herlev and Gentofte Hospital, Copenhagen University Hospital, Copenhagen, Denmark. This study was registered with PROSPERO (International Prospective Register of Systematic Reviews; registration number CRD42018106484). No funding was received for conducting this study. Jonas Bloch Thorlund reports a grant from Pfizer outside the submitted work. Milena Simic is supported by the Sydney University Research Accelerator (SOAR) fellowship. Richard Day has received support for his participation as "Community Expert" from Reckitt Benckiser ending in 2018. The other authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Jonas Bloch Thorlund, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, 5230 Odense M, Denmark. E-mail: jthorlund@health.sdu.dk

limited because no direct statistical comparison can be made between all relevant interventions. We used network meta-analysis to provide more valid estimates of the comparative effectiveness of opioids, NSAIDs, and exercise therapy for knee OA pain. Such information is important for musculoskeletal rehabilitation clinicians when supporting patients to make decisions about treatment for knee OA. Important knowledge gaps may also be identified.

METHODS

We preregistered the protocol in the PROSPERO (International Prospective Register of Systematic Reviews) database (registration number CRD42018106484; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=106484) and performed the study according to the guidelines for network meta-analysis using Stata.[40] We report the findings of our study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension statement for reporting systematic reviews incorporating network meta-analyses.[24]

Eligibility Criteria

We included randomized controlled trials comparing exercise therapy with NSAIDs, exercise therapy with opioids, or NSAIDs with opioids for knee OA pain. We excluded trials where patients had a knee replacement surgery and trials where patients suffered from conditions other than OA, unless separate data were available for patients with knee OA. Trials including mixed populations of both knee and hip OA were included, as the majority of these patients typically have knee OA (this was a deviation from the protocol registered at PROSPERO). We defined exercise therapy as a regimen or plan of physical activities designed and prescribed for a specific therapeutic goal (ie, to reduce knee OA pain or improve muscle function), as defined by the Medical Subject Headings term in PubMed. We excluded trials that involved combined interventions in which exercise therapy constituted less than 50% of the intervention. We included all trials on drugs classified as NSAIDs or opioids according to the Anatomical Therapeutic Chemical codes.

Literature Search and Study Selection

We carried out systematic searches in MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials. Citation tracking was performed in Web of Science. The final search was performed on April 15, 2021. We developed the search strategies in MEDLINE and adjusted to the other databases using a combination of key words (ie, Medical Subject Headings) and text words (**SUPPLEMENTAL FILE 1**). We had no restrictions for publication year and language. Two members of the study team (K.P. and D.B.B.) independently assessed all titles and abstracts of the identified reports for eligibility. If at least 1 of the reviewers judged a trial eligible, we obtained the full text and had 2 members independently evaluate for study inclusion eligibility. To identify additional trials, we reviewed reference lists of included trials and reviews published within the last 5 years. We resolved disagreements on inclusion by consensus.

Because we identified a low number of trials with direct comparison between treatments and no trials investigating exercise therapy vs opioids, we created an external anchor for the comparison in the network meta-analysis from trials comparing NSAIDs to placebo, opioids to placebo, and exercise therapy to control interventions. We used the same search (filtered for systematic reviews) to identify the most recent and relevant meta-analyses in Cochrane reviews. When no suitable Cochrane review was found, we extracted data from the most recent/relevant systematic review and meta-analysis we could identify (this approach was a deviation from the protocol registered with PROSPERO).

Data Extraction

The prespecified outcome of interest was pain. When a report provided data on more than 1 pain scale, a published hierarchy for the selection of patient-reported outcomes was used.[26] We extracted data on outcomes for each of the intervention groups for the longest follow-up assessment reported in the included trials. As recommended in the *Cochrane Handbook for Systematic Reviews of Interventions*, we extracted the standard deviation (SD) of the outcome measurements, and when the SD was not described, it was estimated from standard errors, confidence intervals (CIs), *P* values, or interquartile ranges.[21] If data were reported in graphical form only, mean values and measures of dispersion were extrapolated using the WebPlotDigitizer software. In crossover trials, both phases of the trial were included, as the crossover effect of opioids and NSAIDs are minimal. For each intervention group, we also extracted the number of participants who were randomized, distribution of sex, mean age at baseline, body mass index at baseline, pain (intensity) at baseline, details about the interventions, number of AEs, type of AEs, and number of withdrawals due to AEs. Furthermore, we extracted definition criteria of OA for each trial. A customized data extraction form was used to independently extract all data by 2 of the study authors (M.S. and J.B.T.).

Similarity of Study Populations

A qualitative assessment of the clinical similarity of the treatment populations was performed based on mean age, sex distribution, OA severity (eg, Kellgren-Lawrence score), and baseline pain.

Data Synthesis

The effects from individual trials were expressed as the standardized mean difference (SMD) with a 95% CI. The SMD was estimated as the mean difference at the end of follow-up between treatment groups divided by the pooled SD. This estimate of the treatment effect size has a slight bias especially in smaller studies

overestimating the effect, and a correction factor was applied to convert the effect size to Hedges's *g*.[19] In trials reporting the effect as the number of participants reaching a predefined level of pain, the effect was estimated as an odds ratio (with a 95% CI) and transformed into an SMD using the formula proposed by Chinn.[9] First, we pooled the results from the individual trials with direct comparisons between the 3 interventions (ie, opioids, NSAIDs, and exercise). Then, we performed frequentist network meta-analysis based on the direct comparison of interventions (ie, opioids, NSAIDs, and exercise) to estimate the effect of pairwise comparisons (ie, based on direct and indirect comparisons).

We identified few trials with direct comparison between treatments and no trials investigating exercise therapy vs opioids. Therefore, we created an external anchor for the comparison in the network meta-analysis by extracting data from trials included in previous systematic reviews comparing NSAIDs to placebo, opioids to placebo, and exercise therapy to control interventions. The final network meta-analysis was performed including the extracted trials comparing the 3 interventions to placebo/control interventions.

Assessing Inconsistency

The heterogeneity in the analyses including direct comparisons between treatments was estimated using the I^2 statistic,[22] which measures the proportion of variation in the combined estimates attributable to between-study heterogeneity.[23] We checked the overall model for consistency using the command "network meta inconsistency" in Stata applying an F test for evaluating consistency. Side-split tables were produced to identify the source of inconsistency. The relative rank of the interventions along with the surface under the cumulative ranking curve was estimated.

Interpreting the Results

The SMD is often poorly understood.[25] To facilitate interpretation of the results, we also converted the estimated SMD in the final network meta-analysis into pain scores on a visual analog scale (VAS) using previously published methods. The converted VAS pain score (from 0 to 100 mm) was calculated by multiplying the SMD with an SD equal to 16.9 mm for pain.[6] The SDs used to convert the SMD into millimeters were based on a cohort of 914 patients with knee OA.[47] We considered that a difference in change in VAS pain between interventions had to be at least 15 mm to be clinically important. Finally, the network meta-analysis was repeated, stratified by OA classification (only patients with knee OA or patients with mixed knee and hip OA), age (over or under the median age), and the percentage of female participants in the study (over or under the median percentage of female participants). Risk of publication bias was investigated using funnel plots. All analyses were performed with Stata (Version 17.0; StataCorp LLC, College Station, TX), using a restricted maximum likelihood method to estimate the combined effect size and the between-study variance.

Assessing the Risk of Bias

Two reviewers (M.S. and J.B.T.) independently assessed risk of bias for trials with direct comparisons using the Cochrane Collaboration's risk-of-bias tool,[21] including the following domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and other bias. Other bias addressed the source of funding. We rated each domain as "low," "high," or "unclear" risk of bias. Disagreements between the 2 reviewers were resolved by consensus.

For the trials identified to create the external anchor for the network meta-analysis, we extracted the results of the risk-of-bias assessments performed in those trials (reported in the supplemental files). We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) adjusted for the network meta-analysis framework to assess the overall quality of the evidence, evaluating the certainty of the estimates on (1) study limitations, (2) indirectness and intransitivity, (3) statistical heterogeneity and statistical inconsistency, (4) imprecision, and (5) publication bias (using the GRADE approach was not registered in the PROSPERO protocol) (**SUPPLEMENTAL TABLE 1**).[39]

RESULTS

Our literature search identified 7719 independent references after excluding duplicates. Of these, 58 were considered for full-text review, and 13 trials including 1398 patients met the inclusion criteria.[1,5,7,11,13,14,27,29,34-36,48] Reasons for exclusion after the full-text review are reported in the flowchart (**FIGURE 1**) and in **SUPPLEMENTAL FILE 2**.

Study Characteristics

Of the included trials, 11 compared NSAIDs to opioids[1,5,7,11,13,27,29,33-36] and 2 compared NSAIDs to exercise therapy.[14,48] We did not identify any trials comparing opioids to exercise therapy. Five of the trials comparing NSAIDs to opioids were crossover trials, meaning that 181 of the 1398 patients were exposed to both treatments and acted as their own control.[7,13,27,35,36] The most common pain outcome was Western Ontario and McMaster Universities Osteoarthritis Index pain (n = 5), followed by VAS pain (n = 4) and numeric rating scale pain (n = 2); the remaining 2 trials used other pain outcomes. In 2 trials, pain data were extrapolated from figures.[33,36] Mean patient age ranged from 53 to 69 years, and mean baseline pain ranged from 34 to 74 mm on a 0- to 100-mm VAS. Follow-up for the primary endpoint in the trials ranged from 2 days to 52 weeks; most trials (n = 8) had their primary endpoint within 4 to 14 weeks (**SUPPLEMENTAL TABLE 2**).

The NSAIDs provided (n = 13 trials) were diclofenac (n = 3),[5,35,48] naproxen (n = 3),[27,33,36] ibuprofen (n = 1),[13] celecoxib (n = 2),[7,11] etoricoxib (n = 1),[1] and a mix of

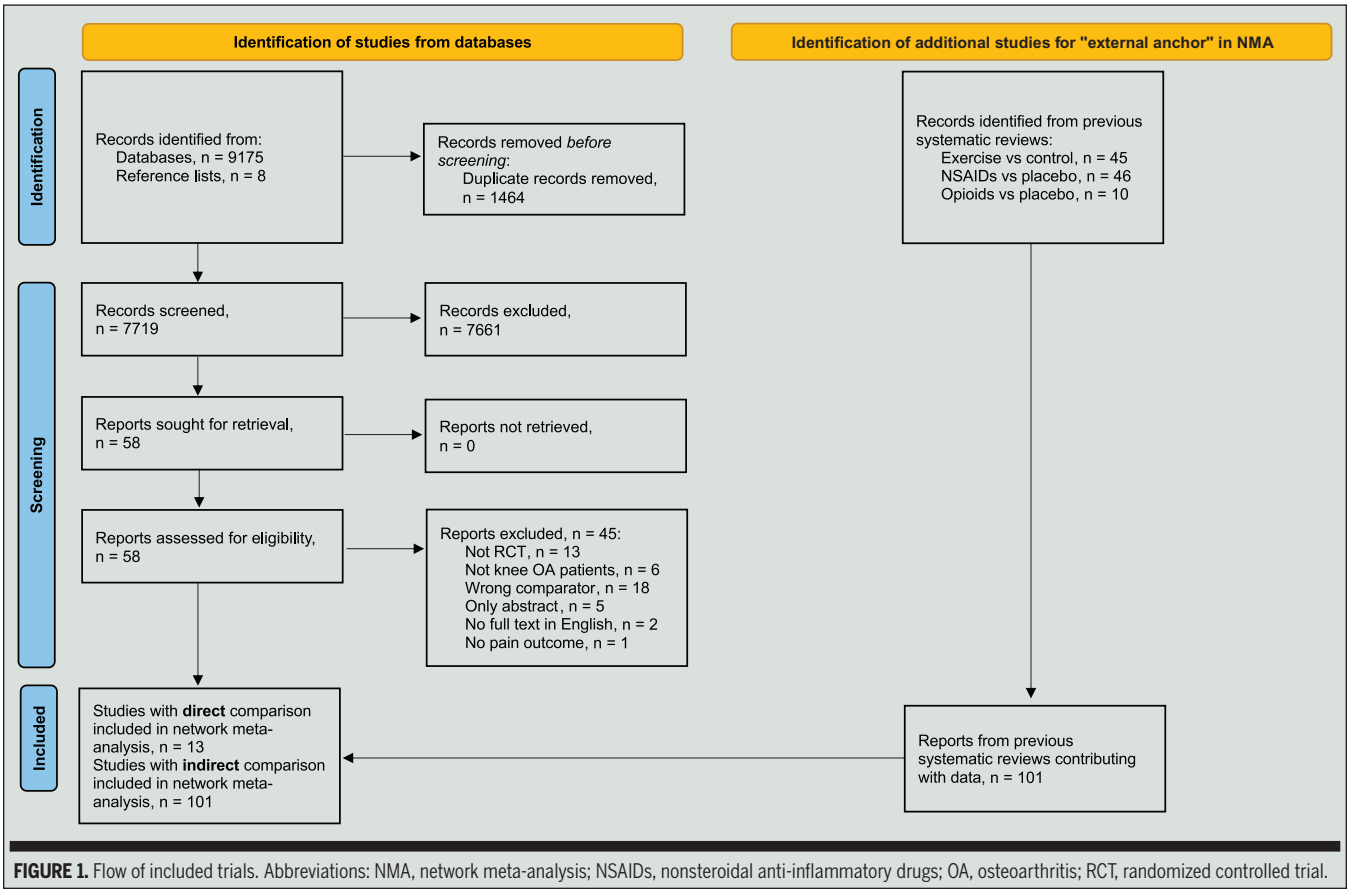

FIGURE 1. Flow of included trials. Abbreviations: NMA, network meta-analysis; NSAIDs, nonsteroidal anti-inflammatory drugs; OA, osteoarthritis; RCT, randomized controlled trial.

NSAIDs (n = 3),[14,29,34] which were delivered orally (n = 11),[1,5,7,11,13,14,27,33-36] topically (n = 1),[48] or both orally and topically (n = 1).[29] The opioids provided (n = 11 trials) were tramadol (n = 7),[5,11,27,33-36] codeine (n = 1),[13] oxycodone (n = 1),[7] tapentadol (n = 1),[1] and a mix of opioids (n = 1),[29] which were delivered orally in all trials except for 1 trial[29] that used both oral and transdermal delivery. In the trials including exercise therapy (n = 2), 1 trial delivered a quadriceps home exercise program,[14] and the other delivered quadriceps and hamstrings isokinetic exercises using a seated dynamometer[48] (**SUPPLEMENTAL TABLE 2**).

To increase precision in the network meta-analysis, we established an external anchor for the 3 interventions by including trials identified from systematic reviews and meta-analyses comparing NSAIDs and opioids (ie, tramadol) to a placebo comparator and exercise therapy to control interventions (such as no intervention, wait-list control, patient education, ultrasound, etc).[15,41,45] From these sources, 45 trials compared exercise therapy to control interventions, 46 compared NSAIDs to placebo, and 10 compared opioids (ie, tramadol) to placebo (**SUPPLEMENTAL TABLE 3**).

Similarity of Study Populations

The populations in trials across treatments were similar with respect to mean age, sex distribution, knee OA severity, and baseline pain. However, knee OA severity and baseline pain were only reported in a limited number of trials.

Results of the Network Meta-analysis

All active treatments (ie, exercise therapy, NSAIDs, and opioids) showed small-to-moderate treatment effects (SMD, 0.27-0.45) compared with placebo/control treatment (**FIGURE 2** and **TABLE 1**). The treatment effect of NSAIDs on knee OA pain was similar to that of opioids (SMD, 0.02; 95% CI, −0.14 to 0.18; corresponding to 0.3 mm on a 0- to 100-mm VAS pain scale), with low confidence in the estimate. Exercise showed a larger effect compared with NSAIDs (SMD, 0.54; 95% CI, 0.19 to 0.89; corresponding to 9.1-mm VAS pain), with very low confidence in the estimates due to study limitations, inconsistency, and indirectness. All estimated SMDs were mixed estimates (ie, a combination of direct and indirect comparisons). Exercise had the highest probability of ranking as the "best" intervention in the network meta-analysis, followed by NSAIDs and opioids, and control intervention ranked "worst," with low confidence in the ranking (**TABLE 2**).

Pairwise Comparisons

In the 11 trials with a direct comparison between opioids and NSAIDs, we found

no difference in the treatment effect on knee OA pain (SMD, 0.03; 95% CI, −0.13 to 0.18; I^2 = 31.8%) (**FIGURE 3**). In the 2 trials investigating a direct comparison between NSAIDs and exercise therapy, we found a large SMD in favor of exercise, but with wide CIs crossing the line of no effect and considerable heterogeneity (SMD, 0.80; 95% CI, −0.19 to 1.79; I^2 = 90.7%) (**FIGURE 4**).

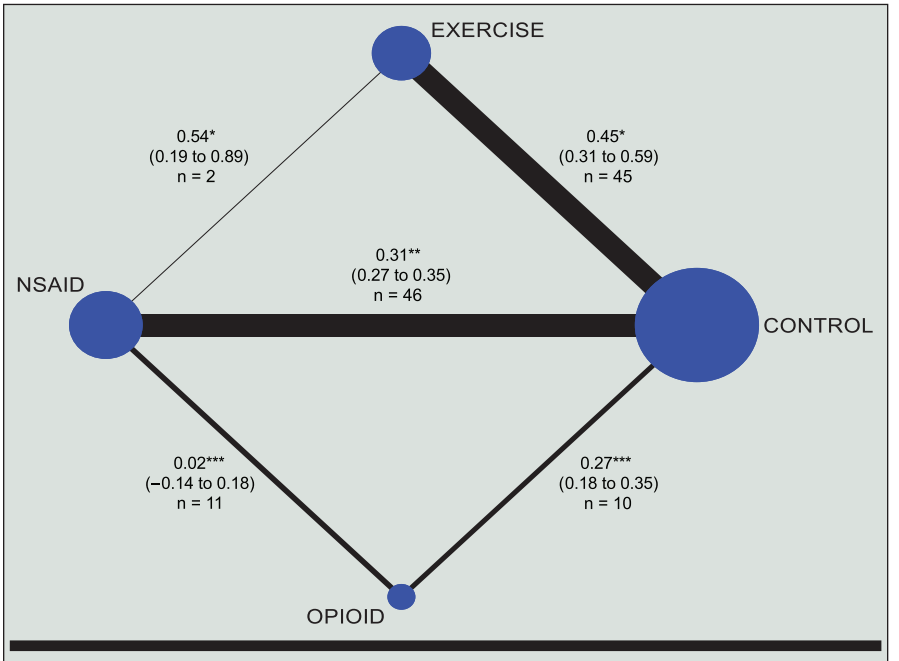

FIGURE 2. Graphic presentation of comparisons between control interventions (placebo or control), exercise therapy, nonsteroidal anti-inflammatory drugs (NSAIDs), and opioids with standardized mean differences, 95% confidence intervals, and number of trials (n). No estimate available for exercise therapy vs opioids. *, favor exercise therapy; **, favor NSAIDs; ***, favor opioids.

Inconsistency in the Network Meta-analysis

The network meta-analysis did not provide a valid estimate for the comparison between exercise and opioids, as no trials with direct comparison were found and consistency was not reached (**FIGURE 2** and **TABLE 1**). Side-split tables revealed the largest difference between direct and indirect estimates in comparisons with exercise. The estimates in the network meta-analysis were dominated by trials comparing to placebo/control (**SUPPLEMENTAL FIGURE 1**), and the network meta-analysis showed considerable inconsistency (F[2, 109] = 4.78, P = .010). When we excluded 1 trial comparing exercise therapy with NSAIDs, which had extreme results from the network meta-analysis,[48] we observed overall model consistency (P = .709) while the estimates remained essentially the same.

Subgroup Analysis

In additional analyses, stratified by OA classification (only patients with knee

TABLE 1 Results From Network Meta-analysis Comparing Placebo/Control Interventions, Exercise Therapy, NSAIDs, and Opioids for Knee Osteoarthritis Pain[a]

Comparison	SMD (95% CI)	SMD Converted to mm VAS Pain (95% CI)	No. of Trials With Direct Comparison	SMD Favors	Nature of Evidence	Confidence	Downgrading Due to
Exercise vs control	0.45 (0.31 to 0.59)	7.6 (5.2 to 10.0)	45	Exercise	Mixed	Low	Study limitation,[b] inconsistency[c]
NSAIDs vs placebo	0.31 (0.27 to 0.35)	5.2 (4.6 to 5.9)	46	NSAIDs	Mixed	Moderate	Inconsistency[c]
Opioids vs placebo	0.27 (0.18 to 0.35)	4.6 (3.0 to 5.9)	10	Opioids	Mixed	Low	Study limitation,[b] inconsistency[c]
Exercise vs NSAIDs	0.54 (0.19 to 0.89)	9.1 (3.2 to 15.0)	2	Exercise	Mixed	Very low	Study limitation,[b] inconsistency,[c] imprecision[d]
Exercise vs opioids	NA	NA	NA	NA	None	NA	NA
NSAIDs vs opioids	0.02 (−0.14 to 0.18)	0.3 (−2.4 to 3.0)	11	Opioids	Mixed	Low	Study limitation,[b] imprecision[d]
Ranking of the treatment						*Low*	*Study limitation,[e] inconsistency[f]*

Abbreviations: CI, confidence interval; NSAIDs, nonsteroidal anti-inflammatory drugs; SMD, standardized mean difference; VAS, visual analog scale.
[a]*Overall network meta-analysis heterogeneity, F test: F(2, 109) = 4.78, P = .01 and I2 = 6.6%.*
[b]*Evidence mainly from trials with high risk of bias.*
[c]*Inconsistency (ie, difference between direct and indirect evidence or difference in estimates of the included trials).*
[d]*Confidence interval includes values favoring either treatment.*
[e]*Study limitation (ie, almost 60% of the included trials were judged high risk of bias).*
[f]*Inconsistency (ie, inconsistency in the overall network meta-analysis).*

TABLE 2 Relative Ranking of Individual Treatments Estimated From the Network Meta-analysis

	Treatment			
Ranking	**Exercise**	**NSAIDs**	**Opioids**	**Control[a]**
Best	100.0	0.0	0.0	0.0
2nd	0.0	85.5	14.5	0.0
3rd	0.0	14.5	85.5	0.0
Worst	0.0	0.0	0.0	100.0
Mean rank	1.0	2.1	2.9	4.0
SUCRA	1.0	0.6	0.4	0.0

Abbreviations: NSAIDs, nonsteroidal anti-inflammatory drugs; SUCRA, surface under the cumulative ranking curve.
[a]*Control: Placebo or control interventions.*

OA or patients with mixed knee and hip OA), age (over or under 61 years), and percentage of female participants in the study (over or under 70%), the estimates remained essentially the same (**SUPPLEMENTAL TABLE 4**).

Adverse Events

AEs were not consistently reported in the included trials, precluding meaningful summary measures. In trials comparing NSAIDs with opioids, a larger proportion of patients who received opioids reported experiencing AEs, and more patients dropped out for this reason. For an overview of the number and types of AEs, please refer to **SUPPLEMENTAL TABLE 5.**

Risk of Bias

One trial[33] had low risk of bias for all domains, and 2 trials[7,36] were considered low on all domains except "other bias," as these trials were sponsored by pharmaceutical companies or had authors who were employed by pharmaceutical companies. Many trials were scored as having high risk of bias or unclear risk of bias for "blinding of participants and personnel" (6/13 trials),[1,14,27,29,34,48] "blinding of outcome assessment" (6/13 trials),[1,14,27,29,34,48] "incomplete outcome data" (6/13 trials),[1,5,11,13,34,48] and "selective outcome reporting" (9/13 trials).[1,5,11,13,14,27,34,35,48] Between 5 and 8 trials had high risk of bias or uncertain risk on the remaining domains (**SUPPLEMENTAL TABLE 6**). Risk-of-bias assessment of the trials included to provide the external anchor is reported in **SUPPLEMENTAL TABLES 7-9.** An inspection of the funnel plot did not indicate a severe risk of publication bias for the overall network meta-analysis and pairwise comparisons (**SUPPLEMENTAL FIGURE 2**).

DISCUSSION

We were unable to identify any trials with direct comparison between exercise therapy and opioids for knee OA pain and could not provide any valid estimate for this comparison. Eleven trials investigated

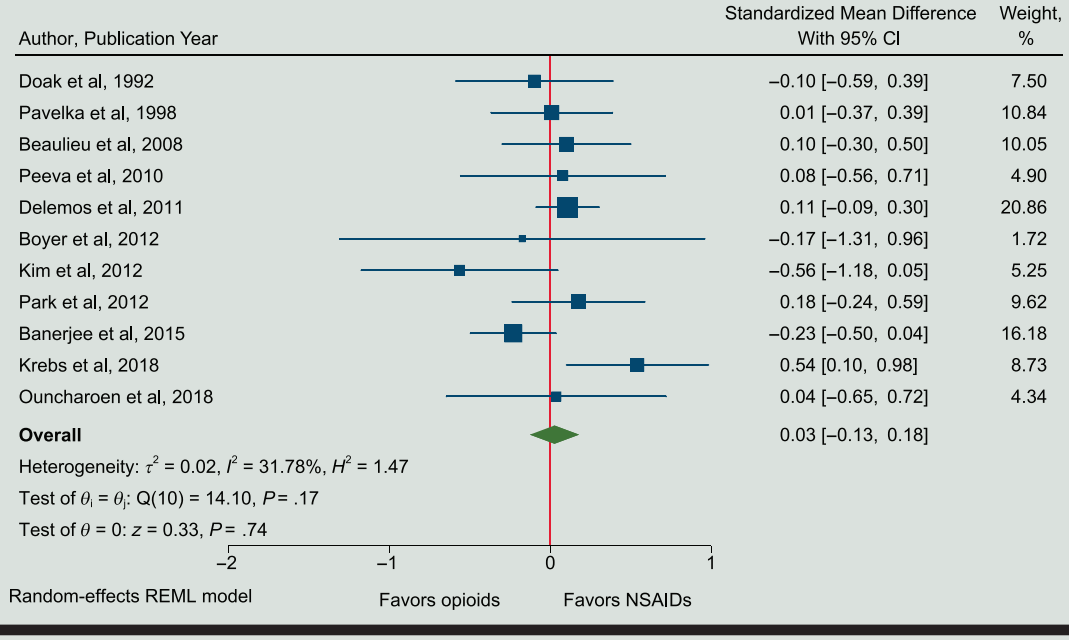

FIGURE 3. Results from the analysis of direct comparison of included trials comparing nonsteroidal anti-inflammatory drugs (NSAIDs) with opioids. Abbreviations: CI, confidence interval; REML, restricted maximum likelihood.

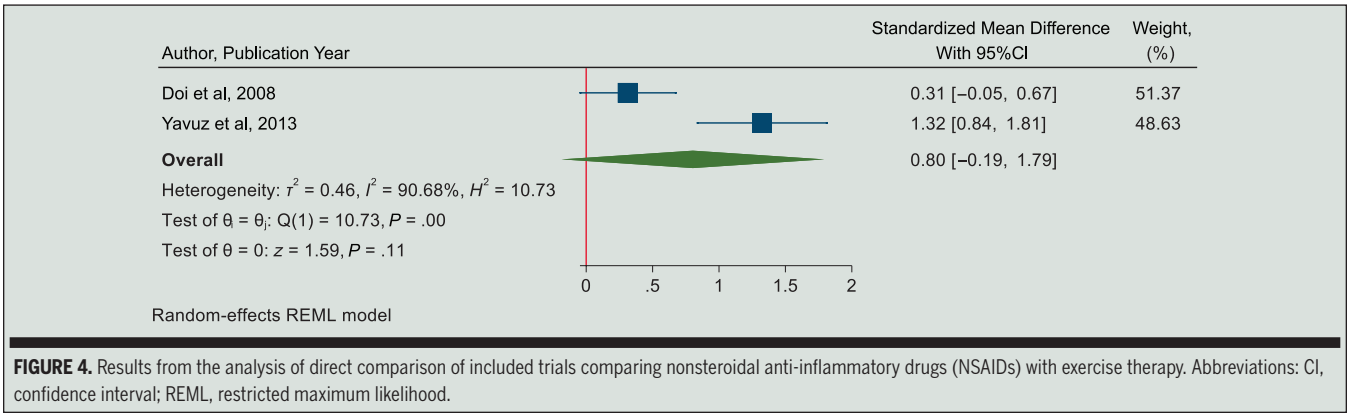

FIGURE 4. Results from the analysis of direct comparison of included trials comparing nonsteroidal anti-inflammatory drugs (NSAIDs) with exercise therapy. Abbreviations: CI, confidence interval; REML, restricted maximum likelihood.

NSAIDs vs opioids for knee OA pain, and the estimates from our network meta-analysis suggested similar pain-relieving effects, but with low confidence due to study limitations and imprecision. Exercise appeared superior for pain relief than NSAIDs, corresponding to 9.1 mm on a 0- to 100-mm VAS pain scale. However, this is unlikely to represent a clinically meaningful difference.[17,46] We have very low confidence in this estimate due to trial limitations (ie, only 2 high-risk-of-bias trials with direct comparison), inconsistency, and imprecision (ie, CIs overlapping the line of no effect). All 3 interventions showed small-to-moderate treatment effects when compared with placebo or control interventions. All estimates should be interpreted with caution, as estimates were driven by indirect comparisons, which highlights the need for trials comparing exercise therapy with NSAIDs and opioids.

A previous review attempting to compare the treatment effect of exercise therapy with analgesics for knee OA pain generally found no difference between interventions.[20] However, all estimates were based on indirect comparisons. We build on these findings by including trials with direct comparison between interventions. We initially aimed to only include trials with direct comparison between the 3 investigated interventions, but we had to adjust this strategy as we only found a limited number of trials with direct comparisons. To create an external anchor for the comparison in the network meta-analysis, we included data from relevant systematic reviews comparing opioids and NSAIDs to placebo as well as exercise therapy to control interventions (ie, active non-exercise intervention or no treatment including wait list). Thus, all estimates in this study are based on a mix of direct and indirect comparisons, and we could not provide any estimate comparing exercise therapy with opioids.

Comparison of Estimates From Network Meta-analyses With and Without External Anchor

Estimates from the analysis of NSAIDs vs opioids based on the network meta-analysis (SMD, 0.02) and the meta-analysis including direct comparisons alone (SMD, 0.03) for knee OA pain were similar. We found a considerable difference in the size of the estimates in the network meta-analysis with only direct comparisons for exercise therapy vs NSAIDs (SMD, 0.80) compared with the network meta-analysis including the external anchor (SMD, 0.54). The main reason for this discrepancy was that only 2 high-risk-of-bias trials were included in the direct comparison between exercise therapy (involving lower limb strengthening) and NSAIDs, with extreme results reported in 1 of these trials, resulting in high heterogeneity.

Risk of Harm

Use of NSAIDs for treating OA pain is associated with risk of harm. The risk of harm is greater with opioids, which also have a substantial addiction potential.[10,12,37,45] On the contrary, exercise therapy for knee OA pain has minimal or no risk of AEs[31,38] and is therefore unanimously recommended by international clinical guidelines as first-line treatment for all patients with knee OA.[30] However, quality-of-care utilization studies report that exercise therapy as treatment is greatly underprescribed.[32,4] Too many patients with OA are missing out on guideline-recommended first-line treatment, as they are directed to second-line pharmacological treatment.

Ranking Treatments

Our analyses ranked exercise first (low confidence) and suggested that exercise therapy may yield superior treatment effects to NSAIDs (compared head-to-head) and better treatment effects than NSAIDs and opioids when compared with placebo/control interventions for knee OA pain. This ranking suggests that a potential to reverse from second-line pharmacological care to first-line treatment with exercise therapy may exist for patients using NSAIDs and opioids, particularly those who initially missed out on proper first-line care. This is important for clinicians when considering treatment options for patients with knee OA pain who are using analgesics.

Limitations

We identified only a limited number of trials reporting direct comparisons between the investigated interventions.

Thus, we supplemented the network meta-analysis with data from trials comparing the interventions to placebo/control interventions. The majority of trials contributing to the network meta-analysis therefore provided data from indirect comparisons. We identified no trials comparing exercise therapy to opioids for knee OA pain, and as a result, we could not provide a valid estimate of this comparison despite addition of the external anchor. Similar to previous attempts to compare different analgesic interventions with exercise,[20] our review also suffers from the limitation that patients in the included trials comparing treatments (exercise, NSAIDs, and opioids) to placebo/control may have varying disease severity, and to a large extent, we rely on data from indirect comparisons.

Trials involving pharmacological interventions included a variety of NSAIDs and opioids. As the majority of trials with direct comparison of opioids used tramadol (7 out of 11), we extracted data from a recent Cochrane review on tramadol for OA.[45] For NSAIDs, a broader range of drugs was used in the included trials; thus, we opted to include data from the most recent systematic review that best matched this variation.[2]

We observed considerable inconsistency in the network meta-analysis. This was mainly due to an extreme effect size in 1 trial that compared exercise therapy with NSAIDs.[48] When we excluded the trial, the overall model reached consistency (P = .709), and estimates remained essentially the same. However, the heterogeneity in the comparison between NSAIDs and exercise remained high.

Ten of the 13 trials with direct comparisons between the 3 investigated interventions had unclear or high risk of bias on 2 or more domains. Some trials included both patients with knee and hip OA. To increase the number of included trials with direct comparisons, we decided to include these trials as most patients had knee OA, and in the analysis stratified by OA classification, results remained similar (**SUPPLEMENTAL TABLE 4**). Participants in crossover trials contributed with data to both the NSAID and opioid arms of these trials. This may result in the overestimation of the precision of the reported estimates.

Reporting of AEs varied in definition and reporting, and it was therefore not possible to pool these data in a meaningful way. We did not include disability as an outcome, as we considered pain was the main and most comparable domain targeted by the 3 interventions investigated.

There were some deviations from the registered protocol, as follows: (1) We decided to include trials with mixed populations as the majority of these patients typically have knee OA, (2) we created an external anchor in the network meta-analysis including trials comparing the 3 investigated interventions with placebo/control interventions, (3) we used the GRADE approach to assess the overall quality of evidence, and (4) a number of preplanned subgroup analyses could not be performed due to limited data availability.

CONCLUSION

Our network meta-analysis suggested exercise therapy as the best treatment for knee OA pain, followed by NSAIDs and opioids. Differences between treatments were likely not clinically relevant (<10-mm VAS pain), and the overall confidence in the ranking of treatments was low, with few trials reporting direct comparisons of exercise therapy, NSAIDs, and opioids and opioids for knee OA pain. A potential to reverse from second-line pharmacological care to first-line treatment with exercise therapy may exist for patients using NSAIDs and opioids, particularly those who initially missed out on proper first-line care.

CAUTION: Only a small number of trials had direct comparison between exercise therapy and NSAIDs and between NSAIDs and opioids. No trials directly compared exercise therapy with opioids. The confidence in the estimates from the network meta-analysis was generally low.

KEY POINTS

FINDINGS: Our network meta-analysis ranked exercise therapy as the best treatment for knee OA pain, followed by NSAIDs and opioids, although differences between treatments are likely not clinically relevant.

IMPLICATIONS: More high-quality studies with direct comparison are needed to better ascertain the comparative effectiveness of exercise therapy, NSAIDs,

STUDY DETAILS

AUTHOR CONTRIBUTIONS: Drs Thorlund, Simic, Day, Koes, and Juhl participated in the conception and design of the study. Drs Thorlund, Simic, Pihl, and Juhl and Dorthe Bang Berthelsen were responsible for the acquisition of data. Dr Juhl performed the analysis, and Drs Juhl and Thorlund interpreted the data. Dr Thorlund drafted the manuscript. All authors critically revised the manuscript for important intellectual content and approved the final version of the manuscript.

DATA SHARING: Data are available upon request. Extracted data from trials included in the network meta-analysis can be shared upon request to the corresponding author. There are no restrictions on reuse of data.

PATIENT AND PUBLIC INVOLVEMENT: No patients were involved in setting the research question or the conduct of this study.

REFERENCES

1. Banerjee M, Mondal S, Sarkar R, Mondal H, Bhattacharya K. Comparative study of efficacy and safety of tapentadol versus etoricoxib in mild to moderate grades of chronic osteorthritis of knee. *Ind J Rheumatol*. 2016;11:21-25. https://doi.org/10.1016/j.injr.2015.12.001
2. Bannuru RR, McAlindon TE, Sullivan MC, Wong JB, Kent DM, Schmid CH. Effectiveness and implications of alternative placebo treatments: a systematic review and network

meta-analysis of osteoarthritis trials. *Ann Intern Med*. 2015;163:365-372. https://doi.org/10.7326/M15-0623
3. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthr Cartil*. 2019;27:1578-1589. https://doi.org/10.1016/j.joca.2019.06.011
4. Basedow M, Williams H, Shanahan EM, Runciman WB, Esterman A. Australian GP management of osteoarthritis following the release of the RACGP guideline for the non-surgical management of hip and knee osteoarthritis. *BMC Res Notes*. 2015;8:536. https://doi.org/10.1186/s13104-015-1531-z
5. Beaulieu AD, Peloso PM, Haraoui B, et al. Once-daily, controlled-release tramadol and sustained-release diclofenac relieve chronic pain due to osteoarthritis: a randomized controlled trial. *Pain Res Manag*. 2008;13:103-110. https://doi.org/10.1155/2008/903784
6. Bliddal H, Christensen R. The treatment and prevention of knee osteoarthritis: a tool for clinical decision-making. *Expert Opin Pharmacother*. 2009;10:1793-1804. https://doi.org/10.1517/14656560903018911
7. Boyer KA, Angst MS, Asay J, Giori NJ, Andriacchi TP. Sensitivity of gait parameters to the effects of anti-inflammatory and opioid treatments in knee osteoarthritis patients. *J Orthop Res*. 2012;30:1118-1124. https://doi.org/10.1002/jor.22037
8. Cheatle MD. Prescription opioid misuse, abuse, morbidity, and mortality: balancing effective pain management and safety. *Pain Med*. 2015;16:S3-S8. https://doi.org/10.1111/pme.12904
9. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med*. 2000;19:3127-3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M
10. da Costa BR, Nuesch E, Kasteler R, et al. Oral or transdermal opioids for osteoarthritis of the knee or hip. *Cochrane Database Syst Rev*. 2014:CD003115. https://doi.org/10.1002/14651858.CD003115.pub4
11. DeLemos BP, Xiang J, Benson C, et al. Tramadol hydrochloride extended-release once-daily in the treatment of osteoarthritis of the knee and/or hip: a double-blind, randomized, dose-ranging trial. *Am J Ther*. 2011;18:216-226. https://doi.org/10.1097/MJT.0b013e3181cec307
12. Derry S, Wiffen P, Moore A. Topical nonsteroidal anti-inflammatory drugs for acute musculoskeletal pain. *JAMA*. 2016;315:813-814. https://doi.org/10.1001/jama.2016.0249
13. Doak W, Hosie J, Hossain M, James IGV, Reid I, Miller AJ. A novel combination of iboprofen and codeine phosphate in the treatment of osteoarthritis: a double-blind placebo controlled study. *J Drug Dev*. 1992;4:179-187.
14. Doi T, Akai M, Fujino K, et al. Effect of home exercise of quadriceps on knee osteoarthritis compared with nonsteroidal antiinflammatory drugs: a randomized controlled trial. *Am J Phys Med Rehabil*. 2008;87:258-269. https://doi.org/10.1097/PHM.0b013e318168c02d
15. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev*. 2015;1:CD004376. https://doi.org/10.1002/14651858.CD004376.pub3
16. Furlan AD, Sandoval JA, Mailis-Gagnon A, Tunks E. Opioids for chronic noncancer pain: a meta-analysis of effectiveness and side effects. *CMAJ*. 2006;174:1589-1594. https://doi.org/10.1503/cmaj.051528
17. Grant MD, Samson D. Special report: measuring and reporting pain outcomes in randomized controlled trials. *Technol Eval Cent Assess Program Exec Summ*. 2006;21:1-2.
18. Han B, Compton WM, Blanco C, Crane E, Lee J, Jones CM. Prescription opioid use, misuse, and use disorders in U.S. adults: 2015 National Survey on Drug Use and Health. *Ann Intern Med*. 2017;167:293-301. https://doi.org/10.7326/M17-0865
19. Hedges LV, Olkin I. *Statistical Methods for Meta-analysis*. New York: Academic Press; 1985.
20. Henriksen M, Hansen JB, Klokker L, Bliddal H, Christensen R. Comparable effects of exercise and analgesics for pain secondary to knee osteoarthritis: a meta-analysis of trials included in Cochrane systematic reviews. *J Comp Eff Res*. 2016;5:417-431. https://doi.org/10.2217/cer-2016-0007
21. Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011). Oxford, UK: The Cochrane Collaboration; 2011.
22. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539-1558. https://doi.org/10.1002/sim.1186
23. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-560. https://doi.org/10.1136/bmj.327.7414.557
24. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162:777-784. https://doi.org/10.7326/M14-2385
25. Johnston BC, Alonso-Coello P, Friedrich JO, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ*. 2016;188:25-32. https://doi.org/10.1503/cmaj.150430
26. Juhl C, Lund H, Roos EM, Zhang W, Christensen R. A hierarchy of patient-reported outcomes for meta-analysis of knee osteoarthritis trials: empirical evidence from a survey of high impact journals. *Arthritis*. 2012;2012:136245. https://doi.org/10.1155/2012/136245
27. Kim SY, Ryou JW, Hur J. Comparison of effectiveness and safety of tramadol/acetaminophen and non-steroidal anti-inflammatory drugs (NSAIDs) for treatment of knee osteoarthritis in elderly patients. *J Rheum Dis (Korean)*. 2012;19:25-29. https://doi.org/10.4078/jrd.2012.19.1.25
28. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. *Arthritis Rheumatol*. 2020;72:220-233. https://doi.org/10.1002/art.41142
29. Krebs EE, Gravely A, Nugent S, et al. Effect of opioid vs nonopioid medications on pain-related function in patients with chronic back pain or hip or knee osteoarthritis pain: the SPACE randomized clinical trial. *JAMA*. 2018;319:872-882. https://doi.org/10.1001/jama.2018.0899
30. McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthr Cartil*. 2014;22:363-388. https://doi.org/10.1016/j.joca.2014.01.003
31. Niemeijer A, Lund H, Stafne SN, et al. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. *Br J Sports Med*. 2020;54:1073-1080. https://doi.org/10.1136/bjsports-2018-100461
32. Osteras N, Jordan KP, Clausen B, et al. Self-reported quality care for knee osteoarthritis: comparisons across Denmark, Norway, Portugal and the UK. *RMD Open*. 2015;1:e000136. https://doi.org/10.1136/rmdopen-2015-000136
33. Ouncharoen T. Tramadol versus naproxen for pain relief in knee osteoarthritis: a pragmatic randomized controlled trial. *Clin Academia*. 2018;42:53-64.
34. Park KS, Choi JJ, Kim WU, Min JK, Park SH, Cho CS. The efficacy of tramadol/acetaminophen combination tablets (Ultracet®) as add-on and maintenance therapy in knee osteoarthritis pain inadequately controlled by nonsteroidal anti-inflammatory drug (NSAID). *Clin Rheumatol*. 2012;31:317-323. https://doi.org/10.1007/s10067-011-1818-y
35. Pavelka K, Peliskova Z, Stehlikova H, Ratcliffe S, Repas C. Intraindividual differences in pain relief and functional improvement in osteoarthritis with diclofenac or tramadol. *Clin Drug Investig*. 1998;16:421-429. https://doi.org/10.2165/00044011-199816060-00002
36. Peeva E, Beals CR, Bolognese JA, et al. A walking model to assess the onset of analgesia in osteoarthritis knee pain. *Osteoarthr Cartil*. 2010;18:646-653. https://doi.org/10.1016/j.joca.2009.12.008
37. Puljak L, Marin A, Vrdoljak D, Markotic F, Utrobicic A, Tugwell P. Celecoxib for osteoarthritis. *Cochrane Database Syst Rev*. 2017;5:CD009865. https://doi.org/10.1002/14651858.CD009865.pub2
38. Quicke JG, Foster NE, Thomas MJ, Holden MA. Is long-term physical activity safe for older adults with knee pain?: a systematic review. *Osteoarthr Cartil*. 2015;23:1445-1456. https://doi.org/10.1016/j.joca.2015.05.002

39. Salanti G, del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682. https://doi.org/10.1371/journal.pone.0099682
40. Shim S, Yoon BH, Shin IS, Bae JM. Network meta-analysis: application and practice using Stata. *Epidemiol Health*. 2017;39:e2017047. https://doi.org/10.4178/epih.e2017047
41. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E. Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review. *Osteoarthr Cartil*. 2016;24:962-972. https://doi.org/10.1016/j.joca.2016.01.135
42. Solomon DH, Rassen JA, Glynn RJ, et al. The comparative safety of opioids for nonmalignant pain in older adults. *Arch Intern Med*. 2010;170:1979-1986. https://doi.org/10.1001/archinternmed.2010.450
43. Thorlund JB, Turkiewicz A, Prieto-Alhambra D, Englund M. Inappropriate opioid dispensing in patients with knee and hip osteoarthritis: a population-based cohort study. *Osteoarthr Cartil*. 2020;28:146-153. https://doi.org/10.1016/j.joca.2019.10.004
44. Thorlund JB, Turkiewicz A, Prieto-Alhambra D, Englund M. Opioid use in knee or hip osteoarthritis: a region-wide population-based cohort study. *Osteoarthr Cartil*. 2019;27:871-877. https://doi.org/10.1016/j.joca.2019.01.005
45. Toupin April K, Bisaillon J, Welch V, et al. Tramadol for osteoarthritis. *Cochrane Database Syst Rev*. 2019;5:CD005522. https://doi.org/10.1002/14651858.CD005522.pub3
46. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis*. 2005;64:29-33. https://doi.org/10.1136/ard.2004.022905
47. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis*. 2005;64:34-37. https://doi.org/10.1136/ard.2004.023028
48. Yavuz M, Ataoğlu S, Baki AE, Özşahin M, Içmeli C. Compared effects and effectiveness in applications of isokinetic exercise, laser, and diclophenac iontophoresis in primary osteoarthritis of knee. *Duzce Med J*. 2013;15:15-21.