Draft Comparative Effectiveness Review

Number XX

Physical Therapy for Knee Pain Secondary to Osteoarthritis

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
http://www.ahrq.gov

This information is distributed solely for the purposes of predissemination peer review. It has not been formally disseminated by the Agency for Healthcare Research and Quality. It does not represent and should not be construed to represent an Agency for Healthcare Research and Quality or Department of Health and Human Services (AHRQ) determination or policy.

Contract No. XXX-XXXX-XXXXX-X

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by the XXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-XXXX-XXXXX-X). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the Children's Health Insurance Program (CHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting comparative effectiveness reviews (CERs) of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see http://www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that CERs will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input from are essential to the Effective Health Care Program. Please visit the Web site (http://www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

We welcome comments on this CER. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang M.D., M.P.H.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Suchitra Iyer, Ph.D.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Physical Therapy for Knee Pain Secondary to Osteoarthritis

Structured Abstract

Objectives: To assess the association between intermediate and patient-centered outcomes and harms with physical therapy interventions in community-dwelling adults with chronic pain secondary to knee osteoarthritis and to examine validity and minimal clinically important differences of the tools for outcome measurement.

Data Sources: We searched major electronic bibliographic databases and trial registries up to September 1, 2011.

Review Methods: We performed a systematic review of randomized and nonrandomized studies published in English to synthesize rates or means of measured pain, function, and quality of life with physical therapy interventions. Observational studies provided evidence of the association between changes in knee joint functional tests and patient-centered outcomes and minimal clinically important differences in validated tools for outcome measures. We performed meta-analyses using random effects models to synthesize the evidence.

Results: From a total of 4,214 retrieved references, 405 therapeutic and 151 diagnostic references were eligible for the review. Pooled analyses failed to show consistent significant, strong, or clinically important changes in outcomes. Aerobic and aquatic exercise improved disability (low strength of evidence); aerobic and strengthening exercise reduced pain (low strength of evidence) and improved function (low strength of evidence). Proprioception exercise reduced pain (low strength of evidence), and ultrasound improved function (low strength of evidence). Specific education programs, Tai Chi, diathermy, orthotics, and magnetic stimulation demonstrated no benefits in pooled analyses. Individual randomized controlled trials (RCTs) failed to show significant, strong, or clinically important changes in outcomes. Direct evidence of comparative effectiveness was limited. Disability measures did not differ between aerobic, aquatic, and strengthening exercise. Pain did not differ between electrical stimulation and exercise in pooled analyses. Evidence from individual RCTs was insufficient to conclude the best treatment option between physical therapy interventions or to conclude differences in effects by patient characteristics. Patients with high compliance tended to have a better treatment response with exercise. We found no association between the duration of examined interventions or followup times and better intermediate/patient-centered outcomes. Adverse events were uncommon and not severe enough to deter participants from continuing treatment. Gait, mobility restrictions, muscle strength, and range of motion measures were associated with disability measures in individual studies. Minimal clinically important differences in scales were determined for 26 tools but have not been used to examine clinical importance of improvements. The definition of the Patient Acceptable Symptom State that accounted for patient satisfaction was available for Western Ontario McMaster Universities Osteoarthritis Index, Visual Analog Scale for pain, and for Patient Global Assessment Scale.

Conclusions: A low level of evidence suggested that only a few physical therapy interventions, including exercise (aerobic, aquatic, strengthening, and proprioception) and ultrasound, were effective, and no single physical therapy intervention improved all outcomes. Study quality and heterogeneity in populations and treatments, including concomitant treatments, downgraded strength of evidence to low or moderate in most cases. Studies focused on a single modality of physical therapy rather than the combinations typically used in practice. Benefits with physical therapy interventions were not consistently evaluated according to clinical importance of improvement in scales and tests. Adverse events were uncommon and not severe enough to deter participants from continuing treatment. Evidence about long-term adherence to and benefits of available physical therapy interventions is lacking.

Contents

Executive Summary ES-1
Introduction 1
Measuring Outcomes of Physical Therapy Intervention 2
Key Question 1 2
Key Question 2 3
Key Question 3 3
Methods 4
Search Strategy 4
Results 12
Key Question 1. What are the effectiveness and comparative effectieness of available physical therapy interventions (without drug treatment) for adult patients with chronic knee pain due to OA on intermediate and patient-centered outcomes when compared to no active tretment or another active physical therapy modality? 12
Key Question 1a. Role of patient characteristics on outcomes 17
Key Question 1b. Association between the dose/intensity/frequency of examined interventions and intermediate/patient-centered outcomes 19
Key Question 1c. Association between the duration of examined interventions and intermediate/patient-centered outcomes 20
Key Question 2. What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after physical therapy interventions? 33
Key Question 2a. What is the validity of the testss and measures used to determine intermediate outcomes of physical therapy on osteoarthritis in association with patient-centered outcomes? 37
Key Question 2b. Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes? 38
Key Question 2c. What are minimal clinically important differences of the tests and measures used to determine intermediate outcomes? 38
Key Question 3. What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls? 47
Key Question 3a. Which patient characteristics are associated with the harms of examined physical therapy interventions? 47
Key Questin 3b. Do harms differ by the duration of the treatment and time of followup? 47
Summary and Discussion 50
Key Findings 50
Limitations 53
Future Research 54
References 66
Abbreviations 80

Tables

Table 1. Physical therapy interventions eligible for review 5
Table 2. Outcomes with physical therapy interventions from controlled clinical trials (pooling with random effects models standardized mean differences) 22

Table 3. Narrative evidence summary of effectiveness of physical therapy interventons for knee osteoarthritis 26
Table 4. Narrative evidence summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis....30
Table 5. Association between intermediate and clinical outcomes (insufficient evidence from individual studies)....40
Table 6. Regression association between intermediate and clinical outcomes....41
Table 7. Relative measure of association between intermediate outcomes and functional disability....42
Table 8. Regression association between psychological and clinical outcomes 46
Table 9. Adverse events reported with physical therapy for osteoarthritis of the knee....48
Table 10. Conclusions about efficacy of physical therapy interventions 55
Table 11. Summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis....58
Table 12. Reported physical therapy modalities as components of the examined physical therapy interventions....59
Table 13. Future research recommendations 62

Figures

Figure 1. Risk of developing disability in activities of daily living in compliance subgroups....32
Figure 2. Pain at the longest time of followup with physical therapy interventions vs. no active treatment (pooled with random effects model at 99% CI standardized mean differences from head-to-head randomized trials)....56
Figure 3. Ratio of means in pain measurements at the longest time of followup with physical therapy interventions vs. no active treatment (pooled with random effects model at 95% CI from head-to-head randomized trials)....57
Figure 4. Physical therapists' involvement in administering modalities from trials that examined the effectiveness of physical therapy interventions....60
Figure 5. Physical therapists' involvement in trials that examined the effectiveness of physical therapy interventions 61

Appendixes

Appendix A. Literature Search Strings
Appendix B. Excluded Studies
Appendix C. Ongoing Studies of Physical Therapy Interventions
Appendix D. Review Questions According to PICOTS Framework
Appendix E. Data Abstraction Forms
Appendix F. Evidence Tables and Figures

Executive Summary

Background

Osteoarthritis (OA), the most common form of arthritis,[1] is a progressive joint disorder caused by gradual loss of cartilage.[2] Osteoarthritis of the knee afflicts 28 percent of adults over age 45[3] and 37 percent of adults over age 65 in the United States.[4] OA may disproportionately affect African Americans and women.[3-6] It is a leading cause of disability among non-institutionalized adults,[4] and its prevalence, health impact, and economic consequences will increase dramatically during the next few decades as the population ages.[7]

OA treatments aim to reduce or control pain, improve physical function, prevent disability, and enhance quality of life.[8] Conservative treatment options include pain relievers, anti-inflammatory drugs, weight loss, general physical exercise, and physical therapy.[9, 10] Optimal OA management combines pharmacologic treatments with physical therapy interventions[9-12] and, when conservative treatments fail, surgery.[9, 10] Surgical treatments for knee OA include realignment osteotomy and knee replacements.[13] In the United States, about 556,400 knee replacement surgeries are performed annually.[13] By 2030 that number is projected to increase 600 percent.[14]

Comprehensive, up-to-date guidelines are available from the Osteoarthritis Research Society International (OARSI), the American Academy of Orthopedic surgeons, and the National Institute for Health and Clinical Excellence. These guidelines recommend exercise (including local muscle strengthening and general aerobic fitness) as a core treatment for symptomatic osteoarthritis, irrespective of patient age, comorbidity, pain severity, or disability.[9, 10, 15] Effectiveness has not been clearly established for other nonpharmacologic physical therapy interventions as adjunct to core treatment (e.g. thermal, manipulation, electrical nerve stimulation, and orthotics).[9]

Patients' functional status, pain, and quality of life constitute patient-centered clinical outcomes.[10] Consumers use improvement in patient-centered outcomes to judge the success of physical therapy.[16, 17]Thus, some consensus exists that clinical trials for symptomatic knee OA should examine outcomes such as pain, physical function, patient global assessment, and joint imaging.[18] However, published studies inconsistently define improvement and treatment success.[19-22] In practice, physical therapists often use intermediate outcomes to evaluate treatment effectiveness, including instrumental measurements of gait, balance, and range of motion. Likewise, reimbursement is currently driven by functional outcomes, including gait, transfers, and activities of daily living. Yet, we are not certain whether these outcomes predict pain, disability, and quality of life.

This report synthesizes published evidence about effectiveness of physical therapy for pain secondary to knee OA in adults. We focused on community-dwelling adults in ambulatory care settings and on interventions applicable to physical therapy practice. Our systematic review is intended to help clinicians, consumers, and policymakers make informed decisions based on synthesized evidence and other relevant factors.

Input from Stakeholders

We developed research questions and an analytic framework after discussions with key informants. Research questions were posted for public comment. Key informants recommended that we focus on patient-centered outcomes and physical therapy interventions relevant for

clinical practice in the United States. Key informants also recommended that we review intermediate outcomes used by physical therapists to judge treatment success. Comprehensive information about mono- and combined interventions at different times of followup will help patients, clinicians, and health care providers make evidence-based decisions about physical therapy interventions for adults with pain secondary to knee OA.

Candidates to serve as key informants, technical experts, and peer reviewers were approved by the Task Order Officer from the Agency for Healthcare Research and Quality (AHRQ) after disclosure of conflicts of interest. We developed the protocol and analytical framework (Figure A) with input from experts who served on the Technical Expert Panel (TEP).

Objectives

For the topic of physical therapy interventions for adults with knee OA, we aimed to conduct (1) a comprehensive review of the literature about the association between intermediate and patient-centered outcomes; and (2) a comprehensive synthesis of evidence of the clinical efficacy and comparative effectiveness of the interventions. We followed the principles from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews from AHRQ (www.effectivehealthcare.ahrq.gov). We examined the following questions:

Key Question 1

What are the effectiveness and comparative effectiveness of available physical therapy interventions (without drug treatment) for adult patients with chronic knee pain due to OA on intermediate and patient-centered outcomes when compared to no active treatment or another active physical therapy modality?

a. Which patient characteristics are associated with the benefits of examined interventions of physical therapy on intermediate and patient-centered outcomes?
b. Do changes in intermediate and patient-centered outcomes differ by the dose, duration, intensity, and frequency of examined interventions of physical therapy?
c. Do changes in intermediate and patient-centered outcomes differ by the time of followup?

Key Question 2

What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after physical therapy interventions?

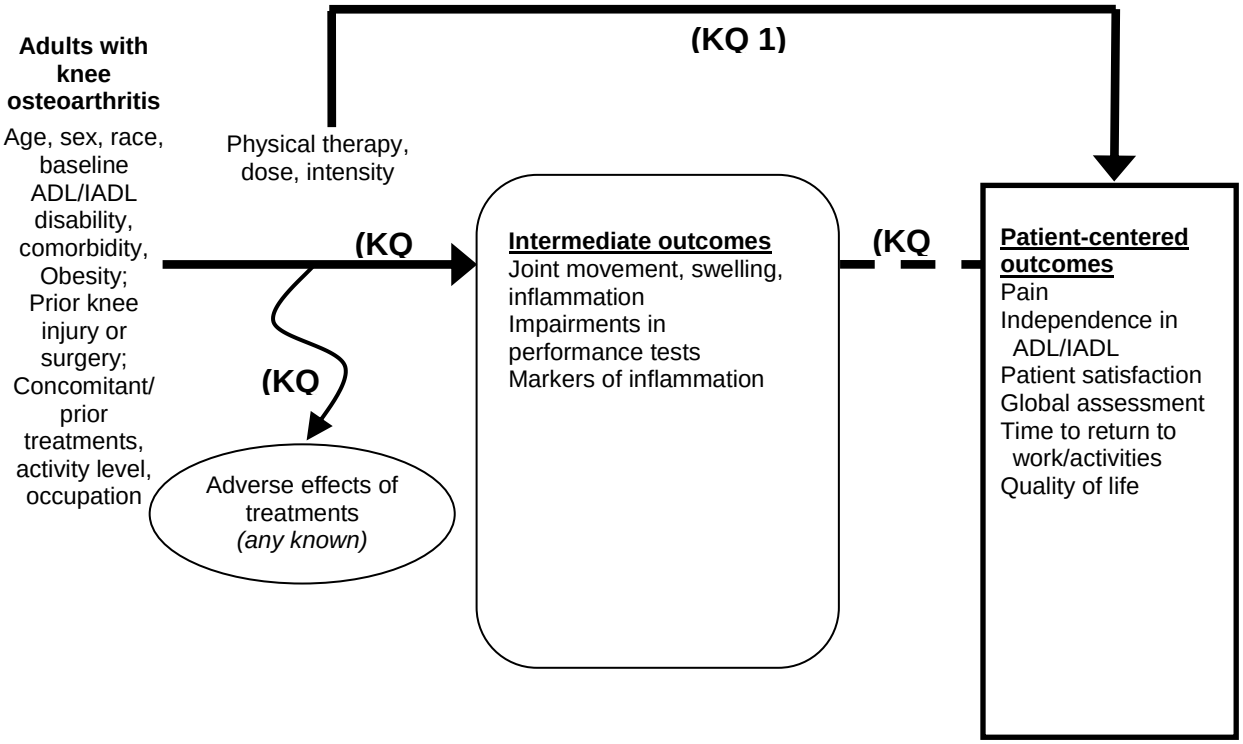
a. What is the validity of the tests and measures used to determine intermediate outcomes of physical therapy on OA in association with patient-centered outcomes?
b. Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes?
c. What are minimal clinically important differences of the tests and measures used to determine intermediate outcomes?

Key Question 3

What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls?

a. Which patient characteristics are associated with the harms of examined physical therapy interventions?
b. Do harms differ by the duration of the treatment and time of followup?

Figure A. Analytic Framework



Methods

Data Sources

We sought studies from a wide variety of sources, including MEDLINE® (via OVID and PubMed®), the Cochrane Library, the Physiotherapy Evidence Database (PEDro), SCIRUS, Allied and Complementary Medicine (AMED), and the Health and Psychosocial Instruments bibliography database up to September 1, 2011. We conducted manual searches of reference lists from systematic reviews and eligible studies. Grey literature search included regulatory documents, conducted clinical trials, and abstracts presented in scientific meetings.

Study Selection

At least two investigators independently evaluated each study for eligibility. We defined the target population, eligible independent and dependent variables, outcomes, time, and setting following the PICOTS framework developed in the protocol. We included original studies of adults with knee OA published in English after 1970. We defined and listed eligible interventions as those within the scope of physical therapy practice but not necessarily administered by physical therapists or physical therapist assistants in a given study. In clinical practice, physical therapists typically use multiple interventions for knee OA, often in combination with other nonphysical therapy interventions for knee OA. To minimize risk of bias and obtain valid estimates of treatment benefits and harms, we focused on randomized controlled trials (RCTs). While randomization may distribute the effects of other treatments equally, their impacts must still be taken into account. Moreover, some nonphysical therapy treatments, such

as pain relievers, may in part mask the benefits of physical therapy, especially for pain. We also reviewed observational studies with multivariate adjustment for concomitant treatments and confounding factors.[23] We applied the best-available evidence approach to include poor-quality observational studies when evidence was not available from RCTs or high-quality observational studies.[24] We included unpublished RCTs presented at peer-reviewed scientific meetings only if they provided data on clinical outcomes of eligible treatments not available in published articles.

We selected the interventions and methods to assess the outcomes in accordance with the Practice Pattern 4E: Impaired Joint Mobility, Motor Function, Muscle Performance, and Range of Motion Associated with Localized Inflammation from the Guide to Physical Therapist Practice.[25]

For Key Question 2 we included any observational studies that reported the association between intermediate and patient-centered outcomes.

We defined the target population as community-dwelling adults with knee pain secondary to knee OA. We excluded studies involving children, adolescents, hospitalized patients, or patients in long-term care facilities; studies that included patients with knee or hip OA that did not separately report the outcomes; and studies of surgical treatments or pharmacological treatments. We also excluded studies of adults with knee OA who had knee arthroplasty within 6 months before the study. For Key Question 2 we did not review validation of tests in populations with diseases other than knee OA.

We defined harms as a totality of all possible adverse consequences of an intervention.[26] We included published and unpublished evidence of adverse effects with eligible interventions, regardless of how authors perceived causality of treatments.[26] We did not contact the investigators of the primary studies.

Data Extraction

Four researchers used standardized forms to extract data. We conducted a double independent quality control for the data extracted from RCTs. We abstracted minimum data sets for therapeutic studies. For categorical variables, we abstracted a number of events among treatment groups. We abstracted means and standard deviations of continuous variables. For RCTs, we abstracted the number randomized to each treatment group. We abstracted the time when the outcomes were assessed as weeks from randomization and the time of followup after treatments. For observational studies, we extracted relative measures of the association (relative risk, hazard ratio, odds ratio) with standard error or 95 percent confidence interval (CI), and reported adjustments for patient characteristics. For diagnostic studies for Key Question 2, we abstracted the number of positive (true and false) and negative (true and false) after index diagnostic tests when compared to reference standard.

We abstracted baseline patient characteristics, including eligible and mean age; mean body mass index; proportion of women and minorities; proportion of disabled; proportions with severe knee OA comorbidities and multi-joint OA; baseline physical activity level; occupation; and concomitant drug and physical therapy interventions. We abstracted settings and physical therapist supervision of the treatments. We abstracted dose, length, and intensity of physical therapy interventions when reported by the authors.

Quality Assessment and Strength of Evidence

We evaluated the quality of studies and classified them by their designs. Using predetermined criteria from Cochrane tool of bias,[27] we evaluated the quality of therapeutic studies to assess risk

of bias that included randomization, adequacy of randomization and allocation concealment, masking of the treatment status, masking of outcome assessment, and intention-to-treat principles. We examined sponsorship and conflict of interest but did not downgrade quality using this information. Well-designed RCTs are believed to have low risk of bias. We defined studies as having medium risk of bias if they were susceptible to some bias, but it was not sufficient to invalidate the results (e.g., RCTs with unclear allocation concealment or without intention-to-treat principle). We defined studies as having high risk of bias if they had significant flaws that imply biases of various types that may invalidate the results, including nonrandom treatment allocation, no strategies to reduce bias, or ignoring randomization in analysis. We evaluated diagnostic studies for Key Question 2 using criteria from the Quality Assessment of Diagnostic Accuracy Studies (QUADAS).[28]

We estimated the applicability of the population by evaluating the selection of adults in observational studies and clinical trials. We also examined setting (including involvement of physical therapists or physical therapist assistants) and exclusion criteria for each study. We assessed strength of evidence for each major outcome according to risk of bias, consistency, directness, and precision.[29] When appropriate, we included dose-response association, presence of confounders that would diminish an observed effect, or strength of association.[29] We focused on direct evidence from head-to-head RCTs. We downgraded strength of evidence if risk of bias was moderate or high, if heterogeneity was significant, or if estimates were inconsistent or imprecise. We defined treatment effect estimates as precise when pooled estimates had reasonably narrow 95 percent CIs. Because single studies cannot give a definitive answer about treatment effects, we defined evidence from single studies as insufficient.[29]

Data Synthesis and Analysis

We have synthesized and presented the evidence according to the classification of physical therapy interventions from the Guide to Physical Therapist Practice by the American Physical Therapy Association.

For categorical variables we calculated rates, relative risk, and absolute risk differences. For continuous variables we calculated mean differences with 95 percent CI. Because pain was integral to the definition of our target population, we calculated mean differences at 99 percent confidence intervals for that outcome to have a greater confidence that the target population difference in pain is within a confidence interval. We also calculated ratios of means that describe percentage differences in pain with active versus control PT interventions.[30] We calculated estimates by applying intention-to-treat principles. If we found more than one study from a particular trial, the results from the latest published papers were used.

We examined and synthesized evidence of other nonsurgical treatments for knee OA if reported in the studies. We then compared effects of the examined physical therapy interventions across the studies according to reported concomitant drug treatments. We conducted sensitivity and subgroup analyses according to concomitant drug treatments when the available data were suitable for pooling.

Using a standard preplanned algorithm, we explored heterogeneity by clinical diversity.[31] Such characteristics included age, sex, race, baseline activities of daily living (ADL), instrumental activities of daily living (IADL), comorbidity, obesity, and significant skeletal abnormality. We explored heterogeneity by treatment dose (when applicable) and duration, as well as by whether the control treatment included education or exercise. We explored

heterogeneity by quality criteria of individual studies and by disclosed conflict of interest[31] rather than using a global quality score.[32, 33]

We focused on patient-centered outcomes including pain, disability, and quality of life.[34] We categorized intermediate outcomes as measurements of gait, strength, balance, transfers, endurance, joint function, or composite measure of functional performance. We reviewed validity and reliability of the tests within the scope of physical therapy practice. Evidence of the association between intermediate and patient-centered outcomes of physical therapy interventions was synthesized from observational studies that adjusted for treatments and confounding factors. We synthesized evidence from the studies that reported diagnostic values of intermediate outcomes to predict clinical outcomes. In a separate analysis, we synthesized the evidence from linear, logistic, or Cox regression models of the association between intermediate and clinical outcomes.

Using Meta-analyst[35] and STATA[36] software at a 95 percent CI, we calculated differences in relative risk and absolute risk from the abstracted events, and we calculated nonstandard mean differences in continuous variables from the reported means and standard deviations. We used correction coefficients, forced intention to treat, and calculations for missing data that were recommended by guidelines.[27] We defined magnitude of the effect using Cohen's criteria as: small, middle, and large effect corresponding to >0.2, >0.5, and >0.8 standardized mean differences.[37] Pooling criteria for Key Questions 1 and 3 required that interventions and outcomes be similarly defined. We calculated standardized mean differences for different measures of the same outcome with Cohen and Hodges methods. We categorized treatment effects from the studies by the clinical importance of differences in intermediate outcomes, using definitions of minimal clinically important differences from published studies and evidence-based reports.[38] We categorized the results from each tested hypothesis as nonsignificant differences in continuous outcomes, or significant with <20, 20-50, or >50 percent difference from control interventions.[39]

We tested consistency of the results by comparing the direction and strength of the association[29] and assessed heterogeneity of results using Chi square and I square tests.[40, 41] We also explored heterogeneity with meta-regression and sensitivity analysis. We reported the results from random effects models only[42] and chose the random effects model to incorporate in the pooled analysis differences across trials in patient populations, concomitant treatments, and definitions of interventions and outcomes.[31]

We qualitatively synthesized the evidence from poorly reported RCTs and observational studies. For studies that included knee and hip OA, we evaluated results but did not include them in pooled analyses because treatment effects may differ in patients with hip versus knee OA.

For Key Question 2, we summarized results of individual studies in evidence tables to analyze sensitivity, specificity, predictive values, diagnostic odds ratios, and predictive likelihood ratios, with a focus on the latter.[43, 44] Ratios of 1 indicated that the tests do not provide a likelihood of accurate diagnosis.[44] Ratios of more than 10 provided large and often conclusive increases in the likelihood of accurate diagnosis.[44] We tabulated each article for results of index diagnostic tests and reference standards. We evaluated validation and the proposed minimal important differences in total scores when this information was available. To judge validity from the studies that reported correlation coefficients between index and reference methods, we categorized correlation into the following categories: weak correlations as <20 percent, medium correlation as 20-50 percent, strong correlation as 50-75 percent, and very strong correlation as >75 percent.[37] To answer the question of which intermediate outcomes meet the criteria of

surrogates for patient-centered outcomes, we used Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) Criteria for Surrogate Endpoints.[39, 45] We examined whether randomized trials of physical therapy interventions evaluated the association between intermediate outcome change and patient-centered outcome change.[45]

Results

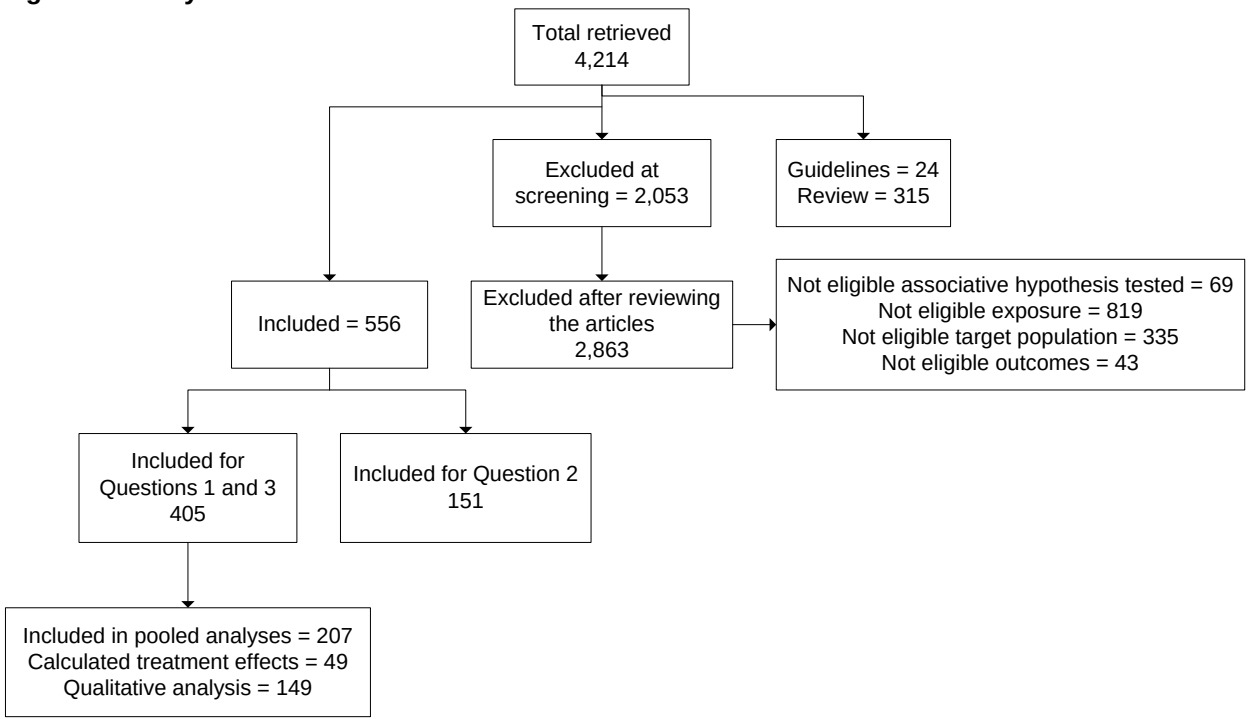
We identified and retrieved 4,214 references. We included 556 references for this review (Figure B). We excluded 3,319 articles. For Key Questions 1 and 3, we synthesized evidence from 405 articles. Overall, 256 studies contributed to synthesis of evidence and conclusions. We included into pooled analyses 207 studies and calculated treatment effects estimates without pooling in 49 studies. We examined evidence from but did not include in the synthesis an additional 149 studies.

Definitions of physical therapy interventions and outcomes varied dramatically among studies; thus, only a small proportion of comparisons met pooling criteria. We prioritized pooled analyses and results at longest time of followup over nonpooled results and short followups. Most studies lasted 4 to 6 weeks with a followup of 6 months.

Because the studies used different tools to measure the same outcomes, we used standardization in all pooled analyses. The studies examined continuous measures of the outcomes and rarely categorize the patients according to clinical importance of the changes.

Thirty-three percent of studies (85) had low risk of bias, 47 percent of studies (121) had medium risk of bias; and 16 percent of studies (42) had high risk of bias. We could not evaluate risk of bias in poorly reported studies (9 studies, 3.5 percent).

Figure B. Study flow



Key Question 1. Effectiveness of Physical Therapy Interventions

We found very few significant differences in outcomes between active and control treatments. Summary Tables A and B show how many studies examined each outcome, estimated effect sizes, and our level of confidence that the evidence reflects a true estimate of the treatment effect that is not likely to be changed by future research. No single physical therapy intervention improved all outcomes. Pooled analyses demonstrated that aerobic and aquatic exercise improved disability measures, aerobic and strengthening exercise reduced pain and improved function, proprioception exercises reduced pain, and ultrasound improved function (Table C). We observed no benefits from specific education programs, Tai Chi, diathermy, orthotics, or magnetic stimulation. Individual (nonpooled) RCTs failed to show consistent significant, strong, or clinically important changes in outcomes. Individual small RCTs may fail to show statistically significant effects due to low statistical power. Most trials had moderate risk of bias. We downgraded strength of evidence due to risk of bias or inconsistency in pooled estimates.

Education Programs

Results of three articles from two RCTs that examined the effects of specific education programs offered a moderate level of evidence suggesting no significant effect on pain relief.[46-48]

Proprioception Exercises

Results from four RCTs offered a low level of evidence that proprioception exercises led to pain relief but did not improve composite function or gait function.[49-52]

Aerobic Exercises

Twelve studies[46-48, 53-61]offered a low level of evidence that aerobic exercises (1) reduced long-term pain and disabilities compared to no aerobic exercise and (2) improved composite function, and gait function within 3 months of initiation.[47,54-56,58-60,62-66] At 12 months followup, the benefits of aerobic exercise continued for gait function[56, 62] but not for composite function.[47, 56, 59] Aerobic exercise resulted in clinically important improvement in pain, disability, and joint function in the majority of RCTs.

Aquatic Exercises

Three RCTs examined the effects of aquatic exercises. [63-65]Two studies showed that aquatic exercises reduced disability (low level of evidence) but had no significant effects on pain relief or quality of life (moderate level of evidence). [63, 64]

Strengthening Exercises

Nine studies offered a low level of evidence that strengthening exercises had no effect on disability.[49, 56, 64, 66-71] Analyzing somewhat overlapping studies, we observed a persistent improvement in pain relief, composite function, and gait function at 3 month through 12 months followup,[49, 50, 56, 64, 66-70, 72-82] but strengthening exercises did not improve health perception[67, 68] or quality of life[64, 68] (low level of evidence).

Tai Chi

A low level of evidence from three small trials[83-85] demonstrated no significant effects of Tai Chi on pain, disability, or sustained effects on composite function measures.

Massage

Three RCTs offered a low level of evidence that massage somewhat improved composite function[86, 87] but not joint function.[86, 88]

Joint Mobilization

Evidence from individual RCTs was insufficient to conclude the impact of joint mobilization.[89-91]

Orthotics

Seven RCTs demonstrated that orthotics had no effect on short-term outcomes of composite function (low level of evidence) or gait function (high level of evidence).[92-98] Evidence from three Japanese studies offered a low level of evidence that elastic subtalar strapping improved composite function around 3 months.[99-101]

Taping

Two RCTs examined the effects of taping, but different reporting formats precluded pooled analyses, which lowered the strength of evidence to insufficient. Pain, disability, composite function, and gait function did not differ with therapeutic taping.[102, 103]

Electrical Stimulation

Seven RCTs[76, 104-109] provided moderate evidence that electrical stimulation offered no improvement for disability or other functional measures. Pooled analyses offered a low level of

evidence that electrical stimulation significantly improved pain short-term and around 3 months after starting the intervention[75, 104-108, 110-114] but worsened pain at 6 months.[104, 108] Additionally, electrical nerve stimulation combined with exercises did not significantly improve pain. Global assessment[106, 107] and muscle strength (measured at 60 degree extension)[76, 105]improved significantly at 3 months with electrical stimulation.

Pulsed Electromagnetic Fields (PEMF)

Four RCTs provided a moderate level of evidence that PEMF neither reduced pain nor improved composite function.[115-118]

Ultrasound

Five RCTs[77, 78, 105, 119, 120] offered a low level of evidence that ultrasound reduced pain and improved function. Pain reduction was significant with a moderate effect size.[77, 78, 105] A low level of evidence also demonstrated that ultrasound had no effect on disability but significantly improved composite function and gait function with a large effect size.[77, 78, 105, 119, 120]

Diathermy

Five RCTs[105, 121-124] offered a low level of evidence that diathermy significantly decreased pain at 1 month followup, but the effect was insignificant at 3 months.[105, 122-125] A moderate level of evidence demonstrated no effect with diathermy on disability, composite function, joint function, or gait function.

Heat

Evidence from individual RCTs was insufficient to draw definitive conclusions about heat.[126 105]

Cryotherapy

Evidence for cryotherapy was limited to individual studies that showed no significant effects.[114, 126]

Comparative Effectiveness of Physical Therapy Interventions

Direct evidence of comparative effectiveness of physical therapy interventions from single studies was insufficient for the majority of comparisons (Tables D and E). Aerobic and aquatic exercises offered similar benefits for disability,[64, 127] a finding consistent with the similar effect sizes we observed for aerobic and aquatic exercises compared to control groups (Table F). Direct comparisons between aerobic and strengthening exercises on disability and composite function found no significant difference between the two.[56, 68]

We found no significant difference between laterally and neutrally wedged insoles on composite function,[101, 128] or between orthotics and brace on composite function.[129] Pain, disability, global assessment, quality of life, and joint function did not differ between laterally and neutrally wedged insoles.[130] In comparisons between electrical stimulation and exercises, several small studies found no significant difference on pain relief and gait function.[75, 76, 104, 131]

Key Question 1a. Role of Patient Characteristics on Outcomes

Evidence for the majority of subgroup analyses was insufficient for determining how patient characteristics affect outcomes.

Adherence

We found a moderate level of evidence from three RCTs that subgroups with high compliance tended to have better outcomes for exercise (aerobic, aquatic, and strengthening).[63, 132-135] The higher exercise compliance group had the lowest risk of incident ADL disability,[132] a lower average depression score,[133] a higher mean Quality of Well-Being Scale score,[63] and greater improvements in both 6-minute walking distance and disability.[135]

Age

Evidence was inconclusive for how age differences affect treatment outcomes, because studies were inconsistent with active and control treatments, outcomes, and definitions of age subgroups.[56, 116]

Malalignment

Two RCTs provided inconclusive evidence for how malalignment affects treatment outcomes. The RCTs found greater benefit in patients with the genu varus group[136] and in those without malalignment.[81]

Body Mass Index (BMI)

Evidence from two RCTs was inconsistent in assessing the role of BMI in predicting treatment effects.[56, 137] Improvement in function by lateral wedge insoles was better in adults of normal weight;[137] whereas very obese participants (defined by the top tertile) experienced similar benefits from aerobic exercise interventions and resistance training programs.[56]

Comorbidity

Evidence from individual studies was not sufficient to conclude differences in benefits with or without comorbidity.[133, 138]

Sex

Evidence from individual studies was not sufficient to conclude differences in benefits between men and women. Five studies reported clinical outcomes in male and female subgroups for exercise and orthotics[56, 93, 101, 139, 140] and demonstrated no significant differences in outcomes.

Race

Evidence from a single study was inconclusive for how racial differences modify treatment effects of exercise.[56]

Severity

Baseline OA severity may affect the impact of physical therapy interventions on clinical outcomes. However, findings were inconsistent and varied across studies depending on the treatments, outcomes, and/or cut-off grades. Furthermore, RCTs reported post-hoc analyses of changes from baseline in functional measures among patients with different baseline severity scores. Clinical outcomes in severity subgroups were reported in seven RCTs involving brace,[129, 136] insole,[96, 97, 129, 137] exercise (strengthening or ROM),[140] and weight reduction and/or electrical stimulation.[111] Three RCTs found no consistent modification effect of baseline severity.[96, 97, 137]

Key Question 1b. Association Between Dose/Duration/Intensity/ Frequency of Examined Interventions and Intermediate/Patient-Centered Outcomes

For the majority of comparisons, evidence was insufficient to make conclusions about the association between the dose/duration/intensity/frequency of examined interventions and outcomes.

Exercise

Studies inconsistently defined intensity; however, RCTs did indicate that low- and high-intensity exercises were equally beneficial.[79, 141] Patient outcomes demonstrated a positive dose response association with exercise frequency, with better benefits from 30-minute exercise sessions three times a week. [134] Home-based exercise therapy supervised by a visiting physical therapist offered similar benefits over 24 weeks with six or two visits weekly.[142]

Orthotics

For patients with genu varus deformity from OA, medium duration (5-10 hours each day) of insole with subtalar strapping wear was better than short duration (less than 5 hours) and long duration (more than 10 hours).[92] Two studies showed that a less elevated wedged insole seemed more comfortable than a more elevated wedge.[96, 143]

Electrical Stimulation

Evidence from individual RCTs was insufficient to draw valid conclusions about the effects of electrical stimulation.[144-147]

Ultrasound

Evidence from two RCTs showed that pulsed ultrasound was better than continuous ultrasound in improving disability, gait, and composite function measures.[77, 120]

Key Question 1c. Association Between Time of Followup and Intermediate/Patient-Centered Outcomes

The associations between followup times and outcomes varied by treatments and outcomes of interest. The effects of aerobic, aquatic, and strengthening exercises and ultrasound did not differ at shorter versus longer followup times. In combining aerobic, aquatic, strengthening, proprioception, and Tai Chi exercises, changes in intermediate and patient-centered outcomes did not differ by followup time (all p-values greater than 0.05). Results held consistent with or without inclusion of Tai Chi. Outcomes of pain, gait, and composite function after ultrasound did not differ by followup time, but evidence was insufficient for disability. Electrical stimulation improved pain at short-term followup but significantly worsened pain at longer followup times (p-value <0.001). In contrast, we observed that diathermy's benefits for disability increased with longer followup times (p-value = 0.009).

Association Between Duration of Examined Interventions and Intermediate/ Patient-Centered Outcomes

The duration of examined interventions varied broadly. For example, exercise programs ranged from 2 to 72 weeks. We found no significant association between the duration of examined interventions and intermediate/patient-centered outcomes. In combined results for aerobic, aquatic, strengthening, proprioception, and Tai Chi exercises, changes in intermediate

and patient-centered outcomes did not differ by the duration of examined interventions with all p-values greater than 0.05.

Key Question 2. Association Between Intermediate and Patient-Centered Outcomes

Evidence for the association between intermediate and clinical outcomes was limited to individual studies, and thus insufficient to draw valid conclusions. We found substantial variability in definitions of index and reference methods, definitions of outcomes, and methods of examining diagnostic values and associations between intermediate and clinical outcomes.

We synthesized the evidence of association between intermediate and clinical outcomes from 43 studies that included 25,799 adults with knee OA. Disability measures were associated with gait, mobility restrictions, muscle strength, and range of motion measures, but the magnitude and clinical importance of the association were unclear.

Key Question 2a. Validity of the Tests and Measures Used to Determine Intermediate Outcomes of Physical Therapy on OA in Association with Patient-Centered Outcomes

Validation of the tests and measures used to determine intermediate outcomes of physical therapy on OA was reported in 65 studies of 15,734 adults. The studies used a variety of reference methods to judge validity according to significant correlation coefficients. Only a small proportion of the studies demonstrated a strong (more than 50 percent) correlation between index and reference method measurements. Strength of correlation varied across validity types.

Key Question 2b. Which Intermediate Outcomes Meet the Criteria of Surrogates for Patient-Centered Outcomes?

None of the intermediate outcomes met surrogate criteria for patient-centered outcomes as defined by the OMERACT Criteria for Surrogate Endpoints.[39, 45] Gait was proposed by the TEP as a feasible candidate for a surrogate endpoint. However, no study analyzed the association between gait and patient-centered outcomes of physical therapy for adults with knee OA. One RCT did conclude that knee pain and self-efficacy mediated the effects of exercise on stair climb time.[148] A single longitudinal study of elderly adults demonstrated that impaired gait and the Physical Performance Test were independent predictors of nursing home placement.[149] Three cohort studies, including The Einstein Aging study,[150] Chinese Elderly Cohort,[151] and the Women's Health and Aging Study,[152] examined the association between gait and nursing home placement. However, the studies included adults with any etiology of gait problem, including neurological diseases or heart failure. Further, the definitions of impaired gait and magnitude of the association were inconsistent across the studies.

Key Question 2c. What are Minimal Clinically Important Differences of the Tests and Measures Used to Determine Intermediate Outcomes?

No RCTs of physical therapy interventions determined minimal clinically important differences (MCIDs). MCIDs in outcome measurements were, however, reported in 30 observational studies of 13,162 adults. The studies used the anchor method, which compares patient perception of improvement with absolute change in scale score or with percentage

difference from baseline levels. The percent difference from baseline levels incorporated baseline severity of the diseases. MCIDs were available for 26 tools.

Few studies determined Patient Acceptable Symptom State (PASS) for knee OA. PASS is defined as the highest level of symptom patients can tolerate and still be satisfied with their treatment. The studies used the same anchor method for determining PASS as they did for determining MCIDs. The difference is in anchoring questions. PASS involves asking patients whether they are satisfied with their functional status in relation to daily activities and quality of life. PASS was determined for three scales (WOMAC, VAS, and Patient Global Assessment).

Key Question 3. Harms from Physical Therapy Interventions Available for Adult Patients with Chronic Knee Pain Due to Osteoarthritis

Adverse events were uncommon and varied across interventions. Skin irritation was reported with brace, insole, tape, and electrical stimulation; swelling with brace, diathermy, and exercise; muscle soreness with electrical stimulation; or throbbing sensation with diathermy, electrical stimulation, and PEMF; increased pain with diathermy, exercise, insole, and PEMF; falls with insole; and need for surgery with diathermy. Rates of adverse events did not differ significantly among treatment groups. Adverse events were not severe enough to deter participants from continuing treatment.

Discussion

Our report analyzes patient-centered outcomes including pain, disability, and function with physical therapy interventions available in the United States for adults with knee OA. Our findings generally agree with previously published guidelines[10, 15] and systematic reviews[19, 21, 153] that recommend exercise for adults with symptomatic knee OA. Few physical therapy interventions demonstrated any significant effectiveness, and no single intervention improved all outcomes. Pooled analyses demonstrated that diathermy, Tai Chi, orthotics, and magnetic stimulation failed to show any benefits.

Our review was complicated by the discrepancy between the actual practice of physical therapy and the design of research evaluating these interventions. Published studies have focused on individual physical therapy interventions. In practice, however, physical therapists or physical therapist assistants most often treat their patients with a combination of modalities. In addition, clinical care for adults with knee OA includes pharmacological interventions,[154-156]while our review was limited to nonpharmacological treatments, further complicating our effort. We tried to examine how pain relievers (a common concomitant treatment) may influence physical therapy outcomes for pain, but rare and inconsistent reporting of drug treatments impeded synthesis of evidence. To address such complexity, our review focused on randomized trials since these equally distribute concomitant treatments among treatment groups and thus provided valid estimates of effects of the examined interventions. While randomization might equally distribute the effect of pain relievers, for example, it would not prevent the dampening of potential effects from physical therapy interventions. The trials we examined rarely provided information about all other treatments patients may have received. Nor did the trials analyze outcomes separately in patient subgroups by concomitant treatments. Finally, heterogeneity in populations, treatments, and definitions of the outcomes downgraded strength of evidence to low or moderate in most cases.

Low strength of evidence was in large part due to risk of bias: frequent exclusion of patients from the analyses, inadequate allocation concealment, and unmasked outcome assessment. Few

studies reported that researchers who assessed outcomes were unaware of the treatment status of the patients.[65, 82, 83, 157-159] We excluded several studies with poorly reported results and did not report evidence based on individual studies with high risk of bias. We also excluded trials with poor applicability for the target population of adults with symptomatic knee OA. Thus, we excluded trials that enrolled patients with knee or hip OA without reporting those outcomes separately. Many trials failed to provide sufficient details about the nature and intensity of specific interventions or about the involvement of the physical therapists themselves, further impeding our ability to make valid conclusions.[160, 161]

Variability in definitions and measurements of the outcomes also hampered synthesis of evidence. Validated measurements of functional impairments relevant to physical therapy practice are listed in the *Guide to Physical Therapist Practice*;[25] however, the guide recommends neither clinically important thresholds for such measures nor monitoring treatment effects according to patient-centered outcomes. Most trials reported outcomes as average scores for all patients in each treatment group with no evaluation of the clinical importance of the averages. Average scores do not reveal how many or which types of patients develop disability or experience clinically meaningful improvement in pain, function, or quality of life. Due to variability in definitions of the outcomes, we had to calculate standardized mean differences. Statistically significant differences in this construct do not necessarily reflect the clinical importance of improvement in outcomes. More analyses need to address predetermined subgroups. OARSI has recommended evaluating treatment success according to patient-centered outcomes and clinically important differences in the WOMAC scale.[39, 162] In addition, many studies have determined minimal clinically important differences of 26 validated tests using the anchor method, which compares changes in scales with patient perception of improvement.[163, 164] Yet, published studies of physical therapy interventions have not categorized patients according to meaningful improvements in pain, disability measures, or quality of life.

Benefits with treatments should be judged by patient satisfaction, not just measured improvement in scales or performance tests. PASS is gaining increasing recognition as a valid and reliable approach across many areas of medical practice, including rheumatology.[165] PASS is used to identify the level of symptom state patients can tolerate while still considering their health states satisfactory and their treatment successful. PASS is available for three scales (WOMAC, VAS for pain, and the Patient Global Assessment). Expanded use of PASS would help to improve the quality of physical therapy practice and to increase the impact of studies examining physical therapy interventions.

Our report has implications for future research. First, consensus is needed regarding methods to judge benefits of physical therapy interventions.[166] Benefits should be defined as clinically important improvements in pain, independence in ADL, and quality of life. Treatment success should be estimated using rates of patient-centered outcomes. Through meta-analysis of individual patient data from previously conducted RCTs, researchers would be able to categorize patients according to the clinical importance of any changes they experienced. They would also be able to analyze rates of patient-centered outcomes. This would require that principal investigators of RCTs be willing to share their data. Individual patient data meta-analyses may also provide good estimates of treatment effects in patient subpopulations by age, comorbidity, severity of knee OA, and concomitant treatments. Future RCTs should examine comparative effectiveness of combined physical therapy treatments. Fully powered trials should examine comprehensive and multi-modal interventions that more closely resemble physical therapy

practice. Future studies should also analyze the effects of concomitant treatments such as pain relievers on pain and function.

Key Messages

Key Question 1

- Effectiveness of physical therapy (PT) interventions
  - Few PT interventions were shown to be effective.
  - No single PT intervention was shown to improve all outcomes.
  - Research focused on individual PT interventions, in contrast with common PT practice of combining interventions.
  - Pooled analyses demonstrated that:
    - Aerobic and aquatic exercise improved disability measures.
    - Aerobic and strengthening exercise reduced pain and improved function.
    - Proprioception exercise reduced pain.
    - Ultrasound improved function.
  - Pooled analyses also demonstrated that the following PT interventions failed to show any benefits:
    - Specific education program
    - Tai Chi
    - Diathermy
    - Orthotics
    - Magnetic stimulation (PEMF)
  - Individual (non-pooled) RCTs failed to show consistent significant, strong, or clinically important changes in outcomes.
- Comparative effectiveness of PT interventions
  - Evidence about comparative effectiveness of PT interventions was limited.
  - Pooled analyses demonstrated that:
    - Disability measures did not differ between aerobic, aquatic, and strengthening exercises.
    - Pain did not differ between electrical stimulation and exercise in pooled analyses.
  - Individual RCTs of other treatment comparisons found no consistent clinically important differences in outcomes and provided insufficient evidence to conclude the best treatment option.
- Which patient characteristics are associated with the benefits of examined PT interventions on intermediate and patient-centered outcomes?
  - Evidence from individual RCTs was insufficient to conclude differences in effects by patient characteristics.
  - Patients with high compliance tended to have better treatment response with exercise interventions.
- Do changes in intermediate and patient-centered outcomes differ by the dose, duration, intensity, and frequency of examined PT interventions?
  - The duration of examined interventions was not associated with better intermediate or patient-centered outcomes.

    - Evidence regarding the association between the dose/intensity/frequency of examined interventions and outcomes was insufficient for the majority of comparisons.
- Do changes in intermediate and patient-centered outcomes differ by the time of followup?
    - The effects of the treatments that significantly improved outcomes, including exercise (aerobic, aquatic, and strengthening) and ultrasound, did not differ at shorter versus longer followup times.
    - Electrical stimulation improved pain short term but significantly worsened pain at longer followup.
    - Study quality and heterogeneity in populations and treatments including concomitant treatments hampered strength of evidence to low or moderate in most cases.

Key Question 2

- What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after PT interventions?
    - Gait, mobility restrictions, muscle strength, and range of motion measures were associated with disability measures.
    - Individual studies offered insufficient evidence for determining which intermediate outcomes strongly and consistently predict patient-centered outcomes.
- What is the validity of the tests and measures used to determine intermediate outcomes of PT on OA in association with patient-centered outcomes?
    - Validation was reported in many articles but few demonstrated a strong (more than 50 percent) correlation between index and reference method measurements.
    - Original studies concluded that tests are valid based on significance, not strength of correlation.
- Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes?
    - None of the intermediate outcomes met surrogate criteria for patient-centered outcomes.
- What are minimal clinically important differences (MCIDs) of the tests and measures used to determine intermediate outcomes?
    - MCIDs of the tests were determined using the anchor method, which compares changes in scales with patient perception of improvements. MCIDs were available as absolute change in score or as relative change as a percent difference from baseline levels, the latter accounting for baseline severity of the disease.
    - The definition of PASS that accounted for patient satisfaction was available for WOMAC, VAS for pain, and for the Patient Global Assessment Scale. PASS defines the highest level of symptom state patients can tolerate and still be satisfied with their treatment.
    - Validated tools defined threshold values of clinical importance for evaluating treatment success in adults with knee OA. However, more often studies used continuous measures of the outcomes providing an average score for all patients in each treatment group with no evaluation of the clinical importance of these averages. Average scores do not reveal how many or which patients develop disability or experience clinically meaningful improvement in pain, function, or quality of life.

Key Question 3

- What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls?
  - Adverse events were uncommon, varied across interventions, and included skin irritation with brace/insole/tape/electrical stimulation, swelling with brace/diathermy/exercise, muscle soreness with electrical stimulation, warming/throbbing sensation with diathermy/electrical stimulation/PEMF, increased pain with diathermy/exercise/insole/PEMF, and falls with insole. Adverse events were not severe enough to deter participants from continuing treatment.

Table A. Effectiveness of physical therapy intervention on patient-centered outcomes (standardized mean differences pooled with random effects models)

Physical therapy intervention	Pain # of studies Mean difference/ Strength of evidence	Disability # of studies Mean difference/ Strength of evidence	Psychological disability # of studies Mean difference/ Strength of evidence	Global assessment # of studies Mean difference/ Strength of evidence	Health perception # of studies Mean difference/ Strength of evidence	Quality of life # of studies Mean difference/ Strength of evidence
Education program	2 NS / M	1 -0.36 / I	1 NS / I		1 -0.29 /I	
Proprioception exercises	3 -0.72 / L					
Aerobic exercises	6 -0.21 / L	4 -0.26 / L	4 NS / L	1 NS / I	3 NS / L	1 NS / I
Aquatic exercises	2 NS / M	2 -0.26 / L	1 NS / I			2 NS / M
Strengthening exercises	3 -0.68 / L	2 NS / L	2 NS / L	1 OR 20* / I	2 NS / L	2 NS / L
Tai Chi	2 NS / L	2 NS / L	1 -0.79 / I			1 NS / I
Massage	1 -1.03 / I	1 C -16.59* / I	1 NS / I		1 C -9.82* / I	
Joint mobilization	1 NS / I	1 -2.7* / I		1 -3.1* / I		
Joint mobilization + exercise		1 -0.41 / I				
Orthotics	1 NS / I	1 NS / I		1 NS / I		
Elastic subtalar strapping						
Taping	1 NS / I	1 NS / I				
E-stim	2 0.6/L	2 NS / M		2 -0.43 / L	1 NS / I	
PEMF	2 NS / M	1 NS / I		1 OR 2.83* / I		
Ultrasound	3 -0.61 / L	2 NS / L				
Diathermy	3 NS / M	2 NS / M	1 NS / I	1 NS / I	1 NS / I	1 NS / I
Heat	1 NS / I	1 C -5.5* / I				1 C -0.41* / I
Cryotherapy	1 NS / I	1 NS / I				1 NS / I

Table B. Effectiveness of physical therapy intervention on intermediate outcomes

Physical therapy intervention	Composite function	Joint function	Gait function	Strength	Transfers
Education program	1 NS / I		1 NS / I	1 NS / I	
Proprioception exercises	3 NS / L	1 NS / I	3 NS / L	2 Improvement†	
Aerobic exercises	3 NS / L	1 NS / I	2 -0.56 / L	3 Improvement (1/8) / I	1 -0.52 / I
Aquatic exercises	1 NS / I			2 Inconsistent #	
Strengthening exercises	**4** **-1.18 / L**	1 -0.29* / I	**4** **-1.4 / L**	3 Improvement (17/20) / I	1 -0.44 / I
Tai Chi	2 NS / L	2 NS / L	1 NS / I		
Massage	2 -0.57 / L	2 NS / L	1 C -2.01* / I		
Joint mobilization			1 NS / I		
Joint mobilization + exercise			1 NS / I		
Orthotics	2 NS / L	1 NS / I	4 NS / H	2 NS / I	
Elastic subtalar strapping	3 -0.27 / L				
Taping	1 NS / I		1 NS / I	1 NS / I	
E-stim	3 NS / M	2 NS / M	3 NS / L	2 Improvement‡	1 NS / I
PEMF	2 NS / M	1 NS / I	1 NS / I	1 NS / I	
Ultrasound	**2** **-1.17 / L**	1 NS / I	**2** **-1.49 / L**	1 Improvement (8/24) / I	
Diathermy	3 NS / M	2 NS / M	3 NS / M	1 Improvement(4/12) / I	
Heat	1 NS / I	1 NS / I	1 NS / I	1 Improvement (7/12) / I	
Cryotherapy	1 NS / I			1 NS / I	

Results were presented as: Number of trials; Pooled with random effects model standardized mean difference; Level of evidence was determined according to four domains: risk of bias, consistency, directness, and precision.
NS: Nonsignificant result; OR: Odds ratio; C: Changes, as reported from the original study; *: nonstandardized estimate; ^: Only limited results could be pool-analyzed because measures of strength varied among studies, such as measured at different degrees of extension or flexion, concentric or eccentric contraction,

and right or left knee. Results were presented as numbers of measures with significant improvement over total numbers of measures unless otherwise indicated. †: Pooled analyses showed significant improvement of extensor and flexor at 60 degree, but insignificant improvement of both at 180 degree. #: Aquatic exercises significantly reduced muscle strength in one study but significant increase muscle strength in another. ‡: Pooled analyses showed significant improvement of extensor at 60 degree, but insignificant improvement at 120 degree. Negative values for mean differences or changes, and OR larger than 1 reflect improvements in knee symptoms compared with control groups. Bold: large magnitude of the effect (standardized mean difference >0.8).

I	Insufficient evidence results from individual RCTs
L	Low level of evidence
M	Moderate level of evidence
H	High level of evidence
Empty cells	No studies examined the association

Table C. Summary of effectiveness of physical therapy interventions for knee osteoarthritis

Physical therapy intervention	Studies/subjects	Conclusions/strength of evidence
Education program	Studies=3/Subjects=435	An education program improved disability (I) and health perception measures (I) but did not improve pain (M), gait (I) and composite measures of function (I)
Proprioception exercises	Studies=4/Subjects=225	Proprioception exercises improved pain (L) but did not improve joint (I), gait (L) and composite measures of function (L)
Aerobic exercises	Studies=12/ Subjects=1,256	Aerobic exercises improved pain (L), disability (L), gait (L), and transfer (L) measures of function but did not improve psychological disability (L), global assessment (I), health perception (L), quality of life (I), joint (I) and composite measures of function (L)
Aquatic exercises	Studies=3/Subjects=314	Aquatic exercises improved disability (L) but did not improve pain (M), psychological disability (I), quality of life (M), and composite measures of function (I)
Strengthening exercises	Studies=9/ Subjects=1,840	Strengthening exercises improved pain (L), global assessment (I), gait (L), joint (I), transfer (I), and composite (L) function measures but did not improve disability (L), psychological disability (L), health perception (L), and quality of life (L)
Tai Chi	Studies=3/Subjects=140	Tai Chi improved psychological disability (I) but did not improve pain (L), disability (L), gait (I), joint (L) and composite (L) function measures
Massage	Studies=3/Subjects=162	Massage improved pain (I), disability (I), health perception (I), gait (I) and composite (L) function measures but did not improve psychological disability (I) and joint (L) function measures
Joint mobilization	Studies=2/Subjects=119	Joint mobilization improved disability (I) and global assessment (I) but did not improve pain (I) and gait (I) function measures
Joint mobilization with exercise	Studies=1/Subjects=120	Joint mobilization with exercise improved disability (I) but did not improve gait (I) function measures
Orthotics	Studies=7/Subjects=510	Orthotics did not improve pain (I), disability (I), global assessment (I), joint (I), gait (H) and composite (L) function measures
Elastic subtalar strapping	Studies=3/Subjects=235	Elastic subtalar strapping improved composite function measures (L)
Taping	Studies=2/Subjects=94	Taping did not improve pain (I), disability (I), gait (I) and composite (I) function measures
E-stim	Studies=7/Subjects=345	E-stim improved global assessment (L), but worsened pain (L), and did not improve disability (M), health perception (I), and gait (L), joint (M), transfer (I), and composite (M) function measures,
PEMF	Studies=4/Subjects=263	PEMF improved global assessment (I) but did not improve pain (M), disability (I), and gait (I), joint (I) and composite (M) function measures
Ultrasound	Studies=5/Subjects=353	Ultrasounds improved pain (L), gait (L) and composite (L) function measures but did not improve disability (L), and joint function measures (I)
Diathermy	Studies=5/Subjects=342	Diathermy did not improve pain (M), disability (M), psychological disability (I), global assessment (I), health perception (I), quality of life (I), and joint (M), gait (M) and composite (M) function measures
Heat	Studies=3/Subjects=185	Heat improved disability (I) and quality of life (I), but did not improve pain (I), gait (I), joint (I), and composite (I) function measures
Mud pack	Studies=1/Subjects=60	Mud pack improved pain (I), disability (I) and global assessment measures (I)
Cryotherapy	Studies=2/Subjects=90	Cryotherapy did not improve pain (I), disability (I), quality of life (I), and composite function measures (I)

E-stim=Electrical stimulation; PEMF=Pulsed electromagnetic fields; strength of evidence as L: Low; M: Moderate; H: High; I: Insufficient; Level of evidence was determined according to four domains.

Table D. Comparative effectiveness of physical therapy intervention on patient-centered outcomes (standardized mean differences pooled with random effects models)

Active vs. control physical therapy intervention	Pain	Disability	Psychological disability	Global assessment	Health perception	Quality of life
Aerobic exercises vs. strengthening exercises	2 NS / L	2 NS / L				
Aquatic exercises vs. aerobic exercises	2 NS / L	1 NS / I				1 NS / I
Proprioception exercises vs. strengthening exercises	1 NS / I					
Tai Chi vs. stretching exercises	1 NS / I	1 C -0.96* / I	1 C -8.9* / I	1 NS / I		
Lateral wedged insole vs. neutrally wedged insole	2 NS / L	1 NS / I		1 NS / I		1 NS / I
Orthotics vs. brace	1 NS / I					
E-stim vs. exercises	2 NS / L					
E-stim vs. ultrasound	1 NS / I					

Table E. Comparative effectiveness of physical therapy intervention on intermediate outcomes

Active vs. control physical therapy intervention	Composite function	Joint function	Gait function	Strength^	Transfers
Aerobic exercises vs. strengthening exercises	2 NS / L		1 -0.51 / I	1 NS / I	1 NS / I
Aquatic exercises vs. aerobic exercises	1 NS / I		1 NS / I		
Proprioception exercises vs. strengthening exercises	1 0.5 / I		1 NS / I		
Tai Chi vs. stretching exercises	1 NS / I	1 NS / I	1 NS / I		1 C -5.98* / I
Lateral wedged insole vs. neutrally wedged insole	2 NS / M	1 NS / I	1 NS / I	1 NS / I	
Orthotics vs. brace	1 NS / I				
E-stim vs. exercises	1 -1.09 / I	1 -1.27 / I	2 NS / L	1 NS / I	
E-stim vs. ultrasound	1 NS / I		1 NS / I	1 NS / I	

Results were presented as: [Number of trials/N=number of patients] Pooled with random effects model standardized mean difference; Level of evidence was determined according to four domains: risk of bias, consistency, directness, and precision.
NS: Nonsignificant result; C: Changes, as reported from the original study; *: nonstandardized estimate; ^: Only limited results could be pool-analyzed because measures of strength varied among studies, such as measured at different degrees of extension or flexion, concentric or eccentric contraction, and right or left knee. Negative values for mean differences or changes reflect improvements in knee symptoms in the first intervention, compared with the second intervention group. Bold: large magnitude of the effect (standardized mean difference >0.8).

I	Insufficient evidence - results from individual RCTs
L	Low level of evidence
M	Moderate level of evidence
H	High level of evidence
Empty cells	No studies examined the association

Table F. Summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis

Active vs. control physical therapy intervention	Studies/subjects	Conclusions/strength of evidence
Aerobic exercises vs. strengthening exercises	studies=2/ subjects=345	Aerobic exercises improved gait function measures (I) but did not improve pain (L), disability (L), transfer (I), and composite (I) function measures, compared to strengthening exercises
Aquatic exercises vs. aerobic exercises	studies=2/ subjects=116	Aquatic exercises did not improve pain (L), disability (I), quality of life (I), gait (I) and composite (I) function measures, compared to aerobic exercises
Proprioception exercises vs. strengthening exercises	studies=1/ subjects=72	Proprioception exercises worsened composite function measures (I) but did not improve pain (I), gait function (I), compared to strengthening exercises
Tai Chi vs. stretching exercises	studies=1/ subjects=40	Tai Chi improved disability (I), psychological disability (I), and transfer function (I) but did not improve pain (I), global assessment (I), gait (I), joint (I), and composite (I) function measures, compared to stretching exercise
Laterally vs. neutrally wedged insole	studies=5/ subjects=734	Laterally wedged insole did not improve pain (L), disability (I), global assessment (I), quality of life (I), gait (I), joint (I), and composite function measures (M), compared to neutrally wedged insole
Orthotics vs. brace	studies=1/ subjects=91	Orthotics did not improve pain (I) and composite function measures (I), compared to brace
E-stim vs. exercises	studies= 3/ subjects=112	E-stim improved joint (I) and composite (I) measures of function but did not improve pain (L) and gait (I) function, compared to exercises
E-stim vs. ultrasound	studies=1/ subjects=35	E-stim did not improve pain (I), gait (I) and composite (I) function measures, compared to ultrasound

E-stim=Electrical stimulation; strength of evidence as L: Low; M: Moderate; I: Insufficient; Level of evidence was determined according to four domains.

Introduction

Osteoarthritis (OA), the most common form of arthritis,[1] is a progressive disorder in which gradual cartilage loss causes bony spurs and cysts to develop at the surface and margins of the joints. Inflammation, pain, stiffness, limited movement, and possible deformity of the joint may result.[2] OA of the knee afflicts 28 percent of adults over age 45[3] and 37 percent of adults over age 65 in the United States.[3-6]Already, OA is a leading cause of disability among noninstitutionalized adults;[4] those affected by it have slower gait velocities and use more assistive walking devices and nonsteroidal anti-inflammatory drugs and narcotics than those not affected. Further, the Centers for Disease Control and Prevention anticipates that the prevalence, health impact, and economic consequences of OA will surge during the next few decades as the population ages.[7]

Treatments for OA aim to reduce or control pain, improve physical function, prevent disability, and enhance quality of life—all of which constitute clinical outcomes of importance to patients.[8, 10] Treatment options include pain relievers, anti-inflammatory drugs, weight loss, general physical exercise, physical therapy, and, when conservative treatments fail, surgery.[9, 10]

Surgical treatments for knee OA include realignment osteotomy and total knee arthroplasty revisions (knee replacements).[13] In the United States, about 556,400 knee replacement surgeries take place annually,[13] a figure that increased by nearly 300 percent between 1990 and 2004.[7, 13, 14] By 2030, the annual number of total knee arthroplasty revisions in the United States is projected to increase 600 percent.[14]

The Osteoarthritis Research Society International (OARSI) asserts that optimal OA management combines nonpharmacologic and pharmacologic modalities.[9-12] However, scant evidence exists for the efficacy of adjunct therapies for knee OA other than exercise.[167, 168] Based on the findings of one systematic review,[10, 15] OARSI and the American Academy of Orthopaedic Surgeons recommend a variety of physical therapy interventions, including low-impact aerobic fitness exercises, range of motion/flexibility exercises, quadriceps strengthening, and patellar taping for short-term pain relief.[10, 15] OASRI and the Academy were unable to recommend for or against acupuncture as an adjunctive therapy for pain relief. The National Institute for Health and Clinical Excellence guidelines[9] agree that exercise (including local muscle strengthening and general aerobic fitness) should be a core treatment for osteoarthritis irrespective of patient age, comorbidity, pain severity, or disability. The National Institute for Health and Clinical Excellence also suggests other nonpharmacologic physical therapy interventions as adjunct to core treatment.[9]

Many systematic reviews, including three from the Cochrane Collaboration,[19, 21, 153] have synthesized data on physical therapy interventions for adult patients with knee OA. However, each published review examines a specific intervention instead of examining and comparing a range of available physical therapies. Meanwhile, many physical therapies for knee OA have yet to be evaluated. Most existing studies focus on exercise therapy; however, studies currently underway include physical therapy components such as insole treatment, knee bracing, wedged orthoses, manual therapy, weight loss, home-based exercises, strength training, knee stability training, electrical stimulation, and ultrasound. Publication of substantial new research evidence may alter the calculated risk-benefit ratio for some OA physical therapies and thus necessitate the updating of research evidence.[11, 169]

Measuring Outcomes of Physical Therapy Interventions

Measurement of physical therapy benefits should address patient-centered outcomes rather than the results of instrumental tests.[170] Additionally, clinicians and policymakers should consider patient-centered outcomes when making treatment and reimbursement decisions. Research based on patient-centered outcomes provides patients and clinicians the necessary information for effective and informed decisions about physical therapy and other health care services.[34]

Patient-centered outcomes for adults with pain secondary to knee OA include reduction in pain and improvement in functional disability and quality of life.[171] Some consensus exists that clinical trials for knee OA should examine pain, physical function, patient global assessment, and joint imaging.[18] However, published studies have inconsistently interpreted and defined improvement and treatment success.[19-22] Clinical trials have estimated the benefits and harms of physical therapy with validated scales of pain, function, and quality of life.[16, 17] Clinicians and researchers have used statistically significant changes in scale scores to define treatment success, without accounting for whether these score changes have clinical importance. Score changes that amount to benefits patients recognize as important—known as minimal clinically important differences, or MCIDs—have been determined by individual studies, but have not been systematically reviewed.

Further, studies of physical therapy interventions have examined intermediate outcomes. These outcomes have been defined as improvement in tests of balance, knee joint range of motion, gait speed, or muscle strength, as measured by a variety of assessment tools. Yet, validation of such measurements of functional impairments has not been systematically reviewed. In addition, clinical trials have concluded benefits with physical therapy interventions according to absolute changes in functional measurements while ignoring the clinical importance of such changes. Likewise, reimbursement for physical therapy services is currently driven by measurements of gait, transfers, and ADLs, irrespective of patient perception of improvement.

No systematic reviews or primary studies of physical therapy for adult patients with knee OA have specifically examined the relationship between changes in intermediate outcomes and meaningful changes in patient-centered outcomes such as disability in activities of daily living, quality of life, or lost work time. Quality of care could be improved by evaluating how clinical effects are measured and documented, as well as by reviewing outcomes information for research.

Our review intends to contribute to evidence-based recommendations by synthesizing published efficacy evidence for physical therapy for knee pain secondary to OA. We systematically reviewed studies that examined physical therapy interventions and that assessed intermediate and patient-centered outcomes.

The key questions used to guide this study are shown below.

Key Question 1

What are the effectiveness and comparative effectiveness of available physical therapy interventions (without drug treatment) for adult patients with chronic knee pain due to OA on intermediate and patient-centered outcomes when compared to no active treatment or another active physical therapy modality?

a. Which patient characteristics are associated with the benefits of examined interventions of physical therapy on intermediate and patient-centered outcomes?

b. Do changes in intermediate and patient-centered outcomes differ by the dose, duration, intensity, and frequency of examined interventions of physical therapy?
c. Do changes in intermediate and patient-centered outcomes differ by duration of examined interventions of physical therapy and the time of followup?

Key Question 2

What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after physical therapy interventions?

a. What is the validity of the tests and measures used to determine intermediate outcomes of physical therapy on OA in association with patient-centered outcomes?
b. Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes?
c. What are minimal clinically important differences of the tests and measures used to determine intermediate outcomes?

Key Question 3

What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls?

a. Which patient characteristics are associated with the harms of examined physical therapy interventions?
b. Do harms differ by the duration of the treatment and time of followup?

Methods

Search Strategy

We sought studies from a wide variety of sources, including MEDLINE® (via OVID and PubMed®), the Cochrane Library, the Physiotherapy Evidence Database (PEDro), SCIRUS, Allied and Complementary Medicine (AMED), and the Health and Psychosocial Instruments bibliography database up to September 1, 2011. We conducted manual searches of reference lists from systematic reviews and eligible studies.[172] We searched for unpublished literature including regulatory documents, conducted clinical trials, and abstracts presented in scientific meetings. We searched clinical trial registries including ClinicalTrials.gov and WHO Clinical Trials for completed trials related to the key questions. To find closed studies of physical therapy interventions for adults with knee OA, we searched the website www.clinicaltrials.gov in March 2011.

Our EPC search strategy used relevant medical subject headings (MeSH) terms, text words, and weighted word frequency algorithms to identify related articles. Members of our Technical Expert Panel (TEP) and peer reviewers suggested additional published studies. We documented each recommended, included, and excluded study. Our evidence search involved several steps: (1) conduct a comprehensive literature search in the databases listed above to retrieve references into the EndNote reference management software; (2) screen abstracts against pre-established inclusion/exclusion criteria; and (3) retrieve and review full articles on eligible studies to determine potential inclusion in the evidence synthesis.

The search strategies for the three research questions are described in Appendix A.

Excluded references are shown in Appendix B. Our analysis of the results from ongoing studies is presented in Appendix C.

Eligibility

At least two investigators evaluated each study for eligibility according to recommendations from the *Cochrane Handbook for Systematic Reviews of Interventions*.[27] We developed an algorithm to define study eligibility for each research question. We followed the guidelines to select evidence from controlled trials and observational studies.[23] We defined the target population, eligible independent and dependent variables, outcomes, time, and setting following the PICOTS framework (Appendix D). We formulated a list of eligible interventions following discussions with key informants and technical experts and after consideration of public comments. Eligible treatments were those within the scope of physical therapy practice but not necessarily administered by physical therapists or physical therapist assistants in a given study. In clinical practice, physical therapists or physical therapist assistants typically use multiple interventions for knee OA, often in combination with other nonphysical therapy treatments. To minimize risk of bias and to obtain valid estimates of benefits and harms, we focused on randomized controlled trials (RCTs). While randomization may distribute the effects of other treatments equally, their impacts still need to be taken into account. Moreover, some nonphysical therapy treatments such as pain relievers and nonsteroidal anti-inflammatory drugs may mask the benefits of physical therapy, especially for pain. Thus, we also reviewed observational studies with multivariate adjustment for concomitant treatments and confounding factors.[23]

We applied the best-available evidence approach to include poor quality observational studies when evidence was not available from RCTs or high-quality observational studies.[24, 173]

We reviewed the evidence of the efficacy and comparative effectiveness of physical therapy for knee pain secondary to OA. We classified interventions and methods to assess the outcomes according to the classifications in the practice pattern "Impaired Joint Mobility, Motor Function, Muscle Performance, and Range of Motion Associated with Localized Inflammation" from the *Guide to Physical Therapist Practice*.[25]

Inclusion Criteria

1. Original epidemiologic studies, including RCTs, nonrandomized multicenter clinical trials, and observational studies that used the strategies to reduce bias (adjustment, stratification, matching, propensity scores).
2. Publication in English after 1970.
3. Target population of community-dwelling adults with knee OA.
4. Eligible intermediate and patient-centered outcomes as listed in the analytic framework (Figure A in the Executive Summary).
5. Eligible interventions as listed in Table 1.

Table 1. Physical therapy interventions eligible for review

General modality	Specific intervention	Definition
Patient/client-related instruction		
Instruction, education, and training of patients/clients and caregivers	Current condition Enhancement of performance Health, wellness, and fitness Plan of care Risk factors for pathology/ pathophysiology, impairments, functional limitations, or disabilities	
Therapeutic exercise	Aerobic capacity/endurance conditioning or reconditioning	Increased workload over time Walking programs Aquatic therapy
	Flexibility exercises	Muscle lengthening Range of motion Stretching
	Gait and locomotion training	Gait training Implement and device training
	Strength, power, and endurance training for limb muscles	Active assistive, active, and resistive exercises Quadriceps strengthening Aquatic programs Standardized, programmatic, complementary exercise approaches Task-specific performance training Body mechanics and postural stabilization Body mechanics training
	Balance, coordination, and agility training	Neuromuscular education or re-education Posture awareness training
	Muscle relaxation technique for pain management	
Functional training in self-care, home management, work, community, and leisure integration or reintegration (including ADL, IADL, work	ADL training	
	Devices and equipment use and training	Assistive and adaptive device or equipment training during ADL and IADL Orthotic, protective, or supportive device or equipment training during ADL and

Table 1. Physical therapy interventions eligible for review (continued)

General modality	Specific intervention	Definition
hardening, and work conditioning)		IADL
	Functional training programs	Simulated environments and tasks Task adaptation
	IADL training	
	Injury prevention or reduction	Injury prevention education during self-care, home management, work, community, and leisure integration or reintegration Injury prevention or reduction with use of devices and equipment Safety awareness training during self-care, home management, work, community, and leisure integration and reintegration
Manual therapy techniques (Including mobilization/ manipulation)	Detailed examination to reveal impaired movements	
	Manual techniques with reinforcing exercise to improve movement	
	Manual traction	
	Massage	Connective tissue massage Therapeutic massage
	Mobilization/manipulation	Soft tissue Knee joint, other joints
	Passive range of motion	
Prescription, application of devices and equipment	Adaptive devices	Raised toilet seats
	Assistive devices	Canes Crutches Walkers Long-handled reachers Power devices Static and dynamic splints Grab bars and tub chairs
	Orthotic devices	Braces Shoe inserts Splints
	Protective devices	Braces Protective patellar taping
	Supportive devices	Supportive taping
Electrotherapeutic interventions	Electrical stimulation	Electrical muscle stimulation Functional electrical stimulation High-voltage pulsed current Neuromuscular electrical stimulation Transcutaneous electrical nerve stimulation
Physical agents and mechanical interventions	Nonthermal agents	Pulsed electromagnetic fields
	Aquatic therapy	Pools
	Sound agents	Ultrasound
	Thermotherapy	Dry heat Hot packs Diathermy Cold modalities
	Cryotherapy	Cold packs Ice massage

Abbreviations: ADL = activities of daily living; IADL = instrumental activities of daily living

For Key Question 2, we also included studies that examined the association between intermediate and patient-centered outcomes of physical therapy interventions, including any observational studies that reported this association.

We included observational studies when trial data were insufficient to estimate treatment benefits and harms.[24] We included unpublished RCTs presented at peer-reviewed scientific meetings only if they provided data on clinical outcomes of eligible treatments not otherwise available in published articles.

We included RCTs with subjects who had both knee and hip OA if outcomes for the two groups were reported separately. For Key Question 2, we included studies of tests and measures of functional outcomes in adults with knee OA.

Exclusion Criteria

1. Studies that involved children, adolescents, hospitalized patients, or patients in long-term care facilities.
2. Studies that included patients with knee or hip OA but did not separately report the outcomes.
3. Studies that involved surgical treatments or pharmacologic treatments for knee OA.
4. Studies of adults with knee OA who had knee arthroplasty within 6 months before the study.
5. Studies that validated tests and measures in populations with other diseases.
6. Studies that reported absolute values of the diagnostic tests in adults with knee OA.
7. Studies that did not test associative hypotheses or that did not provide adequate information on tested hypotheses (e.g., least square means, relative risk).
8. Case series when the evidence was available from RCTs or controlled observational studies.
9. Secondary data analyses, nonsystematic reviews, letters, or comments.

To assess harms of treatments we followed the recommendations from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews from the Agency for Healthcare Research and Quality (AHRQ) (www.effectivehealthcare.ahrq.gov)[26] and included published and unpublished evidence of the adverse effects of eligible interventions including:

- RCTs
- Published nonrandomized controlled trials
- Observational studies
- Observational studies based on patient registries or analyses of large databases
- Case reports and postmarketing surveillance

We defined harms as a totality of all possible adverse consequences of an intervention.[26] We analyzed harms regardless of how authors perceived causality of treatments.

We did not contact the investigators of the primary studies. The Scientific Resource Center requested Scientific Information Packets from appropriate manufacturers per usual procedures.

Quality Assessment

We evaluated the quality of studies according to recommendations from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews from AHRQ

(www.effectivehealthcare.ahrq.gov)[173] and the Cochrane Handbook for Systematic Reviews of Interventions.[27]

We classified the studies by design to distinguish randomized and nonrandomized controlled clinical trials from observational studies. Then we abstracted predefined criteria for critical appraisal of quality. We evaluated risk of bias with criteria of internal validity. We evaluated applicability with criteria of external validity. For interventional studies, we used criteria from the *Methods Guide for Comparative Effectiveness Reviews*[173] and from the Cochrane tool of bias.[27]

Quality criteria for therapeutic studies included randomization, adequacy of randomization and allocation concealment, masking of the treatment status, masking of the outcomes assessment, and intention-to-treat principles.[27] We evaluated disclosure of conflict of interest by the authors of individual studies and funding sources but did not use this information to downgrade quality of individual studies. We did not downgrade methodological quality of poorly reported studies; however, we synthesized evidence from these studies separately.

For observational studies, we evaluated strategies to reduce bias in study design and analysis, including adjustment for confounding and valid outcome measurements. For diagnostic studies, we applied the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria.[28, 174]

Rather than ranking overall quality or using a global score, we incorporated quality of studies into synthesis of evidence using individual quality criteria.[33, 175]

We defined well-designed RCTs with adequate allocation concealment, intention to treat principles in analysis, and adequate randomization as studies with low risk of bias.

We defined studies as having medium risk of bias if they were susceptible to some bias, but it was not sufficient to invalidate the results (e.g., open-label RCTs, RCTs with unclear allocation concealment, or crossover RCTs without assessment of carryover effect).

We defined studies as having high risk of bias if they had significant flaws that imply biases of various types that may invalidate the results, including nonrandom treatment allocation, no strategies to reduce bias, or ignoring randomization in analysis.

Grading the Evidence for Each Key Question

We assessed strength of evidence by following the guidelines from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews from AHRQ.[29] We judged the strength of evidence for each major outcome according to risk of bias, consistency, directness, and precision.[29] We focused on direct evidence from head-to head RCTs. We defined treatment effect estimates as precise when pooled estimates had reasonably narrow 95 percent confidence intervals (CIs). When appropriate, we included dose-response association, presence of confounders that would diminish an observed effect, or strength of association.

We defined high level of evidence on the basis of consistent findings from well-designed RCTs that reflected true effects of the treatments; these are findings for which future research would be very unlikely to change the estimate of effect. We assigned a moderate level of evidence when RCTs with medium risk of bias reported consistent treatment effects or large observational studies reported consistent associations. We assigned a low level of evidence to data from RCTs with serious flaws in design/analysis, and from post hoc subgroup analysis; these are findings for which further research is likely to change the estimate. We defined insufficient evidence when a single study examined treatment effects or associations.

We graded the quality of evidence for primary outcomes across studies as illustrated below.

Grade	Definition
High	**High confidence that the evidence reflects the true effect.** Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	**Moderate confidence that the evidence reflects the true effect.** Further research may change our confidence in the estimate of effect and may change the estimate.
Low	**Low confidence that the evidence reflects the true effect.** Further research is likely to change the confidence in the estimate of effect and is likely to change the estimate.
Insufficient	Evidence either is unavailable or does not permit a conclusion.

Applicability

We estimated applicability of the population by evaluating subject selection in observational studies and clinical trials.[176] Studies of community-dwelling adults with knee OA recruited from the general population had high applicability. Large observational cohorts based on national registries, population-based effectiveness trials, and nationally representative administrative and clinical databases had higher applicability, as did studies of interventions conducted by physical therapists and studies with followup times of 3 months, 6 months, or longer.

Data Extraction

Four researchers independently evaluated the studies and extracted the data. We conducted a double independent quality control for the data extracted from RCTs. We assessed errors in data extractions by comparing established ranges for each variable and the data charts with the original articles. Any detected discrepancies were discussed. We abstracted information relevant to the PICOT framework for each question. We abstracted minimum datasets to reproduce the results presented by the authors. For categorical variables we abstracted a number of events among treatment groups. For continuous variables we abstracted means and standard deviations.

For RCTs, we abstracted the number randomized to each treatment group as the denominator and calculated estimates by applying intention-to-treat principles. We abstracted the time when the outcomes were assessed as weeks from randomization and the time of followup after treatments. For observational studies, we extracted relative measures of the association (relative risk, hazard ratio, odds ratio) with standard error or 95 percent CI and reported adjustments for patient characteristics. For diagnostic studies, we abstracted the number of positive (true and false) and negative (true and false) after index diagnostic tests when compared to gold standard.

Data abstraction forms are shown in Appendix E. We abstracted sponsorship of the studies, sponsor participation in study design and in analysis and presentation of data, and conflict of interest by the authors. We abstracted baseline patient characteristics, including eligible and mean age; mean body mass index; proportion of women and minorities, disabled subjects, and subjects with severe knee OA, comorbidities, and/or multi-joint OA; baseline physical activity level; occupation; and concomitant drug and physical therapy interventions. We abstracted settings and supervision of treatments by physical therapists. We abstracted dose, length, and intensity of interventions when reported by the authors.

Data Synthesis

We synthesized and presented the evidence according to the classification of physical therapy interventions from the Guide to Physical Therapist Practice by the American Physical Therapy Association. We summarized the results into evidence tables (Appendix F).

For categorical variables we calculated rates, relative risk, and absolute risk differences. For continuous variables we calculated mean differences with 95 percent CI. For pain, we calculated mean differences at 99 percent CI. We also calculated ratios of means that describe percentage differences in pain with active versus control interventions.[30] If we found more than one study from a particular trial, the results from the latest published papers were used.

We addressed the role of concomitant treatments in association with patient outcomes and synthesized the evidence according to other nonsurgical treatments for knee OA reported in the studies. We then compared effects of the examined physical therapy interventions across the studies according to reported concomitant drug treatments. We conducted sensitivity and subgroup analyses according to concomitant drug treatments when the available data were suitable for pooling. Using a standard preplanned algorithm, we explored heterogeneity by clinical diversity[31] (age, sex, race, baseline ADL, IADL, comorbidity, and obesity). We explored heterogeneity by dose (when applicable), by duration of the treatments, and by whether the control treatment included education or exercise.

Rather than using the global quality score, we explored heterogeneity by quality criteria of individual studies and by the disclosed conflict of interest.[31-33]

Following guidelines and recommendations from key informants and TEP members, we focused on patient-centered outcomes including pain, disability, and quality of life.[34] We categorized intermediate outcomes as measurements relevant to the practice of physical therapy such as gait, strength, balance, transfers, endurance, joint function, and composite measure of function. We synthesized evidence depending on measurements of the outcomes with validated scales. We reviewed validity and reliability of the tests within the scope of physical therapy practice. Evidence of the association between intermediate and patient-centered outcomes was synthesized from observational studies that adjusted for treatments and confounding factors. We synthesized evidence from the studies that reported diagnostic values of intermediate outcomes to predict clinical outcomes. In a separate analysis, we synthesized the evidence of the association between intermediate and clinical outcomes from linear, logistic, and Cox regression models.

We calculated differences in relative risk and absolute risk from the abstracted events using Meta-analyst[35] and STATA[36] software at a 95 percent CI. We used correction coefficients and forced intention to treat to estimate treatment effects among all randomized patients irrespective of the authors' exclusion of subjects from the analyses.[27] We calculated nonstandard mean differences in continuous variables from the reported means and standard deviations by using Meta-analyst[35] and STATA[36] software at a 95 percent CI. We defined magnitude of the effect using Cohen's criteria of small, medium, and large effect corresponding to >0.2, >0.5, and >0.8 standardized mean differences.[37] We analyzed the adjusted regression coefficients with a standard error of association between intermediate and patient-centered outcomes.

Pooling criteria for Key Questions 1 and 3 included that definitions of physical therapy interventions and outcomes be the same. We calculated standardized mean differences for different measures of the same outcome with Cohen and Hodges methods. We categorized treatment effects by the clinical importance of differences in intermediate outcomes. We used definitions of minimal clinically important differences from published studies and evidence-

based reports.[38] We categorized the results from each tested hypothesis as either nonsignificant differences in continuous outcomes, or as significant with <20, 20-50, or >50 percent differences from the control interventions.[39]

We tested consistency in the results by comparing the direction and strength of the association[29] and used Chi square and I square tests to assess heterogeneity in study results.[40, 41] We explored heterogeneity with meta-regression and sensitivity analysis and reported the results from random effects models only.[42] We chose the random effects model because it incorporates in the pooled analysis differences across trials in patient populations, baseline rates of the outcomes, and definitions of interventions and outcomes.[31]

We assumed the presence of publication bias and did not use statistical tests for bias (defined as the tendency to publish positive results).[177-181] We used several strategies to reduce bias, including a comprehensive literature search of published and unpublished evidence in several databases, reference lists of systematic reviews, proceedings of scientific meetings, contacts with experts for additional references, and agreement on the eligibility status by several investigators. We examined publication rates among registered in ClinicalTrials.gov studies that examined physical therapy interventions in adults with knee osteoarthritis.

The numbers needed to treat to achieve one event of patient-centered outcome were calculated as reciprocals of the absolute risk differences in rates of outcome events in the active and control groups.[36, 182] The number of avoided or excess events (respectively) per population of 1,000 is the difference between the two event rates multiplied by 1,000. We calculated means and 95 percent CI for treatment events per 1,000 treated, multiplying pooled absolute risk difference by 1,000.[183]

For Key Question 2, we summarized results of individual studies in evidence tables to analyze sensitivity, specificity, predictive values, diagnostic odds ratios, and predictive likelihood ratios. We focused on the latter.[43, 44] Ratios of 1 indicated that the tests do not provide likelihood of accurate diagnosis.[184] Ratios of more than 10 provided large and often conclusive increases in the likelihood of accurate diagnosis.[184] We tabulated each article for results of index diagnostic tests and reference standards. We evaluated validation and the proposed minimal clinically important differences in total scores when this information was available. To judge validity from the studies that reported correlation coefficients between index and reference methods, we categorized correlation into the following categories: weak correlation as <20 percent, medium correlation as 20-50 percent, strong correlation as 50-75 percent, and very strong correlation as >75 percent.[37] To answer the question of which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes, we used Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) Criteria for Surrogate Endpoints.[39, 45] We examined whether randomized trials evaluated the association between intermediate and patient-centered outcomes.[45]

Results

We identified and retrieved 4,214references (Figure B in Executive Summary). We excluded 3,319 references (Appendix B), leaving 556 references for this review. For Key Questions 1 and 3, we synthesized evidence from 405 articles. Overall 256 studies contributed to synthesis of evidence and conclusions. We included into pooled analyses 207 studies, calculated treatment effects estimates without pooling in 49 studies. We examined evidence from but did not include an additional 149 studies in the synthesis.

Detailed evidence tables with all included studies are available by request from the authors. Eligible references included published results from individual studies, abstracts presented in scientific meetings, and FDA statistical reviews of several studies (Appendix Table F1). Our search of www.clinicaltrials.gov for completed studies identified 18 publications of 69 relevant studies (26 percent publication rate) (Appendix Table F2). We received no response from manufacturers of physical therapy equipment companies in response to our requests for scientific information packages (Appendix Table F3).

Key Question 1. What are the effectiveness and comparative effectiveness of available physical therapy interventions (without drug treatment) for adult patients with chronic knee pain due to OA on intermediate and patient-centered outcomes when compared to no active treatment or another active physical therapy modality?

We identified 405 studies that reported intermediate or patient-centered outcomes in adults with knee OA. We used in pooled analyses evidence from 207 RCTs. We acknowledged the findings of, but excluded from our synthesis, 149 studies that provided insufficient details about outcomes (Appendix Table F4). We evaluated quality and treatment effects from but ultimately excluded RCTs of adults with knee or hip OA that failed to report those outcomes separately (Appendix Table F5).

Eligible for synthesis of evidence, 256 studies enrolled 27,864 adults. Most, but not all, reported consent of the subjects and ethical approval (Appendix Table F6). Almost half of the studies (100, 39 percent) reported no funding source; 17 (7 percent) were sponsored by industry, and 61 (24 percent) by government. The studies recruited an average of 109 (standard deviation 136) adults (Appendix Table F7).

Regarding risk of bias, 32 percent of studies (82) had low risk, 47 percent (121) had medium risk; and the 18 percent (45) had high risk. We could not evaluate risk of bias in poorly reported studies (Appendix Table F6). Most frequently, high risk of bias was due to exclusion of patients from the analyses and differences among treatment groups at baseline (inadequate randomization) (Appendix Table F8). Among RCTs included in pooled analyses, 64 RCTs (31 percent) had low, 110 (53 percent) medium, 25 (12 percent) high, and eight (4 percent) unclear risk of bias.

We concluded that overall the studies had good applicability to our target population because they primarily recruited older adults with knee OA. On average, women constituted 73±18 percent of the participants. Body mass index (BMI) of participants averaged around 29±3kg/m$^{2.}$ (Appendix Table F7). Most studies did not report race of participants (Appendix Table F9). Adults in 114 studies (44 percent) were taking anti-inflammatory drugs or pain relievers. Half of the studies provided no information about exact pharmacological treatments. Most studies did

not report participants' occupation, knee injury, comorbidity, duration of condition, or proportion of subjects with baseline disability. Few studies explicitly stated that they excluded patients with prior knee surgery (Appendix Table F9). Most studies did not report the proportion of patients who had undergone surgery.

For two reasons, we concluded that the studies overall had poor applicability to the actual practice of physical therapy. First, most examined an isolated intervention, which is inconsistent with physical therapy's standard of practice, which is combination therapy. Second, many of the interventions studied were physical agents/modalities (i.e., orthotics, ultrasound, taping, etc.). Again, this contradicts the practice of physical therapy, in which physical agents/modalities are used infrequently and never in isolation, but rather in combination with other more "active" interventions (i.e., exercises).

Among the studies for pooled analyses, aerobic exercise was the most frequently studied treatment (19 studies of 5,681adults), followed by strength exercise (141 studies of 2,916 adults), aquatic exercise (3 studies of 1,278 adults), and electrical stimulation (20 studies of 1,155 adults) (Appendix Table F10). Outcomes were categorized by comparisons, domains of outcomes, and followup times. Definitions of physical therapy interventions and outcomes varied dramatically among studies; thus, only a small proportion of comparisons met pooling criteria. We prioritized pooled analyses and results at longest time of followup over nonpooled results and short followups (Appendix Table F11).

Effectiveness of Physical Therapy Interventions

We found very few significant differences in outcomes between active and control treatments. The summary table (Table A in the Executive Summary) shows how many studies examined each outcome, estimated effect sizes, and our level of confidence that the evidence reflects the true effect of the treatment and that the estimate is unlikely to be changed by future research (Appendix Table F12). No single physical therapy intervention improved all outcomes (Table 2). Individual small RCTs may fail to show statistically significant effects due to low statistical power.

Pooled analyses demonstrated that aerobic and aquatic exercise improved disability measures; aerobic and strengthening exercise reduced pain and improved function; proprioception exercises reduced pain; and ultrasound improved function (Table 3). Specific education programs, Tai Chi, diathermy, orthotics, and magnetic stimulation failed to show any benefits. Individual (nonpooled) RCTs failed to show consistent significant, strong, or clinically important changes in outcomes (Appendix Table F13).

Education Programs

Results of three articles from two RCTs offered a low level of evidence that specific education programs had no significant effect on pain relief,[46-48] either at 3 months[46, 54, 111, 185] or at longer followups.[46-48] One study suggested that dietary education improved disability and health perception;[48] however, evidence from a single RCT was insufficient for valid conclusions.

Proprioception Exercises

Results from four RCTs offered a low level of evidence that proprioception exercises led to pain relief but did not improve composite function or gait function.[49-52] Magnitude of the effect varied across the studies with significant heterogeneity in pooled estimates. Sensitivity analysis

restricted to two studies with low risk of bias revealed a larger effect size.[49, 50] One study suggested that proprioception exercises improved knee reposition error. [50]

Aerobic Exercises

Twelve studies[46-48, 53-61] offered a low level of evidence that aerobic exercises (1) reduced long-term pain and disabilities compared to no aerobic exercise and (2) improved composite function and gait function within 3 months of initiation.[46, 53-55, 57-59, 186-190]

Both pain and disability outcomes, either measured at 1 month, 3 months, or 12 months, improved significantly with aerobic exercise, although no significant improvement occurred at 6 months.[46, 48, 53-59, 61, 186-191] In contrast, within 3 months, but not at 6 months or with longer followups, aerobic exercise improved composite function.[46, 47, 55-59, 68, 157, 188, 189, 191] The benefits of aerobic exercise on gait function started at 1 month[61, 190, 192] and persisted to 3 months,[53, 55, 61, 157, 188, 190-193] 6 months,[47, 188, 190, 191] and more than 6 months.[56, 62] Aerobic exercise significantly improved the long-term outcome of transfer function.[56, 188]

Magnitude of the effect varied across the studies with significant heterogeneity in pooled estimates. Meta-regression exploring heterogeneity in disability around 3 months after aerobic exercise compared to placebo indicated that larger proportion of female participants and those with larger BMI had significantly better outcomes (Appendix Table F14). We found similar results favoring female participants for pain relief around 3 months, but could find no significant association between pain relief and BMI (Appendix Table F15). Subgroup analyses revealed that the involvement of a physical therapist might explain different effect sizes of gait function around 3 months (Appendix Table F16 –F18).

Aquatic Exercises

Three RCTs examined the effects of aquatic exercises.[63-65] Two studies showed that aquatic exercise reduced disability (low level of evidence) but had no significant effects on pain relief or quality of life (moderate level of evidence).[63, 64]

Strengthening Exercises

Nine studies offered a low level of evidence that strengthening exercises had no significant effect on disability.[49, 56, 64, 66-71] Analyzing somewhat overlapping studies, we observed a persistent improvement in pain relief, composite function, and gait function at 3 month through more than 12 months followup,[49, 50, 56, 66-70, 74-76, 79, 81, 82] but strengthening exercises did not improve health perception[67, 68] or quality of life[64, 68] (low level of evidence). Individual studies also found strengthening exercises improved joint function,[71] transfer function,[56] and global assessment.[70]

Magnitude of the effect varied across the studies with significant heterogeneity in pooled estimates. Meta-regression exploring heterogeneity in pain relief around 3 months after strengthening exercise compared to placebo indicated that younger participants had significantly better outcomes (Appendix Table F19). In meta-regression exploring heterogeneity in gait function around 3 months after strengthening exercise, we found that exercise frequency of two or three times per week improved gait function significantly better than a frequency of five times per week (Appendix Table F20). We explored heterogeneity by the involvement of a physical therapist and study quality and found no consistent association with outcomes (Appendix Tables F21-F25).

Tai Chi

A low level of evidence from three small trials[83-85] demonstrated no significant effects of Tai Chi on pain, disability, or composite function measures. Evidence from a single RCT was insufficient to draw valid conclusions about Tai Chi's affect on psychological outcomes.[83]

Massage

Three RCTs offered a low level of evidence that massage somewhat improved composite function[86, 87] but not joint function.[86, 88] Individual studies showed that massage also improved pain,[86] disability,[88] health perception,[86] and gait function;[88] however, this evidence was insufficient for definitive conclusions.

Joint Mobilization

Evidence was insufficient to conclude the impact of joint mobilization. Individual studies showed that joint mobilization with or without exercise reduced disability.[89 91] However, joint mobilization with or without exercise did not improve gait function.[89, 90]

Orthotics

Seven RCTs demonstrated that short-term outcomes of composite function (low level of evidence) and gait function (high level of evidence) did not differ with orthotics.[92-98] A low level of evidence from three Japanese studies showed that elastic subtalar strapping improved composite function around 3 months.[99-101]

Taping

Two RCTs examined the effects of taping, but different reporting formats precluded pooled analyses, which lowered the strength of evidence to insufficient. The two studies showed that pain, disability, composite function, and gait function did not differ with therapeutic taping.[102, 103]

Electrical Stimulation

Seven RCTs[76, 104-109] offered a low level of evidence that electrical stimulation significantly improved pain short-term and around 3 months after starting the intervention[75, 104-108, 110-114] but worsened pain at 6 months.[104, 108] Global assessment[106, 107] and muscle strength (measured at 60 degree extension)[76, 105] around 3 months improved significantly with electrical stimulation (low level of evidence). These significant findings were consistent without substantial heterogeneity across the studies. Pooled analyses provided moderate evidence of no improvement on disability or other functional measures.[76, 104-107, 147, 194-196]

Pulsed Electromagnetic Fields (PEMF)

Four RCTs offered a moderate level of evidence that PEMF neither reduced pain nor improved composite function.[115-118] One study showed that PEMF significantly increased subjective success,[118] but the evidence was insufficient for definitive conclusions.

Ultrasound

Five RCTs offered a low level of evidence that ultrasound reduced pain and improved function.[77, 78, 105, 119, 120] Pain reduction was significant with a moderate effect size.[77, 78, 105] A low level of evidence also demonstrated that ultrasound did not improve disability but did significantly improve composite function and gait function with a large effect size.[77, 78, 105, 119, 120]

Magnitude of the effect on gait function at 3 months varied across the studies with significant heterogeneity in pooled estimates. We were unable to examine heterogeneity due to the small number of studies (N=3) and inadequate variation of the preplanned criteria.

Diathermy

Five RCTs[105, 121-124] offered a low level of evidence that diathermy significantly decreased pain at 1 month followup, but the effect was insignificant at 3 months.[105, 122-125] A moderate level of evidence demonstrated that diathermy had no affect on disability, composite function, joint function, or gait function. A single study also demonstrated no beneficial effects on psychological disability, global assessment, or health perception.[121, 124]

Heat

Evidence was insufficient to draw valid conclusions. In one study heat improved disability and quality of life[126] but had no effect on composite function and gait function.[105]

Cryotherapy

Evidence for cryotherapy was limited. Individual studies showed no significant effects.[114, 126]

Clinical Importance of Treatment Effects with Physical Therapy Interventions

Few individual RCTs categorized patients by clinical importance of the changes in measured pain, disability, or joint function and failed to demonstrate consistent improvement with physical therapy interventions (Appendix Table F26). Rates of patient-rated treatment success were greater with a brace,[197] electrical stimulation,[106, 198] mud pack,[199] and PEMF.[118] Evidence from individual RCTs were insufficient for definitive conclusions about clinically important improvement with physical therapy interventions.

As a part of planned synthesis of evidence, we compared the differences in continuous measures of pain and disability reported in trials with the minimal clinically important differences determined in observational studies (Appendix Table F13) and found few clinically important improvements. Aerobic exercise resulted in clinically important improvement in pain, disability, and joint function in the majority of individual RCTs (Appendix Table F27).

Comparative Effectiveness of Physical Therapy Interventions

Evidence was insufficient for the majority of comparisons (Tables E-F in the Executive Summary and Table 4 below). Aerobic and aquatic exercises had the same benefits on disability,[64, 127] a finding consistent with the similar effect sizes demonstrated by these two interventions in efficacy studies. Direct comparisons showed no significant differences between aerobic and strengthening exercises on disability and composite function.[56, 68]One study found aerobic exercise was better than strengthening exercise in gait function.[56] One study demonstrated that Tai Chi was better than stretching exercises for disability, psychological disability, global assessment, and transfer function.[200]

Laterally and neutrally wedged insoles demonstrated similar effects on gait function,[101, 128] as did orthotics and brace on composite function.[129] A recent study showed that pain, disability, global assessment, quality of life, and joint function did not differ between laterally and neutrally wedged insoles.[130] Several small studies found no significant difference between electrical

stimulation and exercise for pain relief and gait function.[75, 76, 104, 131] One study showed insignificant differences between electrical stimulation and ultrasound on composite and gait function.[105]

Key Question 1a. Role of patient characteristics on outcomes

Adherence

We found a moderate level of evidence from three RCTs that subgroups with high compliance tended to have better outcomes for exercise (aerobic, aquatic, and strengthening).[63, 132-135] The higher compliance group had the lowest risk of incident ADL disability,[132] a lower average depression score,[133] a higher mean Quality of Well-Being Scale score,[63] and greater improvements in both 6-minute walking distance and disability.[135]

Three articles came from the Fitness Arthritis and Seniors Trial (FAST), which investigated the effects of two exercise programs (aerobic and strengthening) for adults 60 and older.[132-134] The authors examined dose-response effects between exercise frequency and three outcome variables: knee pain, self-reported difficulties with ADL, and ADL performance.[134] The results indicated that exercise for patients with knee OA should be three times each week with moderate duration (35 minutes). The authors defined exercise compliance for both types of exercise by the percentage attendance at exercise sessions and found the lowest risk of ADL disability and a lower average depression score for those in the highest compliance tertile (Figure 1).[132, 133] One study performed an economic evaluation of aquatic exercise for persons with osteoarthritis; the mean Quality of Well-Being Scale (QWB) score for adherers was significantly higher than for nonadherers or controls.[63] Using multiple linear regression models among an Arthritis, Diet, and Activity Promotion Trial subsample, the authors found that higher exercise adherence was associated with greater improvements in 6-minute walking distance and in disability.[135]

Age

Evidence was insufficient to determine how age differences affect treatment outcomes. Three studies reported clinical outcomes by age subgroup for bracing, exercise (aerobic or strengthening), or PEMF.[56, 116, 136] Heterogeneity across studies (different active and control treatments, outcomes, and definitions of age subgroups) led to inconclusive evidence.

In 117 knee OA patients, explorative subgroup analyses showed that patients younger than 60 experienced a slightly better effect of the brace for knee function (measured by an estimated improvement of 3.38 on the Hospital for Special Surgery or HSS score) than patients 60 years and older (estimated HSS score improvement 2.48).[136] Pain severity with bracing showed a similar modest trend. Using a cut-off age of 70, the FAST trial found that younger and older participants randomized in aerobic or strengthening exercise programs improved in self-reported disability, pain, and 6-minute walk distance compared with the health education group.

While PEMF demonstrated no beneficial symptomatic effect in all patients, those younger than 65 improved significantly after two weeks in stiffness[116] but not in ADL or pain.

Alignment

Two RCTs provided inconclusive evidence for how malalignment affects treatment outcomes. RCTs found greater benefits in patients with the *genu varus* group[136] and in those without malalignment.[81]

Stratified by the alignment, the *genu varus* group (n=95) showed a better and significant effect of the brace for the knee function score (estimated HSS score improvement 4.15; P=0.03) compared to the *genu valgus* group (n=22) (estimated HSS score improvement 0.20; P=0.96).[136] This trend was similar but not as prominent for pain relief.

One study examined the impact of malalignment on the way strengthening exercise affects knee adduction moment, pain, and function.[81] The results indicated that strengthening exercise did not significantly alter the knee adduction moment or function in either the more malaligned or the more neutral group, but the latter experienced significant pain reduction.

Body Mass Index

Two RCTs offered inconsistent evidence for the role of BMI in predicting treatment effects.[56, 137] One study compared treatment with and without a lateral wedge insole and found that those with a BMI of less than 30 kg/m^2 had a 29-point improvement in the WOMAC Pain subscale, compared with an improvement of only 6-points in those whose BMI was more than 30 kg/m^2.[137] In contrast, the very obese participants (defined by the top tertile) who were assigned to the aerobic exercise or resistance training programs improved in self-reported disability, pain, and 6-minute walk distance compared with the health education group.

Comorbidity

Individual studies offered insufficient evidence for determining any modifying effect of comorbidity. The FAST study of 439 older adults with knee OA investigated the effects of comorbidity on the benefits of resistance or aerobic exercise,[138] and the results indicated that comorbidity had no significant impact.

Depression

Individual studies offered insufficient evidence for determining differences in benefits between patients with and without depression. The FAST study investigated the effects of depression on the benefits of exercise.[133] Aerobic or resistance exercise significantly improved disability, pain, and walking speed irrespective of baseline depressive symptoms. In addition, aerobic (but not resistance) exercise significantly lowered depressive symptoms at 128 months of followup compared to the control educational group. The authors concluded that depression had no substantial impact on the benefits of exercise.

Sex

Individual studies offered insufficient evidence for determining differences in benefits between men and women. Five studies that reported clinical outcomes of exercise and orthotics in male and female subgroups[56, 93, 101, 139, 140] demonstrated no significant difference in effects.

Race

A single study offered insufficient evidence for determining differences in benefits between racial groups. One study that performed subgroup analysis between whites and African Americans offered insufficient evidence for prediction of treatment effects by race.[56] Both African Americans and whites assigned to the aerobic exercise interventions or the resistance training program improved in self-reported disability, pain, and 6-minute walk distance compared with the health education group.

Severity

Baseline OA severity may affect the impact of physical therapy interventions on clinical outcomes. However, findings were inconsistent and varied across studies depending on the treatments, outcomes, and/or cut-off grades. Furthermore, RCTs reported *post hoc* analyses of changes from baseline in functional measures among patients with different baseline severity scores. Six RCTs used the Kellgren and Lawrence (K/L) scale to grade severity, and one categorized severity according to the Ahlback score.[136] Clinical outcomes in severity subgroups were reported in seven RCTs involving brace,[129, 136] insole,[96, 97, 129, 137] exercise (strengthening or ROM),[140] and weight reduction and/or electrical stimulation.[111] Three RCTs found no consistent modification effect of baseline severity.[96, 97, 137] In one RCT, pain reduction was greater in patients with severe OA than in those with mild.[136]

These inconsistent findings may be due to inconsistent outcomes and/or cutoff grades in the original studies. In a group of 221 older adults randomized to strength training or range of motion exercises, the WOMAC Pain subscale did not differ between K/L grade 2–3 and grade 0-1.[140] Interestingly, percentage of joint space narrowing >0.5-mm after a 30-month followup was higher in the strength training arm than in the range of motion arm in subgroup of K/L grade 0-1, but not grade 2-3. This finding is unexplained and further confirmation is warranted.

One study evaluated the effect of weight reduction and/or electrical stimulation on patients with knee osteoarthritis and obesity.[111] The study found that subgroups with severity grades 3 and 4 had more pain decrease than the group with grade 2, irrespective of whether the treatment arms received weight reduction, electrical stimulation, or weight reduction plus electrical stimulation. However, the study did not examine whether baseline severity modified benefits with manual therapy. Further, patients whose OA severity varied also had different baseline scores in the VAS pain scale.

Key Question 1b. Association between the dose/intensity/frequency of examined interventions and intermediate/patient-centered outcomes

For the majority of comparisons, we found insufficient evidence for determining any association between the dose/intensity/frequency of examined interventions and outcomes.

Exercise

Although definitions of intensity differed among studies, prior research indicated similar benefits for low- and high-intensity exercise, defined by one study as 40 percent and 70 percent heart rate reserve, respectively.[141] In one study, low- and high-intensity exercises similarly improved function, gait, and pain.[141] Another study found that the effects of high-resistance strength training (>60 percent of one repetition maximum) appear larger than those of low-resistance strength training (10 percent of one repetition maximum), but the differences were statistically insignificant.[79] One study examined exercise compliance in order to determine any dose-response effects between exercise frequency and outcomes.[134] The results indicated that exercise for patients with knee OA should be three times each week with moderate duration (35 minutes). One study compared frequency of physical therapist visits for patients receiving home-based exercise over 24 weeks, and found no significant difference between groups who received six visits or two visits.[142]

Orthotics

For patients with *genu varus* deformity from OA, medium duration (5-10 hours daily) of insole with subtalar strapping wear was better than short duration (less than 5 hours) and long duration (more than 10 hours).[92] One study suggested better comfort and effectiveness for an 8- or 12mm elevated wedged insole than a 16mm elevated wedge.[143] Another study also found better comfort for a 5° than a 10° wedge, although the effect of the 5° wedge was smaller.[96] Future researchers should consider these findings in designing their studies.

Electrical Stimulation

We found no short-term clinical difference between low frequency (2Hz pulse trains) and high frequency (80 Hz) electrical stimulation.[145] However, noxious stimulation decreased pain intensity more than innocuous stimulation.[146] In one study, Burst Mode and High Rate stimulation had similar effects on stiffness and pain.[147] Another study demonstrated that for reducing pain, 40 minutes was the optimal duration of electrical stimulation.[144]

Diathermy

Two studies found no significant differences between high- and low-intensity diathermy on disability and gait function.[123, 125]

Ultrasound

Pulsed ultrasound appeared to be better than continuous ultrasound in improving disability, gait, and composite function.[77, 120]

Key Question 1c. Association between the duration of examined interventions and intermediate/patient-centered outcomes

The duration of examined interventions varied broadly. For example, exercise programs ranged from 2 to 72 weeks. We found no significant association between the duration of examined interventions and intermediate or patient-centered outcomes. In combining aerobic, aquatic, strengthening, proprioception, and Tai Chi exercises, changes in intermediate and patient-centered outcomes did not differ by the duration of interventions, with all p-values greater than 0.05 (Appendix Figure F1). For this analysis we used the longest followup standardized effect size in each study. While these results might seem to suggest that a 2-week exercise program is sufficient, we emphasize that exercises should be continuous and that higher adherence to exercise led to better improvement. Evidence was insufficient for valid conclusions about other treatments.

Association Between Time of Followup and Intermediate/Patient-Centered Outcomes

The association between the time of followup and outcomes differed by examined treatments and outcomes. Outcomes did not differ by followup times for treatments that demonstrated significant benefits (aerobic, aquatic, and strengthening exercises and ultrasound). Nor did intermediate or patient-centered outcomes differ by followup time when the effects of aerobic, aquatic, strengthening, proprioception, and Tai Chi exercises were combined (all p-values greater than 0.05) (Appendix Figure F2). The combined results remained consistent with or without inclusion of Tai Chi. Ultrasound's effects did not differ by time of followup for pain, gait, and composite function, while evidence was insufficient for disability (Appendix Figure F3).

Electrical stimulation improved pain at short followups but significantly worsened pain at longer followups (p-value < 0.001) (Appendix Figure F4). In contrast, we observed that diathermy's benefits for disability increased with longer followups (p-value = 0.009) (Appendix Figure F5).

Table 2. Outcomes with physical therapy interventions (pooling with random effects models standardized mean differences)

Treatment	Outcome (sorted by importance of the outcomes) Weeks of followup	Randomized trials Subjects	Cohen Standard mean difference (95% CI) Hodges SMD not shown
Education program	Pain 6-13 weeks	Studies: 3[46, 54, 185] Subjects: 348	0.09 (-0.44; 0.61) I-squared=0.8, p=0.002
Education program	Pain >26 weeks	Studies: 2[46, 47] Subjects: 435	-0.11 (-0.33, 0.10) I-squared=0.226, p=0.275
Proprioception exercise	**Pain 6-13 weeks**	**Studies: 3[49-51] Subjects: 190 (254 knees)**	**-0.724 (-1.36, -0.135) I-squared=0.794, p=0.008**
Proprioception exercise	Function composite 6-13 weeks	Studies: 3[49-51] Subjects: 190	-1.68 (-2.659, 0.402) I-squared=0.955, p=0
Proprioception exercise	Gait function 6-13 weeks	Studies: 3[50-52] Subjects: 169	-0.958 (-1.992, 0.076) I-squared=0.895, p=0
Aerobic exercise	Pain <6 weeks	Studies: 2[61, 190] Subjects: 137	-1.00 (-2.25, 0.25) I-squared=0.926, p=0
Aerobic exercise	**Pain 6-13 weeks**	**Studies: 13[46, 53-55, 57-59, 61, 186-190] Subjects: 930**	**-0.35 (-0.59, -0.11) I-squared=0.722, p=0**
Aerobic exercise	**Pain 14-26 weeks**	**Studies: 7[47, 57, 59, 68, 188-190] Subjects: 947**	**-0.26 (-0.59, 0.08) I-squared=0.837, p=0**
Aerobic exercise	**Pain >26 weeks**	**Studies: 6[46, 47, 57-59](add ref: 8980206) Subjects: 1005**	**-0.21 (-0.35, -0.07) I-squared=0.205, p=0.273**
Aerobic exercise	**Disability <6 weeks**	**Studies: 2[61, 192] Subjects: 103**	**-1.73 (-3.37, -0.09) I-squared=0.898, p=0.002**
Aerobic exercise	**Disability 6-13 weeks**	**Studies: 9[53, 54, 57, 58, 61, 186, 187, 189, 192] Subjects: 727**	**-0.52 (-1.00, -0.03) I-squared=0.89, p=0**
Aerobic exercise	Disability 14-26 weeks	Studies: 3[57, 68, 189] Subjects: 310	-1.59 (-3.48, 0.31) I-squared=0.976, p=0
Aerobic exercise	**Disability >26 weeks**	**Studies: 4[48, 56-58] Subjects: 676**	**-0.259 (-0.411, -0.107) I-squared=0, p=0.486**
Aerobic exercise	Psychological disability 6-13 weeks	Studies: 4[53-55, 61] Subjects: 246	-0.69 (-1.51, 0.13) I-squared=0.87, p=0
Aerobic exercise	Health perception 6-13 weeks	Studies: 2[55 , 61] Subjects: 55	-1.43 (-3.16, 0.3) I-squared=0.872, p=0.005
Aerobic exercise	Health perception >26 weeks	Studies: 3[48, 57, 58] Subjects: 383	-0.093 (-0.294, 0.107) I-squared=0, p=0.362
Aerobic exercise	**Function composite 6-13 weeks**	**Studies: 3[55, 59, 157] Subjects: 312**	**-0.85 (-1.38; -0.33) I-squared=0.772, p=0.004**
Aerobic exercise	Function composite 14-26 weeks	Studies: 2[59, 68] Subjects: 242	-1.56 (-4.32; 1.20) I-squared=0.977, p=0
Aerobic exercise	Function composite >26 weeks	Studies: 3[47, 56, 59] Subjects: 791	-0.18 (-0.448, 0.087) I-squared=0.716, p=0.014
Aerobic exercise	**Gait function <6 weeks**	**Studies: 3[61, 190, 192] Subjects: 206**	**-0.37 (-0.62, -0.11) I-squared=0, p=0.599**
Aerobic exercise	**Gait function 6-13 weeks**	**Studies: 8[53, 55, 61, 157, 188, 190, 192, 193] Subjects: 582**	**-0.58 (-0.76, -0.4) I-squared=0.202, p=0.257**
Aerobic exercise	**Gait function 14-26 weeks**	**Studies: 3[47, 188, 190] Subjects: 453**	**-0.45 (-0.63; -0.27) I-squared=0, p=0.812**
Aerobic exercise	**Gait function >26 weeks**	**Studies: 2[56, 62] Subjects: 609**	**-0.558 (-0.862, -0.254) I-squared=0.7, p=0.036**
Aerobic exercise	**Transfer function >26 weeks**	**Studies: 2[56, 188] Subjects: 327**	**-0.405 (-0.799, -0.011) I-squared=0.379, p=0.044**

Table 2. Outcomes with physical therapy interventions (pooling with random effects models standardized mean differences) (continued)

Treatment	Outcome (sorted by importance of the outcomes) Weeks of followup	Randomized trials Subjects	Cohen Standard mean difference (95% CI) Hodges SMD not shown
Aquatic exercise	Pain 6-13 weeks	Studies: 2[64, 65] Subjects: 99	-0.25 (-0.646, 0.147) I-squared=0, p=0.376
Aquatic exercise	Pain 14-26 weeks	Studies: 2[63, 64] Subjects: 269	-0.176 (-0.416, 0.065) I-squared=0, p=0.349
Aquatic exercise	Disability 6-13 weeks	Studies: 2[64, 65] Subjects: 99	0.005 (-0.554, 0.563) I-squared=0.493, p=0.16
Aquatic exercise	**Disability 14-26 weeks**	**Studies: 2[63, 64] Subjects: 216**	**-0.257 (-0.495, -0.019) I-squared=0, p=0.353**
Aquatic exercise	QL 14-26 weeks	Studies: 2[63, 64] Subjects: 216	-0.097 (-0.334, 0.139) I-squared=0, p=0.952
Strengthening exercise	**Pain 6-13 weeks**	**Studies: 13[49, 50, 64, 66, 70, 72, 74, 79-82, 201] Subjects: 1317**	**-0.644 (-0.89, -0.398) I-squared=0.768, p=0**
Strengthening exercise	**Pain 14-26 weeks**	**Studies: 5[64, 66-69] Subjects: 616**	**-0.658 (-1.177, -0.139) I-squared=0.87, p=0**
Strengthening exercise	**Pain >26 weeks**	**Studies: 3[49, 56, 66] Subjects: 741 (822 knees)**	**-0.68 (-1.21, -0.15) I-squared=0.925, p=0**
Strengthening exercise	Disability 6-13 weeks	Studies: 4[64, 66, 72, 80] Subjects: 558	-0.14 (-0.60; 0.33) I-squared=0.794, p=0.002
Strengthening exercise	Disability 14-26 weeks	Studies: 4[64, 66-68] Subjects: 508	-0.42(-0.89, 0.06) I-squared=0.756, p=0.006
Strengthening exercise	Disability >26 weeks	Studies: 2[56, 66] Subjects: 660	-0.159 (-0.479, 0.161) I-squared=0.768, p=0.038
Strengthening exercise	Psychological disability 14-26 weeks	Studies: 2[67, 68] Subjects: 91	-0.857 (-4.07, 2.357) I-squared=0.977, p=0
Strengthening exercise	QL 6-13 weeks	Studies: 2[64, 80] Subjects: 173	-0.32 (-0.72, 0.08) I-squared=0.368, p=0.208
Strengthening exercise	QL 14-26 weeks	Studies: 2[64, 68] Subjects: 105	-0.43 (-1.011, 0.151) I-squared=0.551, p=0.136
Strengthening exercise	Health 14-26 weeks	Studies: 2[67, 68] Subjects: 91	-4.26 (-12.11, 3.58) I-squared=0.987, p=0
Strengthening exercise	**Function composite 6-13 weeks**	**Studies: 5[49, 50, 79, 81, 82] Subjects: 446**	**-0.895 (-1.203, -0.587) I-squared=0.641, p=0.007**
Strengthening exercise	**Function composite 14-26 weeks**	**Studies: 3[67-69] Subjects: 199**	**-1.003 (-1.955, -0.051) I-squared=0.918, p=0**
Strengthening exercise	**Function composite >26 weeks**	**Studies: 2[49, 56] Subjects: 376**	**-1.003 (-1.955, -0.051) I-squared=0.918, p=0**
Strengthening exercise	**Gait function 6-13 weeks**	**Studies: 9[50, 66, 70, 74, 76, 79, 81, 82, 201] Subjects: 912**	**-0.485 (-0.813, -0.157) I-squared=0.789, p=0**
Strengthening exercise	**Gait function 14-26 weeks**	**Studies: 2[66, 69] Subjects: 467**	**-0.465 (-0.846, -0.085) I-squared=0.666, p=0.05**
Strengthening exercise	**Gait function >26 weeks**	**Studies: 2[56, 66] Subjects: 660**	**-0.392 (-0.591, -0.193) I-squared=0.397, p=0.198**
Tai Chi	Pain 6-13 weeks	Studies: 2[83, 84] Subjects: 77	-0.442 (-0.96, 0.029) I-squared=0, p=0.953
Tai Chi	Disability 6-13 weeks	Studies: 2[83, 84] Subjects: 77	-0.263 (-0.731, 0.204) I-squared=0, p=0.509
Tai Chi	Disability 14-26 weeks	Studies: 2[84, 85] Subjects: 96	-0.285 (-0.968, 0.398) I-squared=0.601, p=0.113
Tai Chi	Function composite 6-13 weeks	Studies: 2[83, 84] Subjects: 77	-0.403 (-0.873, 0.067) I-squared=0, p=0.742
Tai Chi	Function joint 6-13 weeks	Studies: 2 Subjects: 77	-0.09 (-0.555, 0.376) I-squared=0, p=0.682

Table 2. Outcomes with physical therapy interventions (pooling with random effects models standardized mean differences) (continued)

Treatment	Outcome (sorted by importance of the outcomes) Weeks of followup	Randomized trials Subjects	Cohen Standard mean difference (95% CI) Hodges SMD not shown
Massage	Function composite 6-13 weeks	Studies: 2[86, 87] Subjects: 94	-0.566 (-0.946, -0.187) I-squared=0, p=0.703
Massage	Function joint 6-13 weeks	Studies: 2[86, 88] Subjects: 127	-0.243 (-1.255, 0.769) I-squared=0.884, p=0
Orthotics	Function composite <6 weeks	Studies: 2[92, 93] Subjects: 138	-0.583 (-1.191, 0.024) I-squared=0.75, p=0.06
Orthotics	Gait function <6 weeks	Studies: 4[94-97] Subjects: 262	-0.009 (-0.22, 0.203) I-squared=0, p=1
Elastic subtalar strapping	**Function composite 6-13 weeks**	**Studies: 3[99-101] Subjects: 235**	**-0.272 (-529, -0.014) I-squared=0, p=0.039**
E-stim	**Pain <6 weeks**	**Studies: 7[107, 110, 112, 113 114 147] Subjects: 289**	**-0.741 (-1.026, -0.456) I-squared=0.295, p=0.156**
E-stim	Pain 6-13 weeks	Studies: 7[104-108, 113] Subjects: 283	-0.094 (-0.331, 0.143) I-squared=0, p=0.814
E-stim	**Pain 14-26 weeks**	**Studies: 2[104, 108] Subjects: 72**	**0.597 (0.099, 1.095) I-squared=0.088, p=0.295**
E-stim	Disability 6-13 weeks	Studies: 2[106, 107] Subjects: 95	-0.276 (-0.695, 0.144) I-squared=0, p=0.961
E-stim	Global assessment 6-13 weeks	Studies: 2[106, 107] Subjects: 95	-0.43 (-0.862, -0.006) I-squared=0, p=0.373
E-stim	Function composite 6-13 weeks	Studies: 3[105-107] Subjects: 129	-0.09 (-0.45; 0.26) I-squared=0, p=0.634
E-stim	Function joint <6 weeks	Studies: 2[107, 147] Subjects: 97	-0.253 (-0.617, 0.111) I-squared=0, p=0.815
E-stim	Function joint 6-13 weeks	Studies: 2[106, 107] Subjects: 95	-0.295 (-0.714, 0. 125) I-squared=0, p=0.999
E-stim	Gait function <6 weeks	Studies: 4[76, 194-196] Subjects: 196	-0.21 (-0.70, 0.29)I-squared=0.648, p=0.014
E-stim	Gait function 6-13 weeks	Studies: 3[76, 104, 105] Subjects: 149	0.05 (-0.25; 0.36) I-squared=0, p=0.778
E-stim	Strength, 120 degree extension, 6-13 weeks	Studies: 2[104, 105] Subjects: 102	-0.41 (-0.84; 0.02) I-squared=0.157, p=0.305
E-stim	**Strength, 60 degree extension, 2 weeks**	**Studies: 2[76, 105] Subjects: 134**	**-0.55 (-0.90; -0.20) I-squared=0, p=0.434**
PEMF	Pain <6 weeks	Studies: 2; [115, 116] Subjects: 145	0.013 (-0.417, 0.442) I-squared=0.396, p=0.198
PEMF	Function composite <6 weeks	Studies: 2[115, 116] Subjects: 145	-0.127 (-0.607, 0.354) I-squared=0.513, p=0.152
Ultrasound	**Pain <6 weeks**	**Studies: 2[119, 120] Subjects: 149**	**-0.533 (-1.051, -0.014) I-squared=0.655, p=0.055**
Ultrasound	**Pain 6-13 weeks**	**Studies: 3[77, 78, 105] Subjects: 179 (323 knees)**	**-0.607 (-0.913, -0.3) I-squared=0.502, p=0.111**
Ultrasound	Disability <6 weeks	Studies: 2[119, 120] Subjects: 149	-0.393 (-0.799, 0.013) I-squared=0.45, p=0.162
Ultrasound	**Function composite 6-13 weeks**	**Studies: 3[77, 78, 105] Subjects: 179**	**-0.818 (-1.702, 0.066) I-squared=0.886, p=0**
Ultrasound	**Function composite >26 weeks**	**Studies: 2[77, 78] Subjects: 128**	**-1.169 (-1.628, -0.711) I-squared=0.411, p=0.183**
Ultrasound	Gait function <6 weeks	Studies: 2[119, 120] Subjects: 149	-0.54 (-1.337, 0.256) I-squared=0.85, p=0.001
Ultrasound	**Gait function 6-13 weeks**	**Studies: 3[77, 78, 105] Subjects: 179**	**-1.423 (-2.492, -0.353) I-squared=0.908, p=0**
Ultrasound	**Gait function >26 weeks**	**Studies: 2[77, 78] Subjects: 128**	**-1.491 (-2.086, -0.896) I-squared=0.614, p=0.075**

Table 2. Outcomes with physical therapy interventions (pooling with random effects models standardized mean differences) (continued)

Treatment	Outcome (sorted by importance of the outcomes) Weeks of followup	Randomized trials Subjects	Cohen Standard mean difference (95% CI) Hodges SMD not shown
Diathermy	Disability <6 weeks	Studies: 4[122-125] Subjects: 250	-0.21 (-0.447, 0.027) I-squared=0.085, p=0.364
Diathermy	Disability 6-13 weeks	Studies: 2[123, 124] Subjects: 143	-0.046 (-0.342, 0.251) I-squared=0, p=0.667
Diathermy	**Pain** **<6 weeks**	**Studies: 4[122-125]** **Subjects: 250**	**-0.537 (-0.973, -0.101)** **I-squared=0.701, p=0.003**
Diathermy	Pain 6-13 weeks	Studies: 3[105, 123, 124] Subjects: 178	-0.005 (-0.275, 0.265) I-squared=0, p=0.964
Diathermy	Function composite <6 weeks	Studies: 3[122-124] Subjects: 223	-0.468 (-0.952, 0.015) I-squared=0.742, p=0.004
Diathermy	Function composite 6-13 weeks	Studies: 3[123-125] Subjects: 178	0.007 (-0.263, 0.277) I-squared=0, p=0.998
Diathermy	Function joint <6 weeks	Studies: 2[123-125] Subjects: 143	0.197 (-0.104, 0.499) I-squared=0.026, p=0.358
Diathermy	Function joint 6-13 weeks	Studies: 2[123, 124] Subjects: 143	0.162 (-0.134, 0.459) I-squared=0, p=0.871
Diathermy	Gait function <6 weeks	Studies: 3[123-125] Subjects: 170	-0.098 (-0.368, 0.172) I-squared=0, p=0.953
Diathermy	Gait function 6-13 weeks	Studies: 3[105, 123, 124] Subjects: 178	-0.139 (-0.41, 0.132) I-squared=0, p=0.934
Aerobic exercise vs. strengthening exercise	Disability >26 weeks	Studies: 2[56, 68] Subjects: 345	-2.353 (-0.326, 0.097) I-squared=0.986, p=0
Aerobic exercise vs. strengthening exercise	Function composite >26 weeks	Studies: 2[56, 68] Subjects: 345	-0.115 (-6.94, 2.234) I-squared=0, p=0.497
Aquatic exercise vs. aerobic exercise	Disability 6-13 weeks	Studies: 2[64, 127] Subjects: 118	-0.106 (-0.522, 0.31) I-squared=0.24, p=0.251
Aquatic exercise vs. aerobic exercise	Disability 14-26 weeks	Studies: 2[64, 127] Subjects: 116	-0.206 (-0.714, 0.303) I-squared=0.477, p=0.167
Laterally vs. neutrally wedged insole	Function composite 6-13 weeks	Studies: 2[101, 128] Subjects: 363	-0 (-253, 0.253) I-squared=0.474, p=0.107
E-stim vs. exercise	Pain <6 weeks	Studies: 2[131] Subjects: 81	-1.298 (-2.992, 0.396) I-squared=0.913, p=0.001
E-stim vs. exercise	Gait function <6 weeks	Studies: 2[76, 131] Subjects: 81	0.198 (-1.181, 1.577) I-squared=0.888, p=0.003

QL=Quality of life; E-stim=Electrical stimulation; PEMF=Pulsed electromagnetic fields; Bold indicates significant changes

Table 3. Narrative evidence summary of effectiveness of physical therapy interventions for knee osteoarthritis

Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
Education program	Pain/studies=2, subjects=435	An education program did not improve pain measures/Moderate
	Disability/studies=1, subjects=278	An education program improved disability measures/Insufficient
	Psychological disability/studies= 1, subjects=278	An education program did not improve psychological disability measures/Insufficient
	Health perception/studies=1, subjects=278	An education program improved health perception measures/Insufficient
	Composite function/studies=1, subjects=316	An education program did not improve composite function measures/Insufficient
	Gait function/studies=1, subjects=316	An education program did not improve gait function measures/Insufficient
		An education program improved disability (I) and health perception measures (I) but did not improve pain (M) , gait (I) and composite measures of function (I)
Proprioception exercises	Pain/studies=3, subjects=190 (254 knees)	Proprioception exercises improved pain measures/Low
	Composite function/studies=3, subjects=190	Proprioception exercises did not improve composite function measures/Low
	Gait function/studies=3, subjects=169	Proprioception exercises did not improve gait function measures/Low
		Proprioception exercises improved pain (L) but did not improve gait (L) and composite measures of function (L)
Aerobic exercises	Pain/studies=6, subjects=1005	Aerobic exercises improved pain measures/Low
	Disability/studies=4, subjects=676	Aerobic exercises improved disability measures/Low
	Psychological disability/studies=4, subjects=246	Aerobic exercises did not improve psychological disability measures/Low
	Global assessment/studies=1, subjects=183	Aerobic exercises did not improve global assessment measures/Insufficient
	Health perception/studies=3, subjects=383	Aerobic exercises did not improve health perception measures/Low
	Quality of life/studies=1, subjects=56	Aerobic exercises did not improve quality of life measures/Insufficient
	Composite function/studies=3, subjects=791	Aerobic exercises did not improve composite function measures/Low
	Joint function/studies=1, subjects=21	Aerobic exercises did not improve joint function measures/Insufficient
	Gait function/studies=2, subjects=609	Aerobic exercises improved gait function measures/Low
	Transfer function/studies=2, subjects=327	Aerobic exercises improved transfer function measures/Low
		Aerobic exercises improved pain (L), disability (L), and gait (L), and transfer (L) measures of function but did not improve psychological disability (L), global assessment (I), health perception (L), quality of life (I), joint (I), and composite measures of function (L)
Aquatic exercises	Pain/studies=2, subjects=269	**Aquatic exercises did not improve pain measures/Moderate**
	Disability/studies=2, subjects=216	Aquatic exercises improved disability measures/Low
	Psychological disability/studies=1, subjects=214	Aquatic exercises did not improve psychological disability measures /Insufficient
	Quality of life/studies=2, subjects=216	Aquatic exercises did not improve quality of life measures/moderate
	Composite function/studies=1, subjects=45	Aquatic exercises did not improve composite function measures /Insufficient
		Aquatic exercises improved disability (L) but did not improve pain (M), psychological disability(I), quality of life (M), and composite measures of

Table 3. Narrative evidence summary of effectiveness of physical therapy interventions for knee osteoarthritis (continued)

Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
		function (I)
Strengthening exercises	Pain/studies=3, subjects=741 (822 knees)	Strengthening exercises improved pain measures/Low
	Disability/studies=2, subjects=660	Strengthening exercises did not improve disability measures/Low
	Psychological disability/studies=2, subjects=91	Strengthening exercises did not improve psychological disability measures/Low
	Global assessment/studies=1, subjects=67	Strengthening exercises improved global assessment measures/ Insufficient
	Health perception/studies=2, subjects=91	Strengthening exercises did not improve health perception measures/Low
	Quality of life/studies=2, subjects=105	Strengthening exercises did not improve quality of life measures/Low
	Composite function/studies=2, subjects=376	Strengthening exercises improved composite function measures/Low
	Joint function/studies=1, subjects=786	Strengthening exercises improved joint function measures/Insufficient
	Gait function/studies=2, subjects=660	Strengthening exercises improved gait function measures/Low
	Transfer function/studies=1, subjects=295	Strengthening exercises improved transfer function measures/Insufficient
		Strengthening exercises improved pain (L), global assessment (I), gait (L), joint (I), transfer (I), and composite (L) function measures but did not improve disability (L), psychological disability (L), health perception (L), and quality of life (L)
Tai Chi	Pain/studies=2, subjects=77	Tai Chi did not improve pain measures/Low
	Disability/studies=2, subjects=96	Tai Chi did not improve disability measures/Low
	Psychological disability/studies=1, subjects=44	Tai Chi improved psychological disability measures/Insufficient
	Composite function/studies=2, subjects=77	Tai Chi did not improve composite function measures/Low
	Joint function/studies=2, subjects=77	Tai Chi did not improve joint function measures/Low
	Gait function/studies=1, subjects=44	Tai Chi did not improve gait function measures/Insufficient
		Tai Chi improved psychological disability (I) but did not improve pain (L), disability (L), gait (I), joint (L) and composite (L) function measures
Massage	Psychological disability/studies=1, subjects=59	Massage did not improve psychological disability measures/Insufficient
	Health perception/studies=1, subjects=59	Massage improved health perception measures/Insufficient
	Composite function/studies=2, subjects=94	Massage improved composite function measures/Low
	Joint function/studies=2, subjects=127	Massage did not improve joint function measures/Low
	Gait function/studies=1, subjects=68	Massage improved gait function measures/Insufficient
		Massage improved health perception (I), gait (I) and composite (L) function measures but did not improve psychological disability (I) and joint (L) function measures
Joint mobilization	Pain/studies=1, subjects=43	Joint mobilization did not improve pain measures/Insufficient
	Disability/studies=1, subjects=43	Joint mobilization improved disability measures/Insufficient
	Global assessment/studies=1, subjects=43	Joint mobilization improved global assessment measures/Insufficient
	Gait function/studies=1, subjects=76	Joint mobilization did not improve gait function measures/Insufficient
		Joint mobilization improved disability (I) and global assessment (I), but did not improve pain (I) and gait function measures(I)
Joint mobilization with exercise	Disability/studies=1, subjects=120	Joint mobilization with exercise improved disability measures/Insufficient
	Gait function/studies=1, subjects=120	Joint mobilization with exercise did not improve gait function measures/ Insufficient

Table 3. Narrative evidence summary of effectiveness of physical therapy interventions for knee osteoarthritis (continued)

Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
		Joint mobilization with exercise improved disability (I) but did not improve gait (I) function measures
Orthotics	Global assessment/studies=1, subjects=119	Orthotics did not improve global assessment measures/Insufficient
	Composite function/studies=2, subjects=138	Orthotics did not improve composite function measures/Low
	Joint function/studies=1, subjects=119	Orthotics did not improve joint function measures/Insufficient
	Gait function/studies=4, subjects=262	**Orthotics did not improve gait function measures/High**
		Orthotics did not improve global assessment (I), joint (I), gait (H) and composite (L) function measures
Elastic subtalar strapping	Composite function/studies=3, subjects=235	Elastic subtalar strapping improved composite function measures/Low
Taping	Disability/studies=1, subjects=58	Taping did not improve disability measures/Insufficient
	Composite function/studies=1, subjects=58	Taping did not improve composite function measures/Insufficient
	Gait function/studies=1, subjects=36	Taping did not improve gait function measures/Insufficient
		Taping did not improve disability (I), gait (I) and composite (I) function measures
E-stim	Pain/studies=2, subjects=72	E-stim worsened pain measures/Low
	Disability/studies=2, subjects=95	**E-stim did not improve disability measures/Moderate**
	Global assessment/studies=2, subjects=95	E-stim improved global assessment measures/Low
	Health perception/studies=1, subjects=37	E-stim did not improve health perception measures/Insufficient
	Composite function/studies=3, subjects=129	**E-stim did not improve composite function measures/Moderate**
	Joint function/studies=2, subjects=95	**E-stim did not improve joint function measures/Moderate**
	Gait function/studies=3, subjects=149	E-stim did not improve gait function measures/Low
	Strength/studies=2, subjects=134	E-stim improved strength measures/Low
	Transfer function/studies=1, subjects=34	E-stim did not improve transfer function measures/Insufficient
		E-stim improved global assessment (L) and strength (L) measures, but worsened pain (L), and did not improve disability (M), health perception (I), and gait (L), joint (M), transfer (I), and composite (M) function measures,
PEMF	Pain/studies=2, subjects=145	**PEMF did not improve pain measures/Moderate**
	Disability/studies=1, subjects=86	PEMF did not improve disability measures/Insufficient
	Global assessment/studies=1, subjects=32	PEMF improved global assessment measures/Insufficient
	Composite function/studies=2, subjects=145	**PEMF did not improve composite function measures/Moderate**
	Joint function/studies=1, subjects=90	PEMF did not improve joint function measures/Insufficient
	Gait function/studies=1, subjects=32	PEMF did not improve gait function measures/Insufficient
		PEMF improved global assessment (I) but did not improve pain (M), disability (I), and gait (I), joint (I) and composite (M) function measures
Ultrasound	Pain/studies=3, subjects=204 (323 joints)	Ultrasound improved pain measures/Low
	Disability/studies=2, subjects=149	Ultrasound did not improve disability measures/Low
	Composite function/studies=2, subjects=128	Ultrasound improved composite function measures/Low
	Joint function/studies=1, subjects=67	Ultrasound did not improve joint function measures/Insufficient
	Gait function/studies=2, subjects=128	Ultrasound improved gait function measures/Low
		Ultrasounds improved pain (L), gait (L) and composite (L) function measures but did not improve disability (L), and joint function measures (I)

Table 3. Narrative evidence summary of effectiveness of physical therapy interventions for knee osteoarthritis (continued)

Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
Diathermy	Pain/studies=3, subjects=178	**Diathermy did not improve pain measures/Moderate**
	Disability/studies=2, subjects=143	**Diathermy did not improve disability measures/Moderate**
	Psychological disability/studies=1, subjects=40	Diathermy did not improve psychological disability measures/Insufficient
	Global assessment/studies= 1,subjects=113	Diathermy did not improve global assessment measures/Insufficient
	Health perception/studies=1, subjects=40	Diathermy did not improve health perception measures/Insufficient
	Composite function/studies=3, subjects=178	**Diathermy did not improve composite function measures/Moderate**
	Joint function/studies=2, subjects=143	**Diathermy did not improve joint function measures/Moderate**
	Gait function/studies=3, subjects=178	**Diathermy did not improve gait function measures/Moderate**
		Diathermy did not improve pain (M), disability (M), psychological disability (I), global assessment (I), health perception (I), and joint (M), gait (M) and composite (M) function measures
Heat	Disability/studies=1, subjects=34	Heat improved disability measures/Insufficient
	Quality of life/studies=1, subjects=34	Heat improved quality of life measures/Insufficient
	Composite function/studies=1, subjects=34	Heat did not improve composite function measures/Insufficient
	Gait function/studies=1, subjects=34	Heat did not improve gait function measures/Insufficient
		Heat improved disability (I) and quality of life (I), but did not improve gait (I) and composite (I) function measures
Mud pack	Disability studies=1, subjects=60	Mud pack improved disability measures/Insufficient
	Global assessment/studies=1, subjects=60	Mud pack improved global assessment measures/Insufficient
		Mud pack improved disability (I) and global assessment measures (I)
Cryotherapy	Disability/studies=1, subjects=68	Cryotherapy did not improve disability measures/Insufficient
	Quality of life/studies=1, subjects=68	Cryotherapy did not improve quality of life measures/Insufficient
	Composite function/studies=1, subjects=68	Cryotherapy did not improve composite function measures/Insufficient
	Pain/studies=1, subjects=22	Cryotherapy did not improve pain measures/Insufficient
		Cryotherapy did not improve pain (I), disability (I), quality of life (I), and composite function measures (I)

E-stim=Electrical stimulation; PEMF=Pulsed electromagnetic fields; **Bold** indicates findings with moderate or high strength of evidence; strength of evidence as L: Low; M: Moderate; H: High; I: Insufficient; Level of evidence was determined according to four domains.

Table 4. Narrative evidence summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis

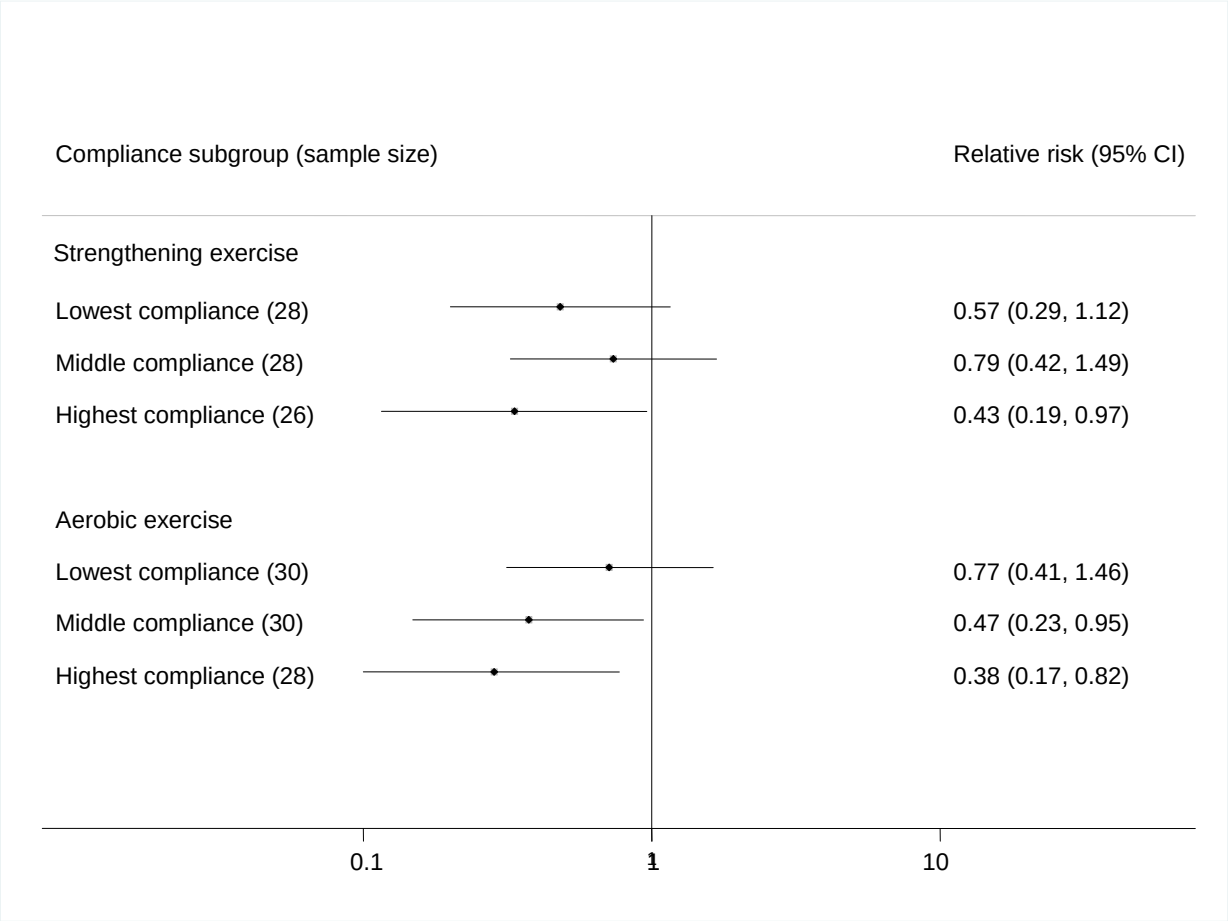
Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
Aerobic exercises vs. strengthening exercises	Disability/studies=2, subjects=345	Aerobic exercises did not improve disability measures, compared to strengthening exercises/Low
	Composite function/studies=2, subjects=345	Aerobic exercises did not improve composite function measures, compared to strengthening exercises/Low
	Gait function/studies=1, subjects=290	Aerobic exercises improved gait function measures, compared to strengthening exercises/Insufficient
	Transfer function/studies=1, subjects=290	Aerobic exercises did not improve transfer function measures, compared to strengthening exercises/Insufficient
		Aerobic exercises improved gait function measures (I) but did not improve disability (I), transfer (I), and composite (I) function measures, compared to strengthening exercises
Aquatic exercises vs. aerobic exercises	Disability/studies=2, subjects=116	**Aquatic exercises did not improve disability measures, compared to aerobic exercises/Moderate**
	Quality of life/studies=1, subjects=52	Aquatic exercises did not improve quality of life measures, compared to aerobic exercises/Insufficient
	Composite function/studies=1, subjects=64	Aquatic exercises did not improve composite function measures, compared to aerobic exercises/Insufficient
	Gait function/studies=1, subjects=64	Aquatic exercises did not improve gait function measures, compared to aerobic exercises/Insufficient
		Aquatic exercise did not improve disability (M), quality of life (I), gait (I) and composite (I) function measures, compared to aerobic exercise
Proprioception exercise vs. strengthening exercise	Composite function/studies=1, subjects=72	Proprioception exercises worsened composite function measures, compared to strengthening exercises/Insufficient
	Gait function/studies=1, subjects=72	Proprioception exercises did not improve gait function measures, compared to strengthening exercise/Insufficient
		Proprioception exercises worsened composite function measures (I) but did not improve gait function (I), compared to strengthening exercises
Tai Chi vs. stretching exercises	Disability/studies=1, subjects=40	Tai Chi improved disability measures, compared to stretching exercise/Insufficient
	Psychological disability/studies=1, subjects=40	Tai Chi improved psychological disability measures, compared to stretching exercise/Insufficient
	Global assessment/studies=1, subjects=40	Tai Chi improved global assessment measures, compared to stretching exercise/Insufficient
	Composite function/studies=1, subjects=40	Tai Chi did not improve composite function measures, compared to stretching exercise/Insufficient
	Joint function/studies=1, subjects=40	Tai Chi did not improve joint function measures, compared to stretching exercise/Insufficient
	Gait function/studies=1, subjects=40	Tai Chi did not improve gait function measures, compared to stretching exercise/Insufficient
	Transfer function/studies=1, subjects=40	Tai Chi improved transfer function measures, compared to stretching exercise/Insufficient

Table 4. Narrative evidence summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis (continued)

Comparison	Outcomes at the longest time of followup	Conclusions/strength of evidence
		Tai Chi improved disability (I), psychological disability (I), global assessment (I), and transfer function (I) but did not improve gait (I), joint (I), and composite (I) function measures, compared to stretching exercise
Laterally vs. neutrally wedged insole	Pain/studies=1, subjects=200	Laterally wedged insole did not improve pain measures, compared to neutrally wedged insole/Insufficient
	Disability/studies=1, subjects=200	Laterally wedged insole did not improve disability measures, compared to neutrally wedged insole/Insufficient
	Global assessment/studies=1, subjects=200	Laterally wedged insole did not improve global assessment measures, compared to neutrally wedged insole/Insufficient
	Quality of life/studies=1, subjects=200	Laterally wedged insole did not improve quality of life measures, compared to neutrally wedged insole/Insufficient
	Composite function/studies=2, subjects=363	**Laterally wedged insole did not improve composite function measures, compared to neutrally wedged insole/Moderate**
	Gait function/studies=1, subjects=45	Laterally wedged insole did not improve gait function measures, compared to neutrally wedged insole/Insufficient
	Joint function/studies=1, subjects=200	Laterally wedged insole did not improve joint function measures, compared to neutrally wedged insole/Insufficient
		Laterally wedged insole did not improve pain (I), disability (I), global assessment (I), quality of life (I), joint (I), gait (I), and composite (M) function measures, compared to neutrally wedged insole
Orthotics vs. brace	Composite function/studies=1, subjects=91	Orthotics did not improve composite function measures, compared to brace/Insufficient
E-stim vs. exercises	Pain/studies=2, subjects=81	E-stim did not improve pain measures, compared to exercise/Low
	Composite function/studies=1, subjects=50	E-stim improved composite function measures, compared to exercise/Insufficient
	Joint function/studies=1, subjects=50	E-stim improved joint function measures, compared to exercise/Insufficient
	Gait function/studies=2, subjects=81	E-stim did not improve gait function measures, compared to exercise/Low
		E-stim improved joint (I) and composite (I) measures of function but did not improve pain (L) and gait (I) function, compared to exercise
E-stim vs. ultrasound	Composite function/studies=1, subjects=35	E-stim did not improve composite function measures, compared to ultrasound/Insufficient
	Gait function/studies=1, subjects=35	E-stim did not improve gait function measures, compared to ultrasound/ Insufficient
		E-stim did not improve gait (I) and composite (I) function measures, compared to ultrasound

E-stim=Electrical stimulation; **Bold** indicates findings with moderate or high strength of evidence; strength of evidence as L: Low; M: Moderate; H: High; I: Insufficient; Level of evidence was determined according to four domains.

Figure 1. Risk of developing disability in activities of daily living in compliance subgroups



The attention control group is used as the reference.

Key Question 2. What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after physical therapy interventions?

Evidence for the association between intermediate and clinical outcomes was limited to individual studies and thus was insufficient to draw valid conclusions (Table 5). Substantial variability occurred between index and reference methods, definitions of outcomes, methods of examining diagnostic values, and associations between intermediate and clinical outcomes. Delineating between patient-centered and intermediate outcomes was somewhat artificial. For example, pain (a patient-centered outcome) is an explanatory factor for several intermediate outcomes including gait, range of motion, and balance. Likewise, patient-centered outcomes such as disability, self-reported pain, and observed IADL dependency were determined by composite measures of objective tests including WOMAC or Arthritis Impact Measurement Scale (Appendix Table F28) (intermediate outcomes). Finally, certain associations between patient-centered outcomes are clinically important for predicting treatment effects—for instance, pain or function may predict disability.

We synthesized the evidence of association between intermediate and clinical outcomes from 43 studies of 25,799 adults with knee OA. Mean age averaged 65 and ranged from 55 to 80 (Appendix table F29). Women constituted 70 percent of participants. Sample size of the studies varied with a median of 149 and mean of 600 participants. Half of all published studies were American or British (13 and 7 respectively). Minorities were included only in the American studies, which did not report those results separately.

The studies used different statistical concepts to examine the relationship between outcomes. First, the studies examined sensitivity, specificity, or positive predictive likelihood of index tests for correct identification of the outcomes according to reference tests. When the outcomes—for example, functional disability—were measured with scales, the studies defined thresholds in numeric score to categorize patients as disabled or not. Studies of diagnostic value had only fair quality (Appendix Table F30).

Second, the studies used linear regression to examine association as changes in measurements of patient-centered outcomes corresponding to changes in measurements of intermediate outcomes. Regression coefficients presented a magnitude of the change in continuous measures of patient-centered outcomes (such as pain) corresponding to one unit change in continuous measures of intermediate outcomes (such as muscle strength). However, it is not clear whether such estimates of the association between outcomes have clinical importance.

Third, the studies used logistic or Cox regression to examine association as rates or odds of patient-centered outcomes corresponding to rates or odds of functional impairments. We judged a magnitude of such association as high when relative risk or odds ratio were more than 2 or less than 0.5. Few studies used adjustment in regression models for lowering risk of bias. Some studies failed to distinguish patients with definitive diagnoses of knee OA from those with self-reported OA or knee pain.

Some studies examined the association between outcomes at the same time points without collecting followup data. These cross-sectional analyses could point out the association between functional impairments and pain or disability at one time point but could not predict future changes in the outcomes. In contrast, prospective studies examined the association between baseline functional impairments and patient-centered outcomes at future time of followup. Those

studies predicted patient-centered outcomes based on the association with intermediate outcomes after adjustment for confounding factors.

Fourth, some cross sectional studies calculated correlation between continuous measurements in the outcomes. Correlation coefficients ranged from -1 (negative correlation) to 0 (no correlation) to 1 (positive correlation). Correlation reflected simply the same linear direction of the changes in intermediate and clinical outcomes, with no consideration of units of measured outcomes or of the clinical importance of the measures.

Diagnostic Value of Outcomes

Few studies reported diagnostic values of intermediate outcomes. The Cohort Hip and Cohort Knee study (CHECK) found that an SF-36 Physical Function score <60 had conclusive diagnostic value for work limitations as determined by the Functional Capacity Evaluation.[202] Adults with SF-36 Physical Function score <60 had a large (>10) positive likelihood ratio of work limitations. In one prospective cohort study, a Pain Numeric Rating Scale score of <4 had a conclusive diagnostic value for patient perception of clinically significant improvement (Appendix Table F31).[203] Another prospective study, Clinical Assessment Study of the Knee (CAS(K)), demonstrated that bilateral knee pain, duration of morning stiffness, and inactivity gelling (stiffness after inactivity) had conclusive diagnostic value for poor WOMAC function at 18 months of followup (positive likelihood ratio = 42 and the area under the ROC curve = 0.73).[204]

Association Between Intermediate and Clinical Outcomes Examined with Regression Models

Measures of functional impairment were associated with poor patient-centered outcomes in individual studies (Table 6). Studies were inconsistent in defining intermediate and patient-centered outcomes and in adjusting for confounding factors. The studies examined the association with logistic or Cox regression reporting hazard rate ratios or odds ratios of categorical patient-centered outcomes (Table 7). Patient-centered outcomes were categorized according to clinically important thresholds in scales. Most studies examined the association with linear regression and reported differences in continuous measures of the outcomes corresponding to 1 unit increase in the measures of intermediate outcomes. No clinical importance of such changes was evident unless the studies proposed regression models estimating quality-of-life index or other patient-centered outcomes based on WOMAC scores. [205]

Gait

Gait measurements were associated with pain and poor functional outcomes (Appendix Table F32). Baseline stance time on stairs was positively associated with time to climb stairs at followup in a randomized trial of older adults with knee OA.[148] In one cross-sectional study, gait speed was positively associated with maximal activity profile (the highest oxygen-demanding activity the participant is still able to perform).[206] Adults with pain due to mild to moderate, clinically diagnosed medial-compartment knee OA had impaired walking speed.[207]

Muscle Strength

Muscle strength was positively associated with better function, but the significance and magnitude of the association differed depending on measures of strength and outcomes (Appendix Table F33). The strongest association was reported in one large cross-sectional study

of more than 6,000 older adults with knee OA.[208] Patients with maximal isometric quadriceps femoris muscle strength (force-generating capacity) of 10-20kg had 137 percent relative risk increase of poor function compared to those with >30 kg of force after controlling for age, sex, and BMI.[208] The association was dose responsive with a greater than 417 percent relative risk increase of poor function in adults with weaker muscles (≤10 kg of Force).[208] The CAS(K) prospective cohort found that adults with weaker quadriceps or hamstring muscle strength had a 50 percent higher relative risk of poor WOMAC functional outcome.[208] Another prospective cohort study, Mechanical Factors in Arthritis of the Knee (MAK), found no significant association between quadriceps and hamstring muscle strength after adjustment for age, BMI, pain intensity, and disease severity.[209] This study defined a physical function outcome as a clinically important change in chair-stand performance.[209] One cross-sectional study found muscle strength to be negatively associated with both observed and self-reported disability measures.[210] Functional disability, as assessed by total WOMAC score, was negatively associated with greater muscle strength.[211] The association between muscle strength and functional disability was stronger in patients with high knee joint laxity.[211] A single study found significant association between quadriceps torque and balance performance (center of pressure path length) after controlling for disease severity, symptom bother, and WOMAC pain.[212]

Impaired Balance Measurements

Impaired balance measurements demonstrated inconsistent association with functional status. A single prospective cohort study of older adults (CAS(K)) found no significant association between the impaired single-leg standing balance test and poor WOMAC functional outcomes after adjustment for age, BMI, knee pain, and stiffness (Appendix Table F34).[204] In contrast, a cross-sectional study of older women found a significant association between balance and outcomes of a functional test consisting of walking for 20 meters, climbing up and down nine stairs, and going from sitting to standing for five repetitions.[213] Physical performance time improved in association with better balance.[213] We cannot be certain whether study design, population, or balance measurements contributed to different conclusions in the studies.

Range of Motion

Among other intermediate outcomes, increased range of motion was negatively associated with observed and self-reported disability (Appendix Table F35).[210] A small cross-sectional study of 86 adults with knee OA found a significant interaction between joint range of motion, muscle strength, and walking speed.[211] A prospective cohort, the CAS(K), demonstrated that morning stiffness of 1-30 minutes predicted a 47 percent increase in relative risk of poor function at 18 months followup.[204] Another prospective cohort study, the MAK, demonstrated a 58 percent increase in relative risk of poor function at 3 years followup per 3º increase in joint laxity after controlling for age, BMI, pain intensity, and disease severity.[209]

Knee Mobility and Stability

Knee mobility and stability were weak predictors of functional performance. Joint mobility measures were not associated with poor WOMAC function (Appendix Table F36).[204]Joint stability measures demonstrated a weak but significant association with poor WOMAC function (Appendix Table F37).[204, 209, 211, 212, 214] Adults with a positive Giving Way Test had a 33 percent relative increase in having a poor WOMAC Function Score.[204] Knee instability or laxity was associated with a 58 percent relative increase in having a poor WOMAC Function Score.[209, 214]

Patient-centered outcomes were associated with a variety of tests (Table 5). Pain was associated with impaired functional status in adults with knee OA (Appendix Table F38). Adults with bilateral knee pain had a 46 percent relative risk increase of having a poor WOMAC Function Score.[204] An increased WOMAC Pain Score was also associated with a poor WOMAC Function Score.[214] At followup, VAS pain intensity was associated with a 48 percent relative increase in risk of having a poor WOMAC Function Score (Appendix Table F39).[209] In contrast, self-reported knee pain[215] or local tender point counts were not associated with poor function (Appendix Table F40).[204]

Disability measures were associated with gait, mobility restrictions, muscle strength, and range of motion (Table 5). Several mobility restriction tests (VAS, TUG Test, Step Test, Gait Speed) were associated with impaired adjusted daily activity scores (Appendix Table F41). Mobility restrictions and the TUG Test were negatively associated with Adjusted Activity Score.[206] Increased gait speed and step test were positively associated with adjusted daily activity score.[206] Greater muscle strength and range of motion of the affected knee(s) were negatively associated with self-reported or observed disability (Appendix Table F35).[210] Patients with more severe knee OA had impaired role function with physical limitations.[216]

We found no studies that reported time to return to work or activities. Patients with self-reported disability had increased risk of total joint replacement within a year of followup (Appendix Table F42). Patients with severe OA according to their Lequesne score had a 137 percent relative risk increase of having knee surgery within 1 year of followup (Adjusted OR 2.37, 95 percent CI, 1.71 to 3.25).[217] Patients who considered themselves disabled had a 57 percent relative risk increase of total joint replacement within 1 year of consultation.[217]

Several studies examined the importance of self-efficacy and mental health for adults with knee OA (Tables 8). Self-efficacy was defined as how patients perceive their ability to manage chronic arthritis, and it was measured using the Arthritis Self-Efficacy Scale.[209] The Arthritis Self-Efficacy Scale function subscale is a self-rating of degree of certainty in one's ability to walk 100 feet on flat ground in 20 seconds, walk down 10 steps in 7 seconds, and get out of an armless chair without using hands for support.[209] Higher values correspond to better self-efficacy.[209] Adults with good self-efficacy had an 11 percent relative decrease in risk of poor WOMAC function (Table 8).[209]

Self-reported health outcomes, including anxiety, depression, and fatigue, were associated with poor functional status. Older adults with knee pain and anxiety had a 29 percent relative risk increase of having poor WOMAC function compared to adults without anxiety (Appendix Table F43).[204] Adults with good mental health had a 42 percent relative risk reduction of having poor WOMAC function.[209] Adults with greater social support had a 15 percent relative risk reduction of having poor WOMAC function.[209] Several studies found that depression and frustration demonstrated strong positive correlation with poor functional status (Appendix Table F44).[210, 215, 218-223]

In summary, disability measures were associated with gait, mobility restrictions, muscle strength, and range of motion measures, but the magnitude and clinical importance of the associations remain unclear. Individual studies offered insufficient evidence for determining which intermediate outcomes strongly and consistently predict patient-centered outcomes.

Key Question 2a. What is the validity of the tests and measures used to determine intermediate outcomes of PT on OA in association with patient-centered outcomes?

Many articles reported validation, but few demonstrated a strong (more than 50 percent) correlation between index and reference method measurements (Appendix Table F45). The studies used a variety of reference methods and judged validity on the basis of significant correlation coefficients. Strength of correlation varied across validity types (Appendix Table F46).

We synthesized the evidence of correlation between intermediate and patient-centered outcomes from 43 studies of 15,734 adults with knee OA. Mean age averaged around 64 years and ranged from 29 to 67 (Appendix Table F47). Women constituted 64 percent of the participants. Sample size of the studies varied with a median of 109 and a mean of 254 participants. Half of all studies were American (20 studies), Canadian (nine studies), or British (eight studies). Only American studies included minorities (but did not separately report those results). Some studies did not distinguish patients with a definitive diagnosis of knee OA from those with self-reported OA or knee pain. The studies analyzed correlation coefficients between index and reference methods and did not use strategies to reduce bias.

Correlation strength varied across measurements and reference standards for intermediate outcomes (Appendix Table F48). Balance measures with Standing Balance Test correlated with radiographic degenerative changes.[224] The Knee Proprioception Test (quantified as the ability to replicate target knee joint angles using a computerized dynamometer) did not correlate with radiographic degenerative changes.[224] Knee range of motion was assessed as self-reported morning stiffness,[225]Arthritis Impact Measurement Scale dexterity subscale,[226] Knee Patient-Specific Index,[227] Lequesne index,[228] or WOMAC stiffness subscale[228] (Appendix Table F48). A strong correlation was reported for the Knee Patient-Specific Index with the WOMAC Stiffness scale[227] and for the Lequesne Index with the WOMAC Stiffness subscale.[228] A strong correlation was demonstrated for range of motion to ipsilateral hip abduction with knee flexion on the affected side.[229] Other tests demonstrated very weak or no correlation with patient-centered outcomes (Appendix Table F49). Measurements of pain and function in relation to symptom bother were validated with the Short Form Health Questionnaire,[230] The Influence of Rheumatic Disease on General Health and Lifestyle mobility subscale,[231] or the Global Functional Rating[232] (Appendix Table F50).

Many studies used validated WOMAC subscales as a reference standard (Appendix Table F51). The WOMAC scale was recommended to measure clinical outcomes in trials involving adults with knee OA.[39, 162] WOMAC is a validated instrument with different subscales for pain and stiffness as well as for physical, social, and emotional function.[233-236] Several studies examined whether the WOMAC satisfied the Rasch model (Rasch Item Response Theory) (Appendix Table F52). To satisfy Rasch criteria, subscales should be unidimensional by measuring the anticipated concept of pain and function, and not have redundant items counting repeatedly toward the overall score. The first study of 655 patients with osteoarthritis of the knee or hip concluded that pain and function subscales were unidimensional and did not collect redundant information.[237] The second study of 158 patients with knee or hip OA found that pain and function items may represent the same construct and introduce redundancy for calculation of overall score.[238] A prospective cohort followed 1,151 adults with knee OA or hip OA for 1 year after arthroplasty.[239] The authors concluded that prospectively collected responses evaluate changes in functional status. The study suggested that when monitoring treatment effects with

WOMAC, certain items should be omitted, including night pain and pain on standing, heavy domestic duties, getting in and out of the bath, and getting on and off the toilet.[239] All studies recommended using WOMAC to measure patient-centered outcomes in adults treated for knee OA.

Key Question 2b. Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes?

In order to be considered surrogates for patient-centered outcomes, measurements in functional impairments are expected to predict patient-centered outcomes. However, none of the intermediate outcomes met this criterion. Gait was proposed by the TEP as a feasible candidate for a surrogate endpoint; however, no studies examined the association between gait and patient-centered outcomes in adults with knee OA treated with physical therapy interventions. While a single RCT concluded that knee pain and self-efficacy mediated the effects of exercise on stair climb time,[148] no evidence supported an association or predictive power between gait change and patient-centered outcome change. A single longitudinal study of older adults demonstrated that impaired gait and Physical Performance Test were independent predictors of nursing home placement.[149] Three cohort studies, including The Einstein Aging Study,[150] the Chinese Elderly Cohort, [151]and the Women's Health and Aging Study,[152] examined the association between gait and nursing home placement (Appendix Table F53). However, these studies included adults with any etiology of gait problem, including neurological diseases or heart failure. Moreover, definitions of impaired gait and magnitude of the association were inconsistent across the studies.

Key Question 2c. What are minimal clinically important differences of the tests and measures used to determine intermediate outcomes?

MCIDs refer to thresholds of change in outcomes measurements that result in significant changes in clinical outcomes. Such thresholds were determined comparing the changes in performance measure with patient perception of improvement. Establishing accurate MCIDs helps to clarify whether statistically significant changes in outcome measures actually equate with patient opinions about treatment success and improved quality of life. MCIDs are critically necessary for evaluating whether changes in commonly used outcomes measurements or scales are of actual clinical importance to patients.

Thirty studies of 13,162 adults reported MCIDs. The studies used the anchor method, which compares changes in scales with patient perception of improvements. MCIDs were available for 26 tools as absolute change in score or relative change as a percent difference from baseline levels. The latter method incorporated baseline severity of the diseases (Appendix Table F54). Only a few studies defined MCID with the distribution method, which is based on distribution of changes in outcomes measurements and defined MCID as an upper quartile of the distribution.

We identified 16 studies that determined (with slight variation) MCIDs in WOMAC scales and subscales (Appendix Table F54). The Osteoarthritis Research Society International Standing Committee for Clinical Trials determined that patient perception of “high” improvement in pain corresponded to at least a 40 percent relative change in WOMAC with a minimum absolute improvement of 20 to 30 NU (normalized units).[240] Patients noticed improvement when WOMAC subscales changed by a margin of 17 to 22 percent of baseline scores.[241]

Few studies determined PASS for knee OA. PASS is defined as the highest level of symptom patients can tolerate and still be satisfied with their treatment. The studies that determined PASS

used the same anchor method as they did for determining MCIDs. The difference is in anchoring questions that asked patients whether they are satisfied with their functional status in relation to daily activities and quality of life. PASS was determined for three scales (WOMAC, VAS for pain, and the Patient Global Assessment) (Appendix Table F55).[242]

MCIDs in 36-Item Short-Form Health Survey (SF-36) were determined for patients with severe knee OA before or after surgery (Appendix Table F56).[243-245] In the SF-36, MCIDs were 12.83 for pain, 0.11 for general health, 0.76 for mental health, and 10.04 for physical functioning.[244] At 6 months followup, mean changes of 22 in an SF-36 bodily pain score and of 38 in physical functioning equated to patient reports of feeling "a great deal better."[245] A variety of other tests and scales proposed MCIDs to judge clinical effectiveness of treatments (including the Timed Up and Go,[246, 247] the 6-Minute Walk Test,[246, 247] and the Short Physical Performance Battery)[247] (Appendix Table F57).

Summary

In individual studies, muscle strength, range of motion, mobility restrictions, and gait were associated with patient-centered outcomes, but individual studies are insufficient for definitive conclusions. Many articles reported validation, but few demonstrated a strong (more than 50 percent) correlation between index and reference method measurements. Original studies concluded that tests are valid based on significance, not strength of correlation. None of the intermediate outcomes met surrogate criteria for patient-centered outcomes. Validated tools defined threshold values of clinical importance for evaluating treatment success. However, studies more often used continuous measures of the outcomes providing an average score for all patients in each treatment group with no evaluation of clinical importance. Average scores, however, cannot reveal how many patients develop disability or experience clinically meaningful improvement in pain, function, or quality of life.

Table 5. Association between intermediate and clinical outcomes (insufficient evidence from individual studies)

Clinical outcome	Intermediate outcome	Conclusion	Strength of evidence
Disability	Physical performance	Physical function assessed with SF-36 and WOMAC tools have conclusive diagnostic value for work limitations	Insufficient
Disability	Physical performance	Several functional tests (VAS: restriction; TUG test; Step test; Walking speed) were associated with impaired adjusted daily activity score	Insufficient
Disability	OA severity	Severity of OA assessed with Osteoarthritis of the Knee Severity Index was negatively associated with role functioning	Insufficient
Disability	Range of motion	Increased range of motion was negatively associated with observed and self reported disability	Insufficient
Disability	Strength	Muscle strength was negatively associated with observed and self reported disability	Insufficient
Physical performance	Balance	Impaired balance was associated with poor physical performance	Insufficient
Pain	Gait	Gait speed was associated with WOMAC pain severity	Insufficient
Pain	Function	Joint laxity (knee instability) was not associated with walking time	Insufficient
Pain	OA severity	Severity of OA assessed with Osteoarthritis of the Knee Severity Index was negatively associated with body pain assessed using SF-36	Insufficient
Function	Balance	Impaired single-leg standing balance was not associated with poor WOMAC function	Insufficient
Function	Disease severity	Self-reported swelling was associated with poor WOMAC function scale	Insufficient
Function	Joint alignment	Fixed flexion deformity was not associated with poor WOMAN function	Insufficient
Function	Joint alignment	Intercondylar and intermalleolar gap in standing were not associated with poor functional outcome	Insufficient
Function	Joint mobility	Duration of morning stiffness was associated with poor WOMAC function	Insufficient
Function	Joint mobility	Hip rotation was not associated with poor WOMAC function	Insufficient
Function	Joint mobility	Knee flexion Range of movement was not associated with poor WOMAC function	Insufficient
Function	Joint mobility	Locking: Pseudo-locking was not associated with poor WOMAC function	Insufficient
Function	Joint stability	Anteroposterior instability was not associated with poor WOMAC function	Insufficient
Function	Joint stability	Laxity (knee instability) was associated with poor WOMAC function	Moderate
Function	Joint stability	Positive Giving Way Test was associated with poor WOMAC function	Insufficient
Function	OA severity	Severity of OA assessed with Osteoarthritis of the Knee Severity Index was negatively associated with physical functioning assessed using SF-36	Insufficient
Function	Proprioception	Proprioceptive inaccuracy was not associated with poor WOMAC function	Insufficient
Function	Strength	Force (quadriceps femoris muscle strength) <20kg was associated with poor WOMAC function	Insufficient
Function	Strength	Hamstring strength(mm Hg): ≤100 vs. ≥185 was associated with poor WOMAC function	Low
Function	Strength	Muscle strength and laxity (knee instability) were associated with reduced walking time	Insufficient
Function	Strength	Muscle strength but not laxity (knee instability) were associated with poor WOMAC function	Insufficient
Function	Strength	Quadriceps strength(mm Hg): ≤140 vs. ≥300 was associated with poor WOMAC function	Insufficient

Table 6. Regression association between intermediate and clinical outcomes

Intermediate outcomes	Disability studies/ N=patients Estimate	Function Studies/patients Estimate	Pain Studies/patients Estimate
Balance		Harrison, 2004[213]/ N=50 NS * Thomas, 2008[204] N=621 NS HR	
Gait speed	**Bennell, 2004[206]/ N=259***	Nebel, 2009[248] N=179 NS *	**Astephen Wilson, 2011[207]/ N=40***
Range of motion	**van Baar, 1998[210]/ N=185***	Thomas, 2008[204] N=621 NS HR	**Van Der Esch, 2006[211]/ N=86***
Strength	**van Baar, 1998[210]/ N=185***	Thomas, 2008[204] N=62 **1.5 HR** Wood, 2008[208] N=741 **5.2 OR** Sharma, 2003[209] N=257 NS OR O'Reilly, 1998[249] N=300 **7.1 OR**	**O'Reilly, 1998[249]** **N=300 18.8 OR**
Swelling		**Thomas, 2008[204]/ N=621** **1.3 HR**	

* linear regression; bold-significant association at 95% CI
HR - hazard rate ratio; OR - odds ratio

Table 7. Relative measure of association between intermediate outcomes and functional disability

Author, year Design Months of followup	Intermediate outcome	Definition of intermediate outcome	Adjustment	Estimate Mean (95% CI)	Area under the receiver operating characteristic curve (ROC)
Thomas*, 2008[204] prospective cohort Months of followup: 72	Balance	Single-leg standing balance(s): <4 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count	HR 1.21 (0.85; 1.72)	0.77
	Balance	Single-leg standing balance(s): <4 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count, prevalent knee radiographic OA	HR 1.21 (0.85; 1.73)	0.77
	Balance	Single-leg standing balance(s): <4 vs. 30	Local tender point count	**HR 1.49 (1.09; 2.04)**	0.68
	Balance	Single-leg standing balance(s): 10-29 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count	HR 1.12 (0.8; 1.55)	0.77
	Balance	Single-leg standing balance(s): 10-29 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count, prevalent knee radiographic OA	HR 1.1 (0.79; 1.54)	0.77
	Balance	Single-leg standing balance(s): 10-29 vs. 30	Local tender point count	HR 1.27 (0.92; 1.74)	0.68
	Balance	Single-leg standing balance(s): 4-9 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count	HR 1.22 (0.88; 1.67)	0.77
	Balance	Single-leg standing balance(s): 4-9 vs. 30	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, inactivity gelling, local tender point count, prevalent knee radiographic OA	HR 1.22 (0.89; 1.68)	0.77
	Balance	Single-leg standing balance(s): 4-9 vs. 30	Local tender point count	**HR 1.5 (1.12; 2.01)**	0.68
Thomas, 2008[204] prospective cohort Months of followup: 72	Range of motion	Duration of morning stiffness (min): ≤30 vs. none	Bilateral knee pain, inactivity gelling	**HR 1.47 (1.13; 1.89)**	0.69
	Range of motion	Duration of morning stiffness (min): >30 vs. none	Bilateral knee pain, inactivity gelling	HR 1.55 (0.99; 2.43)	0.69
	Range of motion	Duration of morning stiffness (min): ≤30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age	HR 1.32 (1.01; 1.73)	0.76
	Range of motion	Duration of morning stiffness (min): ≤30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age, local tender point count, single-leg standing balance	HR 1.25 (0.95; 1.65)	0.77

Table 7. Relative measure of association between intermediate outcomes and functional disability (continued)

Author, year Design Months of followup	Intermediate outcome	Definition of intermediate outcome	Adjustment	Estimate Mean (95% CI)	Area under the receiver operating characteristic curve (ROC)
	Range of motion	Duration of morning stiffness (min): ≤30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age, local tender point count, single-leg standing balance, prevalent knee radiographic OA	HR 1.25 (0.95; 1.65)	0.77
	Range of motion	Duration of morning stiffness (min): ≥30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age	HR 1.22 (0.75; 2)	0.76
	Range of motion	Duration of morning stiffness (min): ≥30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age, local tender point count, single-leg standing balance	HR 1.15 (0.7; 1.89)	0.77
	Range of motion	Duration of morning stiffness (min): ≥30 vs. none	BMI, anxiety, inactivity gelling, bilateral knee pain, age, local tender point count, single-leg standing balance, prevalent knee radiographic OA	HR 1.16 (0.7; 1.91)	0.77
	Range of motion	Inactivity gelling: Yes vs. no	Bilateral knee pain, duration of morning stiffness	HR 1.34 (0.98; 1.83)	NR
	Range of motion	Inactivity gelling: Yes vs. no	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age	HR 1.23 (0.89; 1.71)	0.69
	Range of motion	Inactivity gelling: Yes vs. no	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, local tender point count, single-leg standing balance	HR 1.19 (0.86; 1.66)	0.76
	Range of motion	Inactivity gelling: Yes vs. no	BMI, anxiety, duration of morning stiffness, bilateral knee pain, age, local tender point count, single-leg standing balance, prevalent knee radiographic OA	HR 1.19 (0.85; 1.65)	0.77
Thomas, 2008[204] prospective cohort Months of followup: 72	Strength	Hamstring strength (mm Hg): ≤100 vs. ≥185	Unadjusted	**HR 1.51 (1.12; 2.02)**	NR
	Strength	Hamstring strength (mm Hg): 101-139 vs. ≥185	Unadjusted	HR 1.31 (0.97; 1.76)	NR
	Strength	Hamstring strength (mm Hg): 140-184 vs. ≥185	Unadjusted	HR 1.1 (0.81; 1.5)	NR
	Strength	Quadriceps strength (mm Hg): 141-200 vs. ≥300	Unadjusted	HR 1.27 (0.93; 1.73)	NR
	Strength	Quadriceps strength (mm Hg) ≤140 vs. ≥300	Unadjusted	**HR 1.52 (1.12; 2.06)**	NR
	Strength	Quadriceps strength (mm Hg):200-299 vs. ≥300	Unadjusted	HR 1.08 (0.79; 1.47)	NR

Table 7. Relative measure of association between intermediate outcomes and functional disability (continued)

Author, year Design Months of followup	Intermediate outcome	Definition of intermediate outcome	Adjustment	Estimate Mean (95% CI)	Area under the receiver operating characteristic curve (ROC)
Wood*, 2008[208] cross-sectional Months of followup: NA	Strength	≤10 kg of force (quadriceps femoris muscle strength) vs. >30 kg	NR	**OR 5.17 (3.01; 8.86)**	NR
	Strength	10-20 kg of force (quadriceps femoris muscle strength) vs. >30 kg	NR	**OR 2.37 (1.57; 3.59)**	NR
	Strength	20-30 kg of force (quadriceps femoris muscle strength) vs. >30 kg	NR	OR 1.29 (0.83; 2.01)	NR
Sharma, 2003[209] prospective cohort Months of followup: 72	Strength	Quadriceps strength, ft-lbs	Age, BMI, knee pain intensity, and disease severity (higher K/L grade of the 2 knees)	OR 0.88/20 ft-lbs (0.7; 1.11)	NR
	Strength	Hamstring strength, ft-lbs	Age, BMI, knee pain intensity, and disease severity (higher K/L grade of the 2 knees)	OR 0.86/20 ft-lbs (0.6; 1.23)	NR
O'Reily**,1998[249] nested case-control Months of followup: NA	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): 20-30 vs. >30	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	OR 1.48 (0.37; 5.93)	NR
	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): 10-20 vs. >30	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	**OR 4.88 (1.18; 20.14)**	NR
	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): ≤10 vs. >30	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	**OR 8.23 (1.53; 44.38)**	NR
	Strength	MVC (Maximum voluntary contraction) (kgF): 30-40 vs. >40	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	OR 3.04 (0.86; 10.71)	NR
	Strength	MVC (Maximum voluntary contraction) (kgF): 20-30 vs. >40	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	OR 3.77 (1.02; 13.91)	NR
	Strength	MVC (Maximum voluntary contraction) (kgF): ≤20 vs. >40	Age, sex, BMI, depression, anxiety, activation (percent), radiographic score	**OR 4.98 (1.08; 22.97)**	NR
O'Reilly, 1998[249] nested case-control Months of followup: NA	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): 30-40 vs. >40	Age, sex, BMI, depression, anxiety, activation, radiographic score	OR 1.49 (0.56; 3.96)	NR
	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): 20-30 vs. >40	Age, sex, BMI, depression, anxiety, activation, radiographic score	**OR 3.17 (1.22; 8.26)**	NR

Table 7. Relative measure of association between intermediate outcomes and functional disability (continued)

Author, year Design Months of followup	Intermediate outcome	Definition of intermediate outcome	Adjustment	Estimate Mean (95% CI)	Area under the receiver operating characteristic curve (ROC)
	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): 10-20 vs. >40	Age, sex, BMI, depression, anxiety, activation, radiographic score	**OR 7.1 (2.43; 20.68)**	NR
	Strength	MVC (Maximum voluntary contraction of quadriceps) (kgF): ≤10 vs. >40	Age, sex, BMI, depression, anxiety, activation, radiographic score	**OR 18.83 (4.79; 74.08)**	NR
Thomas, 2008[204] prospective cohort Months of followup: 72	Swelling	Self-reported dramatic swelling: Yes vs. no	Unadjusted	HR 1.09 (0.83; 1.44)	NR
	Swelling	Self-reported swelling in past month: Yes vs. no	Unadjusted	**HR 1.27 (1.03; 1.56)**	NR

NR=Not reported
* Thomas, 2008[204]; Wood, 2008[208]- functional disability was defined using WOMAC function scale;
** O'Reilly, 1998[249] functional disability was defined using WOMAC functional score >19 or SF-36 functional score <90

Table 8. Regression association between psychological and clinical outcomes

Intermediate outcomes	Disability Studies/N=patients Estimate	Function Studies/patients Estimate	Pain Studies/patients Estimate	Quality of life Studies/patients Estimate
Anxiety	O'Reilly, 1998[249] N=300 NS OR	**Thomas, 2008[204]/ N=621 1.3 HR** O'Reilly, 1998[249] **N=300 ***	O'Reilly, 1998[249] N=300 NS OR	
Baseline Self-efficacy		**Rejeski, 1998[148]/ N=439*** **Maly, 2006**[246] **N= 54 ***		
Depression	**O'Reilly, 1998**[249] **N=300 6.2 OR**	**Wolfe, 1999[234]/ N=2115 * WOMAC function** **Wolfe, 1999[234]/ N=2115 * WOMAC stiffness** **O'Reilly, 1998**[249] **N=300 ***	**Wolfe, 1999[234]/ N=2115*** **O'Reilly, 1998**[249] **N=300 2.4 OR**	
Fatigue		**Wolfe, 1999[234]/ N=2115 * WOMAC function** **Wolfe, 1999[234]/ N=2115 * WOMAC stiffness**	**Wolfe, 1999[234]/ N=2115***	
Mental health score		**Sharma, 2003[209]/ N=257 0.6 OR**		
Role functioning emotional score		Sharma, 2003[209]/ **N=**257 NS OR		
Psychological well being: Cheerfulness			**van Baar, 1998[210]/ N=185***	
Self-efficacy score		**Sharma, 2003[209]/ N=257 0.8 OR**		
Functional self-efficacy				Harrison, 2004[213]/ N=50 NS*

* linear regression; bold - significant association at 95% CI
HR - hazard rate ratio; OR - odds ratio

Key Question 3. What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls?

a. Which patient characteristics are associated with the harms of examined physical therapy interventions?

b. Do harms differ by the duration of the treatment and time of followup?

Adverse events were uncommon and varied across interventions. Skin irritation was reported with brace, insole, tape, and electrical stimulation; swelling with brace, diathermy, and exercise; muscle soreness with electrical stimulation; grumbling, warming, or throbbing sensation with diathermy, electrical stimulation, and PEMF; increased pain with diathermy, exercise, insole, and PEMF; falls with insole; and need for surgery with diathermy. Rates of adverse events did not differ significantly between treatment groups (Table 9). Nor were adverse events severe enough to deter participants from continuing treatment. Only four significant findings were reported. Lund et al. found that adverse events were six times more likely following land-based exercise than aquatic exercise, yet insignificantly led to discontinuation.[64] Use of a 16-mm lateral wedged insole resulted in more pain than an 8-mm lateral wedge.[143] Compared to neutrally wedged insoles, laterally wedged insoles led to more back pain, foot pain, and other discomfort.[130] Hinman et al. reported that skin irritation was more likely with therapeutic tape than control tape.[102] Two studies compared skin reactions with active electrical stimulation and inactive sham stimulation. The pooled analysis[106, 198] showed that electrical stimulation did not increase risk of skin irritation.

Table 9. Adverse events reported with physical therapy for OA of knee

Treatments	Definition of adverse events	Studies	Patients	Relative risk Number needed to treat to harm	Strength of evidence
Diathermy vs. placebo	Adverse event including mild pain, mild swelling, feeling of vasodilatation, deterioration of pain, or needed operation	1[121]	113	1.13 (0.30, 4.31)	Insufficient
E-stim vs. placebo	Mild skin reaction	2[106, 198]	136	1.02 (0.53, 1.97)	Low
Interferential (IF) and patterned muscle stimulation vs. low-current transcutaneous electrical nerve stimulation	Adverse event including skin irritation, skin burns, muscle soreness, electrical shock, and unanticipated adverse events	1[250]	109	0.57 (0.20, 1.58)	Insufficient
Aquatic exercise vs. land-based exercise	**Increased pain during and after exercise, or swollen knees**	**1[64]**	**52**	0.25 (0.08, 0.80) **3 (2; 10)**	**Insufficient**
Aquatic exercise vs. land-based exercise	Discontinuation due to adverse effects	1[64]	52	0.14 (0.01, 2.54)	Insufficient
Home based progressive exercise vs. home based control exercise	Adverse event, not specified	1[251]	179	0.60 (0.21, 1.78)	Insufficient
8 mm laterally wedged insole vs. 12 mm laterally wedged insole	Popliteal pain, low back pain, or foot sole pain	1[143]	41	0.7 (0.13, 3.76)	Insufficient
8 mm laterally wedged insole vs. 16 mm laterally wedged insole	**Popliteal pain, low back pain, or foot sole pain**	**1[143]**	**41**	**0.23 (0.06, 0.95)** **3 (2; 13)**	**Insufficient**
12 mm laterally wedged insole vs. 16 mm laterally wedged insole	Popliteal pain, low back pain, or foot sole pain	1[143]	42	0.33 (0.10, 1.06)	Insufficient
Strapped insole vs. Inserted insole	Popliteal pain, low back pain, or foot sole pain	1[99]	90	5.74 (0.72, 45.77)	Insufficient
5 degree lateral wedge insole vs. neutrally wedged insole	Musculoskeletal symptoms	1[137]	180	0.6 (0.231, 1.58)	Insufficient
5 degree lateral wedge insole vs. neutrally wedged insole	Blisters	1[137]	180	0.2 (0.02, 1.68)	Insufficient
5 degree lateral wedge insole vs. neutrally wedged insole	Falls	1[137]	180	1.33 (0.31, 5.79)	Insufficient
5 degree laterally wedged insole vs. neutrally wedged insole	**Self reported problems with insoles**	**1[130]**	**179**	**2.02 (1.31, 3.12)** **4(3;10)**	**Insufficient**
5 degree laterally wedged insole vs. neutrally wedged insole	**Back pain**	**1[130]**	**179**	**9.10 (1.18, 70.35)** **11 (6; 42)**	**Insufficient**
5 degree laterally wedged insole vs. neutrally wedged insole	**Foot pain**	**1[130]**	**179**	**2.31 (1.33, 4.03)**	**Insufficient**
5 degree laterally wedged insole vs. neutrally wedged insole	**Uncomfortable or difficulty fitting in shoes**	**1[130]**	**179**	**3.79 (1.31, 10.99)** **5 (3;13)**	**Insufficient**
5 degree laterally wedged insole vs. neutrally wedged insole	Increased knee pain	1[130]	179	0.40 (0.08, 2.031)	Insufficient
5 degree laterally wedged insole vs. neutrally wedged insole	Felt unstable	1[130]	179	0.34 (0.01, 8.16)	Insufficient

Table 7. Relative measure of association between intermediate outcomes and functional disability (continued)

Treatments	Definition of adverse events	Studies	Patients	Relative risk Number needed to treat to harm	Strength of evidence
5 degree laterally wedged insole vs. neutrally wedged insole	**Any discomfort**	**1**[130]	**163**	**1.79 (1.17, 2.74)**	**Insufficient**
Medial insole vs. neutrally wedged insole	Mild discomfort	1[130]	30	0.29 (0.01, 6.69)	Insufficient
PEMF vs. placebo	Grumbling or throbbing sensation	1[116]	90	1 (0.27, 3.75)	Insufficient
PEMF vs. placebo	Warming sensation	1[116]	90	6 (0.75, 47.85)	Insufficient
PEMF vs. placebo	Aggravation of the osteoarthritic pain in the study knee	1[116]	90	2 (0.19, 21.28)	Insufficient
Therapeutic tape vs. control tape	**Skin irritation**	**1**[102]	**58**	**8 (1.07, 59.95)** **4 (2;15)**	**Insufficient**

Bold- significant association at 95% CI

Summary and Discussion

Key Findings

A number of important findings emerged from this review.

Efficacy of Physical Therapy Interventions

- Few physical therapy interventions showed efficacy (Table 10).
- No single physical therapy intervention improved all outcomes.
- Pain was decreased with aerobic, strengthening, proprioception exercise, massage, and ultrasound (Figure 2). Low strength of evidence suggested a 60 percent reduction in pain intensity with proprioception exercise (Figure 3).
- Research focused on individual physical therapy interventions, whereas typical physical therapy practice uses combined modalities.
- Pooled analyses demonstrated that aerobic and aquatic exercises improved disability measures; aerobic and strengthening exercise reduced pain and improved function, proprioception exercise reduced pain; and ultrasound improved function.
- Pooled analyses also demonstrated that education programs, Tai Chi, diathermy, orthotics, and magnetic stimulation (PEMF) failed to show any benefits.
- Individual (nonpooled) RCTs failed to show consistent significant, strong, or clinically important changes in outcomes.

Comparative Effectiveness of Physical Therapy Interventions

We found limited evidence about comparative effectiveness of physical therapy interventions (Table 11).

- Disability measures did not differ between aerobic, aquatic, and strengthening exercises in pooled analyses.
- Pain did not differ between electrical stimulation and exercise in pooled analyses.
- Individual RCTs of other treatment comparisons found no consistent clinically important differences in outcomes and provided insufficient evidence to conclude the best treatment option.

The Role of Patient Characteristics in Modifying Treatment Effects

- Evidence from individual RCTs was insufficient to conclude differences in effects by patient characteristics.
- Patients with high compliance to exercise tended to have better benefits.

The Role of the Duration or Intensity of the Treatments

- The duration of examined interventions was not associated with better intermediate or patient-centered outcomes.
- Evidence regarding the association between the dose/intensity/frequency of examined interventions and outcomes was insufficient for the majority of comparisons.

- The treatments that demonstrated significant benefits (aerobic, aquatic, and strengthening exercises and ultrasound) did not differ in effect at shorter versus longer time of followup.
- Electrical stimulation improved pain short term but significantly worsened pain at longer followup.

Adverse Effects of Physical Therapy Interventions

- Adverse events were uncommon and not severe enough to deter patients from continuing treatment.
- Adverse events varied across intervention and included: skin irritation with brace/insole/tape/electrical stimulation; swelling with brace/diathermy/exercise; muscle soreness with electrical stimulation; warming/throbbing sensation with diathermy/ electrical stimulation/PEMF, increased pain with diathermy/exercise/insole/PEMF; and falls with insole.

Association Between Intermediate and Patient-Centered Outcomes

- Gait, mobility restrictions, muscle strength, and range of motion measures were associated with disability measures.
- Individual studies offered insufficient evidence for determining which intermediate outcomes strongly and consistently predict patient-centered outcomes.
- Many articles reported validation, but few demonstrated a strong (more than 50 percent) correlation between index and reference method measurements.
- Original studies concluded that tests are valid based on significance, not strength of correlation.
- None of the intermediate outcomes met surrogate criteria for patient-centered outcomes.

Minimal Clinically Important Differences

- Minimal clinically important differences of the tests were determined using the anchor method, which compares changes in scales with patient perception of improvements.
- Minimal clinically important differences were available as absolute change in score or relative change as a percent difference from baseline levels, which accounts for baseline severity of the disease.
- PASS, defined as the highest level of symptom patients can tolerate and still be satisfied with their treatment, was determined for three scales (WOMAC, VAS for pain, and the Patient Global Assessment Scale).
- Validated tools defined threshold values of clinical importance for evaluating treatment success in adults with knee OA. In contrast, more often studies used continuous measures of the outcomes providing an average score for all patients in each treatment group. Clinical importance of such averages was not evaluated. Average scores do not provide information how many patients develop disability or experience clinically meaningful improvements in pain, function, or quality of life.

Our report addresses patient-centered outcomes including pain, disability, and function with physical therapy interventions for adults with knee OA. Our findings agree with previously published guidelines[10, 15] and systematic reviews[19, 21, 153] that recommend exercise as an effective

physical therapy intervention. Overall, however, our analyses demonstrated that few physical therapy interventions were effective and no single intervention improved all outcomes.

Several factors affected the applicability of the research base. This lessened the degree to which our synthesis can fully and accurately address the efficacy and comparative effectiveness of physical therapy interventions for knee OA. Most important, current standards of physical therapy practice involve combined interventions, whereas published studies have examined individual physical therapy interventions. Our focus on randomized clinical trials, which equally distribute concomitant treatments among groups to accurately estimate the effect of an examined intervention, could not mitigate the impact of this discrepancy for several reasons. First, the trials rarely directly tested combinations of therapies or provided information about all other treatments, or reported outcomes separately in patient subgroups by concomitant treatments. Second, clinical care for adults with knee OA includes pharmacological interventions, while our review is limited to nonpharmacologic therapies,[154-156] thus further complicating our efforts. We tried to examine the how pain relievers (an extremely common concomitant treatment) may influence physical therapy outcomes for pain, but rare and inconsistent reporting of drug treatments impeded synthesis of evidence. Finally, heterogeneity in populations, treatments, and definitions of the outcomes hampered strength of evidence to low or moderate in most cases.

Most often, strength of evidence was low due to exclusion of patients from the analyses, inadequate allocation concealment, or unmasked outcome assessment. Few studies reported that the researchers who assessed outcomes were unaware of the treatment status of the patients.[65, 82, 83, 157-159] The majority of trials had moderate risk of bias. We excluded those with poorly reported results or high risk of bias. We also excluded trials that enrolled patients with knee or hip OA without separately reporting the outcomes. Moreover, many trials failed to provide sufficient details about the interventions themselves, their intensity, or the involvement of a physical therapist.[160, 161]

Examined physical therapy interventions included balance and coordination training, biofeedback and muscle relaxation techniques, strength, power, and endurance training, and functional training in self-care (Table 12). We tabulated the number of studies that described individual physical therapy modalities as part of physical therapy interventions. Yet, since very few studies gave exact descriptions of modality type and intensity, we found it difficult to assess how individual modalities contributed to treatment benefits.

Even when original studies did describe individual modalities, they rarely examined or reported the role of physical therapists or physical therapist assistants (Figure 4). Less than half of the studies described patient education or self training, two essential components of physical therapy practice.

Majority of strength exercises reported that physical therapists administered the interventions (Figure 5). In contrast, less than half of the trials with aquatic exercise, and only 28 percent of trials of aerobic exercise, stated that physical therapists administered interventions. Future efforts are needed to improve reporting quality of physical therapy studies. Consolidated Standards of Reporting Trials of nonpharmacologic treatments recommends that studies include a detailed description of the interventional components and, when applicable, individualized treatment recommendations as well as details about standardization of interventions and adherence of care providers with the protocol.[252] Such detailed reporting would shed light on how the direct involvement of physical therapists in treatment may contribute to benefits from exercise in adults with knee OA.

Monitoring Treatment Success

Patients judge treatment success by reduction in pain[253] and improvement in quality of life.[34] Clinical trials of nonpharmacologic treatments focused on pain and various measures of function.[254] Reimbursement for physical therapy practice is currently driven by the validated measures of functional impairments recommended by the *Guide to Physical Therapist Practice*.[25] However, the *Guide* recommends neither clinically important thresholds of change for such measures nor the monitoring of treatment effects according to patient-centered outcomes. Thus, the *Guide* does not direct physical therapists in routine assessment of patient-centered outcomes in clinical practice.

Variability in definitions and measurements of the outcomes affected our synthesis of evidence. Although outcomes were reported as average scores for all patients in each treatment group, the clinical importance of such averages was not evaluated. Average scores do not reflect how many patients develop disability or experience clinically meaningful improvements in pain, function, or quality of life. OARSI, however, has recommended that treatment success be evaluated according to patient-centered outcomes and clinically important differences in the WOMAC scale.[39, 162]

MCIDs refer to thresholds of change in outcomes measurements that result in clinically significant improvements in pain, disability, quality of life, and patient satisfaction with treatment. The studies described clinically important differences of 26 validated tests using the anchor method, which compares changes in scales with patient perceptions of improvement.[163, 164] For three scales (WOMAC, VAS for pain, and the Patient Global Assessment), PASS accounting for patient satisfaction was available. PASS is used to identify the level of symptom state that patients can tolerate while still considering their health state satisfactory and their treatment successful. PASS is gaining increasing recognition as a valid and reliable approach across many areas of medical practice, including rheumatology.[165] Expanded use of PASS would help to both improve the quality of physical therapy practice and the impact of studies examining physical therapy interventions.

In contrast, the studies we examined did not categorize patients according to meaningful improvement in pain, disability measures, or quality of life. Meaningful improvements in patient-centered outcomes should define treatment success in physical therapy practice. Current evidence is insufficient to determine an association between patient-centered outcomes and the measurements of functional impairments that currently drive reimbursement for physical therapy services. Future use of WOMAC in clinical trials, along with routine monitoring of treatment success in physical therapy practice, would produce robust cumulative evidence of the benefits of physical therapy modalities and interventions.

Limitations

Our report has several limitations**.** We relied on published information and did not contact the principal investigators of poorly reported or unpublished studies. Despite an exhaustive literature search, we cannot precisely estimate publication bias. We found that less than 30 percent of eligible studies in ClinicalTrials.gov were published**.**

Further, the research on this topic was also limited by several factors. Evidence to answer research questions was low or insufficient in most cases. Due to variability in definitions of the outcomes, we had to calculate standardized mean differences. Statistically significant differences in this construct do not necessarily reflect the clinical importance of improvement in outcomes.

Information about patient characteristics was not sufficient to conclude individualized treatment recommendations. Minorities and patients with comorbidities are at higher risk of disability and yet were underrepresented in clinical trials.

We did not evaluate adverse effects related to unmet patient expectations, insufficient use of patient and provider time and resources, or treatments that were not cost effective.

Future Research

Our report has implications for future research.[166] Benefits from physical therapy interventions should be defined as clinically important improvement in pain, independence in ADL, and quality of life. Treatment success should be estimated using rates of the patient-centered outcomes. Through meta-analyses of individual patient data from previously conducted RCTs, researchers could categorize patients according to the clinical importance of the changes they experience and analyze rates of patient-centered outcomes. Assuming investigators were willing to share their data, meta-analyses of individual patient data could also provide good estimates of treatment effects in patient subpopulations by age, comorbidity, severity of knee OA, and concomitant treatments. Future RCTs should examine comparative effectiveness of combined physical therapy treatments that consist of effective individual modalities or interventions. Finally, researchers should further examine the extent to which the benefits of exercise for adults with knee OA are affected by the supervision provided by physical therapists or their assistants.

Our report points to areas for future research. Table 13 links a research agenda with each key question.

Table 10. Conclusions about efficacy of physical therapy interventions

Physical therapy intervention	Studies/Subjects	Conclusions/Strength of evidence
Education program	Studies=3/Subjects=435	An education program improved disability (I) and health perception measures (I) but did not improve pain (M), gait (I) and composite measures of function (I)
Proprioception exercises	Studies=4/Subjects=225	Proprioception exercises improved pain (L) but did not improve joint (I), gait (L) and composite measures of function (L)
Aerobic exercises	Studies=12/ Subjects= 1,534	Aerobic exercises improved pain (L), disability (L), gait (L), and transfer (L) measures of function but did not improve psychological disability (L), global assessment (I), health perception (L), quality of life (I), joint (I) and composite measures of function (L)
Aquatic exercises	Studies=3/Subjects=314	Aquatic exercises improved disability (L) but did not improve pain (M), psychological disability (I), quality of life (M), and composite measures of function (I)
Strengthening exercises	Studies=9/ Subjects= 1,840	Strengthening exercises improved pain (L), global assessment (I), gait (L), joint (I), transfer (I), and composite (L) function measures but did not improve disability (L), psychological disability (L), health perception (L), and quality of life (L)
Tai Chi	Studies=3/Subjects=140	Tai Chi improved psychological disability (I) but did not improve pain (L), disability (L), gait (I), joint (L) and composite (L) function measures
Massage	Studies=3/Subjects=162	Massage improved pain (I), disability (I), health perception (I), gait (I) and composite (L) function measures but did not improve psychological disability (I) and joint (L) function measures
Joint mobilization	Studies=2/Subjects=119	Joint mobilization improved disability (I) and global assessment (I) but did not improve pain (I) and gait (I) function measures
Joint mobilization with exercise	Studies=1/Subjects=120	Joint mobilization with exercise improved disability (I) but did not improve gait (I) function measures
Orthotics	Studies=7/Subjects=510	Orthotics did not improve pain (I), disability (I), global assessment (I), joint (I), gait (H) and composite (L) function measures
Elastic subtalar strapping	Studies=3/Subjects=235	Elastic subtalar strapping improved composite function measures (L)
Taping	Studies=2/Subjects=94	Taping did not improve pain (I), disability (I), gait (I) and composite (I) function measures
E-stim	Studies=7/ Subjects=345	E-stim improved global assessment (L), but worsened pain (L), and did not improve disability (M), health perception (I), and gait (L), joint (M), transfer (I), and composite (M) function measures,
PEMF	Studies=4/Subjects=263	PEMF improved global assessment (I) but did not improve pain (M), disability (I), and gait (I), joint (I) and composite (M) function measures
Ultrasound	Studies=5/Subjects=353	Ultrasounds improved pain (L), gait (L) and composite (L) function measures but did not improve disability (L), and joint function measures (I)
Diathermy	Studies=5/ Subjects= 342	Diathermy did not improve pain (M), disability (M), psychological disability (I), global assessment (I), health perception (I), quality of life (I), and joint (M), gait (M) and composite (M) function measures
Heat	Studies=3/ Subjects=185	Heat improved disability (I) and quality of life (I), but did not improve pain (I), gait (I), joint (I), and composite (I) function measures
Mud pack	Studies=1/Subjects=60	Mud pack improved pain (I), disability (I) and global assessment measures (I)
Cryotherapy	Studies=2/Subjects=90	Cryotherapy did not improve pain (I), disability (I), quality of life (I), and composite function measures (I)

E-stim=Electrical stimulation; PEMF=Pulsed electromagnetic fields; strength of evidence as L: Low; M: Moderate; H: High; I: Insufficient; Level of evidence was determined according to four domains.

Figure 2. Pain at the longest time of followup with physical therapy interventions vs. no active treatment (pooled with random effects model at 95% CI standardized mean differences from head-to-head randomized trials)

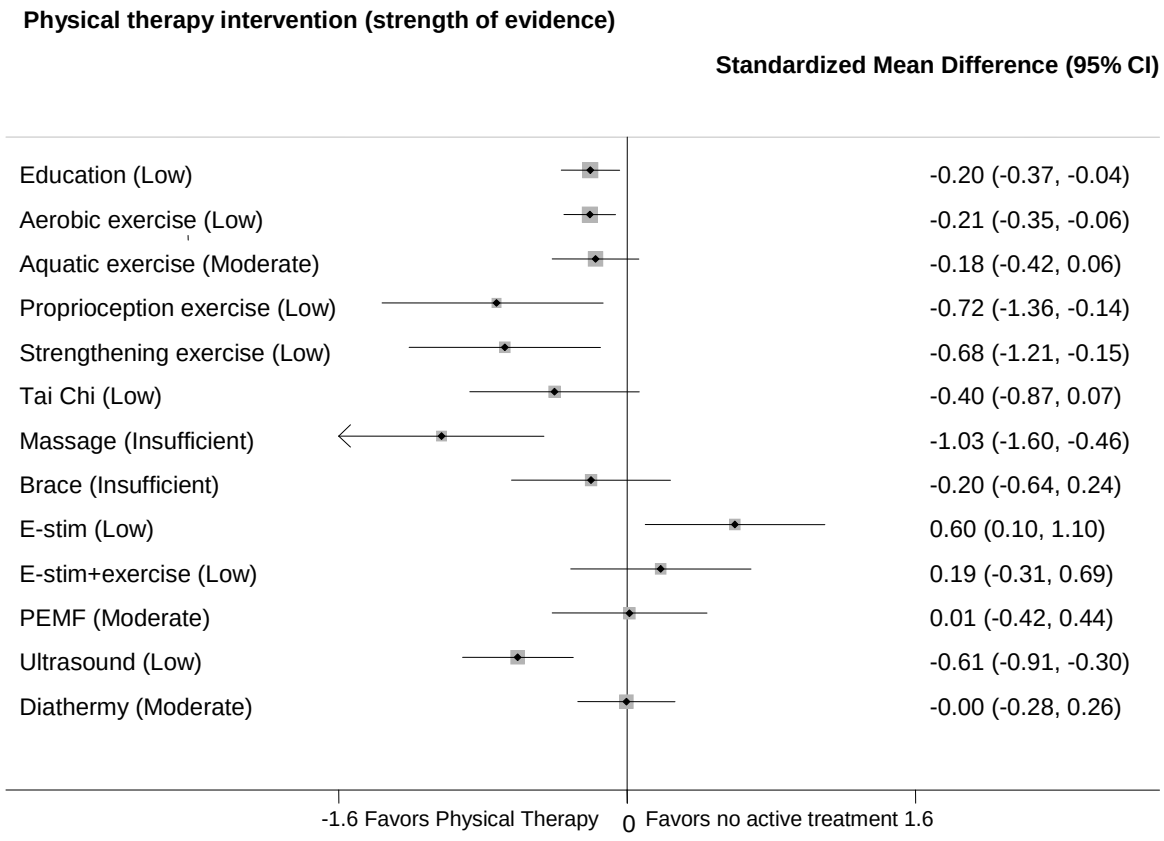


Figure 3. Ratio of means in pain measurements at the longest time of followup with physical therapy interventions vs. no active treatment (pooled with random effects model at 95% CI from head-to-head randomized trials)

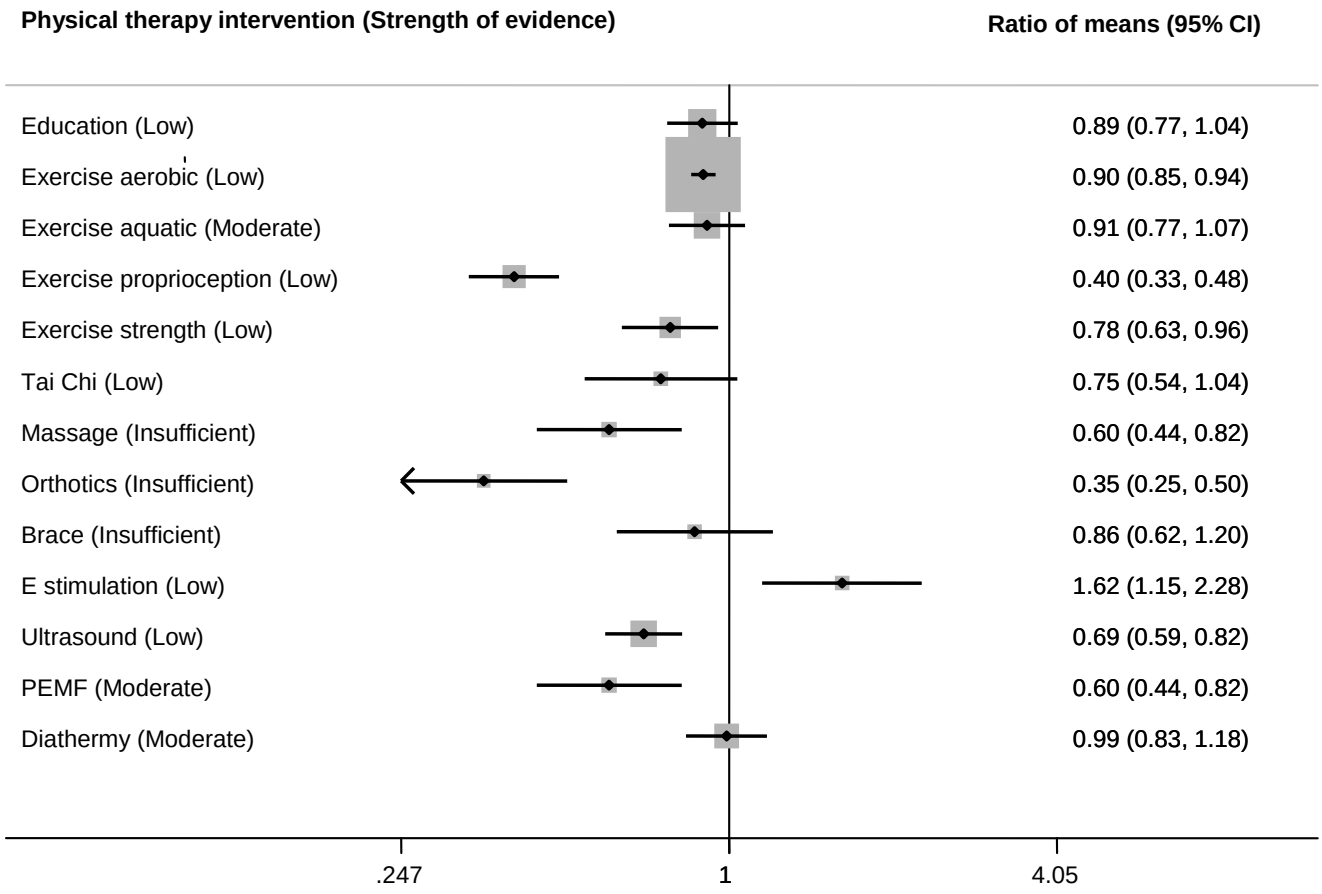


Table 11. Summary of comparative effectiveness of physical therapy interventions for knee osteoarthritis

Active vs. control physical therapy intervention	Studies/subjects	Conclusions/strength of evidence
Aerobic exercises vs. strengthening exercises	studies=2/ subjects=345	Aerobic exercises improved gait function measures (I) but did not improve pain (L), disability (L), transfer (I), and composite (I) function measures, compared to strengthening exercises
Aquatic exercises vs. aerobic exercises	studies=2/ subjects=116	Aquatic exercises did not improve pain (L), disability (I), quality of life (I), gait (I) and composite (I) function measures, compared to aerobic exercises
Proprioception exercises vs. strengthening exercises	studies=1/ subjects=72	Proprioception exercises worsened composite function measures (I) but did not improve pain (I), gait function (I), compared to strengthening exercises
Tai Chi vs. stretching exercises	studies=1/ subjects=40	Tai Chi improved disability (I), psychological disability (I), and transfer function (I) but did not improve pain (I), global assessment (I), gait (I), joint (I), and composite (I) function measures, compared to stretching exercise
Laterally vs. neutrally wedged insole	studies=5 subjects=734	Laterally wedged insole did not improve pain (L), disability (I), global assessment (I), quality of life (I), gait (I), joint (I), and composite function measures (M), compared to neutrally wedged insole
Orthotics vs. brace	studies=1/ subjects=91	Orthotics did not improve pain (I) and composite function measures (I), compared to brace
E-stim vs. exercises	studies=3/ subjects=112	E-stim improved joint (I) and composite (I) measures of function but did not improve pain (L) and gait (I) function, compared to exercises
E-stim vs. ultrasound	studies= 1/ subjects=35	E-stim did not improve pain (I), gait (I) and composite (I) function measures, compared to ultrasound

E-stim=Electrical stimulation; strength of evidence as L: Low; M: Moderate; I: Insufficient; Level of evidence was determined according to four domains.

Table 12. Reported physical therapy modalities as components of the examined physical therapy interventions

Components of the intervention (modalities)	Category	Exercise + education (# of studies)	Exercise aerobic (# of studies)	Exercise aquatic (# of studies)	Exercise proprioception (# of studies)	Exercise strength (# of studies)	Exercise strength (CER) (# of studies)
Aerobic capacity	No	6	3	2	4	19	12
Aerobic capacity	Unclear	3	0	0	0	0	0
Aerobic capacity	Yes	1	29	9	0	6	0
Balance, coordination, and agility training	No	7	29	8	1	22	10
Balance, coordination, and agility training	Unclear	3	0	0	0	0	0
Balance, coordination, and agility training	Yes	0	3	3	3	3	2
Education	No	3	17	10	2	24	11
Education	Yes	7	15	1	2	1	1
Exercise	No	5	1	0	0	2	0
Exercise	Yes	5	31	11	4	23	12
Flexibility exercise	No	7	16	5	4	14	5
Flexibility exercise	Unclear	3	0	0	0	0	0
Flexibility exercise	Yes	0	16	6	0	11	7
Functional training in self-care	No	4	22	11	4	20	9
Functional training in self-care	Yes	6	10	0	0	5	3
Gait training	No	7	31	11	3	23	12
Gait training	Unclear	3	0	0	0	0	0
Gait training	Yes	0	1	0	1	2	0
Manual therapy	No	10	31	11	4	24	12
Manual therapy	Yes	0	1	0	0	1	0
Muscle relaxation technique	No	7	32	11	4	25	12
Muscle relaxation technique	Unclear	3	0	0	0	0	0
Muscle relaxation technique	Yes	0	0	0	0	0	0
Strength, power, and endurance training	No	7	4	1	2	2	0
Strength, power, and endurance training	Unclear	3	0	0	0	0	0
Strength, power, and endurance training	Yes	0	28	10	2	23	12
	Total	10	32	11	4	25	12
		3.89	12.45	4.28	1.56	9.73	4.67

Figure 4. Physical therapists' involvement in administering modalities from trials that examined the effectiveness of physical therapy interventions

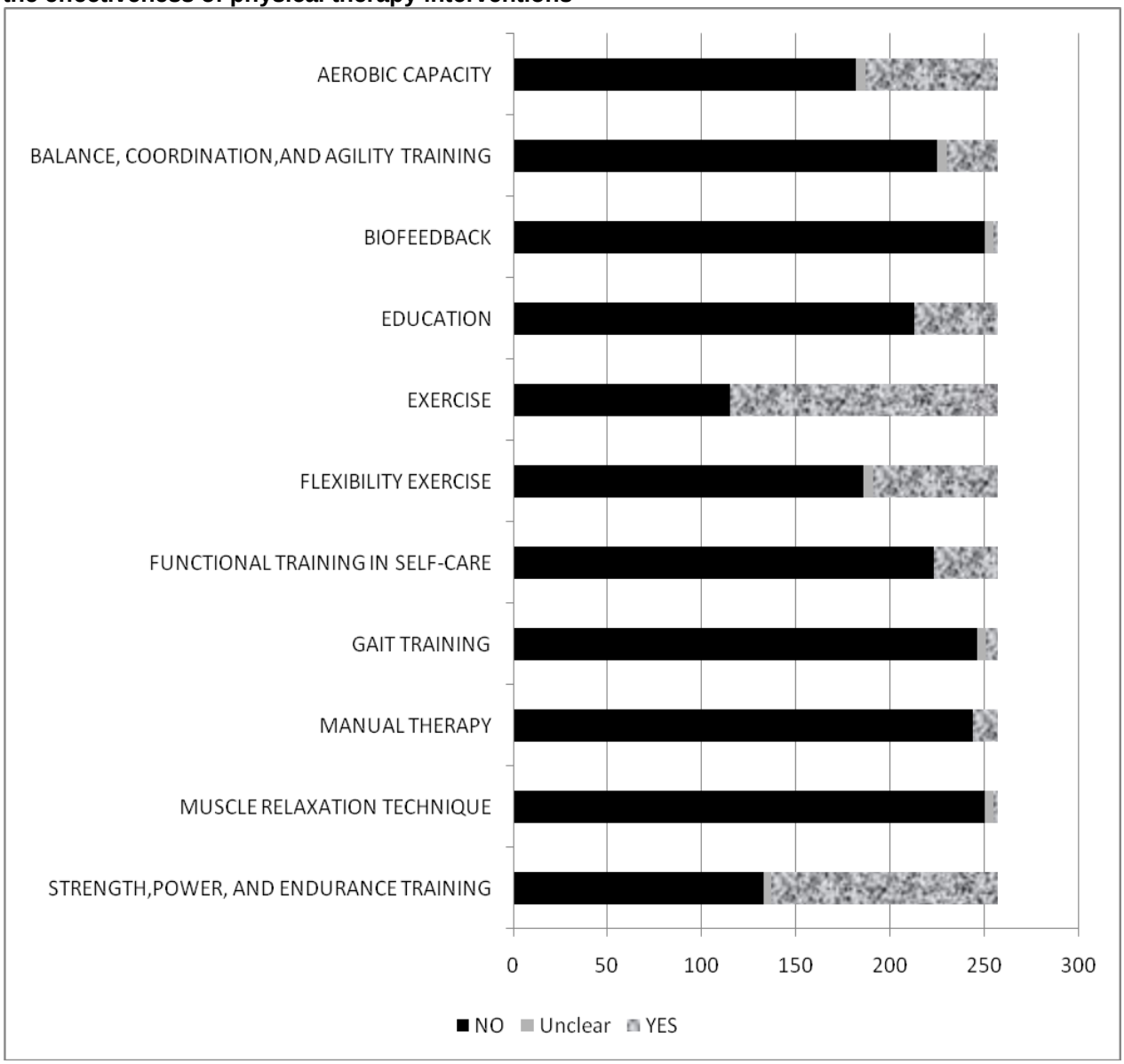


Figure 5. Physical therapists' involvement in trials that examined the effectiveness of physical therapy interventions

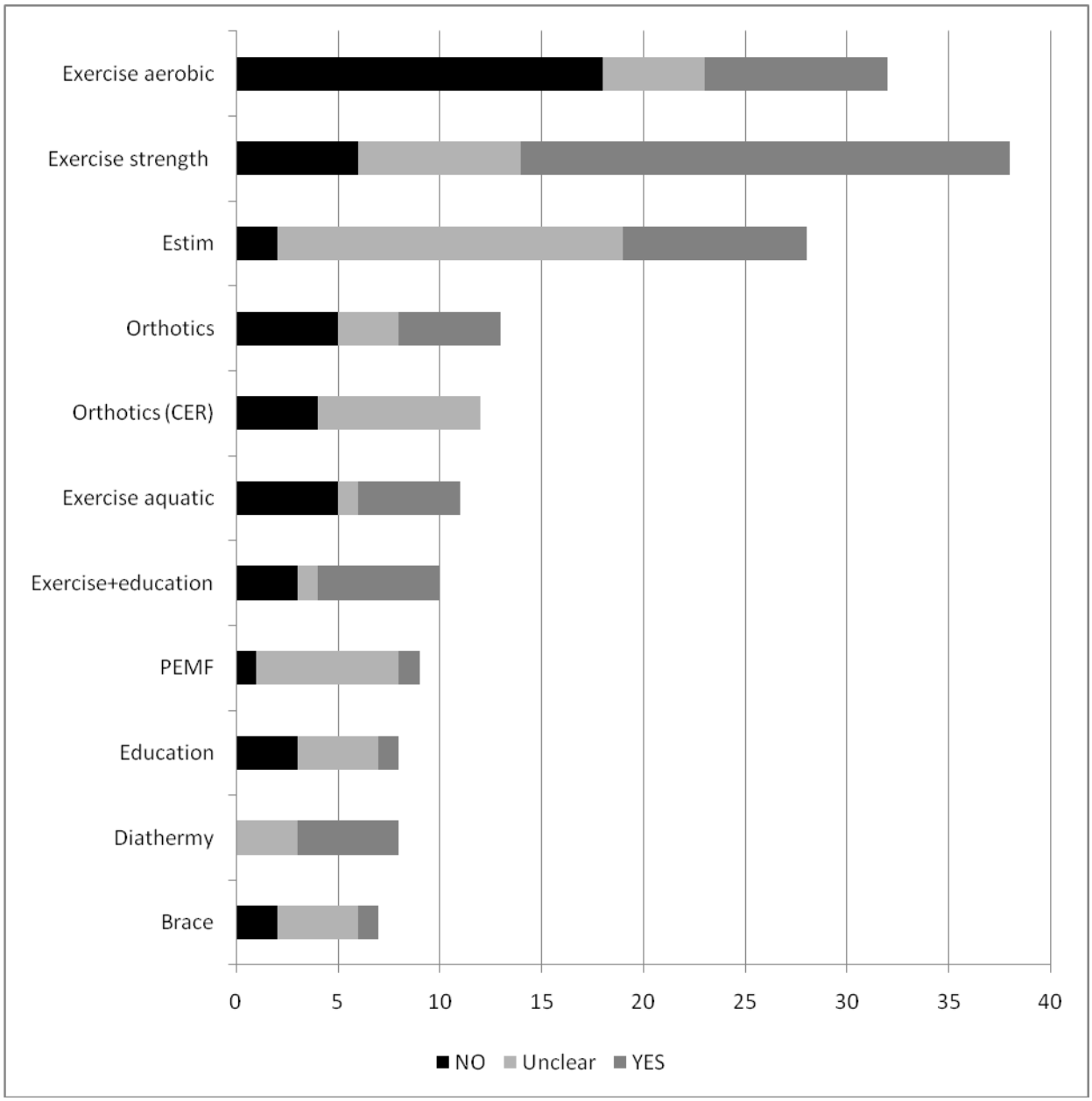


Table 13. Future research recommendations

Key question	Results of literature review	Types of studies needed to answer question	Future research recommendation
What are the effectiveness and comparative effectiveness of available physical therapy interventions (without drug treatment) for adult patients with chronic knee pain due to OA on intermediate and patient-centered outcomes when compared to no active treatment or another active physical therapy modality?	Few physical therapy interventions were shown to be effective. No single physical therapy intervention was shown to improve all outcomes. Research focused on individual physical therapy interventions, in contrast with common physical therapy practice of combining interventions. Pooled analyses demonstrated that: - Aerobic and aquatic exercise improved disability measures - Aerobic and strengthening exercise reduced pain and improved function - Proprioceptive exercise reduced pain -Ultrasound improved function Pooled analyses also demonstrated that the following physical therapy interventions failed to show any benefits: - Specific education program - Tai Chi - Diathermy - Orthotics - Magnetic stimulation Individual (nonpooled) RCTs failed to show consistent significant, strong, or clinically important changes in outcomes. Evidence about comparative effectiveness of physical therapy interventions was limited. Pooled analyses demonstrated that: - Disability measures did not differ with aerobic exercise vs. aquatic or vs. strengthening exercise - Pain did not differ with electrical stimulation vs. exercise in pooled analyses Individual RCTs of other treatment comparisons did not find consistent clinically important differences in outcomes and provided insufficient evidence to conclude the best treatment option.	Meta-analyses of individual patient data Randomized controlled clinical trials.	*Categorize patient outcomes according to clinically important improvement in pain, disability, function, and quality of life*. Examine combined interventions that reflect practice. *Provide detailed information about fidelity of the treatments and involvement of physical therapists and their assistants*. Examine preventive exercise interventions and self management of OA on incidence of disability in community, primary care, and physical therapy settings. *Assess the patient-centered outcomes with robust validated scales (WOMAC) and according to important for the patient's improvement in pain, function, and quality of life.*

Table 13. Future research recommendations (continued)

Key question	Results of literature review	Types of studies needed to answer question	Future research recommendation
Which patient characteristics are associated with the benefits of examined interventions of physical therapy on intermediate and patient-centered outcomes?	Evidence from individual RCTs was insufficient to conclude differences in effects by patient characteristics. Better treatment response was consistently reported in exercise subgroups with high compliance.	Meta-analyses of individual patient data Randomized controlled clinical trials Observational studies	*Subgroup analyses by patient age, severity of OA, multi-joint OA, prior and concomitant treatments success.* The association between patient modifiable risk factors for disability due to knee OA and incidence of pain, disability, and impaired quality of life.
Do changes in intermediate and patient-centered outcomes differ by the dose, duration, intensity, and frequency of examined interventions of physical therapy?	The duration of examined interventions was not associated with better intermediate/ patient-centered outcomes. Evidence regarding the association between the dose/intensity/frequency of examined interventions and outcomes was insufficient for the majority of comparisons.	Meta-analyses of individual patient data Randomized controlled clinical trials Observational studies	*Request from the authors detailed information about dose, included modalities, duration, intensity, and frequency of examined interventions. Re-analyze the conducted studies according to this information.* *Design trials that have enough power to detect differences in patient-centered outcomes according to dose, included modalities, duration, intensity, and frequency of examined interventions.* Use administrative database to analyze treatment utilization (drug, surgery) according to dose, included modalities, duration, intensity, and frequency of physical therapy interventions
Do changes in intermediate and patient-centered outcomes differ by the time of followup?	Among those treatments that demonstrated significant improvement in outcomes, the effects of aerobic, aquatic, and strengthening exercises and ultrasound did not differ at shorter vs. longer time of followup. Electrical stimulation improved pain at short term of followup but significantly worsened pain at longer time of followup	Randomized controlled clinical trials. Observational studies	*Design trials that have enough power to detect differences in long-term patient-centered outcomes* Analyze treatment utilization (drug, surgery) according to time of followup after physical therapy interventions
What is the association between changes in intermediate outcomes with changes in patient-centered outcomes after physical therapy interventions?	Gait, mobility restrictions, muscle strength, and range of motion measures were associated with disability measures. Individual studies offered insufficient evidence for determining which intermediate outcomes strongly and consistently predict patient-centered outcomes	Meta-analyses of individual patient data Randomized controlled clinical trials Observational studies	*Examine the association between responses in intermediate with reasons in patient-centered outcomes. Response must be defined as clinically important changes in measurements and scales.* Use administrative databases in prospective and/pr retrospective analyses of treatment utilization (drug, surgery) and incidence of disability according to response in intermediate outcomes driving reimbursement for physical therapy services.
What is the validity of the tests and measures used to	Validation was reported in many articles but few demonstrated a strong (more than 50%) correlation between index and	Observational validation studies	*Define validity according to strength of the association and diagnostic value of the tests*

Table 13. Future research recommendations (continued)

Key question	Results of literature review	Types of studies needed to answer question	Future research recommendation
determine intermediate outcomes of physical therapy on OA in association with patient-centered outcomes?	reference method measurements. Original studies concluded that tests are valid based on significance, not strength of correlation.		
Which intermediate outcomes meet the criteria of surrogates for patient-centered outcomes?	None of the intermediate outcomes met surrogate criteria for patient-centered outcomes.	Randomized controlled clinical trials	Examine the responses in intermediate measurements for predicting patient-centered outcomes with PT interventions. Response must be defined as clinically important changes in measurements and scales.
What are minimal clinically important differences of the tests and measures used to determine intermediate outcomes?	Minimal clinically important differences of the tests were determined using the anchor method that compared changes in scales with patient perception of improvements. Minimal clinically important differences were available as absolute change in score or relative change as a percent difference from baseline levels counting for baseline severity of the disease. Definition of the PASS that accounted for patient satisfaction was available for WOMAC scale, VAS for pain, and for Patient Global Assessment Scale. Validated tools defined threshold values of clinical importance for evaluating treatment success in adults with knee OA. In contrast, more often studies used continuous measures of the outcomes providing an average score for all patients in each treatment group. Clinical importance of such averages was not evaluated. Average scores did not provide information on how many patients developed disability or experienced clinically meaningful improvement in pain, function, or quality of life.	Randomized controlled clinical trials Observational studies Policy evaluation	Define PASS for intermediate outcomes driving reimbursement for physical therapy services. Examine whether payment decisions based on clinically important improvement in quality of life reduce incidence of disability

Table 13. Future research recommendations (continued)

Key question	Results of literature review	Types of studies needed to answer question	Future research recommendation
What are the harms from physical therapy interventions available for adult patients with chronic knee pain due to osteoarthritis when compared to no active treatment or active controls?	Adverse events were uncommon and varied across interventions, and included: skin irritation with brace/insole/tape/e-stim, swelling with brace/diathermy/exercise, muscle soreness with e-stim, warming/throbbing sensation with diathermy/e-stim/PEMF, increased pain with diathermy/exercise/insole/PEMF, and falls with insole. Adverse events were not severe enough to deter participants from continuing treatment	Randomized controlled clinical trials Observational studies Cost effectiveness analyses	Collect information about all undesirable events patient experienced irrespective of provider opinion about relevance to physical therapy interventions.

References

1. Lawrence RC, Felson DT, Helmick CG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum. 2008 Jan;58(1):26-35. PMID 18163497.
2. Johnson CA. Chapter 12. Approach to the Patient with Knee Pain. CURRENT Rheumatology Diagnosis & Treatment, 2e.
3. Jordan JM, Helmick CG, Renner JB, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol. 2007 Jan;34(1):172-80. PMID 17216685.
4. Dillon CF, Rasch EK, Gu Q, et al. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. J Rheumatol. 2006 Nov;33(11):2271-9. PMID 17013996.
5. Oliveria SA, Felson DT, Reed JI, et al. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. Arthritis Rheum. 1995 Aug;38(8):1134-41. PMID 7639811.
6. Felson DT, Zhang Y, Hannan MT, et al. The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. Arthritis Rheum. 1995 Oct;38(10):1500-5. PMID 7575700.
7. Bernstein AB, Hing E, Moss AJ, et al. Health care in America: Trends in utilization. Hyattsville, Maryland: National Center for Health Statistics; 2003.
8. Imboden J. Chapter 4. Approach to the Patient with Arthritis. CURRENT Rheumatology Diagnosis & Treatment, 2e.
9. National Collaborating Centre for Chronic Conditions. Osteoarthritis: national clinical guideline for care and management in adults. Royal College of Physicians. London; 2008.
10. American Academy of Orthopaedic Surgeons treatment of osteoarthritis of the knee (non-arthroplasty). American Academy of Orthopaedic Surgeons (AAOS). 2008 Dec 6:263.
11. Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage. 2008 Feb;16(2):137-62. PMID 18279766.
12. Jordan KM, Arden NK, Doherty M, et al. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis. 2003 Dec;62(12):1145-55. PMID 14644851.
13. Andersson GBJ, Bouchard J, Bozic KJ, et al. Arthritis and related conditions. The Burden of Musculoskeletal Diseases. Rosemont, IL: Bone and Joint Decade; 2008:71-96.
14. Andersson GBJ, Bouchard J, Bozic KJ, et al. Health care utilization and economic cost of musculoskeletal diseases. The Burden of Musculoskeletal Diseases. Rosemont, IL: Bone and Joint Decade; 2008:195-225.
15. Richmond J, Hunter D, Irrgang J, et al. American Academy of Orthopaedic Surgeons clinical practice guideline on the treatment of osteoarthritis (OA) of the knee. Journal of Bone & Joint Surgery - American Volume. 2010 Apr;92(4):990-3. PMID 20360527.
16. Michael JW, Schluter-Brust KU, Eysel P. The epidemiology, etiology, diagnosis, and treatment of osteoarthritis of the knee. Dtsch Arztebl Int. 2010 Mar;107(9):152-62. PMID 20305774.
17. Felson DT. Clinical practice. Osteoarthritis of the knee. N Engl J Med. 2006 Feb 23;354(8):841-8. PMID 16495396.
18. Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol. 1997 Apr;24(4):799-802. PMID 9101522.
19. Rutjes AW, Nuesch E, Sterchi R, et al. Transcutaneous electrostimulation for osteoarthritis of the knee. Cochrane Database Syst Rev. 2009(4):CD002823. PMID 18 PMID: 19821296.

20. Brouwer RW, Jakma TS, Verhagen AP, et al. Braces and orthoses for treating osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2005(1):004020. PMID 15674927.

21. Rutjes AW, Nuesch E, Sterchi R, et al. Therapeutic ultrasound for osteoarthritis of the knee or hip. Cochrane Database Syst Rev. 2010;20(1):CD003132. PMID 9 PMID: 20091539.

22. Brosseau L, Yonge KA, Robinson V, et al. Thermotherapy for treatment of osteoarthritis. Cochrane Database Syst Rev. 2003(4):CD004522. PMID 14584019.

23. Norris S, Atkins D, Bruening W, et al. Selecting observational studies for comparing medical interventions. Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews. 2010.

24. Norris S, Atkins D, Bruening W, et al. Selecting Observational Studies for Comparing Medical Interventions. 2008PMID 21433401.

25. Impaired Joint Mobility, Motor Function, Muscle Performance, and Range of Motion Associated with Localized Inflammation. Guide to Physical Therapist Practice, 2nd ed. Phys Ther. 2010.

26. Chou R, Aronson N, Atkins D, et al. Assessing harms when comparing medical interventions: AHRQ and the Effective Health-Care Program. J Clin Epidemiol. 2008 Sep 25PMID 18823754.

27. Higgins J, Green S, Cochrane Collaboration. Cochrane handbook for systematic reviews of interventions. Chichester, West Sussex ; Hoboken NJ: John Wiley & Sons; 2008.

28. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006;6:9. PMID 16519814.

29. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID 19595577.

30. Friedrich JO, Adhikari NK, Beyene J. Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods. Journal of Clinical Epidemiology. 2011 May;64(5):556-64. PMID 21447428.

31. Fu R, Gartlehner G, Grant M, et al. Conducting Quantitative Synthesis When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. Methods Guide for Comparative Effectiveness Reviews. 2010.

32. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60. PMID 12958120.

33. Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of meta-analyses on the basis of quality scores should be abandoned. J Clin Epidemiol. 2006 Dec;59(12):1249-56. PMID 17098567.

34. Agency for Healthcare Research and Quality. HHS Awards $473 Million in Patient-Centered Outcomes Research Funding. Agency for Healthcase Research and Quality, Rockville, MD; 2010.

35. Wallace BC, Schmid CH, Lau J, et al. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. BMC Med Res Methodol. 2009;9:80. PMID 19961608.

36. Egger M, Smith GD, Altman DG. Systematic Reviews in Health Care. London: NetLibrary, Inc. BMJ Books; 2001.

37. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Second Edition. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.

38. Chou R, Helfand M, Peterson K, et al. 2006 SepPMID 20704046.

39. Pham T, Van Der Heijde D, Lassere M, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. Journal of Rheumatology. 2003;30(7):1648-54. PMID 12858473.

40. Viechtbauer W. Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med. 2006 Feb 6PMID 16463355.

41. Knapp G, Biggerstaff BJ, Hartung J. Assessing the amount of heterogeneity in random-effects meta-analysis. Biom J. 2006 Apr;48(2):271-85. PMID 16708778.

42. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986 Sep;7(3):177-88. PMID 3802833.

43. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ. 2004 Jul 17;329(7458):168-9. PMID 15258077.

44. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. BMJ. 1994 Jul 16;309(6948):188. PMID 8044101.

45. Lassere MN, Johnson KR, Boers M, et al. Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. J Rheumatol. 2007 Mar;34(3):607-15. PMID 17343307.

46. Farr JN, Going SB, McKnight PE, et al. Progressive resistance training improves overall physical activity levels in patients with early osteoarthritis of the knee: a randomized controlled trial. Phys Ther. 2010 Mar;90(3):356-66. PMID 20056719.

47. Messier SP, Loeser RF, Miller GD, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. Arthritis Rheum. 2004 May;50(5):1501-10. PMID 15146420.

48. Rejeski WJ, Focht BC, Messier SP, et al. Obese, older adults with knee osteoarthritis: weight loss, exercise, and quality of life. Health Psychology. 2002;21(5):419-26. PMID 12211508.

49. Weng MC, Lee CL, Chen CH, et al. Effects of different stretching techniques on the outcomes of isokinetic exercise in patients with knee osteoarthritis. Kaohsiung J Med Sci. 2009;25(6):306-15. PMID 57 PMID: 19560995.

50. Lin DH, Lin CH, Lin YF, et al. Efficacy of 2 non-weight-bearing interventions, proprioception training versus strength training, for patients with knee osteoarthritis: a randomized clinical trial. J Orthop Sports Phys Ther. 2009;39(6):450-7. PMID 60 PMID: 19531879.

51. Tsauo JY, Cheng PF, Yang RS. The effects of sensorimotor training on knee proprioception and function for patients with knee osteoarthritis: a preliminary report. Clin Rehabil. 2008;22(5):448-57. PMID 135 PMID: 18441041.

52. Jan MH, Tang PF, Lin JJ, et al. Efficacy of a target-matching foot-stepping exercise on proprioception and function in patients with knee osteoarthritis. J Orthop Sports Phys Ther. 2008 Jan;38(1):19-25. PMID 18357655.

53. Peloquin L, Bravo G, Gauthier P, et al. Effects of a Cross-Training Exercise Program in Persons with Osteoarthritis of the Knee A Randomized Controlled Trial. J Clin Rheumatol. 1999 Jun;5(3):126-36. PMID 19078371.

54. Keefe FJ, Blumenthal J, Baucom D, et al. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. Pain. 2004 Aug;110(3):539-49. PMID 15288394.

55. An B, Dai K, Zhu Z, et al. Baduanjin alleviates the symptoms of knee osteoarthritis. J Altern Complement Med. 2008 Mar;14(2):167-74. PMID 18315512.

56. Ettinger WH, Jr., Burns R, Messier SP, et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). Journal - American Medical Association. 1997;277(1):25-31. PMID 8980206.

57. Yip YB, Sit JW, Wong DY, et al. A 1-year follow-up of an experimental study of a self-management arthritis programme with an added exercise component of clients with osteoarthritis of the knee. Psychology Health & Medicine. 2008;13(4):402-14. PMID 18825579.

58. Sullivan T, Allegrante JP, Peterson MG, et al. One-year followup of patients with osteoarthritis of the knee who participated in a program of supervised fitness walking and supportive patient education. Arthritis Care Res. 1998 Aug;11(4):228-33. PMID 9791321.

59. Hay EM, Foster NE, Thomas E, et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomised trial. BMJ. 2006 Nov 11;333(7576):995. PMID 17056608.

60. Thorstensson CA, Roos EM, Petersson IF, et al. Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis: a randomized controlled trial [ISRCTN20244858]. BMC Musculoskeletal Disorders. 2005;6(Journal Article):27. PMID 15924620.

61. Aglamis B, Toraman NF, Yaman H. The effect of a 12-week supervised multicomponent exercise program on knee OA in Turkish women. Journal of Back and Musculoskeletal Rehabilitation. 2008;21(2):121-8.

62. Focht BC, Rejeski WJ, Ambrosius WT, et al. Exercise, self-efficacy, and mobility performance in overweight and obese older adults with knee osteoarthritis. Arthritis Rheum. 2005 Oct 15;53(5):659-65. PMID 16208674.

63. Patrick DL, Ramsey SD, Spencer AC, et al. Economic evaluation of aquatic exercise for persons with osteoarthritis. Med Care. 2001 May;39(5):413-24. PMID 11317090.

64. Lund H, Weile U, Christensen R, et al. A randomized controlled trial of aquatic and land-based exercise in patients with knee osteoarthritis. J Rehabil Med. 2008 Feb;40(2):137-44. PMID 18509579.

65. Rooks DS, Huang J, Bierbaum BE, et al. Effect of preoperative exercise on measures of functional status in men and women undergoing total hip and knee arthroplasty. Arthritis Rheum. 2006 Oct 15;55(5):700-8. PMID 17013852.

66. Kuptniratsaikul V, Tosayanonda O, Nilganuwong S, et al. The efficacy of a muscle exercise program to improve functional performance of the knee in patients with osteoarthritis. J Med Assoc Thai. 2002 Jan;85(1):33-40. PMID 12075718.

67. Baker KR, Nelson ME, Felson DT, et al. The efficacy of home based progressive strength training in older adults with knee osteoarthritis: a randomized controlled trial. J Rheumatol. 2001 Jul;28(7):1655-65. PMID 11469475.

68. Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. Rheumatol Int. 2002 Jul;22(3):103-6. PMID 12111084.

69. Topp R, Woolley S, Hornyak J, 3rd, et al. The effect of dynamic versus isometric resistance training on pain and functioning among adults with osteoarthritis of the knee. Arch Phys Med Rehabil. 2002 Sep;83(9):1187-95. PMID 12235596 1187.

70. Borjesson M, Robertson E, Weidenhielm L, et al. Physiotherapy in knee osteoarthrosis: effect on pain and walking. Physiother Res Int. 1996;1(2):89-97. PMID 9238726.

71. Thomas KS, Muir KR, Doherty M, et al. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. British medical journal. 2002;325(7367):752-5. PMID 12364304.

72. Schilke JM, Johnson GO, Housh TJ, et al. Effects of muscle-strength training on the functional status of patients with osteoarthritis of the knee joint. Nursing Research. 1996 Mar-Apr;45(2):68-72. PMID 8604366.

73. Hurley MV, Scott DL. Improvements in quadriceps sensorimotor function and disability of patients with knee osteoarthritis following a clinically practicable exercise regime. Br J Rheumatol. 1998 Nov;37(11):1181-7. PMID 9851266.

74. Gur H, Cakin N, Akova B, et al. Concentric versus combined concentric-eccentric isokinetic training: effects on functional capacity and symptoms in patients with osteoarthrosis of the knee. Arch Phys Med Rehabil. 2002 Mar;83(3):308-16. PMID 11887109.

75. Cheing GL, Hui-Chan CW, Chan KM. Does four weeks of TENS and/or isometric exercise produce cumulative reduction of osteoarthritic knee pain? Pain Reviews. 2002;9(3-4):141-51.

76. Cheing GL, Hui-Chan CW. Would the addition of TENS to exercise training produce better physical performance outcomes in people with knee osteoarthritis than either intervention alone? Clin Rehabil. 2004 Aug;18(5):487-97. PMID 15293483.

77. Huang MH, Lin YS, Lee CL, et al. Use of ultrasound to increase effectiveness of isokinetic exercise for knee osteoarthritis. Arch Phys Med Rehabil. 2005 Aug;86(8):1545-51. PMID 16084806.

78. Huang MH, Yang RC, Lee CL, et al. Preliminary results of integrated therapy for patients with knee osteoarthritis. Arthritis Rheum. 2005 Dec 15;53(6):812-20. PMID 16342083.

79. Jan MH, Lin JJ, Liau JJ, et al. Investigation of clinical effects of high- and low-resistance training for patients with knee osteoarthritis: a randomized controlled trial. Phys Ther. 2008;88(4):427-36. PMID 153 PMID: 18218827.

80. 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(4):258-69. PMID 141 PMID: 18356618.

81. Lim BW, Hinman RS, Wrigley TV, et al. Does knee malalignment mediate the effects of quadriceps strengthening on knee adduction moment, pain, and function in medial knee osteoarthritis? A randomized controlled trial. Arthritis Rheum. 2008;59(7):943-51. PMID 122 PMID: 18576289.

82. Bennell KL, Hunt MA, Wrigley TV, et al. Hip strengthening reduces symptoms but not knee load in people with medial knee osteoarthritis and varus malalignment: a randomised controlled trial. Osteoarthritis Cartilage. PMID 15 - 20175973.

83. Lee HJ, Park HJ, Chae Y, et al. Tai Chi Qigong for the quality of life of patients with knee osteoarthritis: a pilot, randomized, waiting list controlled trial. Clin Rehabil. 2009;23(6):504-11. PMID 71 PMID: 19389743.

84. Brismee JM, Paige RL, Chyu MC, et al. Group and home-based tai chi in elderly subjects with knee osteoarthritis: a randomized controlled trial. Clin Rehabil. 2007 Feb;21(2):99-111. PMID 17264104.

85. Song R, Roberts BL, Lee EO, et al. A randomized study of the effects of t'ai chi on muscle strength, bone mineral density, and fear of falling in women with osteoarthritis. Journal of Alternative & Complementary Medicine. 2010 Mar;16(3):227-33. PMID 20192907.

86. Yip YB, Tam AC. An experimental study on the effectiveness of massage with aromatic ginger and orange essential oil for moderate-to-severe knee pain among the elderly in Hong Kong. Complement Ther Med. 2008 Jun;16(3):131-8. PMID 18534325.

87. Ko T, Lee S, Lee D. Manual therapy and exercise for OA knee: Effects on muscle strength, proprioception, and functional performance. Journal of Physical Therapy Science. 2009;21(4):293-9.

88. Perlman AI, Sabina A, Williams AL, et al. Massage therapy for osteoarthritis of the knee: a randomized controlled trial. Arch Intern Med. 2006 Dec 11-25;166(22):2533-8. PMID 17159021.

89. Deyle GD, Allison SC, Matekel RL, et al. Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. Physical Therapy. 2005;85(12):1301-17. PMID 16305269.

90. Moss P, Sluka K, Wright A. The initial effects of knee joint mobilization on osteoarthritic hyperalgesia. Man Ther. 2007 May;12(2):109-18. PMID 16777467.

91. Pollard H, Ward G, Hoskins W, et al. The effect of a manual therapy knee protocol on osteoarthritic knee pain: a randomised controlled trial. J Can Chiropr Assoc. 2008 Dec;52(4):229-42. PMID 19066697.

92. Toda Y, Tsukimura N, Segal N. An optimal duration of daily wear for an insole with subtalar strapping in patients with varus deformity osteoarthritis of the knee. Osteoarthritis & Cartilage. 2005;13(4):353-60. PMID 15780649.

93. Bar-Ziv Y, Beer Y, Ran Y, et al. A treatment applying a biomechanical device to the feet of patients with knee osteoarthritis results in reduced pain and improved function: a prospective controlled study. BMC Musculoskelet Disord. 2010;11:179. PMID 20698991.

94. Hinman RS, Bowles KA, Bennell KL. Laterally wedged insoles in knee osteoarthritis: do biomechanical effects decline after one month of wear? BMC Musculoskelet.Disord. 2009;10:146.:146. PMID 31 PMID: 19939281.

95. Maly M, Culham E, Costigan P. Static and dynamic biomechanics of foot orthoses in people with medial compartment knee osteoarthritis. Clinical Biomechanics. 2002;17(8):603-10. PMID 12243720.

96. Kerrigan DC, Lelas JL, Goggins J, et al. Effectiveness of a lateral-wedge insole on knee varus torque in patients with knee osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2002;83(7):889-93. PMID 12098144.

97. Kuroyanagi Y, Nagura T, Matsumoto H, et al. The lateral wedged insole with subtalar strapping significantly reduces dynamic knee load in the medial compartment gait analysis on patients with medial knee osteoarthritis. Osteoarthritis & Cartilage. 2007 Aug;15(8):932-6. PMID 17391994.

98. Nigg BM, Emery C, Hiemstra LA. Unstable shoe construction and reduction of pain in osteoarthritis patients. Medicine & Science in Sports & Exercise. 2006;38(10):1701-8. PMID 17019290.

99. Toda Y, Segal N, Kato A, et al. Effect of a novel insole on the subtalar joint of patients with medial compartment osteoarthritis of the knee. Journal of Rheumatology. 2001;28(12):2705-10. PMID 11764221.

100. Toda Y, Tsukimura N. A six-month followup of a randomized trial comparing the efficacy of a lateral-wedge insole with subtalar strapping and an in-shoe lateral-wedge insole in patients with varus deformity osteoarthritis of the knee. Arthritis & Rheumatism. 2004;50(10):3129-36. PMID 15476225.

101. Toda Y, Tsukimura N. Influence of concomitant heeled footwear when wearing a lateral wedged insole for medial compartment osteoarthritis of the knee. Osteoarthritis and Cartilage. 2008;16(2):244-53. PMID 17693101.

102. Hinman RS, Crossley KM, McConnell J, et al. Efficacy of knee tape in the management of osteoarthritis of the knee: blinded randomised controlled trial. BMJ. 2003;327(7407):135. PMID 12869456.

103. Hinman RS, Bennell KL, Crossley KM, et al. Immediate effects of adhesive tape on pain and disability in individuals with knee osteoarthritis. Rheumatology. 2003;42(7):865-9. PMID 12730546.

104. Talbot LA, Gaines JM, Ling SM, et al. A home-based protocol of electrical muscle stimulation for quadriceps muscle strength in older adults with osteoarthritis of the knee. J Rheumatol. 2003 Jul;30(7):1571-8. PMID 12858461.

105. Cetin N, Aytar A, Atalay A, et al. Comparing hot pack, short-wave diathermy, ultrasound, and TENS on isokinetic strength, pain, and functional status of women with osteoarthritic knees: a single-blind, randomized, controlled trial. Am J Phys Med Rehabil. 2008;87(6):443-51. PMID 130 PMID: 18496246.

106. Garland D, Holt P, Harrington JT, et al. A 3-month, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of a highly optimized, capacitively coupled, pulsed electrical stimulator in patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2007 Jun;15(6):630-7. PMID 17303443.

107. Selfe TK, Bourguignon C, Taylor AG. Effects of noninvasive interactive neurostimulation on symptoms of osteoarthritis of the knee: a randomized, sham-controlled pilot study. J Altern Complement Med. 2008;14(9):1075-81. PMID 98 PMID: 19055333.

108. Gaines JM, Metter EJ, Talbot LA. The effect of neuromuscular electrical stimulation on arthritis knee pain in older adults with osteoarthritis of the knee. Appl Nurs Res. 2004 Aug;17(3):201-6. PMID 15343554.

109. Kang RW, Lewis PB, Kramer A, et al. Prospective randomized single-blinded controlled clinical trial of percutaneous neuromodulation pain therapy device versus sham for the osteoarthritic knee: a pilot study. Orthopedics. 2007 Jun;30(6):439-45. PMID 17598487.

110. Taylor P, Hallett M, Flaherty L. Treatment of osteoarthritis of the knee with transcutaneous electrical nerve stimulation. Pain. 1981 Oct;11(2):233-40. PMID 7033891.

111. Huang MH, Chen CH, Chen TW, et al. The effects of weight reduction on the rehabilitation of patients with knee osteoarthritis and obesity. Arthritis Care Res. 2000 Dec;13(6):398-405. PMID 14635316.

112. Law PP, Cheing GL. Optimal stimulation frequency of transcutaneous electrical nerve stimulation on people with knee osteoarthritis. J Rehabil Med. 2004 Sep;36(5):220-5. PMID 15626162.

113. Itoh K, Hirota S, Katsumi Y, et al. A pilot study on using acupuncture and transcutaneous electrical nerve stimulation (TENS) to treat knee osteoarthritis (OA). Chin Med. 2008;3:2. PMID 18312661.

114. Pietrosimone BG, Hart JM, Saliba SA, et al. Immediate effects of transcutaneous electrical nerve stimulation and focal knee joint cooling on quadriceps activation. Med Sci Sports Exerc. 2009 Jun;41(6):1175-81. PMID 19461552.

115. Ay S, Evcik D. The effects of pulsed electromagnetic fields in the treatment of knee osteoarthritis: a randomized, placebo-controlled trial. Rheumatol Int. 2009;29(6):663-6. PMID 100 PMID: 19015858.

116. Thamsborg G, Florescu A, Oturai P, et al. Treatment of knee osteoarthritis with pulsed electromagnetic fields: a randomized, double-blind, placebo-controlled study. Osteoarthritis & Cartilage. 2005;13(7):575-81. PMID 15979009.

117. Trock DH, Bollet AJ, Markoll R. The effect of pulsed electromagnetic fields in the treatment of osteoarthritis of the knee and cervical spine. Report of randomized, double blind, placebo controlled trials. J Rheumatol. 1994 Oct;21(10):1903-11. PMID 7837158.

118. Nicolakis P, Kollmitzer J, Crevenna R, et al. Pulsed magnetic field therapy for osteoarthritis of the knee--a double-blind sham-controlled trial. Wiener klinische Wochenschrift. 2002;114(15-16):678-84. PMID 12602111.

119. Ozgonenel L, Aytekin E, Durmusoglu G. A double-blind trial of clinical effects of therapeutic ultrasound in knee osteoarthritis. Ultrasound Med Biol. 2009 Jan;35(1):44-9. PMID 18829151.

120. Tascioglu F, Kuzgun S, Armagan O, et al. Short-term effectiveness of ultrasound therapy in knee osteoarthritis. Journal of International Medical Research. 2010 Jul-Aug;38(4):1233-42. PMID 20925995.

121. Rattanachaiyanont M, Kuptniratsaikul V. No additional benefit of shortwave diathermy over exercise program for knee osteoarthritis in peri-/post-menopausal women: an equivalence trial. Osteoarthritis Cartilage. 2008 Jul;16(7):823-8. PMID 18178111.

122. Fukuda TY, Alves da Cunha R, Fukuda VO, et al. Pulsed shortwave treatment in women with knee osteoarthritis: a multicenter, randomized, placebo-controlled clinical trial. Phys Ther. 2011 Jul;91(7):1009-17. PMID 21642511.

123. Laufer Y, Zilberman R, Porat R, et al. Effect of pulsed short-wave diathermy on pain and function of subjects with osteoarthritis of the knee: a placebo-controlled double-blind clinical trial. Clin Rehabil. 2005 May;19(3):255-63. PMID 15859526.

124. Akyol Y, Durmus D, Alayli G, et al. Does short-wave diathermy increase the effectiveness of isokinetic exercise on pain, function, knee muscle strength, quality of life, and depression in the patients with knee osteoarthritis? A randomized controlled clinical study. European journal of physical & rehabilitation medicine. 2010 Sep;46(3):325-36. PMID 20926998.

125. Callaghan MJ, Whittaker PE, Grimes S, et al. An evaluation of pulsed shortwave on knee osteoarthritis using radioleucoscintigraphy: a randomised, double blind, controlled trial. Joint Bone Spine. 2005 Mar;72(2):150-5. PMID 15797496.

126. Denegar CR, Dougherty DR, Friedman JE, et al. Preferences for heat, cold, or contrast in patients with knee osteoarthritis affect treatment response. Clin Interv Aging. 2010;5:199-206. PMID 20711439.

127. Silva LE, Valim V, Pessanha AP, et al. Hydrotherapy versus conventional land-based exercise for the management of patients with osteoarthritis of the knee: a randomized clinical trial. Physical Therapy. 2008;88(1):12-21. PMID 17986497.

128. Maillefert JF, Hudry C, Baron G, et al. Laterally elevated wedged insoles in the treatment of medial knee osteoarthritis: a prospective randomized controlled study. Osteoarthritis and Cartilage. 2001;9(8):738-45. PMID 11795993.

129. van Raaij TM, Reijman M, Brouwer RW, et al. Medial Knee Osteoarthritis Treated by Insoles or Braces: A Randomized Trial. Clin Orthop Relat Res. 2010;468(7):1926-32. PMID 13 PMID: 20177839.

130. Bennell KL, Bowles KA, Payne C, et al. Lateral wedge insoles for medial knee osteoarthritis: 12 month randomised controlled trial. BMJ. 2011;342:d2912. PMID 21593096.

131. Durmus D, Alayli G, Canturk F. Effects of quadriceps electrical stimulation program on clinical parameters in the patients with knee osteoarthritis. Clin Rheumatol. 2007 May;26(5):674-8. PMID 16897119.

132. Penninx BW, Messier SP, Rejeski WJ, et al. Physical exercise and the prevention of disability in activities of daily living in older persons with osteoarthritis. Archives of Internal Medicine. 2001 Oct 22;161(19):2309-16. PMID 11606146.

133. Penninx BW, Rejeski WJ, Pandya J, et al. Exercise and depressive symptoms: a comparison of aerobic and resistance exercise effects on emotional and physical function in older persons with high and low depressive symptomatology. J Gerontol B Psychol Sci Soc Sci. 2002 Mar;57(2):P124-32. PMID 11867660.

134. Rejeski WJ, Brawley LR, Ettinger W, et al. Compliance to exercise therapy in older participants with knee osteoarthritis: implications for treating disability. Medicine & Science in Sports & Exercise. 1997;29(8):977-85. PMID 9268953.

135. van Gool CH, Penninx BW, Kempen GI, et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis Rheum. 2005 Feb 15;53(1):24-32. PMID 15696558.

136. Brouwer RW, van Raaij TM, Verhaar JA, et al. Brace treatment for osteoarthritis of the knee: a prospective randomized multi-centre trial. Osteoarthritis & Cartilage. 2006;14(8):777-83. PMID 16563810.

137. Baker K, Goggins J, Xie H, et al. A randomized crossover trial of a wedged insole for treatment of knee osteoarthritis. Arthritis & Rheumatism. 2007;56(4):1198-203. PMID 17393448.

138. Mangani I, Cesari M, Kritchevsky SB, et al. Physical exercise and comorbidity. Results from the Fitness and Arthritis in Seniors Trial (FAST). Aging Clin Exp Res. 2006 Oct;18(5):374-80. PMID 17167301.

139. Fisher NM, Gresham G, Pendergast DR. Effects of a quantitative progressive rehabilitation program applied unilaterally to the osteoarthritic knee. Arch Phys Med Rehabil. 1993 Dec;74(12):1319-26. PMID 8259900.

140. Mikesky AE, Mazzuca SA, Brandt KD, et al. Effects of strength training on the incidence and progression of knee osteoarthritis. Arthritis Rheum. 2006 Oct 15;55(5):690-9. PMID 17013851.

141. Mangione KK, McCully K, Gloviak A, et al. The effects of high-intensity and low-intensity cycle ergometry in older adults with knee osteoarthritis. J Gerontol A Biol Sci Med Sci. 1999 Apr;54(4):M184-90. PMID 10219009.

142. Sen A, Gocen Z, Unver B, et al. The frequency of visits by the physiotherapist of patients receiving home-based exercise therapy for knee osteoarthritis. Knee. 2004 Apr;11(2):151-3. PMID 15066630.

143. Toda Y, Tsukimura N, Kato A. The effects of different elevations of laterally wedged insoles with subtalar strapping on medial compartment osteoarthritis of the knee. Archives of Physical Medicine & Rehabilitation. 2004;85(4):673-7. PMID 15083446.

144. Cheing GL, Tsui AY, Lo SK, et al. Optimal stimulation duration of tens in the management of osteoarthritic knee pain. J Rehabil Med. 2003 Mar;35(2):62-8. PMID 12691335.

145. Jensen H, Zesler R, Christensen T. Transcutaneous electrical nerve stimulation (TNS) for painful osteoarthrosis of the knee. Int J Rehabil Res. 1991;14(4):356-8. PMID 1783484.

146. Defrin R, Ariel E, Peretz C. Segmental noxious versus innocuous electrical stimulation for chronic pain relief and the effect of fading sensation during treatment. Pain. 2005 May;115(1-2):152-60. PMID 15836978.

147. Grimmer K. A controlled double blind study comparing the effects of strong burst mode TENS and high rate TENS on painful osteoarthritic knees. Australian Journal of Physiotherapy. 1992;38(1):49-56.

148. Rejeski WJ, Ettinger WH, Jr., Martin K, et al. Treating disability in knee osteoarthritis with exercise therapy: a central role for self-efficacy and pain. Arthritis Care Res. 1998 Apr;11(2):94-101. PMID 9668732.

149. Reuben DB, Siu AL, Kimpau S. The predictive validity of self-report and performance-based measures of function and health. J Gerontol. 1992 Jul;47(4):M106-10. PMID 1624692.

150. Verghese J, LeValley A, Hall CB, et al. Epidemiology of gait disorders in community-residing older adults. J Am Geriatr Soc. 2006 Feb;54(2):255-61. PMID 16460376.

151. Woo J, Ho SC, Yu AL. Walking speed and stride length predicts 36 months dependency, mortality, and institutionalization in Chinese aged 70 and older. J Am Geriatr Soc. 1999 Oct;47(10):1257-60. PMID 10522962.

152. Jylha M, Guralnik JM, Balfour J, et al. Walking difficulty, walking speed, and age as predictors of self-rated health: the women's health and aging study. J Gerontol A Biol Sci Med Sci. 2001 Oct;56(10):M609-17. PMID 11584033.

153. Fransen M, McConnell S. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev. 2008(4):CD004376. PMID 42 PMID: 18843657.

154. Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines. Arthritis & Rheumatism. 2000 Sep;43(9):1905-15. PMID 11014340.

155. Brand C. Guideline for the non-surgical management of hip and knee osteoarthritis. 2009 July 2009.

156. Mazieres B, Bannwarth B, Dougados M, et al. EULAR recommendations for the management of knee osteoarthritis. Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials. Joint, Bone, Spine: Revue du Rhumatisme. 2001;68(3):231-40. PMID 11394623.

157. Jan MH, Lin CH, Lin YF, et al. Effects of weight-bearing versus nonweight-bearing exercise on function, walking speed, and position sense in participants with knee osteoarthritis: a randomized controlled trial. Arch Phys Med Rehabil. 2009;90(6):897-904. PMID 63 PMID: 19480863.

158. Lund H, Henriksen M, Bartels EM, et al. Can stimulating massage improve joint repositioning error in patients with knee osteoarthritis? J Geriatr Phys Ther. 2009;32(3):111-6. PMID 22 PMID: 20128335.

159. Pisters MF, Veenhof C, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized controlled trial comparing two different physical therapy interventions. Osteoarthritis Cartilage. 2010 Aug;18(8):1019-26. PMID 20488250.

160. Bruckenthal P, Broderick JE. Assessing treatment fidelity in pilot studies assist in designing clinical trials: an illustration from a nurse practitioner community-based intervention for pain. Advances in Nursing Science. 2007;30(1):E72-84. PMID 17299277.

161. Doherty M, Jones A. Design of clinical trials in knee osteoarthritis: practical issues for debate. Osteoarthritis & Cartilage. 1998 Nov;6(6):371-3. PMID 10343768.

162. Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthritis & Cartilage. 2004;12(5):389-99. PMID 15094138.

163. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. Journal of Clinical Epidemiology. 2010 May;63(5):524-34. PMID 19926446.

164. Terwee CB, Roorda LD, Knol DL, et al. Linking measurement error to minimal important change of patient-reported outcomes. Journal of Clinical Epidemiology. 2009 Oct;62(10):1062-7. PMID 19230609.

165. Dougados M, Moore A, Yu S, et al. Evaluation of the patient acceptable symptom state in a pooled analysis of two multicentre, randomised, double-blind, placebo-controlled studies evaluating lumiracoxib and celecoxib in patients with osteoarthritis. Arthritis Research & Therapy. 2007;9(1):R11. PMID 17266764.

166. Fitzgerald GK, Delitto A. Considerations for planning and conducting clinic-based research in physical therapy. Phys Ther. 2001 Aug;81(8):1446-54. PMID 11509074.

167. Mazzuca SA, Brandt KD, Katz BP, et al. Comparison of general internists, family physicians, and rheumatologists managing patients with symptoms of osteoarthritis of the knee. Arthritis Care Res. 1997 Oct;10(5):289-99. PMID 9362595.

168. Ottawa panel evidence-based clinical practice guidelines for therapeutic exercises and manual therapy in the management of osteoarthritis. Physical Therapy. 2005 Sep;85(9):907-71.

169. Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage. 2010;18(4):476-99. PMID 4 PMID: 20170770.

170. Helfand M, Balshem H. AHRQ series paper 2: principles for developing guidance: AHRQ and the effective health-care program. J Clin Epidemiol. 2010 May;63(5):484-90. PMID 19716268.

171. Tallon D, Chard J, Dieppe P. Exploring the priorities of patients with osteoarthritis of the knee. Arthritis Care Res. 2000 Oct;13(5):312-9. PMID 14635301.

172. White CM, Ip S, McPheeters ML, et al. Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews [posted September 2009]. Rockville, MD; 2009.

173. AHRQ. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: 2011.

174. Methods Guide for Medical Test Reviews. Methods Guide for Medical Test Reviews. 2010.

175. Higgins JPT, Green S, Cochrane Collaboration. Cochrane handbook for systematic reviews of interviews. 5.0.2 ed. London: Cochrane Collaboration; 2009.

176. Aschengrau A, Seage GR. Essentials of Epidemiology in Public Health. Sudbury, Mass: Jones and Bartlett; 2003.

177. Egger M, Smith GD. Bias in location and selection of studies. BMJ. 1998 Jan 3;316(7124):61-6. PMID 9451274.

178. Dickersin K, Min YI. NIH clinical trials and publication bias. Online J Curr Clin Trials. 1993 Apr 28;Doc No 50:[4967 words; 53 paragraphs]. PMID 8306005.

179. Higgins J, Green S. The Cochrane Collaboration. The Cochrane handbook for systematic reviews of interventions. Chichester, UK: John Wiley & Sons, Ltd. Cochrane Collaboration; 2005. http://www.cochrane.org/resources/handbook/handbook.pdf2006.

180. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol. 2000 Feb;53(2):207-16. PMID 10729693.

181. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. PMID 21784880.

182. Ebrahim S. The use of numbers needed to treat derived from systematic reviews and meta-analysis. Caveats and pitfalls. Eval Health Prof. 2001 Jun;24(2):152-64. PMID 11523384.

183. Altman DG. Confidence intervals for the number needed to treat. Bmj. 1998 Nov 7;317(7168):1309-12. PMID 9804726.

184. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ. 1994 Jul 9;309(6947):102. PMID 8038641.

185. Shakoor MA, Taslim MA, Hossain MS. Effects of activity modification on the patients with osteoarthritis of the knee. Bangladesh Med Res Counc Bull. 2007 Aug;33(2):55-9. PMID 18481439.

186. Kovar PA, Allegrante JP, MacKenzie CR, et al. Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. Annals of Internal Medicine. 1992 Apr 1;116(7):529-34. PMID 1543305.

187. Bautch JC, Malone DG, Vailas AC. Effects of exercise on knee joints with osteoarthritis: a pilot study of biologic markers. Arthritis Care Res. 1997 Feb;10(1):48-55. PMID 9313390.

188. Talbot LA, Gaines JM, Huynh TN, et al. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc. 2003 Mar;51(3):387-92. PMID 12588583.

189. Yip YB, Sit JW, Fung KK, et al. Impact of an Arthritis Self-Management Programme with an added exercise component for osteoarthritic knee sufferers on improving pain, functional outcomes, and use of health care services: An experimental study. Patient Educ Couns. 2007 Jan;65(1):113-21. PMID 17010554.

190. Messier SP, Thompson CD, Ettinger WH. Effects of long-term aerobic or weight training regimes on gait in an older, osteoarthritic population. Journal of Applied Biomechanics. 1997;13(2):205-25.

191. Hughes SL, Seymour RB, Campbell R, et al. Impact of the fit and strong intervention on older adults with osteoarthritis. Gerontologist. 2004 Apr;44(2):217-28. PMID 15075418.

192. Deyle GD, Henderson NE, Matekel RL, et al. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. Annals of Internal Medicine. 2000;132(3):173-81. PMID 10651597.

193. Peterson MG, Kovar-Toledano PA, Otis JC, et al. Effect of a walking program on gait characteristics in patients with osteoarthritis. Arthritis Care Res. 1993 Mar;6(1):11-6. PMID 8443252.

194. Yurtkuran M, Kocagil T. TENS, electroacupuncture and ice massage: comparison of treatment for osteoarthritis of the knee. Am J Acupunct. 1999;27(3-4):133-40. PMID 10729968.

195. Law PP, Cheing GL, Tsui AY. Does transcutaneous electrical nerve stimulation improve the physical performance of people with knee osteoarthritis? J Clin Rheumatol. 2004 Dec;10(6):295-9. PMID 17043536.

196. Pietrosimone BG, Saliba SA, Hart JM, et al. Effects of disinhibitory transcutaneous electrical nerve stimulation and therapeutic exercise on sagittal plane peak knee kinematics and kinetics in people with knee osteoarthritis during gait: a randomized controlled trial. Clin Rehabil. 2010 Dec;24(12):1091-101. PMID 20713439.

197. Kirkley A, Webster-Bogaert S, Litchfield R, et al. The effect of bracing on varus gonarthrosis. J Bone Joint Surg Am. 1999 Apr;81(4):539-48. PMID 10225800.

198. Zizic TM, Hoffman KC, Holt PA, et al. The treatment of osteoarthritis of the knee with pulsed electrical stimulation. J Rheumatol. 1995 Sep;22(9):1757-61. PMID 8523357.

199. Odabasi E, Turan M, Erdem H, et al. Does mud pack treatment have any chemical effect? A randomized controlled clinical study. J Altern Complement Med. 2008 Jun;14(5):559-65. PMID 18564957.

200. Wang C, Schmid CH, Hibberd PL, et al. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. Arthritis Rheum. 2009;61(11):1545-53. PMID 35 PMID: 19877092.

201. Swank AM, Kachelman JB, Bibeau W, et al. Prehabilitation before total knee arthroplasty increases strength and function in older adults with severe osteoarthritis. Journal of Strength & Conditioning Research. 2011 Feb;25(2):318-25. PMID 21217530.

202. Bieleman HJ, Reneman MF, van Ittersum MW, et al. Self-reported functional status as predictor of observed functional capacity in subjects with early osteoarthritis of the hip and knee: A diagnostic study in the CHECK cohort. Journal of Occupational Rehabilitation. 2009;19(4):345-53. PMID 19557505.

203. Salaffi F, Stancati A, Silvestri CA, et al. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain. 2004 Aug;8(4):283-91. PMID 15207508.

204. Thomas E, Peat G, Mallen C, et al. Predicting the course of functional limitation among older adults with knee pain: do local signs, symptoms and radiographs add anything to general indicators? Annals of the Rheumatic Diseases. 2008 Oct;67(10):1390-8. PMID 18245111.

205. Xie F, Pullenayegum EM, Li SC, et al. Use of a Disease-Specific Instrument in Economic Evaluations: Mapping WOMAC onto the EQ-5D Utility Index. Value Health. 2010 Dec;13(8):873-8. PMID 20667055.

206. Bennell KL, Hinman RS, Crossley KM, et al. Is the Human Activity Profile a useful measure in people with knee osteoarthritis? Journal of Rehabilitation Research and Development. 2004;41(4):621-9. PMID 15558390.

207. Astephen Wilson JL, Deluzio KJ, Dunbar MJ, et al. The association between knee joint biomechanics and neuromuscular control and moderate knee osteoarthritis radiographic and pain severity. Osteoarthritis Cartilage. 2011 Feb;19(2):186-93. PMID 21074628.

208. Wood L, Peat G, Thomas E, et al. Associations between physical examination and self-reported physical function in older community-dwelling adults with knee pain. Physical Therapy. 2008;88(1):33-42. PMID 18029391.

209. Sharma L, Cahue S, Song J, et al. Physical functioning over three years in knee osteoarthritis: role of psychosocial, local mechanical, and neuromuscular factors. Arthritis Rheum. 2003 Dec;48(12):3359-70. PMID 14673987.

210. van Baar ME, Dekker J, Lemmens JA, et al. Pain and disability in patients with osteoarthritis of hip or knee: the relationship with articular, kinesiological, and psychological characteristics. J Rheumatol. 1998 Jan;25(1):125-33. PMID 9458215.

211. Van Der Esch M, Steultjens M, Knol DL, et al. Joint Laxity and the Relationship Between Muscle Strength and Functional Ability in Patients With Osteoarthritis of the Knee. Arthritis Care and Research. 2006;55(6):953-9. PMID 17139642.

212. Hunt MA, McManus FJ, Hinman RS, et al. Predictors of single-leg standing balance in individuals with medial knee osteoarthritis. Arthritis Care & Research. 2010 Apr;62(4):496-500. PMID 20391504.

213. Harrison AL. The influence of pathology, pain, balance, and self-efficacy on function in women with osteoarthritis of the knee. Phys Ther. 2004 Sep;84(9):822-31. PMID 15330695.

214. Kauppila AM, Kyllonen E, Mikkonen P, et al. Disability in end-stage knee osteoarthritis. Disabil Rehabil. 2009;31(5):370-80. PMID 18608423.

215. Jadelis K, Miller ME, Ettinger WH, Jr., et al. Strength, balance, and the modifying effects of obesity and knee pain: results from the Observational Arthritis Study in Seniors (oasis). Journal of the American Geriatrics Society. 2001 Jul;49(7):884-91. PMID 11527479.

216. Clark JA, Spiro A, 3rd, Fincke G, et al. Symptom severity of osteoarthritis of the knee: a patient-based measure developed in the veterans health study. J Gerontol A Biol Sci Med Sci. 1998 Sep;53(5):M351-60. PMID 9754141.

217. Boutron I, Rannou F, Jardinaud-Lopez M, et al. Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year. Osteoarthritis Cartilage. 2008 Sep;16(9):1024-31. PMID 18276169.

218. Summers MN, Haley WE, Reveille JD, et al. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis Rheum. 1988 Feb;31(2):204-9. PMID 3348824.

219. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol. 1986 Jun;5(2):231-41. PMID 3731718.

220. Marks R. Reliability and validity of self-paced walking time measures for knee osteoarthritis. Arthritis Care and Research. 1994;7(1):50-3. PMID 7918728.

221. Marks R, Quinney HA, Wessel J. Proprioceptive sensibility in women with normal and osteoarthritic knee joints. Clin Rheumatol. 1993 Jun;12(2):170-5. PMID 8358974.

222. Creamer P, Lethbridge-Cejku M, Hochberg MC. Determinants of pain severity in knee osteoarthritis: effect of demographic and psychosocial variables using 3 pain measures. J Rheumatol. 1999 Aug;26(8):1785-92. PMID 10451078.

223. Yildiz N, Topuz O, Gungen GO, et al. Health-related quality of life (Nottingham Health Profile) in knee osteoarthritis: correlation with clinical variables and self-reported disability. Rheumatol Int. 2010 Nov;30(12):1595-600. PMID 19823830.

224. Birmingham TB, Kramer JF, Kirkley A, et al. Association among neuromuscular and anatomic measures for patients with knee osteoarthritis. Archives of Physical Medicine & Rehabilitation. 2001;82(8):1115-8. PMID 11494192.

225. Deniston OL, Jette A. A functional status assessment instrument: validation in an elderly population. Health Serv Res. 1980 Spring;15(1):21-34. PMID 7372496.

226. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. The arthritis impact measurement scales. Arthritis Rheum. 1980 Feb;23(2):146-52. PMID 7362665.

227. Wright JG, Santaguida PL, Young N, et al. Patient preferences before and after total knee arthroplasty. J Clin Epidemiol. 2010 Jul;63(7):774-82. PMID 20004554.

228. Faucher M, Poiraudeau S, LefevreColau MM, et al. Algo-functional assessment of knee osteoarthritis: comparison of the test-retest reliability and construct validity of the Womac and Lequesne indexes. Osteoarthritis and Cartilage. 2002;10(8):602-10. PMID 12479381.

229. Steultjens MP, Dekker J, van Baar ME, et al. Muscle strength, pain and disability in patients with osteoarthritis. Clin Rehabil. 2001 Jun;15(3):331-41. PMID 11386405.

230. O'Malley KJ, Suarez-Almazor M, Aniol J, et al. Joint-specific multidimensional assessment of pain (J-MAP): factor structure, reliability, validity, and responsiveness in patients with knee osteoarthritis. J Rheumatol. 2003 Mar;30(3):534-43. PMID 12610814.

231. Steultjens MP, Dekker J, van Baar ME, et al. Internal consistency and validity of an observational method for assessing disability in mobility in patients with osteoarthritis. Arthritis Care Res. 1999 Feb;12(1):19-25. PMID 10513486.

232. Finch E, Kennedy D. The lower extremity activity profile: a health status instrument for measuring lower extremity disability. Physiother Can. 1995 Fall;47(4):239-46. PMID 10153394.

233. Stucki G, Sangha O, Stucki S, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage. 1998 Mar;6(2):79-86. PMID 9692062.

234. Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology. 1999;38(4):355-61. PMID 10378714.

235. Brazier JE, Harper R, Munro J, et al. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology. 1999;38(9):870-7. PMID 10515649.

236. Lingard EA, Katz JN, Wright RJ, et al. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. J Bone Joint Surg Am. 2001 Dec;83-A(12):1856-64. PMID 11741066.

237. Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis. 1999 Sep;58(9):563-8. PMID 10460190.

238. Ryser L, Wright BD, Aeschlimann A, et al. A new look at the Western Ontario and McMaster Universities Osteoarthritis Index using Rasch analysis. Arthritis Care Res. 1999 Oct;12(5):331-5. PMID 11081002.

239. Davis AM, Badley EM, Beaton DE, et al. Rasch analysis of the Western Ontario McMaster (WOMAC) Osteoarthritis Index: results from community and arthroplasty samples. J Clin Epidemiol. 2003 Nov;56(11):1076-83. PMID 14614998.

240. Dougados M, Leclaire P, van der Heijde D, et al. Response criteria for clinical trials on osteoarthritis of the knee and hip: a report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials response criteria initiative. Osteoarthritis & Cartilage. 2000 Nov;8(6):395-403. PMID 11069723.

241. Angst F, Aeschlimann A, Michel BA, et al. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol. 2002 Jan;29(1):131-8. PMID 11824949.

242. 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. Annals of the Rheumatic Diseases. 2005;64(1):34-7. PMID 15130902.

243. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum. 2001 Aug;45(4):384-91. PMID 11501727.

244. Quintana JM, Escobar A, Arostegui I, et al. Health-related quality of life and appropriateness of knee or hip joint replacement. Arch Intern Med. 2006 Jan 23;166(2):220-6. PMID 16432092.

245. Escobar A, Quintana JM, Bilbao A, et al. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis and Cartilage. 2007;15(3):273-80. PMID 17052924.

246. Kennedy DM, Stratford PW, Wessel J, et al. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskeletal Disorders. 2005;6(Journal Article):3. PMID 15679884.

247. Mangione KK, Craik RL, McCormick AA, et al. Detectable changes in physical performance measures in elderly African Americans. Phys Ther. 2010 Jun;90(6):921-7. PMID 20395305.

248. Nebel MB, Sims EL, Keefe FJ, et al. The relationship of self-reported pain and functional impairment to gait mechanics in overweight and obese persons with knee osteoarthritis. Arch Phys Med Rehabil. 2009 Nov;90(11):1874-9. PMID 19887211.

249. O'Reilly SC, Jones A, Muir KR, et al. Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. Annals of the Rheumatic Diseases. 1998;57(10):588-94.

250. Burch FX, Tarro JN, Greenberg JJ, et al. Evaluating the benefits of patterned stimulation in the treatment of osteoarthritis of the knee: a multi-center, randomized, single-blind, controlled study with an independent masked evaluator. Osteoarthritis Cartilage. 2008 Aug;16(8):865-72. PMID 18262443.

251. Petrella RJ, Bartha C. Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial. J Rheumatol. 2000 Sep;27(9):2215-21. PMID 10990236.

252. Boutron I, Moher D, Altman DG, et al. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Annals of Internal Medicine. 2008 Feb 19;148(4):295-309. PMID 18283207.

253. Bellamy N. Measuring beneficial and adverse events in osteoarthritis clinical trials. Inflammopharmacology. 1996;4(2):141-8.

254. Samson DJ, Grant MD, Ratko TA, et al. Treatment of primary and secondary osteoarthritis of the knee. Evidence Report/Technology Assessment. 2007(157):1-157. PMID 18088162.

Abbreviations

ADL	Activities of Daily Living
AHRQ	Agency for Healthcare Research and Quality
AMED	Allied and Complementary Medicine`
BMI	Body Mass Index
CAS(K)	Clinical Assessment Study of the Knee
CHECK	Cohort Hip and Cohort Knee Study
FAST	Fitness Arthritis and Seniors Trial
HSS	Hospital for Special Surgery
IADL	Instrumental Activities of Daily Living
K/L	Kellgren and Lawrence
MAK	Mechanical Factors in Arthritis of the Knee
MCID	Minimal Clinically Important Difference
MeSH	Medical Subject Heading
NU	Normalized Units
OA	Osteoarthritis
OARSI	Osteoarthritis Research Society
OMERACT	Outcomes Measures in Rheumatoid Arthritis Clinical Trials
PASS	Patient Acceptable Symptom State
PEDro	Physiotherapy Evidence Database
PEMF	Pulsed Electromagnetic Fields
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
QWB	Quality of Well Being
RCT	Randomized Controlled Trial
ROC	Receiver Operating Characteristic
ROM	Range of Motion
SF-36	36-Item Short Form Health Survey
TEP	Technical Expert Panel
VAS	Visual Analog Scale