Draft Comparative Effectiveness Review

Number XX

Noninvasive, Nonpharmacological Treatment for Chronic Pain: A Systematic Review

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. To be added for final version.

Prepared by:
To be added for final version.

Investigators:
To be added for final version.

AHRQ Publication No. xx-EHCxxx
<Month Year>

Key Messages

Purpose of Review

To assess evidence on noninvasive, nonpharmacological treatments for common chronic pain conditions, focusing on whether improvements are seen for at least one month post-intervention.

Key Messages

- A number of nonpharmacological interventions may improve function or pain outcomes 1 month to 1 year after the completion of therapy; evidence for some treatments and conditions is limited.
- Exercise, acupuncture, multidisciplinary rehabilitation, mind-body and mindfulness practices and psychological therapies such as cognitive-behavioral therapy may improve function or pain outcomes for specific chronic pain conditions.
- There was no evidence suggesting serious harms from any of the interventions studied, although data on harms were limited.
- Additional comparative evidence with data on the sustainability of effects on function and pain outcomes is particularly needed for chronic tension headache, as well as for osteoarthritis, chronic neck pain, and fibromyalgia.

This report is based on research conducted by an Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. xxx-xxxx-xxxxx). 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.

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

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 periodically be assessed for the currency of conclusions. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program Web site at www.effectivehealthcare.ahrq.gov. Search on the title of the report.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

Suggested citation: <Authors>. <Topic>. Evidence Report/Technology Assessment. No. <#>. (Prepared by <EPC Name> under Contract No. <##>.) AHRQ Publication No. 14-XXXXX>. Rockville, MD: Agency for Healthcare Research and Quality; <Month, Year>. www.effectivehealthcare.ahrq.gov/reports/final/cfm.

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The Centers for Disease Control and Prevention (CDC) requested this report from the EPC Program at AHRQ. AHRQ assigned this report to the following EPC: (INSERT EPC NAME) Evidence-based Practice Center (Contract Number: (INSERT CONTRACT NUMBER).

The reports and assessments provide organizations with comprehensive, evidence-based information on common medical conditions and new health care technologies and strategies. They also identify research gaps in the selected scientific area, identify methodological and scientific weaknesses, suggest research needs, and move the field forward through an unbiased, evidence-based assessment of the available literature. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review and public comment prior to their release as a final report.

AHRQ expects that the EPC evidence reports and technology assessments, when appropriate, will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

If you have comments on this evidence report, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Gopal Khanna, M.B.A.
Director
Agency for Healthcare Research and Quality

Arlene S. Bierman, M.D., M.S.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Elisabeth Kato, M.D., M.R.P.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

To be added for final version.

Key Informants

In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to this report will be added for the final version.

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who participated in reviewing the report will be added for the final version.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers who participated in reviewing the report will be added for the final version.

Noninvasive, Nonpharmacological Treatment for Chronic Pain: A Systematic Review

Structured Abstract

Objectives. To assess the effectiveness of noninvasive, nonpharmacological treatment for selected chronic pain conditions, particularly as alternatives to opioids and other pharmacological treatments, with a focus on evaluating which interventions provide improved function and pain outcomes for at least 1 month post-intervention.

Data sources. Electronic databases (Ovid MEDLINE®, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, no restriction on publication date), reference lists, and ClinicalTrials.gov.

Review methods. Using predefined criteria, we selected randomized controlled trials of noninvasive, nonpharmacological treatments for five common chronic pain conditions (low back pain, neck pain, osteoarthritis of the knee, hip or hand, fibromyalgia, and tension headache) that addressed efficacy or harms compared with usual care, no treatment, waitlist, placebo, or sham intervention; compared with pharmacological therapy; or compared with exercise. The quality of included studies was assessed, data were extracted, and results were summarized quantitatively and qualitatively. Only trials reporting results for at least 1 month post-intervention were included. We focused on evaluating the persistence of effects for therapies beyond the course of treatment at short-term followup (1 to 6 months following completion of treatment), intermediate-term followup (6 to 12 months), and long- term followup (≥12 months).

Results. 205 publications (192 trials) were included in the review. Many included trials were small and the majority of patients were female. In general, there was little followup beyond 1 year after completion of treatment. Most trials enrolled patients who experienced a moderate pain intensity (e.g., >5 on a 0 to 10 point numeric rating scale for pain) and duration of symptoms ranging from 3 months to >15 years.

Chronic low back pain: Function improved slightly in the short term with massage, yoga, and psychological therapies (Strength of evidence [SOE]: Moderate) and with exercise, acupuncture, low-level laser therapy, mindfulness-based stress reduction (MBSR), spinal manipulation, and multidisciplinary rehabilitation (SOE: Low), compared with usual care or inactive controls. Effects on function continued into the intermediate term for yoga, MBSR, spinal manipulation, multidisciplinary rehabilitation (SOE: Low), and psychological therapies (SOE: Moderate). Psychological therapies, were associated with slightly greater improvement than usual care or an attention control on both function and pain at short-term, intermediate-term, and long-term followup (SOE: Moderate). Improvements in pain persisted into the intermediate term for exercise, massage and yoga (moderate effect, SOE: Low), MBSR (small effect, SOE: Low) as well as spinal manipulation, psychological therapies, and multidisciplinary rehabilitation (small effects, SOE: Moderate). For acupuncture there was no difference in pain at intermediate term, but a slight improvement at long-term (SOE: Low). Effects on function were generally smaller than effects on pain. Multidisciplinary rehabilitation slightly improved pain at short and intermediate terms compared with exercise (SOE: Moderate). High-intensity multidisciplinary

rehabilitation (≥20 hours/week or >80 hours total) was not clearly better than non-high-intensity programs.

Chronic neck pain: In the short- and intermediate-term studies, acupuncture and Alexander Technique slightly improved function compared with usual care (both interventions), sham acupuncture, or sham laser (SOE: Low), but no improvement in pain was seen at any time frame (SOE: Low). Short-term, moderate effects on function and pain were seen for low-level laser therapy (SOE: Moderate).

Osteoarthritis: For knee osteoarthritis, only exercise and ultrasound demonstrated small short-term improvements in function compared with usual care, an attention control or sham procedure (SOE: Moderate for exercise, Low for ultrasound). Effects were sustained into the intermediate term only for exercise (SOE: Low), which was also associated moderate improvement in pain (SOE: Low). Long-term, the small improvement in function seen with exercise was sustained, but there was no clear effect on pain (SOE: Low). Evidence was sparse on interventions for hip and hand osteoarthritis. Exercise for hip osteoarthritis was associated with slightly greater function and pain improvement than usual care in the short term (SOE: Low). The effect on function was sustained intermediate term (SOE: Low).

Fibromyalgia: Function improved slightly in the short term with cognitive behavioral therapy (CBT) and tai chi and qigong mind-body practices (SOE: Low) and with acupuncture (SOE: Moderate). Improvements in pain were seen in the short term with exercise (SOE: Moderate) and mind body practices (SOE: Low). Small functional improvement continued into the intermediate term for acupuncture and cognitive behavioral therapy CBT (SOE: Low) and was seen for myofascial release massage and multidisciplinary rehabilitation (SOE: Low). Long term, small improvements in function continued for multidisciplinary rehabilitation but not for exercise or massage (SOE: Low for all) and no clear impact on pain for exercise (SOE: Moderate) or multidisciplinary rehabilitation was seen (SOE: Low).

Chronic tension headache: Evidence was sparse and the majority of trials were of poor quality.

There was no evidence suggesting increased risk for serious treatment-related harms for any of the interventions, although data on harms were limited.

Conclusions. A number of nonpharmacological interventions can provide beneficial effects on function and/or pain that are durable 1 month to 1 year after the completion of therapy. Exercise, multidisciplinary rehabilitation, acupuncture, and mind-body and mindfulness practices may slightly to moderately improve function and pain across multiple chronic pain conditions. Our findings provide some support for clinical strategies that focus on use of nonpharmacological therapies as preferred interventions for chronic pain. Additional comparative research on sustainability of effects beyond the immediate post-treatment period is needed, particularly for conditions other than low back pain.

Contents

Tables

Figures

Appendixes

Evidence Summary

Introduction

Chronic pain substantially impacts physical and mental functioning, productivity, quality of life, and family relationships; it is the leading cause of disability and is often refractory to treatment.[1,2] Chronic pain is defined as pain lasting 12 weeks or longer or persisting past the normal time for tissue healing.[3,4] A monumental public health challenge, chronic pain affects millions of adults in the United States, with a conservative annual cost in personal and health system expenditures estimated at $560 billion to $635 billion.[4]

The National Pain Strategy (NPS) report recommends that pain management be integrated, multimodal, interdisciplinary, evidence-based, and tailored to individual patient needs.[5] In addition to addressing biological factors when known, optimal management of chronic pain must also address psychosocial contributors to pain, while taking into account individual susceptibility and treatment responses.

The NPS points to the "dual crises" of chronic pain and opioid dependence, overdose, and death as providing important context for consideration and implementation of alternative chronic pain management strategies. Nationally, concerns regarding appropriate use, misuse, and diversion of opioids for treatment of chronic pain have been the subject of numerous scientific and news reports and were highlighted in the NPS report[5] and a 2011 Institute of Medicine (IOM, now National Academy of Medicine)[4] report.

Although opioid prescriptions for chronic pain have increased substantially in the past 20 years, evidence shows only modest short-term benefits.[6-8] Lack of evidence on long-term effectiveness[9] and serious safety concerns[10] speak to the need to consider alternative treatments to opioids. The recent evidence-based guidelines on opioid use for chronic pain by the Centers for Disease Control and Prevention (CDC),[11] which include a recommendation on the preferred use of non-opioid treatment over opioid therapy, have prompted additional primary research on alternative methods of managing chronic pain.

Other pharmacological treatments for chronic pain include nonsteroidal anti-inflammatory drugs (NSAIDs), acetaminophen, muscle relaxants, antiseizure medications, antidepressants, and corticosteroids, used alone or in combination with each other or with opioids. Each has potential side effects and contraindications.

Nonpharmacological treatments for chronic pain may include exercise and physical therapy, mind-body practices, psychological therapies, interdisciplinary rehabilitation, mindfulness practices, osteopathic and spinal manipulation, acupuncture, physical modalities, and acupuncture.

Rationale for This Review

Both the IOM report and NPS describe the need for evidence-based strategies for the treatment of chronic pain that address the biopsychosocial nature of this disease, including nonpharmacological treatment. These initiatives, and others, speak to the importance of understanding current evidence on noninvasive, nonpharmacological treatment of chronic pain.

Our review is intended to address some of the needs described in the NPS[5] and IOM[4] reports and others for evidence to inform guidelines and health care policy (including reimbursement policy) related to use of noninvasive, nonpharmacological treatments as alternatives to opioids and other pharmacological treatments. This review also aims to provide additional insights into

research gaps related to use of noninvasive, nonpharmacological alternatives for treating chronic pain.

Scope and Key Questions

This comparative effectiveness review focused on noninvasive, nonpharmacological therapy, with a Key Question (KQ) for each of five common chronic pain conditions:

KQ 1: Chronic low back pain
KQ 2: Chronic neck pain
KQ 3: Osteoarthritis (knee, hip, hand)
KQ 4: Fibromyalgia
KQ 5: Chronic tension headache

For each condition, we addressed the following subquestions:

a. What are the benefits and harms of noninvasive nonpharmacological therapies compared with sham treatment, no treatment, waitlist, attention control, or usual care?
b. What are the benefits and harms of noninvasive nonpharmacological therapies compared with pharmacological therapy (e.g., opioids, NSAIDS, acetaminophen, antiseizure medications, antidepressants)?
c. What are the benefits and harms of noninvasive nonpharmacological therapies compared with exercise or (for headache) biofeedback?

Individual pain management strategies considered in the review include exercise and physical therapy, mind-body practices (yoga, tai chi, qigong), psychological therapies (cognitive-behavioral therapy, biofeedback, relaxation techniques, acceptance and commitment therapy), interdisciplinary rehabilitation, mindfulness practices (meditation, mindfulness-based stress reduction practices), osteopathic and spinal manipulation, acupuncture, and physical modalities (traction, ultrasound, transcutaneous electrical nerve stimulation [TENS], low level laser therapy, interferential therapy, superficial heat or cold, bracing for knee, back or neck, electro-muscular stimulation and magnets), acupuncture, and functional restoration training. We focused on single active interventions and comparators, assessing the persistence of effects for therapies beyond the course of treatment, particularly over the long term. We also asked (KQ 6) whether estimates of benefits and harms differ by age, sex, or presence of comorbidities (e.g., emotional or mood disorders).

Details on the PICOTS (population, interventions, comparators, outcomes, timing, settings) inclusion and exclusion criteria are provided in the full report and in the protocol.

Methods

The methods for this systematic review follow the Agency for Healthcare Research & Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.*[12] See the review protocol (http://effectivehealthcare.ahrq.gov/index.cfm) and the full report of the review for additional details.

Topic Refinement and Review Protocol

The review team developed initial Key Questions and PICOTS with input from the AHRQ Task Order Officer (TOO), representatives from the Centers for Disease Control and Prevention

(CDC) and the Office of the Assistant Secretary for Planning and Evaluation (ASPE), and a group of Key Informants. The EPC review team considered the public comments received on the provisional Key Questions, PICOTS, and analytic framework (posted on the AHRQ Web site), along with input from the AHRQ TOO, CDC and ASPE representatives, and a Technical Expert Panel (TEP) convened for this report. The final version of the protocol for this review was posted on the AHRQ Effective Health Care Program web site (www.effectivehealthcare.ahrq.gov) and registered in the PROSPERO international database of prospectively registered systematic reviews (CRD42017067729).

Literature Search Strategy

A research librarian conducted searches in Ovid® MEDLINE®, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and ClinicalTrials.gov. Searches were conducted without restriction on publication date (see Appendix A for search strategies). A *Federal Register* notice was posted in an effort to identify unpublished data. Reference lists of included articles and the bibliographies of systematic reviews published since 2010 were reviewed for includable literature.

Inclusion and Exclusion Criteria, Study Selection, and Data Abstraction

Inclusion and exclusion criteria were developed *a priori* based on the Key Questions and PICOTS. Abstracts were reviewed by at least two investigators, and full-text articles deemed potentially appropriate for inclusion by at least one of the reviewers were retrieved. Two investigators then independently reviewed all full-text articles for final inclusion. Discrepancies were resolved by discussion and consensus. A list of the included studies can be found in Appendix C of the full report; excluded studies and primary reason for exclusion can be found in Appendix D of the full report. We abstracted data on study characteristics and funding source, populations, interventions, comparators, and results (Appendix E of the full report). Extracted study data were verified for accuracy and completeness by a second team member. The focus of this review was on randomized controlled trials (RCTs) reporting on longer-term outcomes (at least 4 weeks post intervention) that otherwise meet our PICOTS criteria.

Quality Assessment of Individual Studies

Predefined criteria were used to assess the quality (risk of bias) of included studies. RCTs were assessed based on criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions* (Chapter 8.5 Risk of Bias Tool),[13] and precepts for appraisal developed by the Cochrane Back and Neck Group,[14] in conjunction with the approach recommended in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Research*.[12] Two team members independently appraised each included study, with disagreements resolved by consensus. Studies were rated as "good," "fair," or "poor." Detailed assessment of included studies appears in Appendix F of the full report.

Data Analysis and Synthesis

Data were synthesized qualitatively based on ranges and descriptive analysis and quantitatively using meta-analysis where appropriate. Duration of followup post intervention was

reported and categorized as short-term (up to 6 months), intermediate-term (6 to 12 months) and long-term (at least 1 year). A variety of outcome measures were used across trials; we prioritized outcomes of function and pain in reporting results. Based on input from stakeholders, improvement in function was prioritized as the most important outcome.

Results for continuous outcomes as well as dichotomous outcomes were synthesized. Binary outcomes based on the proportion of patients achieving specific thresholds of success for improved function, or other overall measure of success as defined in the trials (e.g., ≥ 30% improvement in pain score) were reported and a risk ratio and 95% confidence interval calculated to evaluate the presence of an association and estimate relative effect size using the Rothman Episheet.[15] For continuous outcomes, mean differences between treatments and 95% confidence intervals were calculated using GraphPad or Stata®/IC 12.1 (StataCorp, College Station, TX) to provide effect sizes and determine presence of a statistical association.

We conducted meta-analysis to get more precise effect estimates for various interventions. To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity. Three continuous outcomes (pain, function and quality of life) provided adequate data for meta-analysis. We assumed random effects across studies and used both the Dersimonian-Laird method[16] and the profile-likelihood model[17] to combine studies. Statistical heterogeneity among the studies was assessed using the standard Cochran's chi-square test and the I^2 statistic.[18] Primary analyses were stratified by disease type, intervention, control group (usual care, exercise or pharmacological treatment) and length of followup (short, intermediate, and long term). Controls included usual care, waitlist, no treatment, placebo, sham treatment, attention control, or other groups that involved at most minimal active treatment. We performed additional sensitivity and subgroup analyses based on specific interventions (e.g., type of acupuncture, type of exercise, intervention intensity etc.) and control types (as described above), and by excluding outlying studies and studies rated poor quality.

To facilitate interpretation of results across trials and interventions, we categorized the magnitude of effects for function and pain outcomes using the system described in our previous review.[19-21] In general we classified effects for measures with a 0-100 scale for pain or function as small/slight (5-10 points)), moderate (>10–20 points), or large/substantial (>20 points). Additional information for specific measures is found in the full report.

Grading the Strength of Evidence for Major Comparisons and Outcomes

The strength of evidence (SOE) was assessed by one researcher, and the initial assessment was independently reviewed by at least one other experienced senior investigator. The overall SOE was determined based on study limitations; consistency of results across studies; the directness of the evidence linking the interventions with health outcomes; effect estimate precision; and reporting bias. Bodies of evidence consisting of RCTs are initially considered as high strength of evidence. All outcomes were considered direct. The SOE for each Key Question and primary outcome (function, pain, harms) was assigned an overall grade of high, moderate, low, or insufficient. When all of the studies for a primary outcome were rated poor quality, we rated the SOE as insufficient. SOE tables for primary outcomes are presented in Appendix G of the full report.

Peer Review and Public Commentary

Peer reviewers with expertise in primary care and management of the included chronic pain conditions have been invited to provide written comments on the draft report. The AHRQ Task Order Officer and an Evidence-based Practice Center Associate Editor will also provide comments and editorial review. The draft report will be posted on the AHRQ Web site for 4 weeks for public comment. A disposition of comments report with authors' responses to the peer and public review comments will be posted after publication of the final comparative effectiveness review (CER) on the public Web site.

Results

Results of Literature Searches

Database searches resulted in 4,470 potentially relevant articles. After dual review of abstracts and titles, 1,091 articles were selected for full-text dual review and 205 publications (192 trials) met inclusion criteria. We included 65 trials (69 publication) on chronic low back pain, 23 trials on chronic neck pain, 51 trials (54 publications) on osteoarthritis, 44 trials (50 publications) on fibromyalgia, and 9 trials on chronic tension headache. One-fourth of the trials excluded at full text did not meet our criteria for followup duration (i.e., a minimum of 1 month of followup after termination of the intervention). Data extraction and quality assessment tables for all included studies are available in Appendixes E and F of the full report.

Thirty-five percent of the included trials were small (fewer than 70 participants). Across studies, the majority of patients were female, with a mean age in most trials of 40 to 45 years old. Exercise was the most common intervention for osteoarthritis and fibromyalgia. Psychological therapies were most common for fibromyalgia, and manual therapies were most common for chronic low back pain. Acupuncture was used in all included conditions. Multidisciplinary rehabilitation was reported primarily for low back pain and fibromyalgia. There were no trials of functional restoration training for any condition. Limited evidence was available for hip or hand osteoarthritis or chronic tension headache. The majority of trials compared nonpharmacological interventions with usual care, waitlist, no treatment, attention control, or placebo/sham, with very few trials employing pharmacological treatments or exercise as comparators. In general, little long-term evidence was available.

The majority of trials (58%) were rated fair quality, and 37 percent were rated as poor, with only 5 percent considered good quality. A primary methodological limitation in many of the trials was the inability to effectively blind participants and, in many cases, providers. Additionally, unacceptable rates of attrition (both overall and differential) and poor reporting of allocation concealment methods were common shortcomings.

Key points are presented in the following sections for interventions and outcomes for which there was low or moderate strength of evidence. Interventions and outcomes with no evidence or insufficient evidence are discussed in the full report.

Key Question 1: Chronic Low Back Pain

Exercise for Low Back Pain

- Exercise was associated with slighter greater effects on short-term function than usual care, an attention control, or a placebo intervention (6 trials, pooled SMD -0.31, 95% CI -0.58 to -0.04, I2=57%); there were no effects on intermediate-term function (3 trials, pooled SMD -0.15, 95% CI -0.48 to 0.18, I2=51%) or long-term function (1 trial, difference 0.00 on the 0 to 100 ODI [Oswestry Disability Index], 95% CI -11.4 to 11.4) (SOE: Low).
- Exercise was associated with slightly to moderately greater effects on pain than controls at short-term (6 trials, pooled difference -0.81 on a 0 to 10 scale, 95% CI -1.26 to -0.,36, I2=0%), intermediate-term (3 trials, pooled difference -1.37, 95% CI -2.10 to -0.65, I2=34%), and long-term (1 trial, difference -1.55, 95% CI -2.378 to -0.32) followup (SOE: Moderate for short-term, low for intermediate-term and long-term).

Psychological Therapies for Low Back Pain

- Psychological therapy was associated with slightly greater effects on function than usual care or an attention control at short-term (3 trials, pooled SMD -0.25, 95% CI -0.38 to -0.12, I2=0%), intermediate-term (3 trials, pooled SMD -0.25, 95% CI -0.37 to -0.13, I2=0%), and long-term followup (2 trials, pooled SMD -0.26, 95% CI -0.39 to -0.12, I2=0%) (SOE: Moderate).
- Psychological therapy was associated with slightly greater effects on pain than controls at short-term (3 trials, pooled difference -0.76 on a 0 to 10 scale, 95% CI -0.99 to -0.53, I2=0%), intermediate-term (3 trials, pooled difference -0.71, 95% CI -0.94 to -0.48, I2=0%), or long-term followup (2 trials, pooled difference -0.53, 95% CI -0.82 to -0.24, I2=3.6%) (SOE: Moderate).

Physical Modalities for Low Back Pain

Ultrasound

- No differences were found between ultrasound versus sham ultrasound in short-term pain (2 trials, SOE: Low).

Low-Level Laser Therapy

- One trial found low-level laser therapy associated with moderately greater effects than sham laser on short-term pain (difference -16.0 on a 0 to 100 scale, 95% CI -28.3 to -3.7) and slightly greater effects on function (difference -8.2 on the 0 to 100 ODI, 95% CI -13.6 to -2.8) (SOE: Low).

Traction

- Two trials found no differences between traction versus sham traction in short-term pain or function (SOE: Low).

Manual Therapies for Low Back Pain

Massage

- There were no differences between massage versus controls in short-term (3 trials, SMD -0.24, 95% CI -0.49 to 0.02, I2=0%) or intermediate-term (2 trials, SMD 0.00, 95% CI -0.22 to 0.22, I2=0%) function (SOE: Moderate for short-term, low for intermediate-term).
- Massage was associated with slighter greater effects on short-term pain than sham massage or usual care (3 trials, pooled difference -0.63 on a 0 to 10 scale, 95% CI -1.00 to -0.26, I2=0%). There was no difference between massage versus controls in intermediate-term pain (2 trials, difference -0.12, 95% CI -0.99 to 0.75, I2=43%) (SOE: Moderate for short-term, low for intermediate-term).

Spinal Manipulation

- Spinal manipulation was associated with slightly greater effects than sham manipulation, usual care, an attention control, or a placebo intervention in short-term function (3 trials, pooled SMD -0.34, 95% CI -0.63 to -0.05, I2=61%) and intermediate-term function (3 trials, pooled SMD -0.40, 95% CI -0.69 to -0.11, I2=76%) (SOE: Low).
- There was no difference between spinal manipulation versus sham manipulation, usual care, an attention control or a placebo intervention in short-term pain (3 trials, pooled difference -0.20 on a 0 to 10 scale, 95% CI -0.66 to 0.26, I2=58%), but manipulation was associated with slightly greater effects than controls on intermediate-term pain (3 trials, pooled difference -0.64, 95% CI -0.92 to -0.36, I2=0%) (SOE: Low for short-term, moderate for intermediate-term).

Mindfulness-Based Stress Reduction for Low Back Pain

- There were no differences between Mindfulness-Based Stress Reduction (MBSR) versus usual care or an attention control in short-term function (3 trials, pooled SMD -0.22, 95% CI -0.53 to 0.10, I2=63%) or intermediate-term function (1 trial, SMD -0.20, 95% CI -0.47 to 0.06) (SOE: Low).
- MBSR was associated with a slightly greater effect than usual care or an attention control on short-term pain, based on the highest quality trials (3 trials, pooled difference -0.73, 95% CI -1.18 to -0.28; I2=45%); MBSR was also associated with slightly greater effects on intermediate-term pain (1 trial, difference -0.75, 95% CI -1.17 to -0.33) (SOE: Moderate for short-term, low for intermediate-term).

Mind-Body Practices for Low Back Pain

Qigong

- One trial found no differences between qigong versus exercise in short-term function (difference 0.9 on the RDQ [Roland-Morris Disability Questionnaire], 95% CI -0.1 to 2.0), though intermediate-term results slightly favored exercise (difference 1.2, 95% CI 0.1 to 2.3) (SOE: Low).

Yoga

- Yoga was associated with slighter greater effects on function than attention or wait list control at short-term followup (5 trials, pooled SMD -0.49, 95% CI -0.75 to -0.23,

I2=59%) and intermediate-term followup (3 trials, pooled SMD -0.33, 95% CI -0.49 to -0.16) (SOE: Moderate for short-term, low for intermediate-term).

- Yoga was associated with moderately greater effects on pain than attention or wait list control at short-term followup (4 trials, pooled difference -1.23 on a 0 to 10 scale, 95% CI -2.08 to -0.39, I2=77%) and intermediate-term followup (2 trials, pooled difference -1.17, 95% CI -1.91 to -0.44, I2=26%) (SOE: Low for short-term, moderate for intermediate-term).

Acupuncture for Low Back Pain

- Acupuncture was associated with slightly greater effects on short-term function than sham acupuncture or usual care (4 trials, pooled SMD -0.22, 95% CI -0.35 to -0.08, I2=44%). There were no differences between acupuncture versus controls in intermediate-term function (3 trials, pooled SMD -0.08, 95% CI -0.36 to 0.20, I2=75%) or long-term function (1 trial, adjusted difference -3.4 on the 0 to 100 ODI, 95% CI -7.8 to 1.0) (SOE: Low).
- Acupuncture was associated with slighter greater effects on short-term pain than sham acupuncture, usual care, an attention control, or a placebo intervention (5 trials, pooled difference -0.55 on a 0 to 10 scale, 95% CI -0.86 to -0.24, I^2=30%). There was no difference in intermediate-term pain (5 trials, pooled mean difference -0.25, 95% CI -0.67 to 0.16, I2=33%); one trial found acupuncture associated with greater effects on long-term pain (mean difference -0.83, 95% CI -1.51 to -0.15) (SOE: Moderate for short-term, low for intermediate-term and long-term).

Multidisciplinary Rehabilitation for Low Back Pain

- Multidisciplinary rehabilitation was associated with slightly greater effects on function than usual care at short-term followup (4 trials, pooled SMD -0.31, 95% CI -0.57 to -0.05, I2=70%) and intermediate-term followup (4 trials, pooled SMD -0.37, 95% CI -0.64 to -0.10, I2=50%); there was no difference in long-term function (2 trials, pooled SMD -0.04, 95% CI -0.31 to 0.24, I2=35%) (SOE: Low).
- Multidisciplinary rehabilitation was associated with slightly greater effects on pain than usual care at short-term followup (4 trials, pooled difference -0.51 on a 0 to 10 scale, 95% CI -0.89 to -0.13, I2=23%) and intermediate-term followup (4 trials, pooled difference -0.63, 95% CI -1.04 to -0.22, I2=0%); the long-term difference was smaller and not statistically significant (2 trials, pooled difference -0.34, 95% CI -0.86 to 0.18, I2=0%) (SOE: Moderate for short-term and intermediate-term, low for long-term).
- Multidisciplinary rehabilitation was associated with slightly greater effects than exercise on short-term pain (6 trials, pooled difference -0.75 on a 0 to 10 scale, 95% CI -1.18 to -0.31, I2=0%) and intermediate-term pain (5 trials [excluding outlier trial], pooled difference -0.55, 95% CI -0.95 to -0.15, I2=0%); there was no effect on long-term pain (2 trials [excluding outlier trial], pooled difference 0.00, 95% CI -0.94 to 0.95) (SOE: Moderate for short-term and intermediate-term, low for long-term).

Comparative Effectiveness of Interventions for Chronic Low Back Pain

- Multidisciplinary rehabilitation was associated with slightly greater effects than exercise on short-term function (6 trials, pooled SMD -0.28, 95% CI -0.54 to -0.01, I2=39%) and intermediate-term function (5 trials [excluding outlier trial], pooled

SMD -0.22, 95% CI -0.40 to -0.03, I2=0%), there was no effect on long-term function (2 trials [excluding outlier trial], pooled SMD -0.06, 95% CI -0.36 to 0.25, I2=0%) (SOE: Moderate for short-term and intermediate-term, low for long-term).

- There were no differences between spinal manipulation versus exercise in short-term function (3 trials, pooled SMD 0.01, 95% CI -0.22 to 0.25; I2=62%) or intermediate-term function (4 trials, pooled SMD 0.02, 95% CI -0.13 to 0.18; I2=48%) (SOE: Low).
- There were no differences between spinal manipulation versus exercise in short-term (3 trials, pooled difference 0.31 on a 0 to 10 scale, 95% CI -0.30 to 0.92; I2=60%) or intermediate-term pain (4 trials, pooled difference 0.22, 95% CI -0.09 to 0.52, I2=9.4%) (SOE: Low).
- One trial found no differences between low-level laser therapy versus exercise therapy in intermediate-term pain or function (SOE: Low).
- One trial found no differences between massage versus exercise in intermediate-term pain, function, or the SF-36 Mental or Physical Component Summary scores (SOE: Low).
- One trial found qigong associated with slightly lower effects on pain versus exercise at short-term followup (difference 7.7 on a 0 to 100 scale, 95% CI 0.7 to 14.7), but the difference at intermediate-term was not statistically significant (difference 7.1, 95% CI -1.0 to 15.2) (SOE: Low).
- There was no statistically significant difference between yoga versus exercise in short-term or intermediate-term function or pain (SOE: Low).

Key Question 2: Chronic Neck Pain

Exercise for Neck Pain

- Across types of exercise, there was no clear improvement in function (3 trials, pooled standardized mean difference (SMD) -0.23, 95% CI -0.71 to 0.15) or pain (3 trials, pooled SMD -0.72, 95% CI -1.49 to 0.06) versus no treatment or advice alone in the short-term (SOE: Low).
- A subgroup of two trials of combination exercises (including 3 of the following 4 exercise categories: muscle performance, mobility, muscle re-education, aerobic) suggests a small benefit in function and pain versus no treatment or advice alone over the short term and function in the long term (SOE: Low).

Psychological Therapies for Neck Pain

- No difference in function (NDI, 0-80 scale) or pain (VAS [Visual Analog Scale for Pain], 0-10 scale) in the short term (adjusted difference 0.1, 95% CI -2.9 to 3.2 and 0.2, 95% CI -0.4 to 0.8, respectively) or intermediate term (adjusted difference 0.2, 95% CI -2.8 to 3.1 and 0.2, 95% CI -0.3 to 0.8, respectively) from one fair quality study comparing relaxation training and no intervention or exercise (SOE: Low for all).

Physical Modalities for Neck Pain

- Low-level laser therapy was associated with a moderate improvement in short-term function (2 trials, pooled difference -14.98 , 95% CI -23.88 to -6.07, I^2=39%, 0-100

scale) and pain (3 trials, pooled difference -1.81 on a 0-10 scale, 95% CI -3.35 to -0.27, I^2=75%) compared with sham (SOE: Moderate for function and pain).

Manual Therapies for Neck Pain

- The effects of massage on function versus self-management attention control were small and not statistically significant in one trial (N=64) in the short term (≥5 point improvement on the NDI, 39% versus 14%, RR 2.7, 95% CI 0.99 to 7.5) and intermediate-term (57% versus 31%, RR 1.8, 95% CI 0.97 to 3.5) (SOE: Low for both time periods).

Mind-Body Practices for Neck Pain

- Alexander Technique resulted in a small improvement in function in the short term (difference -5.56 on a 0-100% scale, 95% CI -8.33 to -2.78) and intermediate-term (difference -3.92, 95% CI -6.87 to -0.97) compared with usual care alone based on one fair-quality trial (SOE: Low).

Acupuncture for Neck Pain

- Acupuncture was associated with slightly greater effects on short term and intermediate-term function versus sham acupuncture, placebo (sham laser) or usual care (short-term, 4 trials, pooled SMD -0.32, 95% CI -0.53 to -0.10, I^2=53.1%; intermediate-term, 3 trials, pooled SMD -0.19, 95% CI -0.35 to -0.02). One trial reported no difference in function in the long term (SMD -0.23, 95% CI -0.61 to 0.16) (SOE: Low for all time periods).
- No difference in pain comparing acupuncture with sham acupuncture, or placebo interventions in in the short term (4 trials, pooled difference -0.2 on a 0-10 scale, 95% CI -0.59 to 0.05, I^2=2%), intermediate term (3 trials, pooled difference 0.45, 95% CI -0.34 to 1.25, I^2=59%) or long term (1 trial, difference -1.8, 95% CI -1.34 to 0.64). (SOE: Low for all time periods).

Comparative Effectiveness of Interventions for Chronic Neck Pain

- There was no clear evidence that massage improved pain in the intermediate-term versus exercise (p>0.05, data not reported) in one fair-quality trial (SOE: Low).
- No clear evidence that basic body awareness therapy improved function in the short-term versus exercise in one fair-quality trial (SOE: Low).

Key Question 3: Osteoarthritis

Exercise for Osteoarthritis of the Knee

- Exercise was associated with slightly greater impact on function than usual care, no treatment or sham intervention short term (7 trials, pooled standardized mean difference (SMD) -0.25, 95% CI -0.4 to -0.09, I^2 = 0%) , at intermediate term (9 trials [excluding outlier trial] pooled SMD -0.78, 95%CI -1.37 to -0.19, I^2 = 91.4% and long term (2 trials, pooled SMD -0.24, 95%CI -0.37 to -0.11 I^2=0%) (SOE: Moderate for short term; Low for intermediate and long term).
- Exercise was associated with a small improvement in pain short term (7 trials, pooled difference -0.44 on a 0 to 10 scale, 95% CI -0.82 to -0.05, I^2= 35%) versus usual care, no treatment or sham intervention (SOE: Moderate), and with moderately greater effect on

pain in the intermediate term (9 trials, pooled difference -1.61 on a 0 to 10 scale, 95% CI -2.51 to -0.72, I^2=91%) compared with usual care, an attention control, or no treatment (SOE: Low). Long term, there was no clear difference between exercise and improvement in pain but data were limited (2 trials, difference -0.24, 95% CI -0.72 to 0.24) (SOE: Low).

Psychological Therapy for Osteoarthritis of the Knee

- One fair quality and one poor quality study of pain coping skills training and cognitive behavioral training versus usual care found no differences in function (WOMAC [Western Ontario and McMaster Universities Osteoarthritis Index] physical, 0-100) or pain (WOMAC pain, 0-100); treatment effects were averaged over short-term to intermediate-term followup (difference -0.3, 95% CI -8.3 to 7.8 for function and -3.9, 95% CI -1.8 to 4.0 for pain) and intermediate-term to long-term followup (mean 35.2, 95% CI 31.8 to 38.6 vs. mean 37.5, 95% CI 33.9 to 41.2, and mean 34.5, 95% CI 30.8 to 38.2 vs. mean 38.0, 95% CI 34.1 to 41.8), respectively (SOE: Low).

Physical Modalities for Osteoarthritis of the Knee

Ultrasound

- One fair-quality trial found continuous and pulsed ultrasound associated with better short-term function (difference of -6.2, 95% CI -8.36 to -4.20, and -5.71, 95% CI -7.72 to -3.70 on a 0-24 scale) and short-term pain intensity (difference -3.3, 95% CI -4.64 to -1.96, and -3.37, 95% CI -4.73 to -2.01 on a 0-10 scale) (SOE: Low).
- One fair-quality trial found no difference between continuous and pulsed ultrasound versus sham in intermediate-term function (difference -2.9, 95% CI -9.19 to 3.39 and 1.6, 95% CI -3.01 to 6.22, on a 0-68 scale) or pain (difference -1.6, 95% CI -3.26 to 0.06 and 0.2, 95% CI -1.34 to 1.74, on a 0-20 scale). There was also no difference between groups for VAS pain during rest or on movement (SOE: Low).

Transcutaneous Electrical Nerve Stimulation

- One trial found no difference between transcutaneous electrical nerve stimulation (TENS) and placebo TENS in intermediate-term function as measured by the WOMAC function subscale (proportion of patients who achieved MCID [minimal clinically important difference] (≥9.1), 38% vs. 39%, RR 1.2 (95% CI 0.6 to 2.2); and difference -1.9 (95% CI -9.7 to 5.9) on a 0-100 scale) or intermediate-term pain (proportion of patients who achieved MCID (≥20) in VAS pain, 56% vs 44%, RR 1.3 (95% CI 0.8 to 2.0); and mean difference –5.6 (95% CI –14.9 to 3.6) on the 0-100 WOMAC pain subscale) (SOE: Low for function and pain).

Electromagnetic Field

- One fair quality trial found pulsed electromagnetic fields associated with small effects on function (difference -3.48, 95% CI -4.44 to -2.51 on a 0-85 WOMAC Activities of Daily Living [ADLs] subscale) and pain (difference -0.84, 95% CI -1.10 to -0.58 on a 0-25 WOMAC pain subscale) versus sham short-term but may not be clinically significant. (SOE: Low).

Acupuncture for Osteoarthritis of the Knee

- There were no clear differences between acupuncture versus control interventions (sham condition, waitlist or usual care) on function in the short-term (5 trials, pooled SMD -0.18, 95% CI -0.55 to 0.20 I^2 =82%) or the intermediate-term (3 trials, pooled SMD -0.12, 95% CI -0/30 to 0.07, I^2 = 41%) (SOE: Low for short term; Moderate for intermediate term).
- There were no clear differences between acupuncture versus control interventions on pain in the short-term (pooled SMD -0.27, 95%CI -0.56 to 0.02, I^2 = 75%) or in the intermediate-term (pooled SMD -0.11, 95% CI -0.30 to 0.07, I^2 = 0%) (SOE: Low for short term; Moderate for intermediate term).

Exercise for Osteoarthritis of the Hip

- Exercise was associated with a small improvement in function versus usual care in the short-term (3 trials, pooled standardized mean difference (SMD) -0.33, 95% CI, -0.53 to -0.12, I^2=0.0%) and intermediate-term (2 trials, pooled SMD -0.28, 95% CI -0.50 to -0.05, I^2=0.0%). (SOE: Low for short and intermediate term).
- Exercise tended toward slightly greater improvement in short-term pain compared with usual care (3 trials, pooled SMD -0.34, 95% CI, -0.63 to -0.04, I^2=48.2%), but the results were no longer significant at intermediate term (2 trials, pooled SMD -0.14, 95% CI -0.37 to 0.08, I2=0%) (SOE: Low for short and intermediate term).

Manual Therapies for Osteoarthritis of the Hip

- Manual therapy was associated with a small short-term effects (mean difference 11.1, 95% CI 4.0 to 18.6, 0-100 scale Harris Hip Score) and small intermediate-term effects (mean difference 9.7, 95% CI, 1.5 to 17.9) on function versus exercise (SOE: Low).

Physical Modalities for Osteoarthritis of the Hand

- One good quality study of low level laser treatment versus sham demonstrated no improvement in terms of function (difference 0.2, 95% CI -0.2 to 0.6) or pain (difference 0.1, 95% CI -0.3 to 0.5) in the short-term (SOE: Low).

Multidisciplinary Rehabilitation for Osteoarthritis of the Hand

- One fair quality trial of multidisciplinary rehabilitation versus waitlist control demonstrated no short-term differences between groups in function (adjusted difference 0.49, 95% CI, -0.09 to 0.37 on 0-36 scale), pain (adjusted difference 0.40, 95% CI, -0.5 to 1.3 on a 0-20 scale) or with regard to the proportion of OARSI OMERACT (Osteoarthritis Research Society International Outcome Measures in Rheumatology) responders (OR 0.82, 95% CI, 0.42 to 1.61) (SOE: Low for all outcomes).

Comparative Effectiveness of Interventions for Osteoarthritis

- Hip Osteoarthritis: Manual therapy was associated with a small effect on pain in the short term (mean differences of -0.72 [95% CI -1.38 to -0.05] for pain at rest) and -1.21 [95% CI -2.29 to -0.25] for pain walking) versus exercise (SOE: Low).

Key Question 4: Fibromyalgia

Exercise for Fibromyalgia

- Exercise was associated with slightly greater effects on function compared with an attention control, no treatment, or usual care in the short-term (7 trials, pooled difference -7.61 on a 0 to 100 scale, 95% CI, -12.78 to -2.43, I^2= 59.9%) (SOE: Low) and intermediate-term (8 trials, pooled difference, -6.04, 95% CI –9.05 to -3.03, I^2= 0%) (SOE: Moderate). There were no clear effects long term (3 trials, pooled difference, -4.33, 95% CI -10.18 to 1.52, I^2= 0%) (SOE: Low).
- Exercise had a slightly greater effect on VAS pain (0-10 scale) compared with usual care, an attention control or no treatment short term (6 trials (excluding outlier trial) pooled difference -0.89, 95% CI -1.32 to -0.46 I^2 = 0%) but there were no clear effects at intermediate term (6 trials, pooled difference -0.31, 95% CI -0.79 to 0.17, I^2= 5.4%) or long term (4 trials, pooled difference -0.18, 95% CI -0.77 to 0.42, I^2= 0%) (SOE: Moderate for all time frames).

Psychological Therapies for Fibromyalgia

- Cognitive behavioral therapy (CBT) was associated with a slightly greater effect on the Fibromyalgia Impact Questionnaire (FIQ) Total Score than usual care or waitlist in the short-term (2 trials, pooled mean difference -10.67, 95%CI -17 to -4.30 I^2 = 0%, 0-100 scale). The pooled estimate at intermediate-term was not statistically significant due to heterogeneity, however individual trials showed a greater effect than usual care and a third trial using 0-10 FIQ Physical Impairment Scale showed a greater effect of CBT than an attention control (mean difference -1.8, 95% CI -2.9 to -0.70) (SOE: Low for short-term and intermediate-term).
- Psychological therapies were associated with a slightly greater improvement in pain compared with usual care, waitlist, or an attention control in the short term (4 trials, pooled mean difference -0.74, 95%CI -1.20 to -0.28, I^2 = 0%) and intermediate term (3 trials, pooled mean difference -0.67, 95% CI -1.21 to -0.31, I^2 = 36.7%) (SOE: Low for short term and intermediate term).

Physical Modalities for Fibromyalgia

- One fair-quality trial showed no differences between magnetic mattress pads compared with sham or usual care in intermediate-term function (difference on the 0-80 scale FIQ -5.0, 95% CI -14.1 to 4.1 vs. sham and -5.5, 95% CI -14.4 to 3.4 vs. usual care) or pain (difference -0.6, 95% CI -1.9 to 0.7 and -1.0, 95% CI -2.2 to 0.2, respectively on a 0-10 Numerical Rating Scale [NRS]) (SOE: Low).

Manual Therapies for Fibromyalgia

- Myofascial release therapy was associated with a slightly greater effect on intermediate-term function as measured by the FIQ (mean 58.6 ± 16.3 vs. 64.1 ± 18.1 on a 100 point scale, p=0.048 for group by repeated measures ANOVA), but not long-term function (mean 62.8 ± 20.1 vs. 65.0 ± 19.8 on the FIQ, 0-100 scale, p=0.329), compared with sham in one fair-quality trial (SOE: Low).
- Myofascial release therapy was associated with slightly greater effects on long-term pain based on the sensory (18.2 ± 8.3 vs. 21.2 ± 7.9 on a 0-33 scale, p=0.038 for group by

repeated measures ANOVA) and evaluative (23.2 ± 7.6 vs. 26.7 ± 6.9 on a 0-42 scale, p=0.036) domains of the McGill Pain Questionnaire (MPQ) in one fair quality trial; there were no differences for the affective domain of the MPQ or for VAS pain (SOE: Low).

Mindfulness-Based Stress Reduction Therapy for Fibromyalgia

- No clear short-term effects of Mindfulness-based Stress Reduction (MBSR) were seen on function compared with waitlist or an attention control (difference 0 to 0.06 on a 0-10 scale) in two trials (one fair and one poor quality) (SOE: Moderate).
- No clear short-term effects of MBSR on pain (difference 0.1 on a 0-100 VAS pain scale in one poor quality trial; difference -1.38 to -1.59 on the affective and -0.28 to -0.71 on the sensory dimension [scales not reported] of the Pain Perception Scale in one fair-quality trial) compared with waitlist or an attention control in two trials (SOE: Moderate). Intermediate and long-term outcomes were not reported.

Mind-Body Therapy for Fibromyalgia

- Over the short-term, two trials of mind-body practices reported slight improvement in function for qigong (mean difference -7.5, 95% CI -13.3 to -1.68) and substantial improvement for tai chi (mean difference -23.5, 95% CI -30 to -17) based on 0-100 scale total FIQ score; heterogeneity may be explained by duration and intensity of intervention and control condition (SOE: Low).
- Qigong and tai chi were associated with moderately greater improvement in pain (0-10 scale) compared with waitlist and an attention control in the short term (2 trials, pooled MD -1.54, 95% CI -2.67, -0.41, I2=75%) (SOE: Low).

Acupuncture for Fibromyalgia

- Acupuncture was associated with slightly greater effects on function based on 0-100 FIQ Total Score in patients with fibromyalgia than sham acupuncture in the short-term (2 trials, pooled difference -8.63, 95% CI =12.12 to -5.13, I^2 = 0%) and intermediate-term (2 trials, pooled difference -9.41, 95% CI -13.96 to -4.85, I^2 = 27.4%) (SOE: Moderate).
- There was no clear effect of acupuncture on pain (0-10 scale) versus sham acupuncture in the short term (3 trials, pooled difference -0.13, 95% CI -1.06 to 0.79, I^2 =72%) or intermediate term (3 trials, pooled difference – 0.53, 95%CI -1.15 to 0.09 I^2 =45.5%) (SOE: Low)

Multidisciplinary Rehabilitation for Fibromyalgia

- There were no clear effects of multidisciplinary treatment for fibromyalgia on function versus usual care based on a 0-100 FIQ total score in the short-term (2 trials, pooled mean difference -5.06, 95% CI -12.38 to 2.25, I^2 =76.2%), however it was associated with a slightly greater effect at the intermediate term (3 trials, pooled difference -7.84, 95% CI -11.43 to -4.25, I^2 = 18.2%) and long term (2 trials, pooled difference -8.42, 95%CI -13.76 to -3.08, I^2 = 24.9%). More multidisciplinary treatment participants experienced a clinically meaningful improvement in FIQ total score compared with usual care at short (OR 3.1, 95% CI 1.6 to 6.2), intermediate (1 trial OR 3.1, 95% CI 1.5 to 6.4) and long term (OR 8.8, 95% CI 2.5 to 30.9) (SOE: Low for short, intermediate and long term).

- There were no clear effects of multidisciplinary treatment for fibromyalgia on pain versus usual care or waitlist in the short-term (2 trials pooled difference on 0-10 scale -0.24, 95% CI -0.63 to 0.15, $I^2 = 0\%$), however multidisciplinary treatment was associated with a slightly greater effect on pain compared with usual care or waitlist at the intermediate term (3 trials, pooled difference -0.68, 95%CI -1.07 to -0.30, $I^2 = 0\%$), but there were no clear differences compared with usual care long term (2 trials, pooled difference -0.25, 95% CI -0.68 to 0.17, $I^2 = 0\%$) (SOE: Low for short, intermediate and long term).

Comparative Effectiveness of Interventions for Fibromyalgia

- There was no clear effect of multidisciplinary pain treatment versus aerobic exercise at long term in one fair-quality trial for function the FIQ total score (difference -1.10, 95% CI -8.40 to 6.20, 0-100 FIQ total score) or pain (difference 0.10, 95% CI -0.67 to 0.87, 0-10 FIQ pain scale) (SOE: Low).
- CBT was associated with a small benefit for function (difference -4.0 on the 0-100 FIQ, 95% CI -7.7 to -0.27), but not for pain (difference 0.2 on a 0-100 VAS, 95% CI -4.0 to 4.4), compared with pregabalin over the intermediate-term in one fair-quality trial (SOE: Low).

Key Question 5: Chronic Tension Headache

Manual Therapies for Tension Headache

- Spinal manipulation therapy was associated with small to moderate improvements, respectively, in function compared with usual care (difference -5.0, 95% CI -9.02 to -1.16 on the Headache Impact Test, scale 36-78 and difference -10.1, 95% CI -19.5 to -0.64 on the Headache Disability Inventory, scale 0-100) and with moderate improvements pain intensity (difference -1.4 on a 0-10 Numerical Rating Scale [NRS] scale, 95% CI -2.69 to -0.16) over the short-term (SOE: Low). Approximately a quarter of the patients had comorbid migraine.

Acupuncture for Tension Headache

- Laser acupuncture was associated with slight improvement in pain intensity (median difference -2, IQR 6.3, on a 0-10 VAS scale) and in the number of headache days per month (median difference -8, IQR 21.5) in one fair-quality trial (SOE: Low).

Comparative Effectiveness of Interventions for Chronic Tension Headache

- No studies compared the interventions of interest to biofeedback and evidence from comparisons with pharmacological interventions was insufficient.

Key Question 6: Differential Efficacy

Evidence was insufficient to determine whether factors such as age, sex or comorbidities modify the effects of treatment.

Harms

Although data on harms were limited, no evidence suggested serious harms for the interventions included in the review. Many trials did not report harms, withdrawals due to adverse events, or differences between compared interventions in risk of harms or withdrawals.

Trials that did report such data found infrequent or rare occurrences of non-serious treatment-related adverse events (e.g., discomfort, soreness, bruising, increased pain, worsening of symptoms), few withdrawals from nonpharmacological treatments due to adverse events, and no differences between comparison groups in frequency of intervention-related adverse events or withdrawals.

Discussion

Key Findings and Strength of Evidence

The key findings of this review, including strength of evidence ratings, are summarized for each chronic pain condition in Tables A–M. (Interventions and comparators with insufficient or no evidence for either function or pain outcomes are not shown.) Domains used to determine the overall strength of evidence are shown in Appendix G of the full report. The strength of evidence was low or insufficient for many interventions and was limited by small numbers of trials available for specific comparisons and for our specified time frames, particularly for long-term followup. We focused on evaluating the persistence of effects for therapies at least 1 month beyond the course of treatment, using the following definitions for post-intervention followup: short-term (1 to 6 months), intermediate-term (6 to 12 months) and long-term (≥12 months). Little long-term evidence was available across conditions and interventions. The majority of trials compared interventions with usual care, and very few trials employed pharmacological treatments or exercise as comparators. In general, effect sizes for most interventions were small, based on mean differences. There tended to be more evidence for the effects of interventions on pain than for function and the effects on function were generally smaller or not clearly present. No trials directly compared interventions with opioids. Our previous reports suggest small to moderate effects of opioids on pain during treatment only (effects would not be expected to persist) with evidence primarily from short-term trials.[9,19,20] Information on adherence to interventions was not well reported; poor adherence may have impacted some of our findings. Harms were poorly reported across interventions. No serious intervention-related adverse events requiring medical attention were identified; reported adverse events were generally minor (e.g., muscle soreness with exercise, bruising with acupuncture) and time-limited (e.g., temporary worsening of pain).

Table A. Nonpharmacological interventions for chronic low back pain compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	x	x	+	++	++
	SOE	✓	✓	✓	✓✓	✓	✓
Psychological Therapies	Effect size	+	+	+	+	+	+
	SOE	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Physical Modalities: Short-Wave Diathermy	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Physical Modalities: Ultrasound	Effect size	?	⊘	⊘	x	⊘	⊘
	SOE	○	⊘	⊘	✓	⊘	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Physical Modalities: Low Level Laser Therapy	Effect size	+	x	⊘	++	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Manual Therapies: Spinal Manipulation	Effect size	+	+	⊘	x	+	⊘
	SOE	✓	✓	⊘	✓	✓✓	⊘
Manual Therapies: Massage	Effect size	+	x	⊘	+	x	⊘
	SOE	✓✓	✓	⊘	✓✓	✓	⊘
Manual Therapies: Traction	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Mindfulness Practices: MBSR	Effect size	+	+	⊘	+	+	⊘
	SOE	✓	✓	⊘	✓✓	✓	⊘
Mind-Body Practices: Yoga	Effect size	+	+	⊘	++	++	⊘
	SOE	✓✓	✓	⊘	✓	✓✓	⊘
Acupuncture	Effect size	+	x	x	+	x	+
	SOE	✓	✓	✓	✓✓	✓	✓
Multidisciplinary Rehabilitation	Effect size	+	+	x	+	+	x
	SOE	✓	✓	✓	✓✓	✓✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
MBSR = mindfulness-based stress reduction; SOE = strength of evidence

Table B. Nonpharmacological interventions for chronic <u>low back pain</u> compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies	Effect size	⊘	?	?	⊘	?	?
	SOE	⊘	○	○	⊘	○	○
Physical Modalities: Low-Level Laser Therapy	Effect size	⊘	x	⊘	⊘	+	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Manual Therapies: Spinal Manipulation	Effect size	x	x	⊘	x	+	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Manual Therapies: Massage	Effect size	⊘	+	⊘	⊘	+	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Mind-Body Practices: Yoga	Effect size	x	x	⊘	+	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Mind-Body Practices: Qi Gong	Effect size	x	+	⊘	+	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Multidisciplinary Rehabilitation	Effect size	+	+	x	+	+	x
	SOE	✓✓	✓✓	✓	✓✓	✓✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table C. Nonpharmacological interventions for chronic <u>neck pain</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Psychological Therapies	Effect size	x	x	⊘	x	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Physical Modalities: Electromagnetic Field	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Physical Modalities: Low Level Laser Therapy	Effect size	++	⊘	⊘	++	⊘	⊘
	SOE	✓✓	⊘	⊘	✓✓	⊘	⊘
Manual Therapies: Traction	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Manual Therapies: Massage	Effect size	x	x	⊘	x	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Mind-Body Practices: Alexander Technique	Effect size	+	+	⊘	⊘	x	⊘
	SOE	✓	✓	⊘	⊘	⊘	⊘
Acupuncture	Effect size	+	+	x	x	x	x
	SOE	✓	✓	✓	✓	✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table D. Nonpharmacological interventions for chronic neck pain compared with pharmacological treatments: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	?	?	?	?	?	?
	SOE	○	○	○	○	○	○
Acupuncture	Effect size	?	⊘	⊘	⊘	⊘	⊘
	SOE	○	⊘	⊘	⊘	⊘	⊘

\+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table E. Nonpharmacological interventions for chronic neck pain compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies	Effect size	X	X	⊘	⊘	⊘	⊘
	SOE	✓	✓	⊘	⊘	⊘	⊘
Manual Therapies: Massage	Effect size	⊘	⊘	⊘	X	⊘	⊘
	SOE	⊘	⊘	⊘	✓	⊘	⊘
Mind-Body Practices: Body Awareness Therapy	Effect size	X	⊘	⊘	⊘	⊘	⊘
	SOE	✓	⊘	⊘	⊘	⊘	⊘

\+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table F. Nonpharmacological interventions for knee osteoarthritis compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	+	+	++	X
	SOE	✓✓	✓	✓	✓✓	✓	✓
Psychological Therapies	Effect size	X	X	X	X	X	X
	SOE	✓	✓	✓	✓	✓	✓
Physical Modalities: Microwave Diathermy	Effect size	+++	⊘	⊘	+++	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Physical Modalities: Ultrasound	Effect size	+++	X	⊘	+++	X	⊘
	SOE	✓	✓	⊘	✓	✓	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Physical Modalities: TENS	Effect size	⊘	x	⊘	⊘	x	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Physical Modalities: Low-Level Laser Therapy	Effect size	?	?	⊘	?	?	⊘
	SOE	○	○	⊘	○	○	⊘
Physical Modalities: Electromagnetic Field	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Physical Modalities: Braces	Effect size	⊘	?	?	?	?	⊘
	SOE	⊘	○	○	○	○	⊘
Manual Therapies: Joint Manipulation	Effect size	⊘	?	⊘	⊘	⊘	⊘
	SOE	⊘	○	⊘	⊘	⊘	⊘
Manual Therapies: Massage	Effect size	⊘	?	⊘	⊘	?	⊘
	SOE	⊘	○	⊘	⊘	○	⊘
Mind-Body Practices: Tai Chi	Effect size	?	?	⊘	?	?	⊘
	SOE	○	○	⊘	○	○	⊘
Acupuncture	Effect size	x	x	⊘	x	x	⊘
	SOE	✓✓	✓	⊘	✓✓	✓	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
TENS = transcutaneous electrical nerve stimulation; SOE = strength of evidence

Table G. Nonpharmacological interventions for <u>hip osteoarthritis</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	?	+	x	?
	SOE	✓	✓	○	✓	✓	○

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table H. Nonpharmacological interventions for <u>hip osteoarthritis compared with exercise</u>: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Manual Therapies	Effect size	++	+	⊘	+	?	⊘
	SOE	✓	✓	⊘	✓	○	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table I. Nonpharmacological interventions for <u>hand osteoarthritis</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Physical Modalities: Low Level Laser Therapy	Effect size	X	⊘	⊘	X	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Multidisciplinary Rehabilitation	Effect size	X	⊘	⊘	X	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table J. Nonpharmacological interventions for <u>fibromyalgia</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	X	+	X	X
	SOE	✓	✓✓	✓	✓✓	✓✓	✓✓
Psychological Therapies: CBT	Effect size	+	+	?	+	+	?
	SOE	✓	✓	○	✓	✓	○
Psychological Therapies: Biofeedback, Imagery	Effect size	?	?	?	+	+	?
	SOE	○	○	○	✓	✓	○
Physical Modalities: Magnetic Pads	Effect size	⊘	X	⊘	⊘	X	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Manual Therapies: Massage (Myofascial Release)	Effect size	⊘	+	X	?	+	X
	SOE	⊘	✓	✓	○	✓	✓
Mindfulness Practices: MBSR	Effect size	X	⊘	⊘	X	⊘	⊘
	SOE	✓✓	⊘	⊘	✓✓	⊘	⊘
Mind-Body Practices: Qigong, Tai Chi	Effect size	+	⊘	⊘	++	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Acupuncture	Effect size	+	+	?	X	X	?
	SOE	✓✓	✓✓	○	✓	✓	○
Multidisciplinary Rehabilitation	Effect size	X	+	+	X	+	X
	SOE	✓	✓	✓	✓	✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
CBT = cognitive behavioral therapy; SOE = strength of evidence

Table K. Psychological therapy for fibromyalgia compared with pharmacological treatments: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
CBT vs. pregabalin; duloxetine	Effect size	⊘	+	⊘	⊘	X	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
CBT = cognitive behavioral therapy; SOE = strength of evidence

Table L. Nonpharmacological interventions for fibromyalgia compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapy	Effect size	?	?	?	?	?	?
	SOE	○	○	○	○	○	○
Multidisciplinary Rehabilitation	Effect size	⊘	⊘	X	⊘	⊘	X
	SOE	⊘	⊘	✓	⊘	⊘	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table M. Nonpharmacological interventions for chronic tension headache compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Manual Therapies: Manipulation	Effect size	+	⊘	⊘	++	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Acupuncture	Effect size	⊘	⊘	⊘	++ (laser) ? (needle)	?	?
	SOE	⊘	⊘	⊘	○	○	○

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; X No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
OTES = occipital transcutaneous electrical stimulation; SOE = strength of evidence

Findings in Relationship to What is Already Known

Many reviews have addressed the effects of interventions for chronic pain management during or immediately following treatments. We focused on evaluating the sustainability of effects for at least one month post-intervention.

This review updates our previous review on low back pain[19] by incorporating new evidence on nonpharmacological treatments for chronic low back pain based on primary literature and devotes more attention to describing effects over short, intermediate and long terms. Consistent with the prior review, small to moderate effects of exercise, yoga, various psychological therapies, acupuncture, spinal manipulation and low-level laser therapy were identified. This review suggests that most effects are at short or intermediate-term followup; long-term data are sparse.

The recent Institute for Clinical and Economic Review (ICER) review[22] on chronic low back pain and neck pain used relevant portions of our previous review for chronic low back pain. New publications listed in the ICER report were included in our current review if they met our inclusion criteria. Our findings are generally consistent with the ICER report relative to chronic low back pain; differences between the reports for chronic neck pain are likely due to differences in inclusion criteria (particularly related to followup post intervention). The ICER report suggests that cognitive and mind-body therapies for treatment of chronic low back pain and chronic neck pain would be cost-effective, resulting in only a small increase ($0.75) per member per month for a hypothetical payer plan covering 1 million members compared with approximately $4.46 per member per month for pain medication.

Our findings indicate that a number of nonpharmacological treatments improve pain and/or function. This is consistent with other reviews including recent reviews on exercise[23] and complementary health approaches[24] for chronic pain management, an AHRQ report on knee osteoarthritis treatment,[25] and a review of chronic pain treatment guidelines on the use of manual and physical therapies.[26]

Applicability

The applicability of our findings may be impacted by a number of factors. Included trials provided limited information on diagnostic criteria, symptom duration, clinical characteristics, comorbid conditions and concomitant treatments, thus it is not clear to what extent this reflects the populations seen in clinical practice or may impact our results. Information on overlapping chronic pain conditions or psychosocial factors was generally not provided in included trials. The extent to which these characteristics were present in trial populations and their impact on our results is not clear. Across conditions, the majority of trial participants were female with trials of fibromyalgia and many in chronic neck pain enrolling female participants exclusively with fewer female participants in trials of low back pain, osteoarthritis, and headache (57%, 68% 75% respectively). Trials also included a broad range of ages, with trials of headache and fibromyalgia generally enrolling younger participants (30 to 50 years old) than those with osteoarthritis (52 to 76 years old). Within each condition, symptom duration varied (e.g., 3 months to 15 years for low back pain, 9 months to 15 years for neck pain). Pain severity for most conditions appeared to be moderate. While we excluded trials where <90 percent of the study sample had the defined condition, there was still likely heterogeneity. For example, some low back pain trials included a small number of patients with radiculopathy and in some trials of chronic tension headache, a large proportion of participants had concomitant migraine headache

and likely medication overuse headache. Most trials excluded persons with serious medical and psychological comorbidities. Our findings are generally most applicable to persons without such comorbidities who have moderate or severe intensity pain that has persisted for >1 year. The heterogeneity in populations across included trials likely is consistent with the heterogeneity seen in clinical practice, so our findings are likely applicable to most primary care clinical settings.

Variability in interventions and comparators may impact applicability. For interventions, there was variability in the numbers of sessions, length of sessions and duration of treatment as well as methods of delivering the intervention. In addition, there was heterogeneity within intervention categories with regard to techniques or methods used. In general, there were no clear differences in effects based on differences in techniques, numbers of sessions, etc. However; conclusions are limited by the relatively small numbers of trials available for stratified analysis. Heterogeneity across and within comparators are another consideration: details of usual care were rarely provided, details of co-interventions were rarely reported and likely varied across trials, and it is assumed that all patients received some sort of "usual care." While the heterogeneity precludes drawing conclusions regarding specific, optimal techniques and their delivery, it likely represents the conditions under which the various interventions are currently delivered and speaks to the need for research to identify what may be optimal.

Implications for Clinical and Policy Decisionmaking

Our review provides evidence that an array of nonpharmacological treatments provide small to moderate improvements in function and pain that are durable for more than 1 month for the five conditions addressed in this review. These encompass the vast majority of chronic pain conditions for which people seek treatment in the United States. The evidence synthesized in this review may help inform guidelines and health care policy (including reimbursement policy) related to use of noninvasive, nonpharmacological treatments as alternatives to opioids for these conditions, and inform policy decisions regarding funding priorities for future research. Recent guidelines from the Centers for Disease Control and Prevention (CDC)[11] in the United States and the Canadian Guidelines for Opioid Use in Chronic Non-Cancer Pain[26] recommend non-opioid treatment as preferred treatment for chronic pain. Further, American College of Physicians guidelines recommend nonpharmacological therapies over medications for chronic back pain.[21] Our findings confirm the feasibility of these guidelines by showing that there are nonpharmacological treatments for chronic pain that have evidence of sustained effectiveness after the completion of therapy. Importantly, some interventions, such as exercise, multidisciplinary rehabilitation, mind-body interventions, and some complementary and integrative medicine therapies such as acupuncture and spinal manipulation also were associated with some sustained effects on function. At the same time, there was no evidence suggesting serious harms, although data on harms were limited.

Evidence reviewed in our report may also help inform decisions regarding prioritization of nonpharmacological therapies by clinicians selecting therapy. Consistent with a biopsychosocial understanding of chronic pain,[4,5] evidence was somewhat more robust for "active" interventions that engage patients in movement and address psychological contributors to pain, particularly at longer-term followup, versus more "passive" treatments focused on symptom relief. Active interventions include exercise, multidisciplinary rehabilitation, cognitive-behavioral therapy, and mind-body interventions. This provides some support for clinical strategies that focus on "active" interventions as primary therapies, with "passive" interventions used in a more adjunctive or supplementary role.

Our review also has policy implications related to treatment access and reimbursement. Given heterogeneity in chronic pain, variability in patient preferences for treatments, and differential responses to specific therapies in patients with a given chronic pain condition, policies that broaden access to a broader array of effective nonpharmacological treatments may have greater impact than those that focus on one or a few therapies. Several considerations could inform policy decisions regarding access to and coverage of nonpharmacological therapies. Efforts could prioritize access to interventions with evidence of persistent effectiveness across different pain conditions, such as exercise, multidisciplinary rehabilitation, mind-body interventions, and acupuncture. Because the level of supporting evidence varies from condition to condition, policymakers may need to consider the degree to which evidence may be reasonably extrapolated across conditions (e.g., effectiveness of psychological therapies for back pain to neck pain). Although the Affordable Care Act has improved access to CAM therapies, variability in reimbursement and authorization procedures remain a potential barrier. Although evidence supports the use of multidisciplinary rehabilitation over exercise therapy or usual care, primarily for low back pain, cost and availability remain important barriers. Our report suggests that less-intensive multidisciplinary rehabilitation may be similarly effective to high-intensity multidisciplinary rehabilitation, which could inform decisions about how such interventions are designed and delivered. In addition, not all patients may require multidisciplinary rehabilitation.[27] Policy efforts that focus on use of multidisciplinary rehabilitation in persons more likely to benefit (e.g., severe functional deficits, failure to improve on standard nonmultidisciplinary therapies, significant psychosocial contributors to pain) could also inform efforts to deliver this modality efficiently.

Limitations of the Evidence Base and the Systematic Review Process

A number of limitations to the evidence base should be noted. First, evidence was relatively sparse for most interventions, particularly with regard to long-term outcomes. Data on outcomes other than pain and function were very limited. Only 5 percent of included trials across conditions were considered to be of good quality; the majority were considered fair (58%).

A number of limitations to the systematic review process should also be noted. To maintain a reasonable scope for this review, inclusion was limited to trials for five common chronic pain conditions: chronic low back pain, chronic neck pain, osteoarthritis of the knee, hip or hand, fibromyalgia, and chronic tension headache, which comprise the vast majority of chronic pain conditions. Our analysis was restricted to trials that reported outcomes after at least 1 month following the end of therapy. We focused on comparisons involving usual care/nonactive therapies and to active treatments of pharmacological interventions and exercise or biofeedback to provide a common point of comparison for individual interventions, based on input from the TEP. Our meta-analyses sometimes included only two or three trials. Meta-analyses based on small numbers of trials must be interpreted with caution.

Research Recommendations

A number of evidence gaps preclude full understanding of the effectiveness, comparative effectiveness and harms of noninvasive, nonpharmacological treatments for chronic low back pain, chronic neck pain, osteoarthritis, fibromyalgia and chronic tension headache. Comparative trials examining effects beyond the immediate post-intervention period, including considerations of attrition and motivation to continue, are needed to better understand whether benefits are sustained over time. Incorporation of pragmatic trial designs that incorporate strategies to

improve participant recruitment, adherence, and continuation could improve retention in trials and facilitate understanding of sustainability of effects. Research to identify optimal techniques and their delivery would help define more standardized interventions to evaluate in future trials. Standardization of protocols and outcomes measures would facilitate comparison of results across trials. Routine collection of common or known harms associated with interventions is needed in future trials.

Outcome measures such as VAS or NRS may not fully capture the impact of pain or allow for accurate classification or evaluation of changes in chronic pain. Inclusion of recommendations for pain assessment[28] that incorporate understanding of pathophysiological mechanisms and address multiple domains of pain, including temporal dimensions, sensory and affective qualities of pain and the location and bodily distribution of pain in trial planning and execution may facilitate more accurate classification and longitudinal tracking of response to interventions. Reporting the proportions of patients achieving a clinically meaningful improvement in pain, function, or quality of life as measures of "success" may provide additional clinical information to complement data on average changes in continuous measures of pain, function, and quality of life for which there is difficulty describing clinically important effects.

There is heterogeneity with regard to research design, execution and outcomes reporting in trials of interventions included in this review compared with well-funded trials of devices or pharmacological agents. Lack of funding to design methodologically sound studies with reasonable sample size of nonpharmacological interventions may have contributed to the general low quality of evidence. Education of researchers examining nonpharmacological approaches to pain management on clinical trial design, execution, and analysis may also assist with improving the quality of the evidence base for many of the interventions.

Conclusions

Our review provides evidence that an array of nonpharmacological treatments provide small to moderate improvements in function and/or pain that are durable for more than one month for the five common chronic pain conditions addressed. Our findings provide some support for clinical strategies that prioritize use of non-pharmacological therapies for chronic pain, including "active therapies" such as exercise, psychological therapies, multidisciplinary rehabilitation, and mind-body interventions. Additional comparative research on the sustainability of effects beyond the immediate post-treatment period is needed for chronic pain conditions other than low back pain.

References

1. Ballantyne JC, Shin NS. Efficacy of opioids for chronic pain: a review of the evidence. Clin J Pain. 2008 Jul-Aug;24(6):469-78. doi: 10.1097/AJP.0b013e31816b2f26. PMID: 18574357.
2. Eriksen J, Sjogren P, Bruera E, et al. Critical issues on opioids in chronic non-cancer pain: an epidemiological study. Pain. 2006 Nov;125(1-2):172-9. doi: 10.1016/j.pain.2006.06.009. PMID: 16842922.
3. Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms. Prepared by the International Association for the Study of Pain, Subcommittee on Taxonomy. Pain Suppl. 1986;3:S1-226. PMID: 3461421.
4. Institute of Medicine. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington, DC: The National Academies Press; 2011.
5. National Pain Strategy Task Force. National Pain Strategy: A Comprehensive Population Health-Level Strategy for Pain. Interagency Pain Research Coordinating Committee (IPRCC), National Institutes of Health (NIH); 1-83. 2015. https://iprcc.nih.gov/National_Pain_Strategy/NPS_Main.htm.
6. Boudreau D, Von Korff M, Rutter CM, et al. Trends in long-term opioid therapy for chronic non-cancer pain. Pharmacoepidemiol Drug Saf. 2009 Dec;18(12):1166-75. doi: 10.1002/pds.1833. PMID: 19718704.
7. Olsen Y, Daumit GL, Ford DE. Opioid prescriptions by U.S. primary care physicians from 1992 to 2001. J Pain. 2006 Apr;7(4):225-35. doi: 10.1016/j.jpain.2005.11.006. PMID: 16618466.
8. Sullivan MD, Edlund MJ, Fan MY, et al. Trends in use of opioids for non-cancer pain conditions 2000-2005 in commercial and Medicaid insurance plans: the TROUP study. Pain. 2008 Aug 31;138(2):440-9. doi: 10.1016/j.pain.2008.04.027. PMID: 18547726.
9. Chou R, Turner JA, Devine EB, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a National Institutes of Health Pathways to Prevention Workshop. Ann Intern Med. 2015 Feb 17;162(4):276-86. doi: 10.7326/M14-2559. PMID: 25581257.
10. Centers for Disease Control and Prevention. Vital signs: overdoses of prescription opioid pain relievers---United States, 1999--2008. MMWR Morb Mortal Wkly Rep. 2011 Nov 4;60(43):1487-92. PMID: 22048730.
11. Dowell D, Haegerich TM, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain--United States, 2016. JAMA. 2016 Apr 19;315(15):1624-45. doi: 10.1001/jama.2016.1464. PMID: 26977696.
12. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
13. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from http://handbook.cochrane.org.; 2011.
14. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015 Nov;40(21):1660-73. doi: 10.1097/BRS.0000000000001061. PMID: 26208232.
15. A collection of R functions supporting the text book Modern Epidemiology, Second Edition, by Kenneth J. Rothman and Sander Greenland. GitHub, Inc.; 2017. https://github.com/epijim/episheet. Accessed October 13, 2017.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986 Sep;7(3):177-88. PMID: 3802833.
17. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. Stat Med. 1996 Mar 30;15(6):619-29. doi: 10.1002/(SICI)1097-

0258(19960330)15:6<619::AID-SIM188>3.0.CO;2-A. PMID: 8731004.

18. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 06;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

19. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatment for Low Back Pain. Comparative Effectiveness Review No. 169. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. HHSA 290-2012-00014-I.) AHRQ Publication No. 16-EHC004-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm. PMID: 26985522.

20. Chou R, Deyo R, Friedly J, et al. Systemic pharmacologic therapies for low back pain: A systematic review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2017 Feb 14;166:[Epub ahead of print]. doi: 10.7326/M16-2458. PMID: 28192790.

21. Chou R, Deyo R, Friedly J, et al. Nonpharmacologic therapies for low back pain: A systematic review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2017 Feb 14;166:[Epub ahead of print]. doi: 10.7326/M16-2459. PMID: 28192793.

22. Tice J, Kumar V, Otunoye I, et al. Cognitive and Mind-Body Therapies for Chronic Low Back and Neck Pain: Effectiveness and Value. Evidence Report. Prepared for The California Technology Assesment Forum. Boston, MA: The Institute for Clinical and Economic Review; 2017. https://icer-review.org/wp-content/uploads/2017/03/CTAF_Chronic_Pain__Evidence_Report_100417.pdf Accessed October 13, 2017.

23. Geneen LJ, Moore RA, Clarke C, et al. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. The Cochrane Library. 2017.

24. Nahin RL, Boineau R, Khalsa PS, et al. Evidence-Based Evaluation of Complementary Health Approaches for Pain Management in the United States. Mayo Clin Proc. 2016 Sep;91(9):1292-306. doi: 10.1016/j.mayocp.2016.06.007. PMID: 27594189.

25. Newberry SJ, FitzGerald J, SooHoo NF, et al. Treatment of Osteoarthritis of the Knee: An Update Review. Comparative Effectiveness Review No. 190. (Prepared by the RAND Southern California Evidence-based Practice Center under Contract No. 290-2015-00010-I.) AHRQ Publication No.17-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2017. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

26. Busse JW, Craigie S, Juurlink DN, et al. Guideline for opioid therapy and chronic noncancer pain. CMAJ. 2017 May 8;198(18):E659-E66. doi: 10.1503/cmaj.170363. PMID: 28483845.

27. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet. 2011 Oct 29;378(9802):1560-71. doi: 10.1016/S0140-6736(11)60937-9. PMID: 21963002.

28. Fillingim RB, Loeser JD, Baron R, et al. Assessment of Chronic Pain: Domains, Methods, and Mechanisms. J Pain. 2016 Sep;17(9 Suppl):T10-20. doi: 10.1016/j.jpain.2015.08.010. PMID: 27586827.

Introduction

Background

Nature and Burden of Chronic Pain

Chronic pain substantially impacts physical and mental functioning, productivity, quality of life, and family relationships; it is the leading cause of disability and is often refractory to treatment.[1,2] A monumental public health challenge, chronic pain affects millions of adults in the United States, with a conservative annual cost in personal and health system expenditures estimated at $560 billion to $635 billion.[3]

Chronic pain is defined as pain lasting 12 weeks or longer or persisting past the normal time for tissue healing.[3,4] Nervous system changes that occur with chronic pain, combined with its psychological and cognitive impacts, have led to conceptualization of chronic pain as a distinct disease entity.[3] This multifaceted disease is influenced by multiple factors (e.g., genetic, central nervous system, psychological, and environmental factors) and complex interactions, making pain assessment and management a challenge. A number of characteristics influence the development of and response to chronic pain, including sex, age, presence of co-morbidities, and psychosocial factors. For example, women report chronic pain more frequently than do men, are at higher risk for some conditions such as fibromyalgia,[3] and may respond to treatment differently than men. Older adults are more likely to have co-morbidities and are more susceptible to polypharmacy, impacting choices and consequences of therapies. Pain is greatly influenced by psychosocial factors, which may predict who will develop chronic disabling pain as well as who will respond to various treatments.

Management of Chronic Pain

The National Pain Strategy (NPS) report recommends that pain management be integrated, multimodal, interdisciplinary, evidence-based, and tailored to individual patient needs.[5] In addition to addressing biological factors when known, optimal management of chronic pain must also address psychosocial contributors to pain, while taking into account individual susceptibility and treatment responses. Self-care is also an important part of chronic pain management.

The NPS points to the "dual crises" of chronic pain and opioid dependence, overdose, and death as providing important context for consideration and implementation of alternative chronic pain management strategies. Nationally, concerns regarding appropriate use, misuse, and diversion of opioids for treatment of chronic pain have been the subject of numerous scientific and news reports and were highlighted in the NPS report[5] and a 2011 Institute of Medicine (IOM)[3] report. Although opioid prescriptions for chronic pain have increased substantially in the past 20 years, evidence shows only modest short-term benefits.[6-8] Lack of evidence on long-term effectiveness[9] and serious safety concerns[10] speak to the need to consider alternative treatments to opioids. The recent evidence-based CDC guidelines on opioid use for chronic pain,[11] which include a recommendation on the preferred use of non-opioid treatment over opioid therapy, has prompted additional primary research on alternative methods of managing chronic pain.

Other pharmacological treatments for chronic pain include nonsteroidal anti-inflammatory drugs (NSAIDs), acetaminophen, muscle relaxants, antiseizure medications, antidepressants, and corticosteroids, used alone or in combination with each other or with opioids. Each has potential side effects and contraindications.

Rationale for This Review

Both the IOM report and the NPS describe the need for evidence-based strategies for the treatment of chronic pain that address the biopsychosocial nature of this disease, including nonpharmacological treatment. These initiatives, and others, speak to the importance of understanding current evidence on noninvasive, nonpharmacological treatment of chronic pain.

The review is intended to address some of the needs described in the NPS[5] and IOM[3] reports and others for evidence to inform guidelines and health care policy (including reimbursement policy) related to use of noninvasive, nonpharmacological treatments as alternatives to opioids and other pharmacological treatments. This review also aims to provide additional insights into research gaps related to use of noninvasive, nonpharmacological alternatives for treating chronic pain.

Scope and Key Questions

This comparative effectiveness review focused on noninvasive, nonpharmacological therapy for five common chronic pain conditions: low back pain, neck pain, osteoarthritis, fibromyalgia, and headache. Individual pain management strategies considered in the review include exercise and physical therapy, mind-body practices (yoga, tai chi, qigong), psychological therapies (cognitive-behavioral therapy, biofeedback, relaxation techniques, acceptance and commitment therapy), interdisciplinary rehabilitation, mindfulness practices (meditation, mindfulness-based stress reduction practices), osteopathic and spinal manipulation, acupuncture, and physical modalities (traction, ultrasound, transcutaneous electrical nerve stimulation, low level laser therapy, interferential therapy, superficial heat or cold, bracing for knee, back or neck, electro-muscular stimulation and magnets), acupuncture, and functional restoration training.

We focused on single active interventions and comparators over the long term. The Key Questions, PICOTS (populations, interventions, comparators, outcomes, timing, settings, and study designs), and analytic framework that guided this review are provided below.

Key Question 1. Adults with chronic low back pain

Key Question 2. Adults with chronic neck pain

Key Question 3. Adults with osteoarthritis-related pain

Key Question 4. Adults with fibromyalgia

Key Question 5. Adults with chronic tension headache

Key Questions 1–5 incorporate the following subquestions.

a. What are the benefits and harms of noninvasive, nonpharmacological therapies compared with sham treatment, no treatment, waitlist, attention control, or usual care?
b. What are the benefits and harms of noninvasive, nonpharmacological therapies compared with pharmacological therapy (e.g., opioids, nonsteroidal anti-inflammatory drugs, acetaminophen, antiseizure medications, antidepressants)?
c. What are the benefits and harms of noninvasive nonpharmacological therapies compared with exercise?

The three-part format for Key Questions 1–5 reflects the following research concepts:

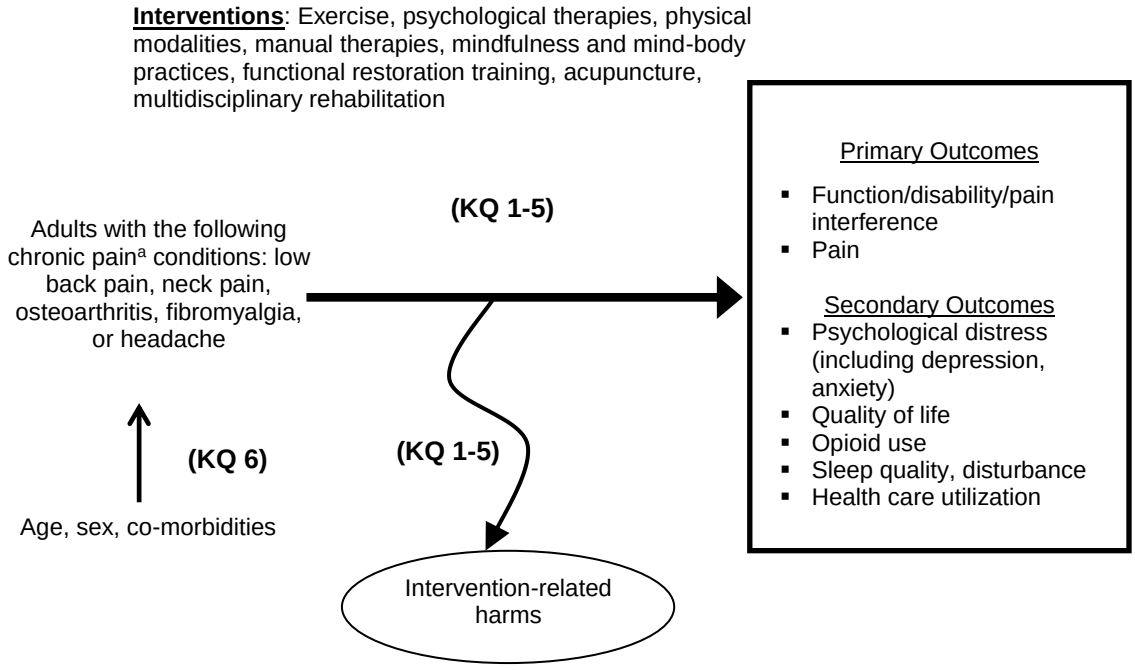
- Part "a" answers the question of whether the various interventions work overall compared with sham, waitlist control, attention control, no treatment, or usual care. For this review, usual care was defined as care that might be provided or recommended by a primary care provider.
- Part "b" answers the question of whether the various interventions work compared with pharmacological alternatives.
- Part "c" answers the question of how outcomes for individual interventions (e.g., acupuncture) compare with a common comparator. Exercise is the most frequent comparison in the literature for many chronic pain conditions, so it provides a common comparator for analysis. It is also recommended in most guidelines for conditions including low back pain, neck pain, fibromyalgia, and osteoarthritis and is widely available. Exercise will serve as common comparator for these conditions. For chronic headache, biofeedback will provide a common comparator for analysis.

Key Question 6. Do estimates of benefits and harms differ by age, sex, or presence of comorbidities (e.g., emotional or mood disorders)?

Analytic Framework

The analytic framework (Figure 1) illustrates the population, interventions, outcomes, and adverse effects that guided the literature search and synthesis.

Figure 1. Analytic framework



KQ = Key Question

[a] Chronic pain is defined as pain lasting ≥ 12 weeks or pain persisting past the normal time for tissue healing.

Methods

The methods for this systematic review follow the Agency for Healthcare Research & Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*[12] and the PRISMA checklist. See the review protocol (http://effectivehealthcare.ahrq.gov/index.cfm) for details.

Topic Refinement and Review Protocol

The review team developed initial Key Questions and PICOTS with input from the AHRQ Task Order Officer (TOO), representatives from the Centers for Disease Control and Prevention (CDC) and the Office of the Assistant Secretary for Planning and Evaluation (ASPE), and a group of Key Informants. The provisional Key Questions, PICOTS, and analytic framework were posted on the AHRQ Web site for public comment from December 27, 2016, to January 23, 2017.

After reviewing public comments, the EPC research team developed the final protocol with input from the AHRQ TOO, CDC and ASPE representatives, and a Technical Expert Panel (TEP) convened for this report. The TEP consisted of nine members with expertise in primary care, rheumatology, pain medicine, behavioral sciences, physical medicine and rehabilitation, and physical therapy. TEP members had expertise in treating patients with one or more of the five conditions included in this report. Many comments from the TEP and public noted that there was substantial heterogeneity within the included chronic pain conditions and within categories of nonpharmacological, noninvasive treatment strategies. Suggestions for including additional chronic pain conditions and additional interventions were made; however, all were considered beyond the scope and resources for this review.

The final version of the protocol for this review was posted on the AHRQ Effective Health Care Program web site (www.effectivehealthcare.ahrq.gov) on April 27, 2017. The protocol was also registered in the PROSPERO database of prospectively registered systematic reviews.

Literature Search Strategy

A research librarian conducted searches in Ovid® MEDLINE®, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and ClinicalTrials.gov to capture both published and gray literature. Searches were conducted without publication date restrictions (see Appendix A for full search strategies). As there are multiple manufacturers/sources for many of the devices examined in this review, a Federal Register notice was posted in an effort to identify unpublished data. We also searched for unpublished studies in ClinicalTrials.gov. Reference lists of included articles and the bibliographies of systematic reviews published since 2010 were reviewed for includable literature. Literature searches will be updated while the draft report is posted for public comment and peer review to capture any new publications. Literature identified during the updated search will be assessed by following the same process of dual review as all other studies considered for inclusion in the report. If any pertinent new literature is eligible for inclusion, it will be incorporated in the final report.

Inclusion and Exclusion Criteria and Study Selection

Inclusion and exclusion criteria were developed *a priori* based on the Key Questions and PICOTS, in accordance with the AHRQ *Methods Guide for Effectiveness and Comparative*

Effectiveness Reviews.[12] Criteria are detailed below in Table 1. Abstracts were reviewed by at least two investigators, and full-text articles were retrieved for all citations deemed potentially appropriate for inclusion by at least one of the reviewers. Two investigators then independently reviewed all full-text articles for final inclusion. Discrepancies were resolved by discussion and consensus. A list of the included studies appears in Appendix B; excluded studies and primary reason for exclusion are listed in Appendix C.

The focus of this review is on randomized controlled trials (RCTs) reporting on longer-term outcomes (at least 4 weeks post intervention) that otherwise meet our PICOTS criteria.

Table 1. Inclusion and exclusion criteria

PICOTS	Inclusion	Exclusion

Patients	**General Inclusion Criteria**
	• Adults with the following chronic pain (defined as pain lasting 12 weeks or longer or pain persisting past the time for normal tissue healing) conditions: low back pain, neck pain, osteoarthritis pain, fibromyalgia, tension or mixed headache.
	
	KQ1: Low back pain
	Adults with chronic, nonradicular low back pain
	
	KQ2: Neck pain
	• Adults with chronic neck pain
	
	KQ3: Osteoarthritis
	• Adults with osteoarthritis-related pain (primary or secondary osteoarthritis) of the hip, knee or hand
	
	KQ4: Fibromyalgia
	• Adults with fibromyalgia
	
	KQ5: Headache
	• Adults with primary chronic tension headache.
	  ○ Primary headaches are attributed to the headache condition itself, not headache caused by another disease or medical condition. Tension headaches are the most common.
	  ○ Chronic headache is defined as 15 or more days each month for at least 12 weeks or history of headache more than 180 days a year.
		General Exclusion Criteria
		• Acute pain
		• Children (<18 years), pregnant or breastfeeding women
		• Patients with chronic pain related to "active" cancer, infection, inflammatory arthropathy,
		• <90% of study sample has the defined condition of interest or <90% received the treatment(s) of interest
		• Treatment for addiction
		• Pain at the end of life
		• Neuropathic pain
		
		KQ1: Low back pain
		• Patients with radiculopathy
		• Low back pain associated with severe or progressive neurological deficits
		• Failed back surgery syndrome
		
		KQ2: Neck pain
		• Patients with radiculopathy or myelopathy
		• Traumatic spinal cord injury
		• Neck pain associated with progressive neurological deficit, loss of strength
		
		KQ3: Osteoarthritis
		• Other types of arthritis (e.g., rheumatoid)
		• Patients with joint replacement
		
		KQ4: Fibromyalgia
		• Conditions with generalized pain not consistent with fibromyalgia
		• Systemic exertion intolerance disease, (myalgic encephalomyelitis/chronic fatigue syndrome)
		• Somatization disorder (Briquet's syndrome)
		
		KQ5: Headache
		• Migraine headache
		• Mixed headache (also known as co-existent tension and migraine headache, chronic daily headache, transformed migraine)
		• Trigeminal neuralgia
		• Cluster headache
		• Secondary headache types as defined in *The International Classification of Headache Disorders*, 3rd edition[13] (i.e., headaches due to an underlying pathology such as cancer, prior medical procedures,

PICOTS	Inclusion	Exclusion
		temporomandibular joint disorders, neck pathology, cervicogenic headache, and medication over-use headache) • Traumatic brain injury
Interventions	**All KQs:** • Exercise (exercise as part of physical therapy, supervised exercise, home exercise, group exercise, formal exercise program) • Psychological therapies (cognitive and/or behavioral therapy, biofeedback, relaxation training) • Physical modalities (traction, ultrasound, transcutaneous electrical nerve stimulation [TENS], low level laser therapy, interferential therapy, electro-muscular stimulation [EMS] diathermy, superficial heat or cold, bracing for knee, back, neck, hand and magnets) • Manual therapies (manipulation, massage) • Mindfulness practices (meditation, mindfulness-based stress reduction practices) • Mind-body practices (yoga, tai chi, qigong) • Acupuncture • Functional restoration training • Multidisciplinary/interdisciplinary rehabilitation[a]	**All KQs:** • Invasive nonsurgical treatments (e.g., injections, nerve block, spinal cord stimulators, parenterally-administered medications) • Surgical interventions (including minimally invasive surgical interventions) • Diet interventions or dietary supplementation • Studies evaluating incremental value of adding a noninvasive, nonpharmacological intervention to another noninvasive, nonpharmacological intervention • Self-management interventions or programs, self-management education programs • Others not listed for inclusion
Comparators	**<u>All KQs, subquestion a</u>** • Sham treatment • Waitlist • Usual care • No treatment • Attention control intended to control for nonspecific effects (e.g., time, attention, expectations); **<u>All KQs subquestion b</u>** • Non-opioid pharmacological therapy (NSAIDS, acetaminophen, anti-seizure medications, antidepressants) • Opioid analgesics **<u>KQs 1-4, 6 subquestion c</u>** • Exercise[b] **<u>KQ 5, 6 subquestion c</u>** • Biofeedback[c]	**All KQs:** • Supplements (e.g., glucosamine, chondroitin, d-ribose, herbal or homeopathic treatments) • Over-the-counter topical agents (e.g., aloe, capsaicin) • Invasive nonsurgical treatments (e.g., injections, nerve block, spinal cord stimulators, parenterally-administered medications) • Surgical interventions (including minimally invasive surgical interventions) • Studies evaluating incremental value of adding a noninvasive, nonpharmacological intervention to another noninvasive, nonpharmacological intervention • Comparisons within nonpharmacological intervention types (e.g., comparisons of different types of exercise with each other, different types of massage with each other) • Others not listed for inclusion
Outcomes	**All KQs:** Primary efficacy outcomes; we will focus on outcomes from validated measures for • Function/disability/pain interference[d] • Pain[d] Harms and Adverse effects	**All KQs:** • Intermediate outcomes (e.g., biomarkers for inflammation) • Other nonclinical outcomes

PICOTS	Inclusion	Exclusion
	Secondary outcomes • Psychological distress (including measures of depression and anxiety) • Quality of life • Opioid use • Sleep quality, sleep disturbance • Health care utilization	
Studies	Randomized controlled trials or high quality systematic reviews of randomized controlled trials published in English; cross-over trials with random assignment of initial treatment will be considered.	**All KQs:** • Studies reporting on intermediate outcomes only • Nonrandomized studies • Abstracts, editorials, letters, conference proceedings • Duplicate publications of the same study that do not report on different outcomes • Single site reports from multicenter trials • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Indirect comparisons • Studies with fewer than 15 patients per treatment arm • Systematic reviews on treatment of chronic neck pain, fibromyalgia, chronic headache, or osteoarthritis that are of low methodological quality. Those that do not report outcomes or time frames of interest may be excluded. Systematic reviews may be excluded based on currency or relevance (e.g., if there is a substantial new body of evidence reflected in a later review).
Setting(s)	Any nonhospital setting or in self-directed care	• Hospital care, hospice care, emergency department care
Timing	Duration of followup: short term (up to 6 months), intermediate term (6-12 months) and long term (at least 1 year); focus on longer term (> 1 year) effects. Trials lasting ≥ 6 months that include a supervised intervention followed by continued home treatment as part of the intervention will be included even though the only followup occurs directly after the intervention.	• Studies with <1 month followup after treatment

NSAID: nonsteroidal anti-inflammatory drug

[a] Multidisciplinary rehabilitation (also known as interdisciplinary rehabilitation, is defined as a coordinated program with both physical and biopsychosocial treatment components (e.g., exercise therapy and cognitive behavioral therapy) provided by professionals from at least two different specialties.

[b] Different forms of exercise will not be compared to each other. Exercise will be compared with nonexercise interventions for low back pain, neck pain, fibromyalgia and osteoarthritis.

[c] Different forms of biofeedback will not be compared to each other. Biofeedback will be compared with the noninvasive interventions for chronic headache.

[d] The magnitude of effects for pain and function will be classified using the same system as in the AHRQ-funded noninvasive treatment for low back pain review recognizing that small effects using this system may not meet standard thresholds for clinically meaningful effects. A small/slight effect was defined for pain as a mean between-group difference following treatment of 5 to 10 points on a 0- to 100-point visual analogue scale (VAS), 0.5 to 1.0 points on a 0- to 10-point numerical rating scale, or equivalent; for function as a mean difference of 5- to 10-point difference on the 0- to 100-point Oswestry Disability Index (ODI) or 1 to 2 points on the 0- to 24-point Roland-Morris Disability Questionnaire (RDQ), or equivalent; and for any outcome as a standardized mean difference (SMD) of 0.2 to 0.5. A moderate effect was defined for pain as a mean difference of 10 to 20 points on a 0- to 100-point VAS, for function as a mean difference of 10 to 20 points on the ODI or 2 to 5 points on the RDQ, and for any outcome as an SMD of 0.5 to 0.8. Large/substantial effects were defined as greater than moderate. We will apply similar

methodology to outcomes measures for the other condition. The clinical relevance of effects classified as small/slight might vary for individual patients depending on preferences, baseline symptom severity, harms, cost, and other factors.

Data Abstraction and Data Management

Using templates, data from included studies were abstracted into categories that included but were not limited to: study design, year, setting, country, sample size, eligibility criteria, attrition, population and clinical characteristics (including age, sex, comorbidities, diagnostic classifications/information), intervention characteristics (including the type, number, intensity, duration of, and adherence to treatments), comparator characteristics, and results (including harms). We also recorded the funding source and role of the sponsor. All abstracted study data were verified for accuracy and completeness by a second team member. Details are further outlined in the protocol. Detailed data abstraction tables appear in Appendix D.

Quality (Risk of Bias) Assessment of Individual Studies

Predefined criteria were used to assess the quality of included studies. We focused on studies with the least potential for bias and the fewest limitations. RCTs were assessed based on criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions* (Chapter 8.5 Risk of Bias Tool),[14] and precepts for appraisal developed by the Cochrane Back and Neck Group.[15] These criteria and methods were used in conjunction with the approach recommended in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Research*.[12] Two team members independently appraised each included study, with disagreements resolved by consensus. Studies were rated as "good," "fair," or "poor" as described in Table 2. Assessments of included studies are in Appendix E.

Table 2. Criteria for grading the quality of individual studies

Rating	Description and criteria
Good	• Least risk of bias, results generally considered valid • Employ valid methods for selection, inclusion, and allocation of patients to treatment; report similar baseline characteristics in different treatment groups; clearly describe attrition and have low attrition; use appropriate means for preventing bias (e.g., blinding of patients, care providers, and outcomes assessors); and use appropriate analytic methods (e.g., intention-to-treat analysis)
Fair	• Susceptible to some bias but not enough to necessarily invalidate results • May not meet all criteria for good quality, but no flaw is likely to cause major bias; the study may be missing information making it difficult to assess limitations and potential problems • Category is broad; studies with this rating will vary in strengths and weaknesses; some fair-quality studies are likely to be valid, while others may be only possibly valid
Poor	• Significant flaws that imply biases of various kinds that may invalidate results; "fatal flaws" in design, analysis or reporting; large amounts of missing information; discrepancies in reporting; or serious problems with intervention delivery • Studies are at least as likely to reflect flaws in the study design or execution as the true difference between the compared interventions • Considered to be less reliable than higher quality studies when synthesizing the evidence, particularly if discrepancies between studies are present

Data Analysis and Synthesis

Data were synthesized qualitatively (based on ranges and descriptive analysis, with interpretation of results) and quantitatively using meta-analysis where appropriate. Results are organized by Key Question (i.e., by condition) and intervention and then organized by comparators for each subquestion (e.g., intervention vs. waitlist or sham for subquestion a). To

the extent that the interventions were distinct, we explored separating them out for analysis and reporting. For example, we categorized various forms of exercise based on their primary mechanisms of action (Appendix F). Interventions with similar characteristics were combined. For example, we combined cognitive behavioral therapy (CBT) and acceptance and commitment therapy (ACT), both of which fall in the broad category of cognitive and behavioral treatments and include many similar methods.[16] Duration of followup post intervention was reported and categorized as short-term (up to 6 months), intermediate-term (6 to 12 months) and long-term (at least 1 year). Prioritized outcomes of function and pain, based on validated measures, are presented first. A variety of outcomes measures were used across trials. We acknowledge that there is overlap between functional outcome measures and quality of life measures. SF-36 and EQ-5D are two such outcome measures that were categorized as quality of life measures for this report. For some conditions, such as osteoarthritis, results are organized by affected region (e.g., knee, hip, hand). Based on input from stakeholders, improvement in function was prioritized as the most important outcome.

Results for continuous outcomes as well as dichotomous outcomes were synthesized. Binary outcomes based on the proportion of patients achieving specific thresholds of success for improved function, or other overall measure of success as defined in the trials (e.g., ≥ 30% improvement in pain score) were reported, and a risk ratio and 95% confidence interval were calculated to evaluate the presence of an association and estimate relative effect size using the Rothman Episheet.[17] For continuous outcomes, mean differences between treatments and 95% confidence intervals were calculated using GraphPad or Stata®/IC 12.1 (StataCorp, College Station, TX) to provide effect sizes and determine presence of a statistical association.

We conducted meta-analysis to get more precise effect estimates for various interventions. To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity. Three continuous outcomes (pain, function, and quality of life) provided adequate data for meta-analysis. Mean difference (MD) was used as the effect measure if the studies reported outcomes using the same scale, or if the outcomes could be converted to the same scale (e.g., pain, converted to 0 to 10 scale, and SF-36); otherwise, standardized mean difference (SMD) was used when the reported outcomes used different scales but measured the same underlying construct (e.g., function). In the primary analysis, MD and SMD were calculated using the followup score, and sensitivity analyses were conducted using the change score from the baseline. When standard deviation (SD) was not reported, or could not be calculated from the reported data, it was imputed using the average SD from the studies of the same meta-analysis, or using the SD value from the baseline if the baseline SD was reported and the followup SD was not.

We assumed random effects across studies and used both the Dersimonian-Laird method[18] and the profile-likelihood model[19] to combine studies. Statistical heterogeneity among the studies was assessed using the standard Cochran's chi-square test and the I^2 statistic.[20] Primary analyses were stratified by disease type, intervention, control group (usual care, exercise or pharmacological treatment) and length of followup (short-, intermediate-, and long-term. Controls included usual care, waitlist, no treatment, placebo, sham treatment, attention control, or other groups that involved at most minimal active treatment. We performed additional sensitivity and subgroup analyses based on specific interventions (e.g., type of acupuncture, type of exercise, intervention intensity etc.) and control types (as described above) and by excluding outlying studies and studies rated poor.

To facilitate interpretation of results across trials and interventions, we categorized the magnitude of effects for function and pain outcomes using the system described in our previous review.[21,22] In general we classified effects for measures with a 0 to 100 scale for pain or function as small/slight (5 to 10 points), moderate (> 10–20 points), or large/substantial (> 20 points).

Grading the Strength of Evidence for Major Comparisons and Outcomes

The strength of evidence for each Key Question and primary outcome (function, pain, harms) was initially assessed by one researcher with experience in determining strength of evidence for each primary clinical outcome in accordance with AHRQ guidance[23,24] and as described in the protocol. The initial assessment was independently reviewed by at least one other experienced senior investigator. The overall strength of evidence was determined based on assessment of study limitations (graded low, moderate, high); consistency of results across trials (graded consistent, inconsistent or for single studies, unknown); the directness of the evidence linking the interventions with health outcomes (graded direct or indirect); effect estimate precision (graded precise or imprecise); and reporting bias (suspected or undetected). Bodies of evidence consisting of RCTs are initially considered high strength. All outcomes were considered direct.

The final strength of evidence grade was assigned by evaluating and weighing the combined results of the above domains and considering the highest quality evidence available. While studies rated as poor quality were not excluded, such studies were considered to be less reliable than higher quality studies when synthesizing the evidence, particularly when discrepancies across studies were noted. The strength of evidence was assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale (Table 3). When all of the studies for a primary outcome were rated poor quality, we rated the SOE as insufficient. Strength of evidence tables for primary outcomes are presented in Appendix G.

Table 3. Description of the strength of evidence grades

Strength of Evidence	Description
High	We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable, i.e., another study would not change the conclusions
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Assessing Applicability

Applicability was assessed using the PICOTS framework by examining the abstracted characteristics of the patient populations for each condition (e.g., demographic characteristics, condition-specific diagnostic criteria, symptoms, presence of medical and psychiatric co-morbidities, other psychosocial factors); the interventions (e.g., availability in the United States; dose, frequency, or intensity of treatment, and methods for administration); and clinical settings (e.g., primary care, specialty setting; developing country vs. developed country) in which the included studies are performed.

The magnitude of effects for pain and function were classified with the system used our previous AHRQ review on noninvasive treatment for low back pain,[21] recognizing that small effects using this system may not meet standard thresholds for clinically meaningful effects. We applied the following definitions:

- Small/slight effect
  - For pain: as a mean between-group difference following treatment of 5 to 10 points on a 0-to-100-point visual analogue scale (VAS), 0.5 to 1.0 point on a 0- to 10-point numerical rating scale, or equivalent
  - For function: as a mean difference of 5 to 10 points on the 0- to 100-point Oswestry Disability Index (ODI) or 1 to 2 points on the 0- to 24-point Roland-Morris Disability Questionnaire (RDQ), or equivalent
  - For any outcome: as a standardized mean difference (SMD) of 0.2 to 0.5

- Moderate effect
  - For pain: as a mean difference of 10 to 20 points on a 0- to 100-point VAS
  - For function: as a mean difference

Information regarding effect size definitions for other outcome measures is available in Appendix H. There is variability across individual patients regarding what may constitute a clinically importance effect, which is influenced by a number of factors such as preferences, duration and type of chronic pain, baseline symptom severity, harms, and costs.

Peer Review and Public Commentary

Peer reviewers with expertise in primary care and management of the included chronic pain conditions have been invited to provide written comments on the draft report. The AHRQ Task Order Officer and an Evidence-based Practice Center Associate Editor will also provide comments and editorial review. The draft report will be posted on the AHRQ Web site for 4 weeks for public comment. A disposition of comments report with authors' responses to the peer and public review comments will be posted after publication of the final CER on the AHRQ Web site.

Results

Introduction

Results are organized by Key Question (i.e., by condition) and intervention and then organized by comparators for each subquestion. We categorized post-intervention followup as short-term (up to 6 months), intermediate-term (6 to 12 months) and long-term (at least 1 year). We prioritized function and pain outcomes based on validated measures, and note there is some overlap between functional outcome measures and quality of life measures. For some conditions (e.g., osteoarthritis), results are organized by affected region.

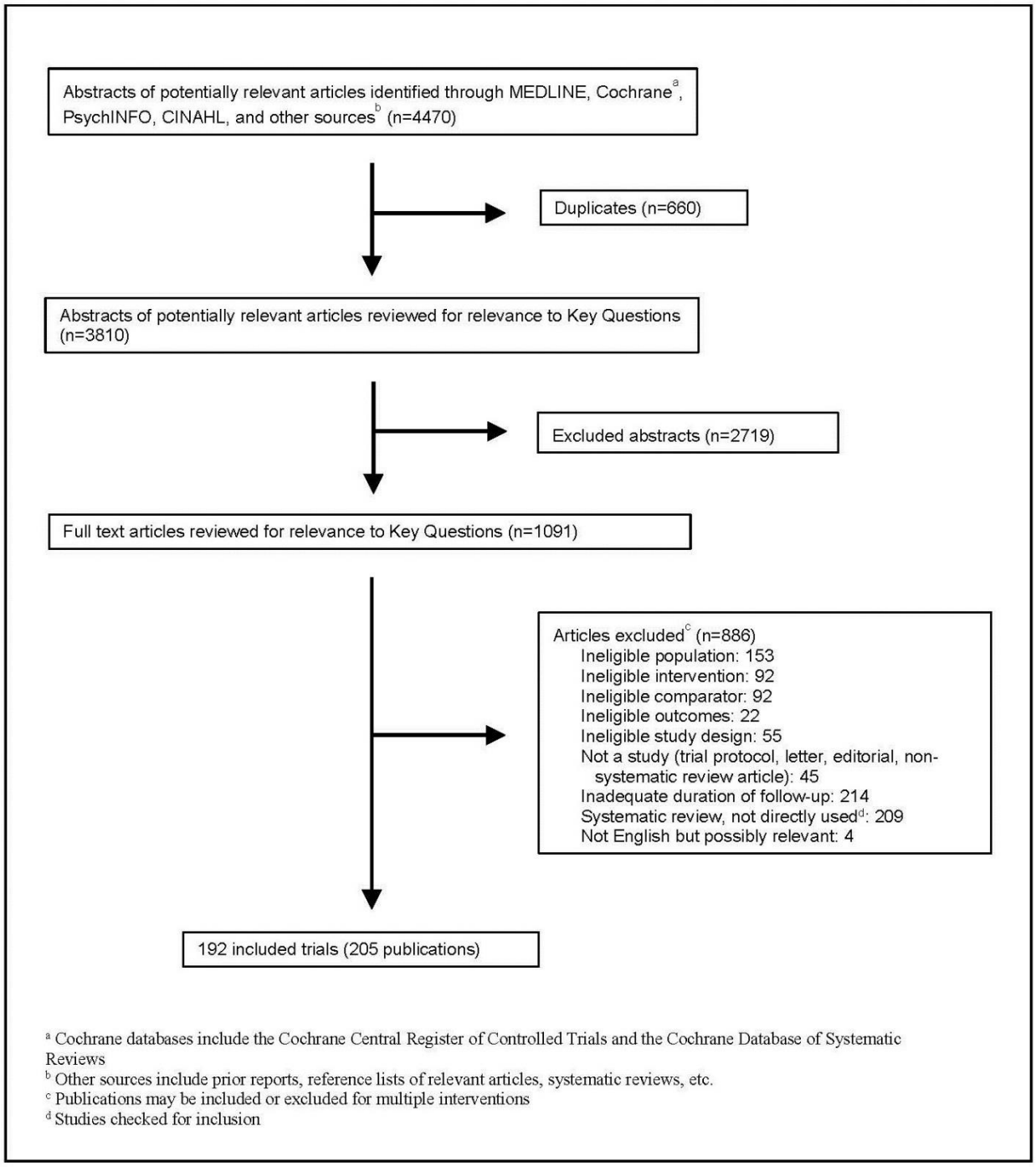
We synthesized data qualitatively and quantitatively, using meta-analysis where appropriate to get more precise effect estimates for various interventions. Three continuous outcomes (pain, function, and quality of life) provided adequate data for meta-analysis. For meta-analyses providing pooled estimates, we report results from heterogeneity testing. I-squared and corresponding p-values describe the degree and statistical significance of heterogeneity across studies; pooled (subtotal) estimates are statistically significant if the confidence interval does not include the value of 0. (See the Methods section of this report and the protocol for additional details on data analysis and synthesis.)

A list of acronyms and abbreviations appears at the end of the report.

Results of Literature Searches

The search and selection of articles are summarized in the literature flow diagram (Figure 2). Database searches resulted in 4,470 potentially relevant articles. After dual review of abstracts and titles, 1,091 articles were selected for full-text dual review, and 205 publications were determined to meet inclusion criteria and were included in this review. One-fourth of the trials excluded at full text did not meet our criteria for followup duration (i.e., a minimum of 1 month of followup after termination of the intervention). Other common reasons for exclusion of primary trials included ineligible population and ineligible intervention or comparator (i.e., combination of treatments or treatments were additive in nature). Data abstraction and quality assessment tables for all included studies are available in Appendixes D and E.

Figure 2. Literature flow diagram



Description of Included Studies

Overall, 192 trials (across 205 publications) were included. For each intervention category, the comparisons evaluated and their respective studies are listed in Table 4.

Table 4. Overview of included studies

Intervention	Comparator	Chronic Low Back Pain: n=65 (69 publications)	Chronic Neck Pain: n=23	Osteoarthritis: n=51 (54 publications)	Fibromyalgia: n=44 (50 publications)	Chronic Tension Headache: n=9
Exercise	Vs. sham, waitlist, no treatment, attention	**6**[25-30]	**5**[31-35]	Knee OA: **18 (21)**[36-56] Hip OA: **4**[36,57-59] Hand OA: **1**[60]	**20 (22)**[61-82]	0
	Vs. pharmaco-logical therapy	0	**1**[83]	0	0	0
Psycho-logical therapies	Vs. sham, waitlist, no treatment, attention	**5**[84-88]	**1**[34]	Knee OA: **2**[89,90] Hip, Hand OA: 0	**10 (11)**[63,81,82,91-98]	**2**[99,100]
	Vs. pharmaco-logical therapy	0	0	0	**3**[91,101,102]	**2**[100,103]
	Vs. exercise (or biofeedback for CTTH)	**1**[104]	**1**[34]	0	**5**[63,81,82,105,106]	0
Physical Modalities	Vs. sham, waitlist, no treatment, attention	**7**[107-113]	**5**[114-118]	Knee OA: **13**[119-131] Hip OA: 0 Hand OA: **2**[132,133]	**1**[134]	**1**[135]
	Vs. pharmaco-logical therapy	0	0	0	0	0
	Vs. exercise (or biofeedback for CTTH)	**1**[136]	0	0	0	0
Manual Therapies	Vs. sham, waitlist, no treatment, attention	**9**[88,113,137-143]	**2**[144,145]	Knee OA: **2**[36,146] Hip OA: **1**[36] Hand OA: 0	**2**[147,148]	**1**[149]
	Vs. pharmaco-logical therapy	0	0	0	0	**1**[150]
	Vs. exercise (or biofeedback for CTTH)	**5**[140,151-154]	**1**[144]	Knee OA: **1**[36] Hip OA: **2**[36,155] Hand OA: 0	0	0
Mindfulness Practices	Vs. sham, waitlist, no treatment, attention	**4**[84,156-158]	0	0	**2 (3)**[159-161]	0
	Vs. pharmaco-logical therapy	0	0	0	0	0
	Vs. exercise (or biofeedback for CTTH)	0	0	0	0	0
Mind-body Practices	Vs. sham, waitlist, no treatment, attention	**6**[162-167]	**1**[168]	Knee OA: **2**[169,170] Hip, Hand OA: 0	**2**[171,172]	0
	Vs. pharmaco-logical therapy	0	0	0	0	0
	Vs. exercise (or biofeedback for CTTH)	**5**[162-164,173,174]	**2**[175,176]	0	0	0

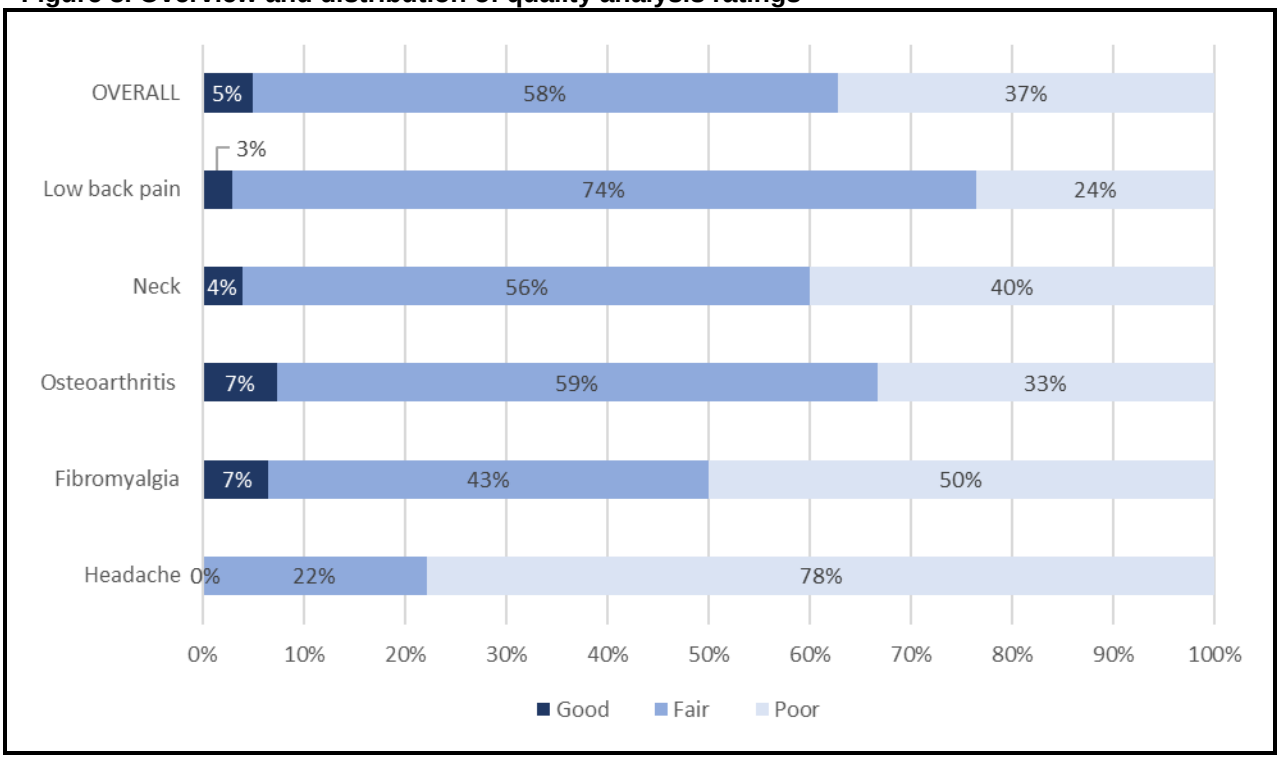
Intervention	Comparator	Chronic Low Back Pain: n=65 (69 publications)	Chronic Neck Pain: n=23	Osteoarthritis: n=51 (54 publications)	Fibromyalgia: n=44 (50 publications)	Chronic Tension Headache: n=9
Acupuncture	Vs. sham, waitlist, no treatment, attention	**8**[142,177-183]	**7**[168,184-189]	Knee OA: **8**[56,190-196] Hip, Hand OA: 0	**3**[197-199]	**3**[200-202]
	Vs. pharmaco-logical therapy	0	**2**[184,203]	0	0	0
	Vs. exercise (or biofeedback for CTTH)	0	0	0	0	0
Function Restoration Training	Vs. sham, waitlist, no treatment, attention	0	0	0	0	0
	Vs. pharmaco-logical therapy	0	0	0	0	0
	Vs. exercise (or biofeedback for CTTH)	0	0	0	0	0
Multi-disciplinary rehab-ilitation	Vs. sham, waitlist, no treatment, attention	**7**[204-209]	0	Knee, Hip OA: 0 Hand OA: **1**[210]	**5 (6)**[80,211-215]	0
	Vs. pharmaco-logical therapy	**1**[216]	0	0	0	0
	Vs. exercise (or biofeedback for CTTH)	**9 (13)**[104,217-228]	0	0	**1**[80]	0

CTTH = chronic tension-type headache; OA = osteoarthritis.

Thirty-five percent of the included trials were small (fewer than 70 participants). Across studies, the majority of patients were female, with a mean ages ranging from 31 to 76 years; patients with osteoarthritis tended to be older in general than those in the other conditions (range, 52 to 76 years). The means of pain duration pain for patients with low back pain, neck pain, and osteoarthritis were similar and varied widely from 6 months to 15 years, while those with fibromyalgia and tension headache had suffered with pain for no fewer than 4 years (up to 22 years). Exercise was the most common intervention for osteoarthritis and fibromyalgia. Psychological therapies were most common for fibromyalgia, and manual therapies were most common for chronic low back pain. Acupuncture was used in all included conditions. Multidisciplinary rehabilitation was reported primarily for low back pain and fibromyalgia. There were no trials of functional restoration training for any condition. Limited evidence was available for hip or hand osteoarthritis or chronic tension headache. The majority of trials compared nonpharmacological interventions with usual care, waitlist, no treatment, attention control, or placebo/sham, with very few trials employing pharmacological treatments or exercise as comparators. In general, little long-term evidence was available across conditions and interventions.

The majority of trials were rated fair quality with only 5 per cent considered good quality (Figure 3). For chronic tension headache, no study was considered good quality. A primary methodological limitation in many of the trials was the inability to effectively blind participants and in many cases providers. Additionally, unacceptable rates of attrition (both overall and differential) and poor reporting of allocation concealment methods were common shortcomings in the included studies.

Figure 3. Overview and distribution of quality analysis ratings



Key Question 1: Chronic Low Back Pain

Exercise for Low Back Pain

Key Points

- Exercise was associated with slighter greater effects on short-term function than controls (6 trials, pooled SMD -0.31, 95% CI -0.58 to -0.04, I2=57%); there were no effects on intermediate-term function (3 trials, pooled SMD -0.15, 95% CI -0.48 to 0.18, I2=51%) or long-term function (1 trial, difference 0.00 on the 0 to 100 ODI, 95% CI -11.4 to 11.4) (SOE: Low).
- Exercise was associated with slightly to moderately greater effects on pain than usual care, an attention control, or a placebo intervention at short-term (6 trials, pooled difference -0.81 on a 0 to 10 scale, 95% CI -1.26 to -0.,36, I2=0%), intermediate-term (3 trials, pooled difference -1.37, 95% CI -2.10 to -0.65, I2=34%), and long-term (1 trial, difference -1.55, 95% CI -2.378 to -0.32) followup (SOE: Moderate for short-term, Low for intermediate-term and long-term).
- No trial evaluated exercise versus pharmacological therapy.
- Comparisons involving exercise versus other nonpharmacological therapies are addressed in the sections for the other therapies.
- Harms were not reported in most trials; one trial did not find an association between exercise and increased pain versus placebo and one trial reported no adverse events (SOE: Low).

Detailed Synthesis

Six trials of exercise therapy for low back pain met inclusion criteria (Table 5 and Appendix D).[25-30] Two trials evaluated neuromuscular re-education exercise (motor control exercises),[25,26] two trials muscle performance exercises (Pilates),[29,30] and two trials combined exercise techniques.[27,28] Sample sizes ranged from 60 to 154 (total sample=553). Three trials compared exercise versus an attention control;[26,27,29] two trials compared exercise versus usual care;[28,30] and one trial compared exercise versus a placebo intervention (detuned diathermy and ultrasound).[25] Four trials were conducted in the United States, Europe, or Australia, and two trials[29,30] were conducted in Brazil. The duration of exercise therapy ranged from 6 to 12 weeks and the number of exercise sessions ranged from 10 to 24. One trial reported outcomes through long-term followup,[26] three trials through intermediate-term followup,[25,27] and the remainder only evaluated short-term outcomes.

Five trials were rated fair-quality and one trial[28] poor-quality (Appendix E). In two fair-quality trials,[25,30] the main methodological limitation was the inability to blind interventions. Limitations in the other trials included unclear randomization and allocation concealment methods, high loss to followup, and baseline differences between intervention groups.

Table 5. Summary of results for low back pain: exercise

Author, Year, Followup,[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Costa, 2009[25] 4 and 10 months Duration of pain: Mean 328 to 335 weeks *Fair*	A: Neuromuscular re-education (motor control exercise) (n=77), 12 sessions over 8 weeks B: Placebo (detuned shortwave diathermy and detuned ultrasound) (n=77) 12 sessions, two sessions/week for 4 weeks, then 1 session/week for 4 weeks	A vs. B Age: 55 vs. 53 years Female: 58% vs. 62% Baseline pain (0-10 VAS): 6.8 vs. 6.6 Baseline RDQ (0-24): 13.1 vs. 13.4	<u>4 months</u> RDQ: 5.3 vs. 4.3, adjusted difference 1.0 (95% CI 0.3 to 1.8) Pain (0-10 VAS): 5.0 vs. 5.6, adjusted difference 1.4 (95% CI 0.3 to 2.4) <u>10 months</u> RDQ: 11.4 vs. 12.3, adjusted difference -1.0 (95% CI -2.8 to 0.8) Pain: 5.0 vs. 6.3, adjusted difference -1.0 (95% CI -1.9 to -0.1)	<u>4 months</u> Global impression of recovery (-5 to +5): 1.5 vs. 0.3, adjusted difference 1.4 (95% CI 0.3 to 1.8) <u>10 months</u> Global impression of recovery: 1.2 vs. -0.3, adjusted difference 1.6 (95% CI 0.6 to 2.6)
Goldby, 2006[26] 3, 6, 12 and 24 months Duration of pain: Mean 11 to 12 years *Fair*	A: Neuromuscular re-education (motor control exercise) (n=84), 10 sessions over 10 weeks B: Attention control (education) (n=40)	A vs. B Age: 43 vs. 41 years Female: 68% vs. 68% Race: 80% vs. 62% Baseline back pain (0-100 NRS): 45.8 vs. 37.6 Baseline ODI (0-100): 40.5 vs. 33.5	A vs. C <u>3 months</u> ODI (0-100): 31.00 vs. 28.1, difference 2.9 (95% CI -3.89 to 9.69) LBO (0-75): 50.92 vs. 54.4, difference -3.48 (95% CI -9.67 to 2.71) Back pain (0-100 NRS): 28.81 vs. 34.4, difference -5.59 (95% CI -17.86 to 6.68) <u>6 months</u> ODI: 25.81 vs. 23.9, difference	A vs. C <u>3 months</u> Nottingham Health Profile: 94.97 vs. 94.32, difference 0.65 (95% CI -36.97 to 38.27) <u>6 months</u> Nottingham Health Profile: 76.3 vs. 77.50, difference -1.20 (95% CI -37.76 to 35.36) <u>12 months</u>

Author, Year, Followup,[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			1.91 (95% CI -6.28 to 10.10) LBO: 55.42 vs. 57.85, difference -2.43 (95% CI -9.14 to 4.28) Back pain: 23.16 vs. 30.25, difference -7.09 (95% CI -20.22 to 6.04) <u>12 months</u> ODI: 24.76 vs. 26.9 difference -2.14 (95% CI -10.14 to 5.86) LBO: 53.86 vs. 50.95, difference 2.91 (95% CI -4.29 to 10.11) Back pain: 29.23 vs. 30, difference -0.77 (95% CI -14.13 to 12.59) <u>24 months</u> ODI: 27 vs. 27; difference 0.00 (95% CI -11.44 to 11.44) LBO: 54.7 vs. 55.2, difference -0.5 (95% CI -9.20 to 8.20) Back pain: 35.4 vs. 50.9, difference -15.50 (95% CI -33.06 to 2.06)	Nottingham Health Profile: 70.06 vs. 87.47 difference -17.41 (95% CI -56.12 to 21.30) <u>24 months</u> Nottingham Health Profile: 82 vs. 83, difference -1.00 (95% CI -60.85 to 58.85)
Kankaaanpaa, 1999[27] 3 and 9 months Duration of pain: Mean 7 to 9 years *Fair*	A. Combined exercise (exercises, stretching, relaxation, muscle function and ergonomic advice) (n=30), 24 sessions over 12 weeks B. Attention Control (n=24) (thermal therapy and minimal massage	A vs. B Age: 40 vs. 39 years Female: 36.6% vs. 33.3% Baseline back pain (0-100 mm VAS): 55.2 vs. 47.0 Baseline Pain and Disability Index (0-70 PDI): 13.2 vs. 9.5	<u>3 months</u> Pain and Disability Index (0-70): 5.7 vs. 12.6, difference -6.9 (95% CI -11.69 to - 2.11) Back pain (0-100 VAS): 26.6 vs. 43.4; difference -16.80 (95% CI -31.12 to -2.47) <u>9 months</u> Pain and Disability Index: 5.7 vs. 11.4, difference -5.7 (95% CI -11.31 to -0.09) Back pain intensity: 23.9 vs. 45.1, difference -21.20 (95% CI -32.69 to -9.71)	
Miyamoto, 2013[29] 4.5 months Duration of pain: Mean 5 to 6 years *Fair*	A. Muscle performance (Pilates) (n=43),12 sessions over 6 weeks B. Attention control (n=43) (education)	A vs. B Age: 41 vs. 38 years Female: 84% vs. 79% Baseline pain (0-10 VAS): 6.6 vs. 6.5 Baseline RDQ: 9.7 vs. 10.5	<u>4.5 months</u> RDQ (0-24): 4.5 vs. 6.7, adjusted difference -1.4 (95% CI -3.1 to 0.03) Patient-Specific Functional Scale (0-10): 6.9 vs. 6.1, adjusted difference 0.2 (95% CI -0.6 to 1.1) Pain (0-10 VAS): 4.5 vs. 5.3, adjusted difference -0.9 (95% CI -1.9 to 0.1)	<u>4.5 months</u> Global impression of recovery (-5 to +5): 2.4 vs. 1.7, adjusted difference 0.7 (95% CI -0.4 to 1.8)
Nassif, 2011[28] 4 months	A. Combined exercise (n=37) (stretching,	A vs. B Age: 45 vs. 45 Female: 11%	<u>4 months</u> RDQ (0-24): 10.0 vs. 10.6, difference -0.6 (95% CI -3.5 to	<u>4 months</u> Dallas Pain Questionnaire anxiety

Author, Year, Followup,[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Duration of pain: NR *Poor*	stability, coordination, and muscle strengthening exercises), 24 sessions over 8 weeks B. Usual care (n=38)	vs. 21% Baseline pain (0-10 VAS): 4.5 vs. 4.9 Baseline RDQ: 13.9 vs. 12.3	2.3) Quebec Back Pain Disability Questionnaire: 27.2 vs. 30.2, difference -3.0 (95% CI -11.7 to 5.7) Pain (0-10 NRS): 3.2 (2.3) vs. 3.5 (2.5), difference -0.3 (95% CI -1.6 to 1.0)	and depression: 31.2 vs. 28.9, difference 2.3 (95% CI -8.2 to 12.8)
Natour, 2014[30] 3 months Duration of pain: >1 year *Fair*	A. Exercise (Pilates) (n=30), 24 sessions over 12 weeks B. Usual care (n=30) (no treatment)	A vs. B Age: 48 vs. 48 Female: 80% vs. 77% Baseline pain (0-10 VAS): 5.5 vs. 5.8 Baseline RDQ: 1.1 vs. 10.6	<u>3 months</u> RDQ (0-24): 7.0 vs. 10.7, difference -3.6, p<0.001 Pain (0-10 VAS): 4.2 vs. 5.8, difference -1.6, p<0.001 SF-36 physical function (0-100): 65.4 vs. 59.6, difference 5.8, p=0.026 SF-36 role physical: 56.4 vs. 40.0, difference 16.4, p=0.086 SF-36 bodily pain: 52.2 vs. 43.9, difference 8.3, p=0.030	<u>3 months</u> SF-36 general health: 65.2 vs. 62.1, difference 3.1, p=0.772 SF-36 mental health: 67.9 vs. 65.3, difference 2.6, p=0.243 SF-36 social functioning: 86.0 vs. 80.4, difference 5.6, p=0.09 No differences on other SF-36 subscales

CI = confidence interval; LBO = Low Back Outcome Score; NHP = Nottingham Health Profile; NR = not reported; ODI = Oswestry Disability Index; RDQ= Roland Morris disability questionnaire; SF-36 = Short-Form 36 questionnaire; VAS = visual analog scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Exercise Compared With Usual Care, an Attention Control, or a Placebo Intervention

Exercise was associated with slightly greater effects on short-term function than controls (6 trials, pooled SMD -0.31, 95% CI -0.58 to -0.04, I2=57%) (Figure 4).[25-30] Four trials that evaluated function using the RDQ (0 to 24 scale) reported a pooled mean difference of -1.96 points (95% CI -3.14 to -0.78).[25,28-30] and one trial that used the Oswestry Disability Index (0 to 100 scale) reported a difference of 2.9 points (95% CI -3.89 to 9.69).[26] There were no clear differences in estimates when analyses were stratified according to the type of exercise (estimates ranged from -0.08 to -0.51 points) or the type of control and when the poor-quality trial was excluded. There were no differences between exercise versus controls in intermediate-term function (3 trials, pooled SMD -0.15, 95% CI -0.48 to 0.18, I2=51%)[25-27] or long-term function (1 trial, difference 0.00, 95% CI -11.4 to 11.4 on the ODI).[26]

Exercise was associated with greater effects on short-term pain than usual care, an attention control, or a placebo intervention (6 trials, pooled difference -0.81 on a 0 to 10 scale, 95% CI -1.26 to -0.36, I^2=0%) (Figure 5).[25-30] There were no clear differences in estimates when analyses were stratified according to the type of exercise (difference -0.52, 95% CI -1.41 to 0.36 in 2 trials of neuromuscular re-education exercises, -1.12, 95% CI -2.28 to -0.14 in 2 trials of muscle performance exercises, and -0.90, 95% I -2.63 to 0.68 in 2 trials of combined exercises) the type of control (usual care, attention control, or placebo intervention) and when the poor-quality trial was excluded. For intermediate-term pain (3 trials, pooled difference -1.37, 95% CI -2.10 to -0.65, I^2=34%)).[25-27] and long-term pain (1 trial, difference -1.55, 95% CI -2.78 to -0.32),[26] effects of exercise on pain were moderate, but findings were based on small numbers of trials.

Data on effects of exercise on quality of life were limited. One trial[26] found no differences between exercise versus an attention control on the Nottingham Health Profile at short-term, intermediate-term, or long-term followup, and one trial[30] found exercise associated with higher scores on the SF-36 physical functioning (difference 5.8 points on to 100 scale, p=0.026), bodily pain (difference 8.3 points, p=0.03), and vitality subscales (difference 5.3 points, p=0.029) at short-term followup; there were no differences on other SF-36 subscales (Table 5).

No trial evaluated effects of exercise on use of opioid therapies or health care utilization. There was insufficient evidence to determine effects of duration of exercise therapy or number of sessions on outcomes.

Exercise Compared With Pharmacological Therapy

No trial of exercise versus pharmacological therapy met inclusion criteria.

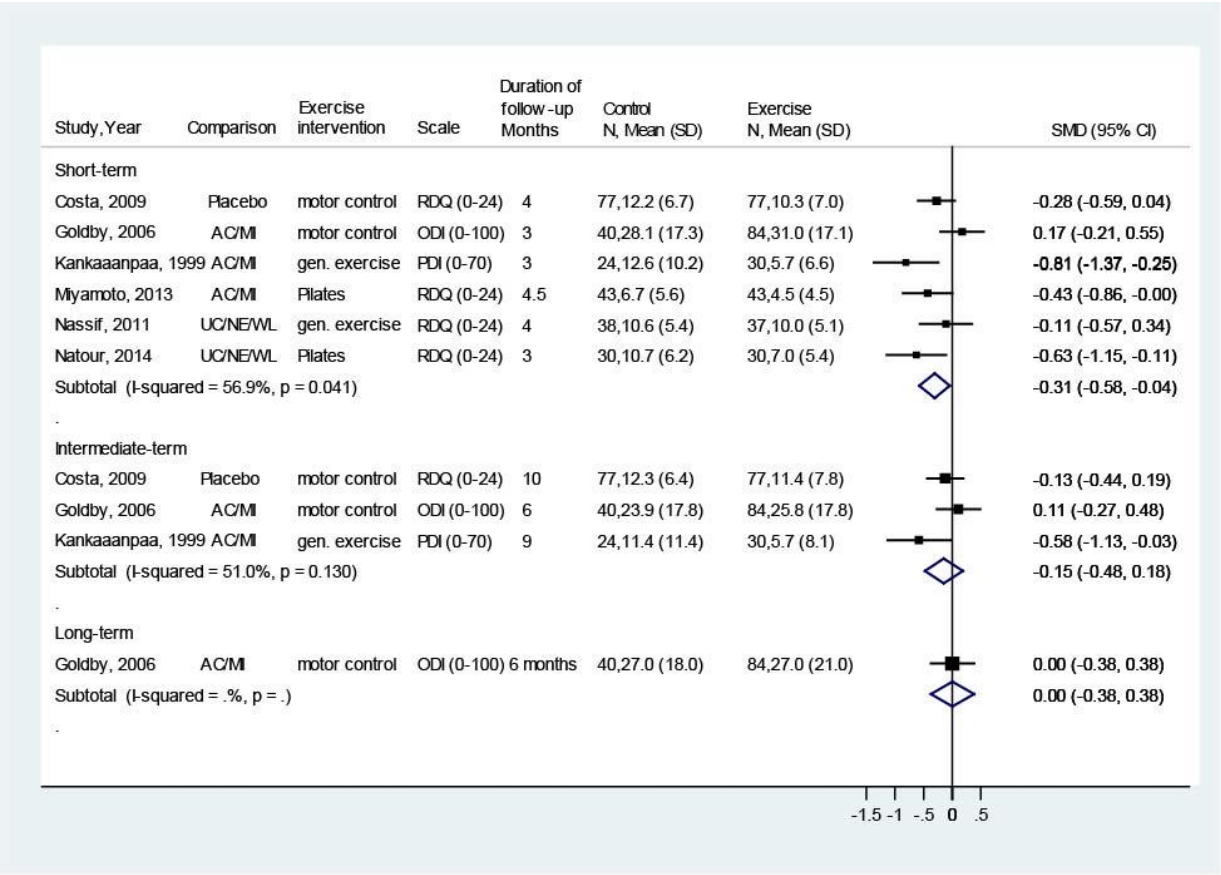
Exercise Compared With Other Non-pharmacological Therapies

Findings for exercise versus other nonpharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

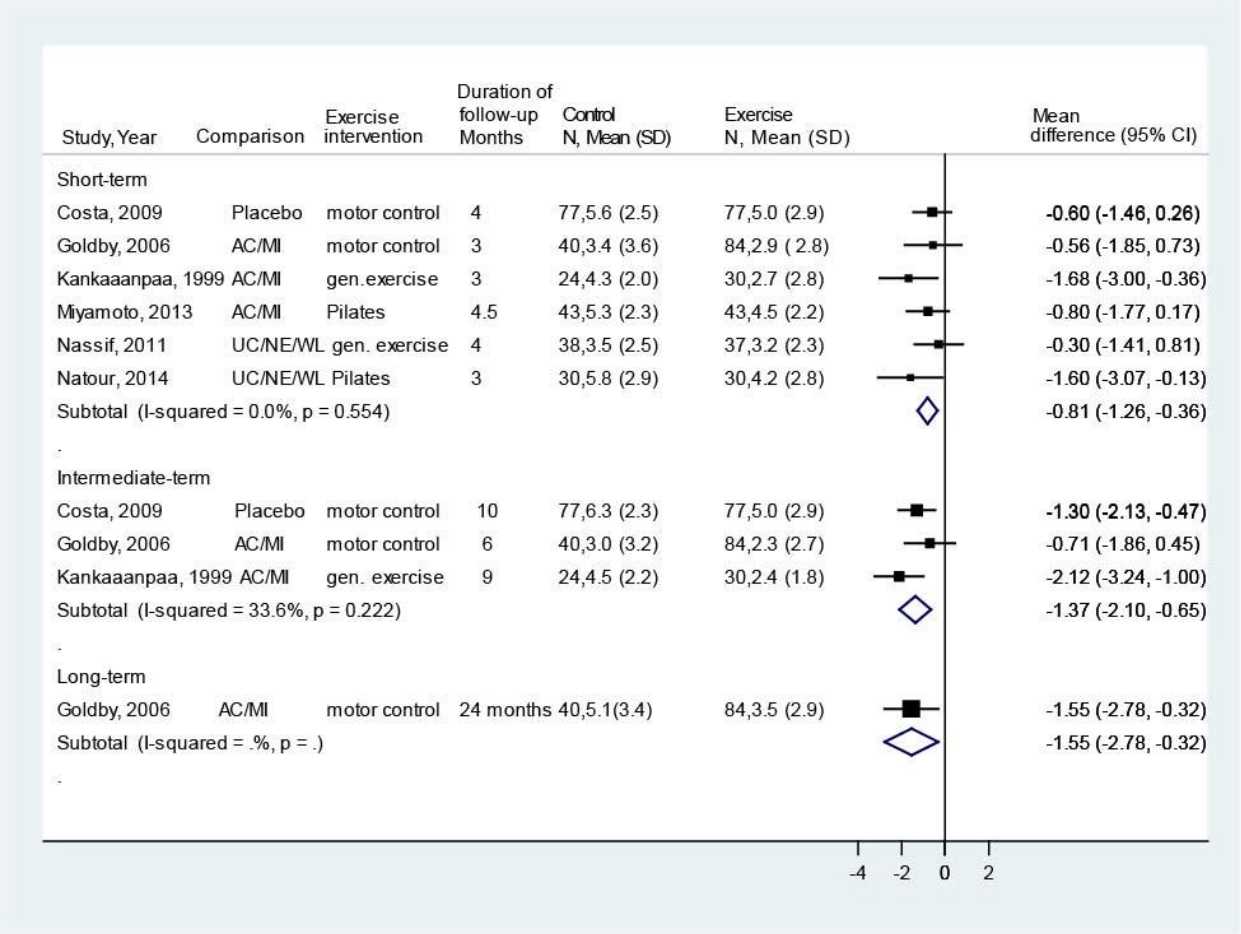
Harms were not reported in most trials. One trial[25] found no difference between exercise and a placebo intervention (detuned diathermy) in likelihood of increased pain, and another trial[29] reported no adverse events (Appendix D).

Figure 4. Exercise versus usual care, an attention control, or a placebo intervention: effects on function



AC/MI=attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; UC/NE/WL=usual care/no exercise/waitlist

Figure 5. Exercise versus usual care, attention control, or a placebo intervention: effects on pain



AC/MI=attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; UC/NE/WL=usual care/no exercise/waitlist

Acupuncture for Low Back Pain

Key Points

- Acupuncture was associated with slightly greater effects on short-term function than sham acupuncture or usual care (4 trials, pooled SMD -0.22, 95% CI -0.35 to -0.08, I2=44%). There were no differences between acupuncture versus controls in intermediate-term function (3 trials, pooled SMD -0.08, 95% CI -0.36 to 0.20, I2=75%) or long-term function (1 trial, adjusted difference -3.4 on the 0 to 100 ODI, 95% CI -7.8 to 1.0) (SOE: Low).
- Acupuncture was associated with slighter greater effects on short-term pain than sham acupuncture, usual care, an attention control, or a placebo intervention (5 trials, pooled difference -0.55 on a 0 to 10 scale, 95% CI -0.86 to -0.24, I^2=30%). There was no difference in intermediate-term pain (5 trials, pooled mean difference -0.25, 95% CI -0.67 to 0.16, I2=33%); one trial found acupuncture associated with greater effects on long-

term pain (mean difference -0.83, 95% CI -1.51 to -0.15) (SOE: Moderate for short-term, Low for intermediate-term and long-term).

- There was no clear difference between acupuncture versus control interventions in risk of withdrawal due to adverse events. Serious adverse events were rare with acupuncture and control interventions (SOE: Low).

Detailed Synthesis

Eight trials of acupuncture for low back pain met inclusion criteria (Table 6 and Appendix D).[142,177-183] All trials evaluated needle acupuncture to body acupoints; one trial also evaluated electroacupuncture.[178] Sample sizes ranged from 40 to 1162 (total sample=2,621). Four trials compared acupuncture versus sham acupuncture,[177,179-181] three trials acupuncture versus usual care,[179,181,183] two trials acupuncture versus a placebo intervention (sham transcutaneous electrical nerve stimulation),[178,182] and one trial acupuncture versus an attention control (self-care education).[142] One trial was conducted in Asia[180] and the rest in the U.S. or Europe. The duration of acupuncture therapy ranged from 6 to 12 weeks and the number of acupuncture sessions ranged from 6 to 15. One trial reported outcomes through long-term followup,[183] four trials through intermediate-term followup.[142,177-179] and the remainder only evaluated short-term outcomes.

One trial was rated good-quality,[177] five trials fair-quality,[142,179-181,183] and two trials[178,182] poor-quality (Appendix E). Limitations in the fair-quality and poor-quality trials included unblinded design, unclear randomization or allocation concealment methods, and high attrition.

Table 6. Summary of results for low back pain: acupuncture

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Brinkhaus, 2006a[177] 4 and 10 months Duration of pain: 14.7 vs. 13.6 years *Good*	A: Needle acupuncture to body acupoints (n=140), 12 sessions over 8 weeks B: Sham acupuncture (n=70)	A vs. B Age: 59 vs. 58 years Female: 64% vs. 75% Baseline pain (0-100 VAS): 63 vs. 66 Baseline Pain Disability Index (0-70): 28.9 vs. 31.5	A vs. B <u>4 months</u> Pain (0-100 VAS): 38.4 vs. 42.1, difference -3.8 (95% CI -12.4 to 4.9) Pain Disability Index (0-70): 19.3 (vs. 21.4, difference -2.1 (95% CI -6.3 to 2.1) SF-36 bodily pain subscale (0-100): 53.6 vs. 49.6, difference 3.9 (95% CI -2.7 to 10.7) FFbH-R (0-100, higher scores indicate better function): 66.0 vs. 64.1, difference 1.9 (95% CI -4.2 to 8.0) Number of days with limited function in past 6 months: 40.9 vs. 59.5, difference -18.6 (95% CI -33.3 to -3.9) <u>10 months</u> Pain (0-100 VAS): 39.2	A vs. B <u>4 months</u> SF-36 PCS (0-100): 39.3 vs. 37.6, difference 1.7 (95% CI -1.3 to 4.7) SF-36 MCS (0-100): 49.9 vs. 46.8, difference 3.1 (95% CI -0.5 to 6.6) Allgemaine Depressionssskala (ADS, t standard): 49.7 vs. 50.3, difference -0.6 (95% CI -2.5 to 3.7) <u>10 months</u> SF-36 PCS: 38.9 vs. 36.1, difference 2.8 (95% CI -0.2 to 5.7) SF-36 MCS: 50.5 vs. 47.2, difference 3.3 (95% CI 0.1 to 6.5) ADS: 48.2 vs. 50.7, difference -2.5 (95% CI -5.3 to 0.4)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			vs. 44.9, difference -5.7 (95% CI -14.4 to 3.0) Pain Disability Index: 19.0 vs. 23.0, difference -4.0 (95% CI -8.1 to 0.1) SF-36 bodily pain subscale: 52.4 vs. 44.0, difference 8.5 (95% CI 1.7 to 15.2) Functional (0-100 FFbH-R): 66.0 vs. 63.1, difference 2.9 (95% CI -3.2 to 9.0) Number of days with limited function in past 6 months: 42.4 vs. 52.9, difference -10.5 (95% CI -27.0 to 6.1)	
Carlsson, 2001[178] 1, 3, 6 months Duration of pain: 6 months or longer *Poor*	A. Needle acupuncture or electroacupuncture (n=34) , 8 sessions over 8 weeks, with followup session at 3 and 6 months B. Placebo (sham transcutaneous electrical nerve stimulation) (n=16)	A vs. B (NR) Age: 50 years Female: 66% Baseline Pain (0-100 VAS): 57 vs. 46 Baseline function: Not reported	A vs. B <u>1 month</u> Pain (0-100 VAS): 50 vs. 60, p not reported Global assessment "pain improved": 47% vs. 13%, RR 3.76 (95% CI 0.98 to 14.4) <u>3 month</u> Pain (0-100 VAS): 42 vs. 56, p not reported Global assessment "pain improved": 44% vs. 13%, RR 6.87 (95% CI 1.87 to 25.1) <u>≥6 months outcomes</u> Pain (0-100 VAS): 41 vs. 50, p not reported Global assessment "pain improved": 41% vs. 13%, RR 3.29 (95% CI 0.85 to 12.8)	A vs. B <u>≥6 months</u> Analgesic intake (tablets per week): 21.4 vs. 21.5 Work full time: 32% vs. 31%
Cherkin, 2001[142] 9.5 months Duration of pain: 3 to 12 months, mean not reported *Fair*	A. Needle acupuncture (n=94),10 sessions over 10 weeks B. Attention control (education) (n=90)	A vs. B Age: 54 vs. 44 years Female: 52% vs. 44% Baseline symptom bothersomeness (0-10): 6.2 vs. 6.1 Baseline modified RDQ (0-23): 12. vs. 12.0	A vs. B <u>9.5 months</u> Symptom bothersomeness (0-10): 4.5 vs. 3.8, adjusted p=0.002 Modified RDQ (0-23): 8.0 vs. 6.4, adjusted p=0.05	A vs. B <u>9.5 months</u> ≥1 work-loss day due to LBP in past month: No difference (data not reported) Medication use: 51% vs. 62%, p<0.05 Provider visits: 1.9 (SD 3.7) vs. 1.5 (SD 4.0) LBP medication fills: 4.4 (SD 8.9) vs. 4.0 (SD 8.6) Imaging studies: 0.2 (SD

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				0.4) vs. 0.1 (SD 0.4) Cost of services (1998 $): 252 (SD 46) vs. 200 (SD 45)
Cherkin, 2009[179] 4.5 and 10.5 months Duration of pain: 3 to12 months, mean not reported *Fair*	A. Needle acupuncture (individualized) (n=157), 10 sessions over 7 weeks B. Needle acupuncture (standardized) (n=158), 10 sessions over 7 weeks C. Sham acupuncture (n=162) D. Usual care (n=161)	A vs. B vs. C vs. D Age: 47 vs. 49 vs. 47 vs. 46 years Female: 68% vs. 56% vs. 60% vs. 64% Baseline pain (0-10 VAS): 5.0 vs. 5.0 vs. 4.9 vs. 5.3 Baseline modified RDQ (0-23): 10.8 vs. 10.8 vs. 9.8 vs. 11.0	A vs. B <u>4.5 months</u> Symptom bothersomeness (0-10): 3.8 (2.5) vs. 3.7 (2.6) vs. 3.5 (2.7) vs. 4.4 (2.6) ≥2 point decrease in symptom bothersomeness: 49% vs. 44% vs. 48% vs. 41% Modified RDQ (0-23): 6.8 (5.5) vs. 6.7 (5.8) vs. 6.4 (6.0) vs. 8.4 (6.0) <u>10.5 months</u> Symptom bothersomeness (0-10): 3.7 (2.6) vs. 3.5 (2.7) vs. 3.4 (2.7) vs. 4.1 (2.6) ≥2 point decrease in symptom bothersomeness: 52% vs. 49% vs. 50% vs. 47% Modified RDQ (0-23): 6.0 (5.4) vs. 6.0 (5.8) vs. 6.2 (5.8) vs. 7.9 (6.5) ≥3 point decrease on RMDQ: 65% vs. 65% vs. 59% vs. 50% >7 days with cutting down on activities due to LBP in the past month: A, B and C 5-7% vs. D 18%, p=0.0005	A vs. B <u>10.5 months</u> SF-36 PCS: No differences, data not provided SF-36 MCS: No differences, data not provided Missed work/school for >1 day in past month: A, B and C 5-10% vs. D 16%, p=0.01 Mean total costs of back-related health services: $160-221 across groups, p=0.65
Cho, 2013[180] 1.5 and 4 months Duration of pain: 3 months *Fair*	A. Needle acupuncture (n=57), 12 sessions over 6 weeks B. Sham acupuncture (n=59)	A vs. B Age: 42 vs. 42 years Female: 83% vs. 86% Baseline pain (0-10 VAS): 6.5 vs. 6.4 Baseline ODI (0-100): 28.2 vs. 24.2	A vs. B <u>1.5 months</u> Pain (0-10 VAS): 2.78 (2.32) vs. 4.06 (2.19) ODI (0-100): 15.5 vs. 15.5, SD not reported Symptom bothersomeness (0-10 VAS): 2.83 (2.34) vs. 3.99 (2.06) <u>4 months</u> Pain (0-10 VAS): 2.79 (2.44) vs. 3.52 (2.53) ODI: 15.3 vs. 15.3, SD not reported Symptom	A vs. B <u>1.5 months</u> Beck Depression Inventory (0-63): 6 vs. 7.5, SD not reported <u>4 months</u> Beck Depression Inventory: 6 vs. 7, SD not reported

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			bothersomeness: 2.85 (2.44) vs. 3.63 (2.37)	
Haake, 2007[181] 1.5 and 4.5 months Duration of pain: Mean 8 years *Fair*	A. Needle acupuncture (n=387), 10-15 sessions over 5 weeks B. Sham acupuncture (n=387) C. Usual care (n=388)	A vs. B vs. C Age: 50 vs. 49 vs. 51 years Female: 57% vs. 64% vs. 58% Baseline Von Korff Chronic Pain Grade Scale (0-100): 67.7 vs. 67.8 vs. 67.8 Baseline Hannover Functional Ability Questionnaire (0-100): 46.3 vs. 46.3 vs. 46.7	A vs. B <u>1.5 months</u> Von Korff Chronic Pain Grade Scale (0-100): 45.4 (19.4) vs. 48.5 (19.5) vs. 54.8 (18.4) Hannover Functional Ability (0-100): 65.4 (22.9) vs. 61.3 (22.7) vs. 56.0 (22.0) <u>4.5 months</u> Von Korff Chronic Pain Grade Scale: 40.2 (22.5) vs. 43.3 (23.0) vs. 52.3 (21.2) Hannover Functional Ability (0-100): 66.8 (23.1) vs. 62.2 (23.0) vs. 55.7 (22.7)	A vs. B <u>1.5 months</u> SF-12 PCS (0-100): 40.3 vs. 39.2 vs. 36.1 SF-12 MCS (0-100): 50.5 vs. 50.2 vs. 48.6 Treatment response (≥33% improvement in pain or ≥12% improvement in function): 55.0% (213/387) vs. 51.9% (201/387) vs. 41.9% (162/387), RR 1.05 (95% CI 0.93 to 1.21) for A vs. B and RR 1.31 (95% CI 1.13 to 1.52) for A vs. C <u>4.5 months</u> SF-12 PCS (0-100): 41.6 vs. 39. vs. 35.8 SF-12 MCS (0-100): 50.7 vs. 50.9 vs. 49.2 Treatment response: 47.6% (184/387) vs. 44.2% (171/387) vs. 27.4% (106/387), RR 1.08 (95% CI 0.92 to 1.25) for A vs. B and RR 1.74 (95% CI 1.43 to 2.11) for A vs. C
Kerr, 2003[182] 4.5 months Duration of pain: Mean 86 vs. 73 months *Poor*	A. Needle acupuncture (n=26), 6 sessions over 6 weeks B. Placebo (sham TENS) (n=20)	A vs. B Age: 43 vs. 43 years Female: 50% vs. 35% Baseline pain (0-100 VAS): 79.7 vs. 76 Baseline function: Not reported	A vs. B <u>4.5 months</u> Pain relief "yes": 91% vs. 75%, RR 1.19 (95% CI 0.89 to 1.60)	
Thomas, 2006[183] 9 and 21 months Duration of pain: Mean 17 weeks *Fair*	A. Needle acupuncture (n=147), 10 sessions over 12 weeks B. Usual care (n=68)	A vs. B Age: 42 vs. 44 Female: 62% vs. 58% Baseline Oswestry Disability Index (0-100): 33.7 vs. 31.4 Baseline McGill Present Pain Index (0-5): 2.64 vs. 2.70	A vs. B <u>9 months</u> Oswestry Disability Index (0-100): 20.6 vs. 19.6, adjusted difference -0.5 (-5.1 to 4.2) McGill Present Pain Index (0-5): 1.43 vs. 1.53, adjusted difference -0.1 (-0.4 to 0.3) SF-36 bodily pain (0-100): 64.0 vs. 58.3,	A vs. B <u>21 months</u> Used medication for LBP in the past 4 weeks: 40% vs. 59%, difference -19% (-35 to -3), p=0.03

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			adjusted difference 5.6 (95% CI -0.2 to 11.4) <u>21 months</u> Oswestry Disability Index: 18.3 vs. 21.0, adjusted difference -3.4 (-7.8 to 1.0) McGill Present Pain Index: 1.42 (1.1) vs. 1.71, adjusted difference -0.2 (-0.6 to 0.1) SF-36 bodily pain: 67.8 vs. 59.5, adjusted difference 8.0 (2.8 to 13.2)	

CI = confidence interval; MCS = Mental Component Summary; ODI = Oswestry Disability Index; PCS = Physical Component Summary; RDQ = Roland Morris disability questionnaire; SF-36 = Short-Form 36 Questionnaire; VAS = Visual Analog Scale; FFbH-R = Funktionsfragebogen Hannover-Rücken

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Acupuncture Compared With Sham Acupuncture, Usual Care, an Attention Control, or a Placebo Intervention

Acupuncture was associated with slightly greater effects on short-term function than sham acupuncture or usual care (4 trials, pooled SMD -0.22, 95% CI -0.35 to -0.08, I2=44%) (Figure 6).[177,179-181] Each trial measured function using a different scale; across trials the SMD ranged from -0.34 to 0.00. Differences were slightly greater in trials that compared acupuncture against usual care (2 trials, SMD -0.42, 95% CI -0.60 to -0.21)[179,181] than against sham acupuncture (4 trials, SMD -0.13, 95% CI -0.24 to 0.01).[177,179-181] None of the trials were rated poor-quality. There were no differences between acupuncture versus controls in intermediate-term function (3 trials, pooled standardized mean difference -0.08, 95% CI -0.36 to 0.20, I2=75%)[142,177,179] or long-term function (1 trial, adjusted difference -3.4 on the 0 to 100 ODI, 95% CI -7.8 to 1.0).[183]

Acupuncture was associated with slighter greater effects on short-term pain than sham acupuncture, usual care, an attention control, or a placebo intervention (5 trials, pooled difference -0.55 on a 0 to 10 scale, 95% CI -0.86 to -0.24, I^2=30%) (Figure 7).[177-181] The pooled estimate was similar when poor-quality trials were excluded. When stratified according to the type of control intervention, acupuncture was associated with greater effects when compared with usual care (2 trials, pooled mean difference -1.00, 95% CI -1.60 to -0.28)[179,181] than when compared with sham acupuncture (4 trials, pooled mean difference -0.20, 95% CI -0.66 to 0.19).[177,179-181] There was no difference between acupuncture versus controls in intermediate-term pain (5 trials, pooled mean difference -0.25, 95% CI -0.67 to 0.16, I2=33%).[142,177-179,183] One trial found acupuncture associated with greater effects on long-term pain than usual care (mean difference -0.83, 95% CI -1.51 to -0.15).[183]

Data on effects of acupuncture on quality of life were limited. In two trials, differences between acupuncture versus sham acupuncture or usual care on short-term or intermediate-term SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) scores

were small (range 0.65 to 3.95 points on a 0 to 100 scale), and most differences were not statistically significant.[177,181] Two trials found no clear effects of acupuncture and controls on measures of depression.[177,180]

Two trials found no clear differences between acupuncture versus an attention control in measures of health care utilization (provider visits, medication fills, imaging studies, costs of services)[142,179] and one trial found no clear differences at intermediate-term followup between acupuncture versus placebo transcutaneous electrical stimulation (TENS) in likelihood of working full time.[178]

One trial found acupuncture associated with a higher likelihood of short-term (4.5 months) treatment response (defined as ≥33% pain improvement and ≥12% functional improvement) versus usual care (48% vs. 27%, RR 1.74, 95% CI 1.43 to 2.11), but there was no difference versus sham acupuncture (RR 1.08, 95% CI 0.92 to 1.25).[181]

No trial evaluated effects of acupuncture on use of opioid therapies or health care utilization. There was insufficient evidence to determine effects of duration of acupuncture or number of acupuncture sessions on findings.

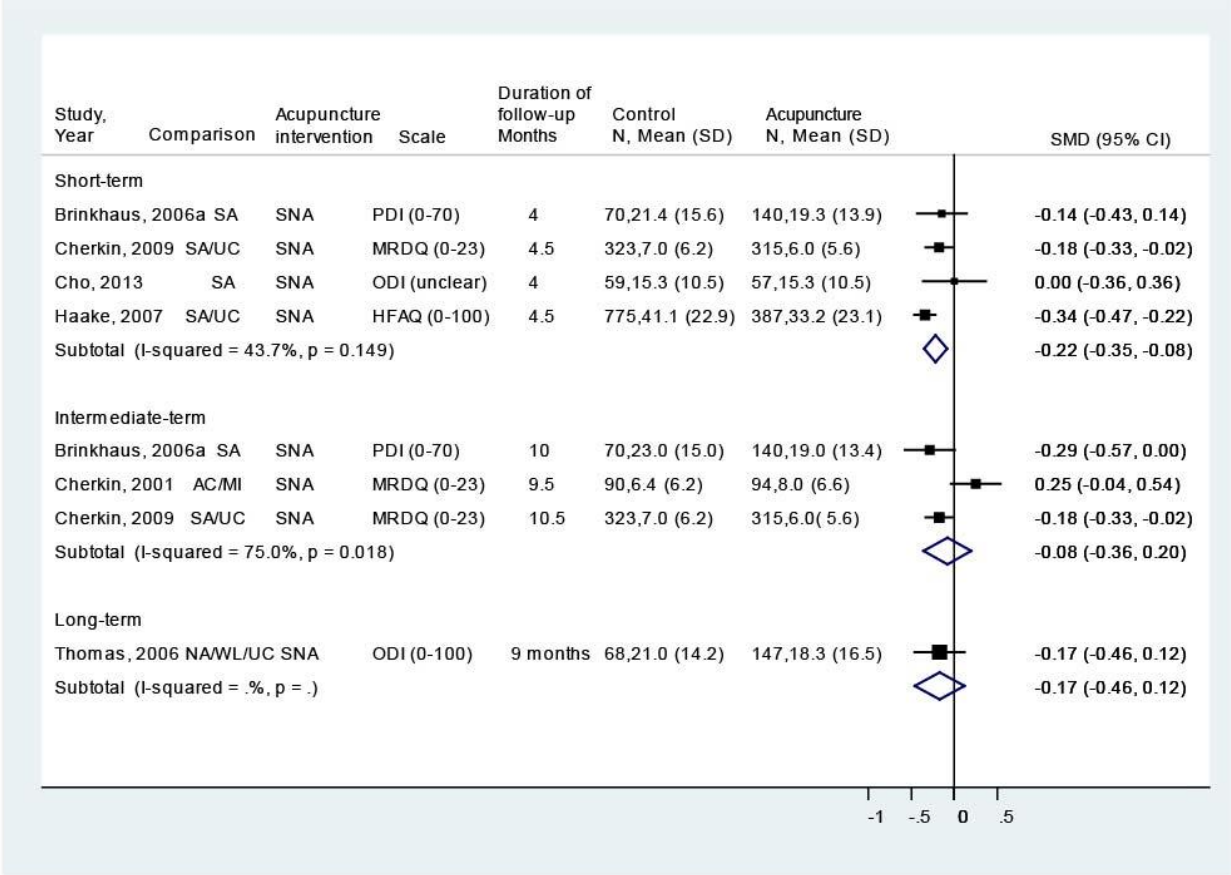
Acupuncture Compared With Pharmacological Therapy or With Exercise

No trial of acupuncture versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

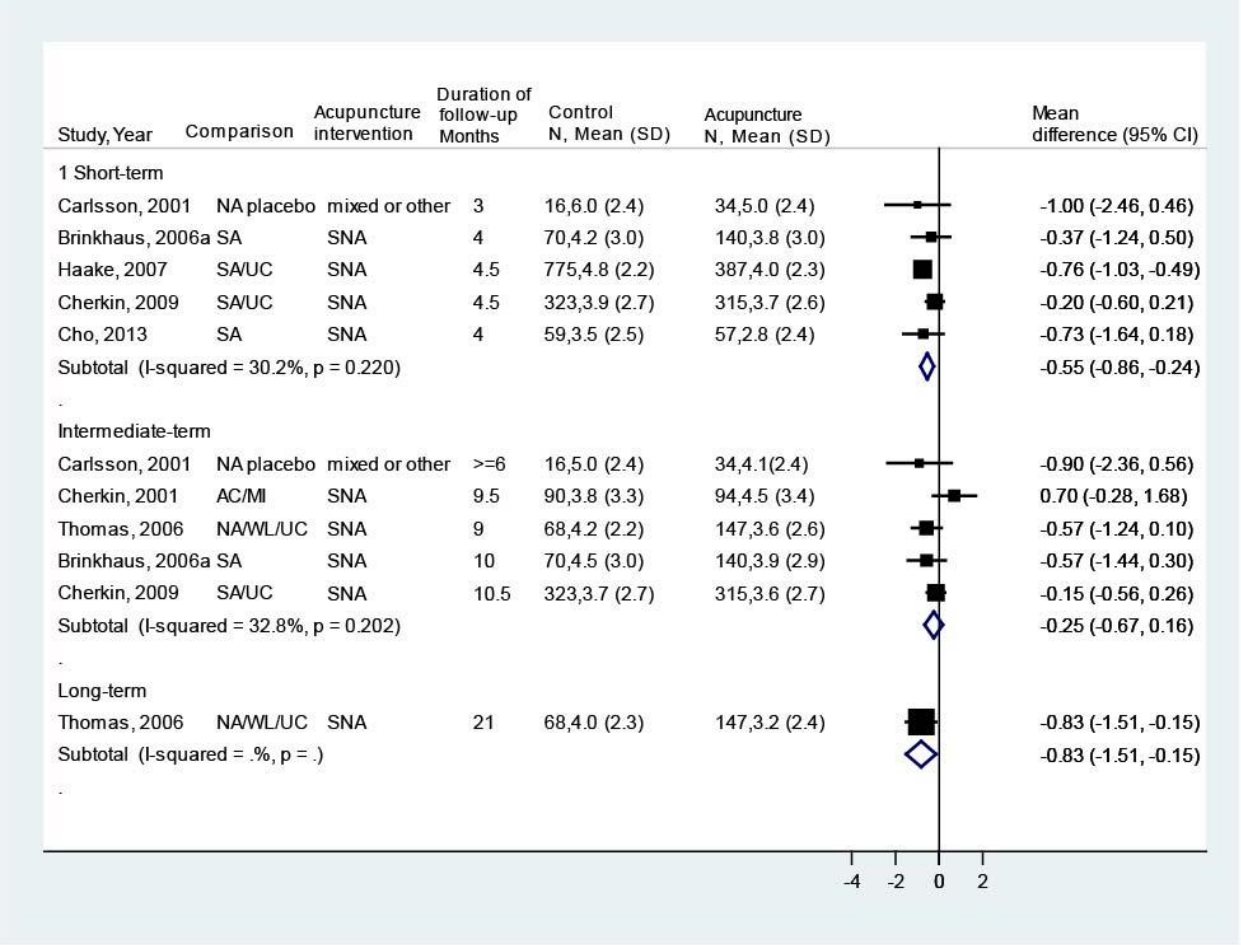
Data on harms were limited, but indicated no clear difference between acupuncture versus control interventions in risk of withdrawal due to adverse events.[179,183] Serious adverse events were rare with acupuncture and control interventions.[142,177,179-181]

Figure 6. Acupuncture versus sham acupuncture, usual care, attention control, or a placebo intervention: effects on function



AC/MI=attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N=number; SA=sham acupuncture; UC/NE/WL=usual care/no exercise/waitlist

Figure 7. Acupuncture versus sham acupuncture, usual care, an attention control, or a placebo intervention: effects on pain



AC/MI = attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N=number; SA = sham acupuncture, UC/NE/WL = usual care/no exercise/waitlist

Short-Wave Diathermy for Low Back Pain

Key Points

- Data from a small, poor-quality trial were insufficient to determine effects of short-wave diathermy versus sham (detuned) diathermy (SOE: Insufficient).

Detailed Synthesis

Data were insufficient from one poor-quality trial (n=68) to evaluate effects of short-wave diathermy (3 times weekly for 4 weeks) versus sham (detuned) diathermy for low back pain (Table 7 and Appendix D).[113] Methodological limitations included unclear randomization and allocation concealment methods, differential attrition, and baseline differences between groups (Appendix E). Although diathermy was associated with worse pain than sham treatment at short-term (8 weeks after completion of therapy) followup (25 vs. 13), statistical significance was not

reported. There were no statistically significant differences in likelihood of using analgesics (7% vs. 22%, RR 0.34, 95% CI 0.08 to 1.50) or being unable to work or having limited activities (7% vs. 19%, RR 0.40, 95% CI 0.09 to 1.80), but estimates were imprecise.

Harms

Adverse events were not evaluated.

Table 7. Summary of results for low back pain: physical modalities (short-wave diathermy)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Gibson, 1985[113] 2 months Duration of pain: 2 to 12 months *Poor*	A. Short wave diathermy (active SWD) (n=34), 12 sessions, 3 session/per week for 4 weeks B. Placebo (detuned SWD) (n=34)	A vs. B Age: 35 vs. 40 years Female: 47% vs. 32% Pain (0-100 VAS): 45 vs. 48	A vs. B 2 months Pain (0-100 VAS, median): 25 vs. 13 (IQR not reported) Unable to work or with limited activities: 7% vs. 19% RR 0.40, 95% CI 0.09 to 1.80	A vs. B 2 months Using analgesics: 7% vs. 22%, RR 0.34, 95% CI 0.08 to 1.50

CI = confidence interval; VAS = Visual Analog Scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Ultrasound for Low back Pain

Key Points

- Two trials found inconsistent effects of ultrasound versus sham ultrasound on short-term function (SOE: Insufficient).Two trials found no differences between ultrasound versus sham ultrasound in short-term pain (SOE: Low).
- One trial found no differences between ultrasound versus sham ultrasound in risk of any adverse events or risk of serious adverse events (SOE: Low)

Detailed Synthesis

Two trials (n=50 and n=455) of ultrasound versus sham ultrasound for low back pain met inclusion criteria (Table 8 and Appendix D).[109,110] The duration of ultrasound therapy was 4 and 8 weeks and the number of sessions was 6 and 10. Both trials evaluated outcomes at short-term (1 month) followup. One good-quality trial[110] was conducted in the United States and one fair-quality trial[109] in Iran (Appendix E). Methodological limitations in the fair-quality trial included failure to blind care providers and unclear blinding of outcome assessors.

Table 8. Summary of results for low back pain: physical modalities (ultrasound)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Ebadi, 2012[109] 1 month Duration of pain: Mean 6 to 8 years *Fair*	A. Ultrasound (n=25), 1.5 W/cm^2 at 1 MHz, 10 sessions over 4 weeks B. Sham ultrasound (n=25)	A vs. B Age: 31 vs. 37 years Female: 25% vs. 50% Pain intensity (mean, 0-100 VAS): 47 vs. 49 Functional Rating Index (mean, 0-100): 41 vs. 44	A vs. B <u>1 month</u> Pain (0-100 VAS): 27.7 vs. 25.5; p=0.48 Functional Rating Index (0-40): 22.8 vs. 30.5; p=0.004	
Licciardone, 2013[110] 3 months Proportion with LBP duration >1 year: 50% *Good*	A. Ultrasound (n=233), 1.2 W/cm^2 at 1 MHz, 6 sessions over 8 weeks B. Sham ultrasound (n=222)	A vs. B Age: 38 vs. 43 years Female: 58% vs. 68% Pain intensity (0-100 VAS): 44 vs. 44 RDQ (0-24): 5 vs. 5	A vs. B <u>1 month, median (IQR)</u> RDQ (0-24): 3 (1-7) vs. 3 (1-7); p=0.93 SF-36 general health (0-100): 72 (52-87) vs. 74 (54-87); p=0.6 Pain improved ≥30%: RR 1.03 (95% CI 0.87 to 1.23) Pain improved ≥50%: RR 1.09 (95% CI 0.88 to 1.35) Pain improved ≥20 mm on 0 to 100 VAS): RR 1.01 (95% CI 0.80 to 1.26) <u>2 months</u> RDQ (0-24): 3 vs. 4; p=0.76 SF-36 general health (0-100): 72 vs.72 (57-85); p=0.53 ≥50% improvement in pain: RR 1.09 (95% CI 0.88 to 1.35) <u>3 months</u> RDQ (0-24): 3 vs. 3; p=0.93	A vs. B <u>1 month</u> Lost 1 or more days work in past 4 weeks because of low back pain: 13% vs. 6%, p=0.11 Prescription drug use for LBP: 16% vs. 18%, p=0.54 SF-36 general health (0-100): 72 (52-87) vs. 74 (54-87), p=0.73 <u>2 months</u> SF-36 general health (0-100): 72 vs. 72, p=0.53 ≥50% improvement in pain: RR 1.09 (95% CI 0.88 to 1.35) <u>3 months</u> SF-36 general health (0-100): 72 vs. 74, p=0.66

CI = confidence interval; NR = not reported; ODI = Oswestry Disability Index; RDQ = Roland Morris disability questionnaire; SF-36 = Short-Form 36 Questionnaire; VAS = Visual Analog Scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Ultrasound Compared With Sham Ultrasound

Limited evidence indicated no clear differences between ultrasound versus sham ultrasound at short-term followup. One good-quality trial (n=455) found no difference between ultrasound versus sham ultrasound in the RDQ (median 3 vs. 3, p=0.93), likelihood for >=50% improvement in pain (RR 1.09, 95% CI 0.88 to 1.35), SF-36 general health (median 72 vs. 74), likelihood of prescription drug use for low back pain (16% vs. 18%, p=0.54), or risk of serious adverse events (1.3% vs. 2.7%, RR 0.48, 95% CI 0.12 to 1.88) or any adverse event (6.0% vs. 5.9%, RR 1.03, 95% CI 0.49 to 2.13).[110] In the smaller (n=50), fair-quality trial, there was no difference between ultrasound versus sham ultrasound in pain (mean 27.7 vs. 25.5 on a 0 to 100

scale, p=0.48), although ultrasound was associated with better function (mean 22.8 vs. 30.5 on the 0 to 40 Functional Rating Index, p=0.004).[109] No trial evaluated longer-term outcomes.

Ultrasound Compared With Pharmacological Therapy or With Exercise

No trial of ultrasound versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

One trial found no differences between ultrasound versus sham ultrasound in risk of any adverse event (RR 1.03, 95% CI 0.49 to 2.13) or serious adverse events (RR 0.48, 95% CI 0.12 to 1.88.[110]

Qigong for Low Back Pain

Key Points

- One trial found no differences between qigong versus exercise in short-term function (difference 0.9 on the RDQ, 95% CI -0.1 to 2.0), although intermediate-term results slightly favored exercise (difference 1.2, 95% CI 0.1 to 2.3) (SOE: Low).
- One trial found qigong associated with slightly lower effects on pain versus exercise at short-term followup (difference 7.7 on a 0 to 100 scale, 95% CI 0.7 to 14.7), but the difference at intermediate-term was not statistically significant (difference 7.1, 95% CI -1.0 to 15.2) (SOE: Low).
- One trial found no difference between qigong versus exercise in risk of adverse events (SOE: Low).

Detailed Synthesis

There was no difference between qigong versus exercise in short-term function (difference 0.9 on the 0 to 24 RDQ, 95% CI -0.1 to 2.0), although intermediate-term results slightly favored exercise (difference 1.2, 95% CI 0.1 to 2.3). One German trial (n=125) compared qigong (weekly sessions for three months) versus exercise therapy (including stretching and strengthening) (Table 9 and Appendix D).[173] It was rated fair-quality due to baseline differences between groups, unblinded design, and suboptimal compliance (Appendix E). Qigong was associated with slightly worse pain versus exercise at short-term followup (mean difference 7.7 on a 0 to 100 scale, 95% CI 0.7 to 14.7) but the difference at intermediate-term was not statistically significant (mean difference 7.1, 95% CI -1.0 to 15.2). There were no differences in sleep, measures of the SF-36 PCS or MCS scores, or in risk of adverse events.

Table 9. Summary of results for low back pain: mind-body practices (qigong)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Blodt, 2015[173] 3 and 9 months Duration of pain: Mean 3	A. Qigong (movement exercises and exercise to change "qi") (n=64)	A vs. B Age (mean): 46 vs. 48 years Female: 91% vs. 70% Baseline pain	A vs. B <u>3 months</u> Average low back pain (0-100 VAS): 35.1 vs. 27.4, difference 7.7 (95% CI 0.7 to 14.7) RDQ (0-24): 4.1 vs. 3.1,	A vs. B <u>3 months</u> SF-36 Physical component score: 45.8 vs. 46.6, difference -0.8 (95% CI –3.4 to 1.9)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
years *Fair*	12 sessions over 12 weeks B. Exercise (strengthening, stretching and relaxation exercises) (n=63)	(0-100 VAS): 55.6 vs. 52.1 Baseline RDQ: 6.2 vs. 5.7	difference 0.9 (95% CI –0.1 to 2.0) SF-36 Bodily pain (0-100): 43.0 vs. 44.6, difference 1.5 (95% CI -1.2 to 4.2) <u>9 months</u> Average low back pain (0-100 VAS): 35.9 vs. 28.8, difference 7.1 (95% CI –1.0 to 15.2) RDQ: 4.3 vs. 3.1, difference 1.2 (95% CI 0.1 to 2.3) SF-36 Bodily pain: 41.4 vs. 43.4, difference -2.0 (95% CI -5.4 to 1.4)	SF-36 Mental component score: 45.4 vs. 46.6, difference 11.2 (95% CI –4.9 to 2.4) Quality of sleep (0-10): 4.6 vs. 4.5, difference 0.0 (95% CI–0.9 to 1.0) Sleep satisfaction (0-10): 5.0 vs. 4.8, difference 0.3 (95% CI –0.6 to 1.1) <u>9 months</u> SF-36 Physical component score: 44.8 vs. 46.5, difference -1.8 (95% CI -4.9 to 1.3) SF-36 Mental component score: 45.0 vs. 45.5, difference -0.5 (95% CI -4.6 to 3.6) Quality of sleep: 4.5 vs. 4.7, difference -0.2 (95% CI -1.0 to 0.7) Sleep satisfaction: 5.1 vs. 5.1, difference -0.1 (95% CI –0.9 to 0.8)

CI = confidence interval; RDQ = Roland Morris Disability Questionnaire; SF-36 = Short-Form 36 Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Massage for Low Back Pain

Key Points

- There were no differences between massage versus controls in short-term function (3 trials, SMD -0.24, 95% CI -0.49 to 0.02, I2=0%) or intermediate-term function (2 trials, SMD 0.00, 95% CI -0.22 to 0.22, I2=0%) (SOE: Moderate for short-term, low for intermediate-term).
- Massage was associated with slighter greater effects on short-term pain than sham massage or usual care (3 trials, pooled difference -0.63 on a 0 to 10 scale, 95% CI -1.00 to -0.26, I2=0%). There was no difference between massage versus controls in intermediate-term pain (2 trials, difference -0.12, 95% CI -0.99 to 0.75, I2=43%) (SOE: Moderate for short-term, Low for intermediate-term).
- One trial found no differences between massage versus exercise in intermediate-term pain, function, or the SF-36 MCS or PCS scores (SOE: Low).
- Two trials of massage reported no serious adverse events; in three trials, the proportion of massage patients who reported increased pain ranged from <1% to 26% (SOE: Low).

Detailed Synthesis

Five trials of massage for low back pain met inclusion criteria (Table 10 and Appendix Table D).[88,141-143,151] Massage techniques varied across trials. Two trials evaluated reflexology,[88,143] one trial myofascial release,[141] and two trials mixed massage techniques that included Swedish massage.[142,151] Sample sizes ranged from 15 to 216 (total sample=625). Two trials compared massage versus sham massage,[141,143] two trials massage versus usual care,[88,151] and one trial compared massage versus an attention control (self-care education).[142] One trial was conducted in India[141] and the rest in the United States or Europe. The duration of massage therapy ranged from 6 to 10 weeks and the number of massage sessions ranged from 6 to 24. Two trials reported outcomes through intermediate-term followup,[142,151] and three only reported short-term outcomes.[88,141,143] No trial reported long-term outcomes.

All of the massage trials were rated fair-quality (Appendix E). Methodological limitations included unclear allocation concealment methods and unblinded design. One trial reported high loss to followup.[88]

Table 10. Summary of results for low back pain: manual therapies (massage)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Ajimsha, 2014[141] 1 month Duration of pain: 2.3 vs. 2.25 years *Fair*	A. Myofascial release (n=38) 24 sessions, 3 session/week for 8 weeks B. Sham myofascial release (n=36)	A vs. B Age: 36 vs. 34 years Female: 76% vs. 78% Baseline pain (0-78 McGill Pain): 23.2 vs. 23.0 Baseline Quebec Back Disability Scale (0-100): 37.1 vs. 35.3	A vs. B <u>1 month</u> McGill Pain Questionnaire (0-78): 13.1 vs. 18.3, mean difference -3.25, p<0.005 Quebec Back Disability Scale (0-100): 28.7 vs. 32.5, mean difference -2.02, p<0.005	NR
Cherkin, 2001[142] 10.5 months Duration of pain >1 year: 64% vs. 62% *Fair*	A. Mixed massage (including Swedish) (n=78) Up to 10 sessions over 10 weeks B. Attention control (self-care education) (n=90)	A vs. B Age: 46 vs. 44 years Female: 69% vs. 56% Baseline modified RDQ (0-23): 11.8 vs. 12.0 Baseline symptom bothersomeness (0-10): 6.2 vs. 6.1	A vs. B <u>10.5 months</u> Modified RDQ (0-23): 6.8 vs. 6.4, p=0.03 Symptom bothersomeness (0-10): 3.2 vs. 3.8, p=0.003	A vs. B <u>10.5 months</u> Low back pain medication: 2.5 vs. 4.0, p=0.69 SF-12 Mental Component Score: no differences, data not shown
Little, 2008[151] 11.5 months Duration of pain: NR *Fair*	A. Mixed massage (including Swedish) (n=72), 6 sessions over 6 weeks/ B: Usual care (n=72)	Age: 45-46 years Female: 64-78% Baseline Deyo troublesomeness: 3.3-3.4	A vs. B <u>10.5 months</u> RDQ (0-24): NR vs. 9.23 (5.3), difference -0.45 (95% CI -2.3 to 1.39) Von Korff disability (0-10): NR vs. 3.32 (2.25), difference 0.46	A vs. B <u>10.5 months</u> Von Korff overall (0-10): NR vs. 4.19, difference 0.31 (95% CI -0.52 to 1.14) SF-36 PCS (0-100): NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	C: Exercise (regular exercise) (n=72) 5 times per week	Baseline RDQ (0-24): 10.8-11.3	(95% CI -0.43 to 1.35) Von Korff pain (0-10): NR vs. 4.74 (2.20), difference 0.29 (95% CI -0.58 to 1.16) Deyo troublesomeness scale (1-5): NR vs. 3.05 (0.80), difference 0.04 (-0.25 to 0.33) A vs. C <u>10.5 months</u> RDQ: -0.45 (-2.3 to 1.39) vs. -1.65 (-3.62 to 0.31) Von Korff disability: 0.46 (-0.43 to 1.35) vs. 0.05 (-0.92 to 1.02) Von Korff pain: 0.29 (-0.58 to 1.16) vs. -0.31 (-1.26 to 0.63) Deyo troublesomeness scale: 0.04 (-0.25 to 0.33) vs. -0.21 (-0.52 to 0.09)	vs. 56.1 (18.6), difference -1.45 (95% CI -9.04 to 6.15) SF-36 MCS (0-100): NR vs. 64.8 (17.5), difference -2.11 (95% CI -9.37 to 5.16) A vs. C <u>10.5 months</u> Von Korff overall: 0.31 (-0.52 to 1.14) vs. -0.19 (-1.09 to 0.72) SF-36 Physical Component Score: -1.45 (-9.04 to 6.15) vs. -2.08 (-10.6 to 6.40) SF-36 Mental Component Score: -2.11 (-9.37 to 5.16) vs. 0.72 (-7.38 to 8.81)
Poole, 2007[88] 4.5 months Duration of pain: 10 vs. 11 vs. 9.5 years *Fair*	A. Reflexology (n=77) 6 sessions over 6-8 weeks B. Usual care (n=75)	A vs. B Age: 47 vs. 47 years Female: 62% vs. 51% Baseline pain (0-100 VAS): 44.5 vs. 40.6 Baseline ODI: 33.0 vs. 36.6	A vs. B <u>4.5 months</u> Pain (0-100 VAS): 39.8 (29.2) vs. 42.7 (28.4) ODI (0-100): 29.0 (20.2) vs. 32.9 (17.6)	A vs. B <u>4.5 months</u> Beck Depression Inventory (0-63): 11.6 (10.9) vs. 12.8 (9.2) SF-36 Physical Functioning : 57.1 (31.8) vs. 52.2 (29.5) SF-36 Social Functioning: 68.1 (31.8) vs. 61.5 (30.8) SF-36 Physical Limitations: 48.2 (46.4) vs. 37.8 (42.5) SF-36 Emotional Limitations: 55.0 (46.5) vs. 62.0 (44.0)
Quinn, 2008[143] 1.5 and 3 months Duration of pain: At least 3 months *Fair*	A. Reflexology (pressure massage stimulation) (n=7) 6 sessions over 6 weeks B. Sham reflexology (n=8)	A vs. B Age (median): 42 vs. 45 Female: 86% vs. 50% Baseline pain (0-10 VAS): 4.7 vs. 3.4 Baseline RDQ: 5 vs. 7.5	A vs. B <u>1.5 months, median (IQR)</u> RDQ: 4 (3 to 4.5) vs. 4.5 (1 to 7) Pain (0-10 VAS): 2.1 (1.5 to 4.9) vs. 4.1 (2.7 to 5.1) McGill Pain Questionnaire (0-77): 11 (6 to 17) vs. 6.5 (5 to 13) <u>3 months, median (IQR)</u> RDQ: 4 (2 to 5) vs. 3.5 (1.8 to 4.8) VAS: 2.2 (1.6 to 3.2) vs. 3.2 (2.6 to 4.6) McGill Pain Questionnaire (0-77): 6 (4 to 13) vs. 7.5 (3.8 to 9.8)	A vs. B <u>1.5 months, median (IQR)</u> SF-36 General health: 52.9 (49 to 54) vs. 42.2 (40 to 51) SF-36 Physical functioning: 48.6 (47 to 50) vs. 43.4 (40 to 50) SF-36 Mental health: 47.2 (43 to 56) vs. 47.2 (42 to 53) <u>3 months, median (IQR)</u> SF-36 General health: 48.2 (46 to 52) vs. 47.0 (38 to 53)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				SF-36 Physical functioning: 50.7 (44 to 51) vs. 45.5 (44 to 50) SF-36 Mental health: 52.8 (39 to 53) vs. 48.6 (44 to 51)

Abbreviations: CI = confidence interval; NR = not reported; MCS = Mental Component Summary; PCS = Physical Component Summary; ODI = Oswestry Disability Index; RDQ =Roland Morris disability questionnaire; SF-36 = Short-Form 36 questionnaire; VAS = visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Massage Compared With Sham Massage, Usual Care, or an Attention Control

Effects of massage on short-term function were small and not statistically significant (3 trials, SMD -0.24, 95% CI -0.49 to 0.02, I2=0%)[88,141,143] and there was no effect on intermediate-term function (2 trials, SMD 0.00, 95% CI -0.22 to 0.22, I2=0%) (Figure 8).[142,151] One trial[151] found no differences between massage versus usual care on the SF-36 MCS or PCS Scores at intermediate-term followup, and one trial[88] found no effects on various SF-36 subscales or the Beck Depression Inventory at short-term followup.

Massage was associated with slighter greater effects on short-term pain than sham massage or usual care (3 trials, pooled difference -0.63 on a 0 to 10 scale, 95% CI -1.00 to -0.26, I2=0%) (Figure 9).[88,141,143] The massage technique was myofascial release in one trial (difference -3.25 on the 0 to 78 McGill Pain Questionnaire, p=0.47)[141]and foot reflexology in the other two (pooled difference -0.53 on a 0 to 10 scale, 95% CI -1.6 to 0.39).[88,143] There was no difference between massage (mixed massage techniques, including Swedish massage) versus an attention control or usual care in intermediate-term pain (2 trials, difference -0.12, 95% CI -0.99 to 0.75, I2=43%).[142,151]

No trial evaluated effects of massage therapy on use of opioid therapies or health care utilization. There was insufficient evidence to determine effects of duration of massage or number of massage sessions on findings.

Massage Compared With Pharmacological Therapies

No trial of massage versus pharmacological therapy met inclusion criteria.

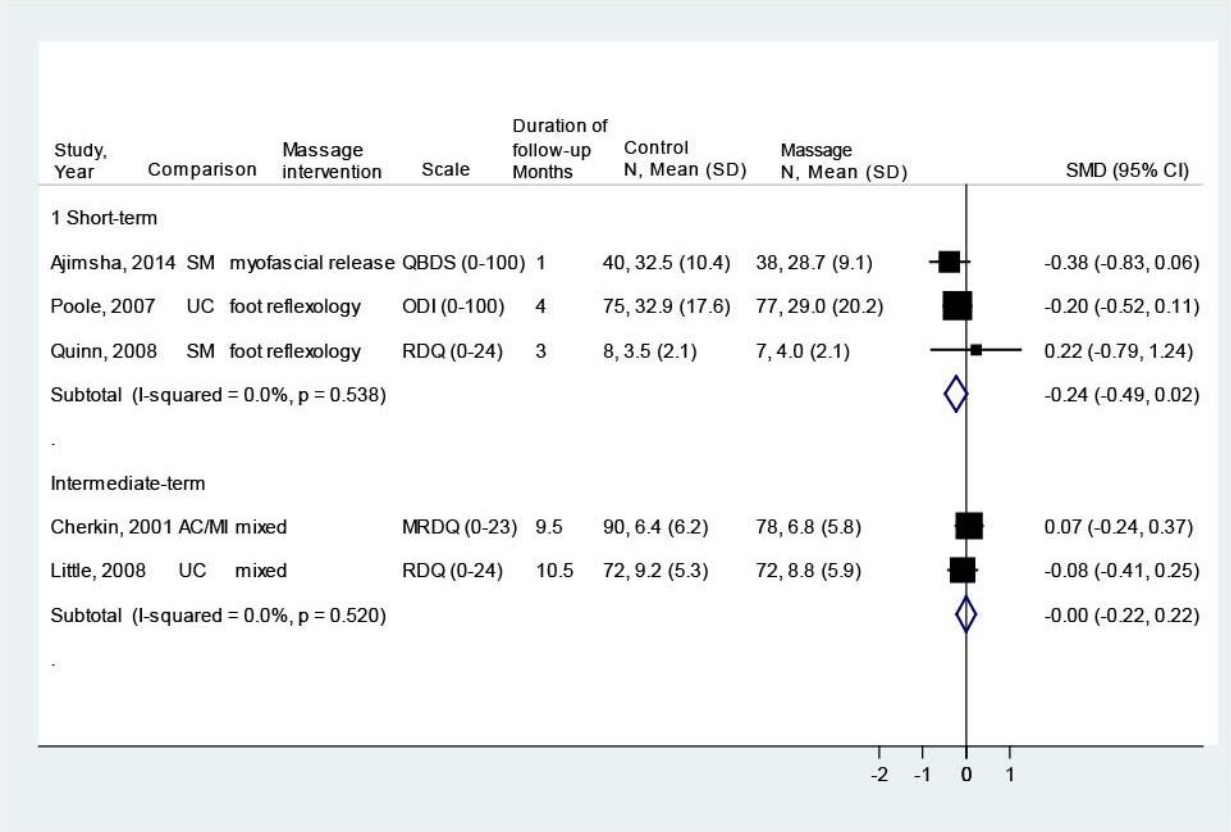
Massage Compared With Exercise

One trial found no differences between massage versus exercise in intermediate-term function (difference 1.2 on the 0 to 24 RDQ, 95% CI -1.47 to 3.87), pain (difference 0.60 on the 0 to 10 Von Korff pain scale, 95% CI -0.67 to 1.87), or the SF-36 MCS or PCS scores (differences 0 to 3 points on 0 to 100 scales, p>0.05).[151]

Harms

Two trials[141,142] reported no serious adverse events, and one trial[143] reported no adverse events. In three trials, the proportion of massage patients who reported increased pain ranged from <1% to 26%.[141,142,151]

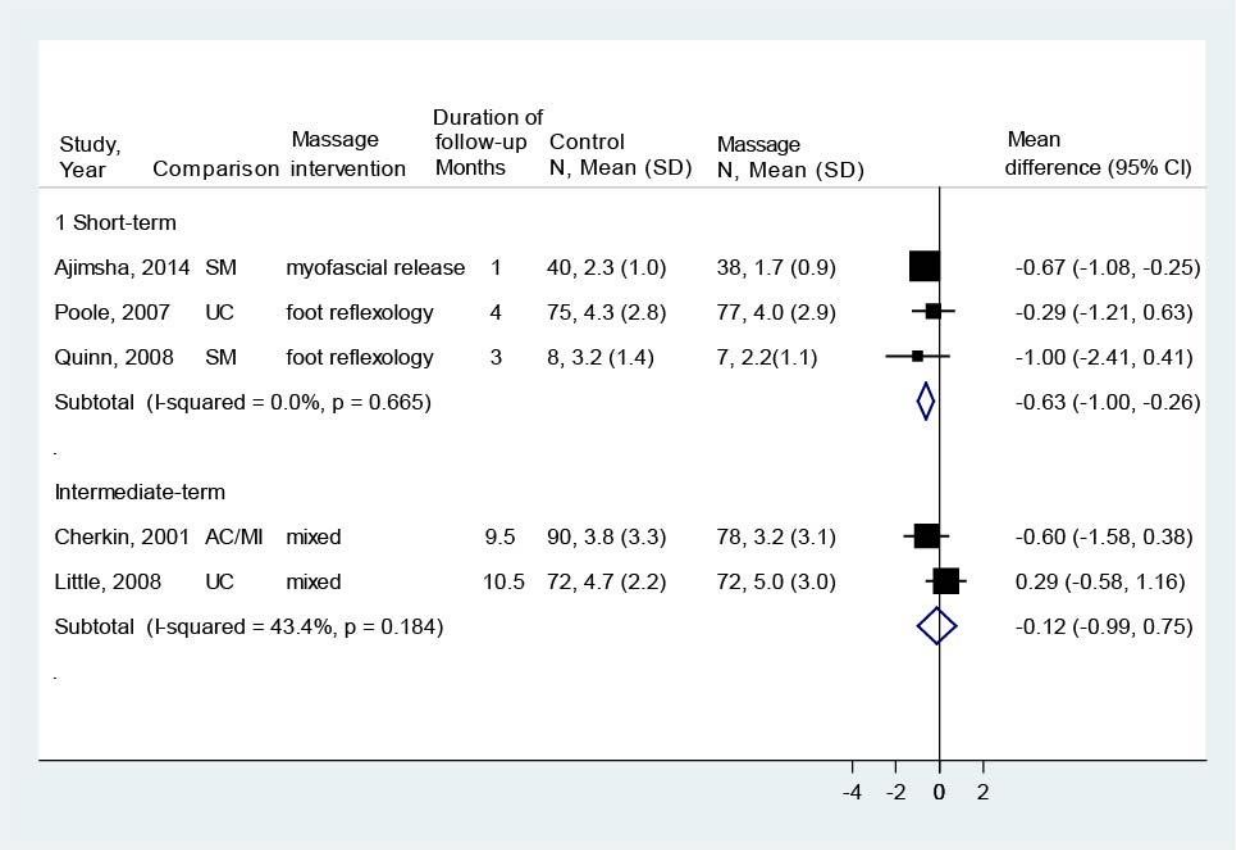
Figure 8. Massage versus sham massage, usual care, attention control, or a placebo intervention: effects on function



AC/MI = attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N=number; SM = sham massage, UC/NE/WL = usual care/no exercise/waitlist

Figure 9. Massage versus sham massage, usual care, attention control, or a placebo intervention: effects on

pain



AC/MI = attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; SM = sham massage, UC/NE/WL = usual care/no exercise/waitlist

Yoga for Low Back Pain

Key Points

- Yoga was associated with slighter greater effects on function than an attention or wait list control at short-term (5 trials, pooled SMD -0.49, 95% CI -0.75 to -0.23, I2=59%) and intermediate-term (3 trials, pooled SMD -0.33, 95% CI -0.49 to -0.16) followup (SOE: Moderate for short-term, Low for intermediate-term).
- Yoga was associated with moderately greater effects on pain than an attention or wait list control at short-term (4 trials, pooled difference -1.23 on a 0 to 10 scale, 95% CI -2.08 to -0.39, I2=77%) and intermediate-term (2 trials, pooled difference -1.17, 95% CI -1.91 to -0.44, I2=26%) followup (SOE: Low for short-term, Moderate for intermediate-term).
- Yoga was associated with no statistically significant differences versus exercise in short-term or intermediate-term pain or function (SOE: Low).
- Yoga was not associated with increased risk of harms versus controls (SOE: Low).

Detailed Synthesis

Seven trials of yoga for low back pain met inclusion criteria (Table 11, Appendix D).[162-167,174] Four trials evaluated Iyengar yoga,[165-167,174] two trials Viniyoga,[163,164] and one trial Hatha yoga.[162] Sample sizes ranged from 60 to 313 (total sample=1,316). Five trials compared yoga

versus an attention control (education),[162-165,167] one trial yoga versus wait list control,[166] and four trials yoga versus exercise.[162-164,174] One trial was conducted in India[174] and the rest in the United States or Europe. The duration of yoga therapy ranged from 4 to 24 weeks and the number of sessions ranged from 4 to 48. In one trial, patients who received 12 weeks of yoga therapy were randomized to ongoing once weekly maintenance sessions or no maintenance.[162] Three trials reported outcomes through intermediate-term followup,[162,165,166] and four only reported short-term outcomes.[163,164,167,174]

All of the trials were rated fair-quality (Appendix E). Trials could not effectively blind patients; other methodological limitations included unclear allocation or randomization methods and high attrition.

Table 11. Summary of results for low back pain: mind-body practices (yoga)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Nambi, 2014[174] 5.5 months Duration of pain: >3 months *Fair*	A. Iyengar yoga (29 poses) (n=30) 5 sessions a week for 4 weeks B. Exercise (stretching exercises for soft tissue flexibility and range of motion) (n=30)	A vs. B Age: 44 vs. 43 years Female: 63% vs. 43% Baseline pain (0-10 VAS): 6.7 vs. 6.7 Baseline function, Physically unhealthy days: 18.0 vs. 17.8	A vs. B <u>5 months</u> Pain (0-10 VAS): 1.8 vs. 3.8, p=0.001 Physically unhealthy days: 2.6 vs. 6.9, p=0.001 Activity limitation days: 2.0 vs. 5.0, p=0.001	A vs. B <u>5.5 months</u> Mentally unhealthy days: 2.1 vs. 5.0, p=0.001 Activity limitation (days): 2.0 vs. 5.0, p=0.001
Saper, 2017[162] 10 months Duration of pain: >3 months *Fair*	A. Hatha yoga (n=127) 12 sessions over 12 weeks, with or without ongoing weekly maintenance sessions B. Exercise (n=129) C. Attention control (education) (n=64)	A vs. B vs. C Age: 46 vs. 46 vs. 44 Female: 57% vs. 70% vs. 66% Baseline pain (0-10 NRS): 7.1 vs. 7.2 vs. 7.0 Baseline modified RDQ: 13.9 vs. 15.6 vs. 15.0	A1 (no maintenance) vs. A2 (maintenance) vs. C, mean (SE) <u>3.5 months</u> Modified RDQ (0-23): 10.1 (0.77) vs. 9.5 (0.77) vs. 11.6 (0.75) Pain (0-10 NRS): 4.3 (0.32) vs. 4.6 (0.32) vs. 5.5 (0.31) <u>9 months</u> Modified RDQ (0-23): 9.2 vs. 8.9 vs. 11.1 Pain (0-10 NRS): 4.3 vs. 4.4 vs. 5.2 A1 vs. A2 vs. B1 vs. B2 <u>3.5 months</u> Modified RDQ (0-23): 10.1 (0.77) vs. 9.5 (0.77) vs. 10.4 (0.84) vs. 10.1 (0.83) Pain (0-10 NRS): 4.3 (0.32) vs. 4.6 (0.32) vs. 4.7 (0.35) vs. 4.8 (0.34) <u>9 months</u> Modified RDQ (0-23): 9.2 (0.88) vs. 8.9 (0.88) vs. 8.9 (0.96) vs.	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			9.4 (0.94) Pain (0-10 NRS): 4.3 (0.36) vs. 4.4 (0.35) vs. 4.0 (0.39) vs. 4.1 (0.37)	
Sherman, 2005[163] 3.5 months Duration of pain: 3 to 15 months *Fair*	A. Viniyoga (n=36) 12 sessions 1 session/week for 12 weeks B. Exercise (n=35) C. Attention control (self-care advice) (n=30)	A vs. B vs. C Age: 44 vs. 42 vs. 45 Female: 69% vs. 63% vs. 67% Baseline symptom bothersomeness (0-10): 5.4 vs. 5.7 vs. 5.4 Baseline RDQ: 8.1 vs. 9.0 vs. 8.0	A vs. B <u>3.5 months</u> Modified RDQ (0-23): 3 vs. 5 (estimated from graph), adjusted difference -1.5 (-3.2 to 0.2)[b] Reduction in RDQ score ≥50%:69% vs. 50%, RR 1.4 (95% CI 0.91 to 2.1) Bothersomeness: 1.8 vs. 3.3 (estimated from graph), adjusted difference -1.4 (95% CI -2.5 to -0.2)[b] Medication use: 21% vs. 50%, RR 0.41 (95% CI 0.20 to 0.87) A vs. C <u>3.5 months</u> Symptom bothersomeness (0-10): 1.8 vs. 4.1, adjusted difference -2.2 (95% CI -3.2 to -1.2) Modified RDQ (0-23): 3 vs. 7, adjusted difference -3.6 (95% CI -5.4 to -1.8) Reduction in RDQ ≥50%: 69% vs. 30%, RR 2.3 (95% CI 1.3 to 4.2)	A vs. B <u>3.5 months</u> Medication use: 21% vs. 59%, RR 0.35 (95% CI 0.15 to 0.73) SF-36: No significant differences (data not provided)
Sherman, 2011[164] 3.5 months Duration of pain: 3 to 6 months *Fair*	A. Viniyoga (n=92) 12 sessions 1 session/week for 12 weeks B. Exercise (n=91) C. Attention control (self-care advice) (n=30)	A vs. B Age: 47 vs. 49 vs. 50 Female: 67% vs. 63% vs. 60% Baseline symptom bothersomeness (0-10): 4.9 vs. 4.5 vs. 4.7 Baseline RDQ: 9.8 vs. 8.6 vs. 9.0	A vs. B <u>3.5 months</u> Symptom bothersomeness (0-10): 3.59 (95 %CI 3.12 to 4.06) vs. 3.34 (95% CI 2.86 to 3.81) Modified RDQ (0-23): 4.49 (95% CI 3.51 to 5.48) vs. 4.26 (95% CI 3.30 to 5.22), adjusted difference -0.35 (95% CI -1.52 to 0.83) Reduction in RDQ score ≥50%: 60% vs. 51%, RR 1.17 (95% CI 0.88 to 1.54) A vs. C <u>3.5 months</u> Symptom bothersomeness (0-10): 3.59 (95% CI 3.12 to 4.06) vs. 3.80 (95% CI 3.14 to 4.46) Modified RDQ (0-23): 4.49 vs. 5.73, adjusted difference -1.81 (95% CI -3.12 to -0.50)	A vs. B <u>3.5 months</u> LBP better, much better, or completely gone: 51% vs. 51%, RR 1.00 (95% CI 0.75 to 1.34) A vs. C LBP better, much better, or completely gone: 51% vs. 20%, RR 2.57, 95% CI 1.39 to 4.78)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			Reduction in RDQ score ≥50%: 60% vs. 31%, RR 1.90 (95% CI 1.21 to 2.99) LBP better, much better, or completely gone: 51% vs. 20%, RR 2.57, 95% CI 1.39 to 4.78)	
Tilbrook, 2011[165] 3 and 6 months Duration of pain: 96 vs. 72 months *Fair*	A. Iyengar yoga (n=156) 12 sessions 1 session/week for 12 weeks B. Attention control (self-care advice) (n=157)	A vs. B Age: 46 vs. 46 Female: 68% vs. 73% Baseline Aberdeen Back Pain Scale (0-100): 25.36 vs. 26.69 Baseline RDQ (0-24): 7.84 vs. 7.75	A vs. B Mean difference in change from baseline (95% CI) <u>3 months</u> RDQ (0-24): -1.48 (-2.62 to -0.33) Aberdeen Back Pain Scale (0 to 100): -1.74 (-4.32 to 0.84) <u>6 months</u> RDQ: -1.57 (-2.71 to -0.42) Aberdeen Back Pain Scale: -0.73 (-3.30 to 1.84)	A vs. B Mean difference in change from baseline (95% CI) <u>3 months</u> SF-12 PCS (0-100): 1.24 (-0.83 to 3.33) SF-12 MCS (0-100): 2.02 (-0.34 to 4.37) <u>6 months</u> SF-12 PCS: 0.80 (-1.28 to 2.87) SF-12 MCS: 0.42 (-1.92 to 2.77)
Williams, 2005[167] 3 months Duration of pain: 11.3 vs. 11.0 years *Fair*	A. Iyengar yoga (n=30), 16 sessions 1 session/week for 16 weeks B. Attention control (education) (n=30)	A vs. B Age: 49 vs. 48 Female: 65% vs. 70% Pain intensity, McGill Pain Questionnaire (0-10 VAS): 2.3 vs. 3.2 Pain Disability Index (7-70): 14.3 vs. 21.2	A vs. B <u>3 months</u> Pain, McGill Pain Questionnaire (0-10 VAS): 0.6 vs. 2.0, p=0.039 Pain Disability Index (7-70): 3.9 vs. 12.7, p=0.009 Present Pain Index (0-5): 0.5 vs. 1.1, p=0.013	A vs. B <u>3 months</u> Stopped or decreased medication use: 50% vs. 33%, p=0.007
Williams, 2009[166] 6 months Duration of pain: 47 vs. 78 months *Fair*	A. Iyengar yoga (n=43) 48 sessions for 24 weeks B. Wait-list (standard medical care) (n=47)	A vs. B Age: 48 vs. 48 years Female: 74% vs. 79% Pain (0-100 VAS): 41.9 vs. 41.2 Oswestry Disability Index (0-100): 25.2 vs. 23.1	A vs. B <u>6 months</u> Pain (0-100 VAS): 22.2 vs. 38.3, p=0.0009 Oswestry Disability Index (0-100): 19.3 vs. 23.5, p=0.001	A vs. B <u>6 months</u> Beck Depression Inventory (0-63): 4.6 vs. 7.8, p=0.0004

CI = confidence interval; EQ = EuroQol; NR = not reported; NRS = Numerical Rating Scale; ODI = Oswestry Disability Index; RDQ = Roland Morris disability questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

[b] Adjusted for baseline scores

Yoga Compared With an Attention Control or Waitlist

Yoga was associated with slightly greater effects on short-term function than controls (5 trials, pooled SMD -0.49, 95% CI -0.75 to -0.23, I2=59%) (Figure 10).[162-165,167] Results were similar when trials were stratified according to whether they evaluated Viniyoga (2 trials, pooled SMD -0.56, 95% CI -1.38 to 0.19), Hatha yoga (1 trial, SMD -0.30, 95% CI -0.61 to 0.00),[162] or Iyengar yoga (2 trials, SMD -0.54, 95% CI -1.41 to 0.14).[167] In four trials that evaluated function using the RDQ or modified RDQ, the mean difference was -1.95 (95% CI -3.06 to -0.85, I2=46%).[162-165] Yoga was also associated with greater effects on intermediate-term function than controls (3 trials, pooled SMD -0.33, 95% CI -0.49 to -0.16, I2=0%).[162,165,166] In two trials that evaluated function with the RDQ or modified RDQ, the mean difference was -1.58 points (95% CI -2.47 to -0.70, I2=0%).[162,165]

Yoga was associated with moderately greater effects on short-term pain than an attention or wait list control (4 trials, pooled difference -1.23, 95% CI -2.08 to -0.39 on a 0 to 10 scale, I^2=77%) (Figure 11).[162-164,167] Estimates were similar from two trials of Viniyoga (pooled difference -1.25, 95% CI -3.78 to 1.27)[163,164] and one trial each of Hatha yoga (difference -1.05, 95% CI -1.81 to -0.29)[162] and Iyengar yoga (difference -1.40, 95% CI -2.27 to -0.53).[167] No trials were rated poor quality. Yoga was also associated with greater effects on intermediate-term pain than controls, based on two trials (pooled mean difference -1.17, 95% CI -1.91 to -0.44, I2=26%).[162,166]

Data on effects of yoga on quality of life were limited. One trial found no difference between yoga versus an attention control on the SF-36 Physical and Mental Component Summaries at short-term or intermediate-term followup (differences 0.42 to 2.02 points on a 0 to 100 scale).[165] One other trial found no differences between yoga versus an attention control on the SF-36, but data were not provided.[163]

One trial found yoga associated with lower (better) scores on the Beck Depression Inventory at intermediate-term followup (mean 4.6 vs. 7.8 on a 0 to 63 scale, p=0.004).[166]

Yoga Compared With Pharmacological Therapy

No trial of yoga versus pharmacological therapy met inclusion criteria.

Yoga Compared With Exercise

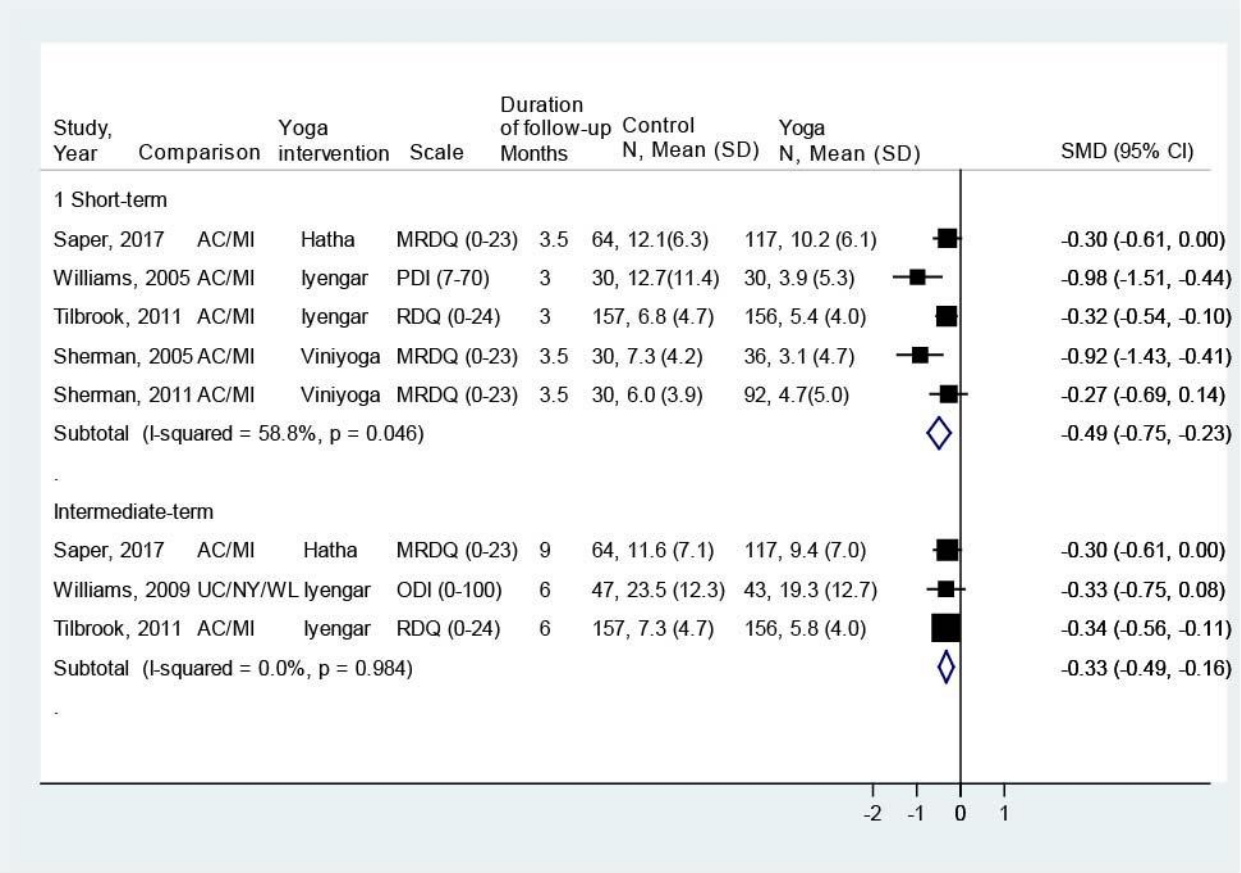
There were no differences between yoga versus exercise in short-term function (3 trials, pooled SMD -0.10, 95% CI -0.34 to 0.13, I2=38%)[162-164] or intermediate-term function (1 trial, SMD -0.01, 95% CI -0.26 to 0.24)[162] (Figure 12). One trial found no difference between yoga versus exercise on the SF-36 at short-term followup (data not provided).[163]

Effects of yoga versus exercise on short-term pain were not statistically significant and there was marked heterogeneity (4 trials, pooled difference -0.89 on a 0 to 10 scale, 95% CI -1.99 to 0.21 I2=92%) (Figure 13).[162-164,174] In one trial of Viniyoga,[164] results favored exercise (difference 0.25, 95% CI -0.41 to 0.91) and in three trials (one each of Viniyoga, Iyengar yoga, and Hatha yoga)[162,163,174] effects favored yoga (mean differences of -0.30 to -2.00). No trials were rated poor quality. One trial found no difference between yoga versus exercise in intermediate-term pain (difference 0.30, 95% CI -0.39 to 0.99).[162]

Harms

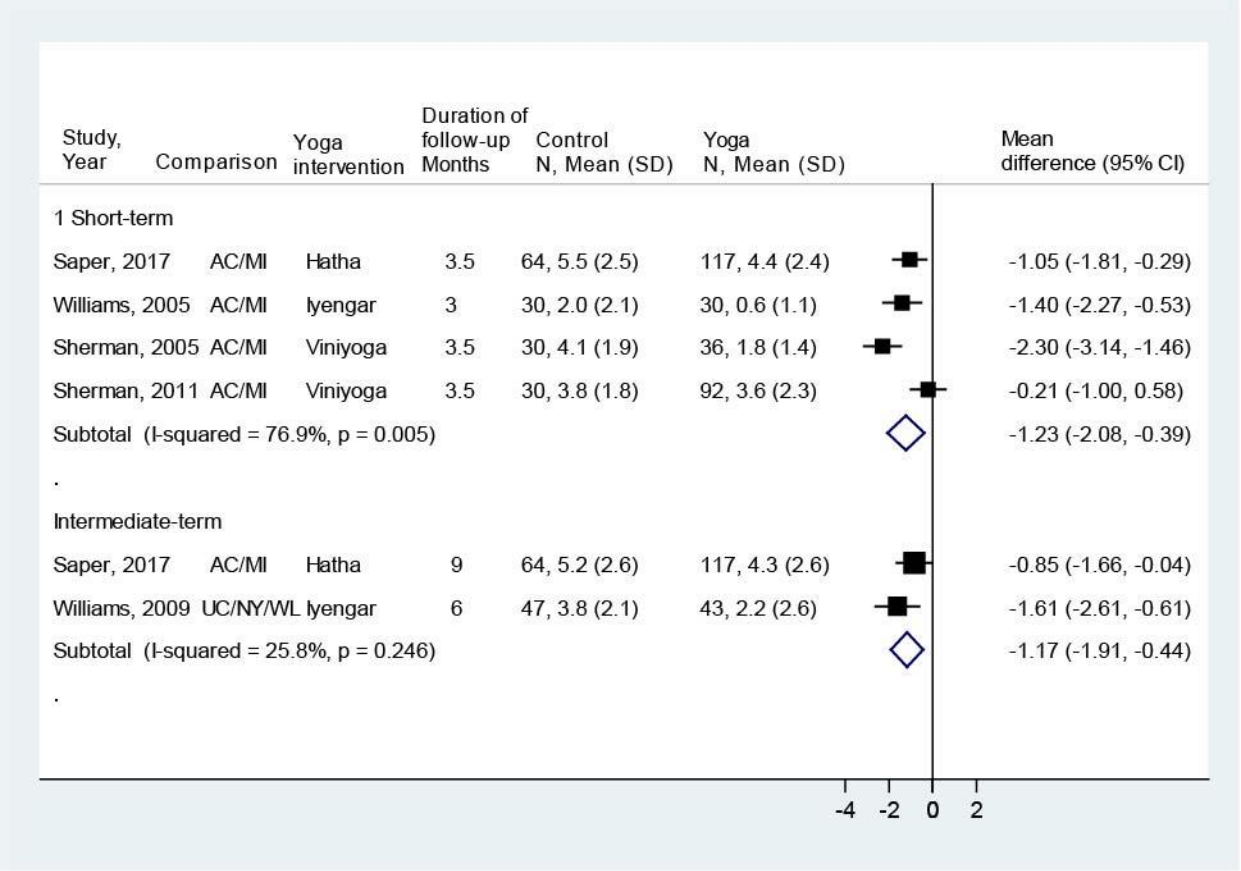
Data on harms were limited, but trials reported no clear difference between yoga versus control interventions in risk of any adverse event.[162,164,165] For serious adverse events, one trial reported a case of cellulitis in a patient randomized to yoga.[162]

Figure 10. Yoga versus attention control or wait list: effects on function



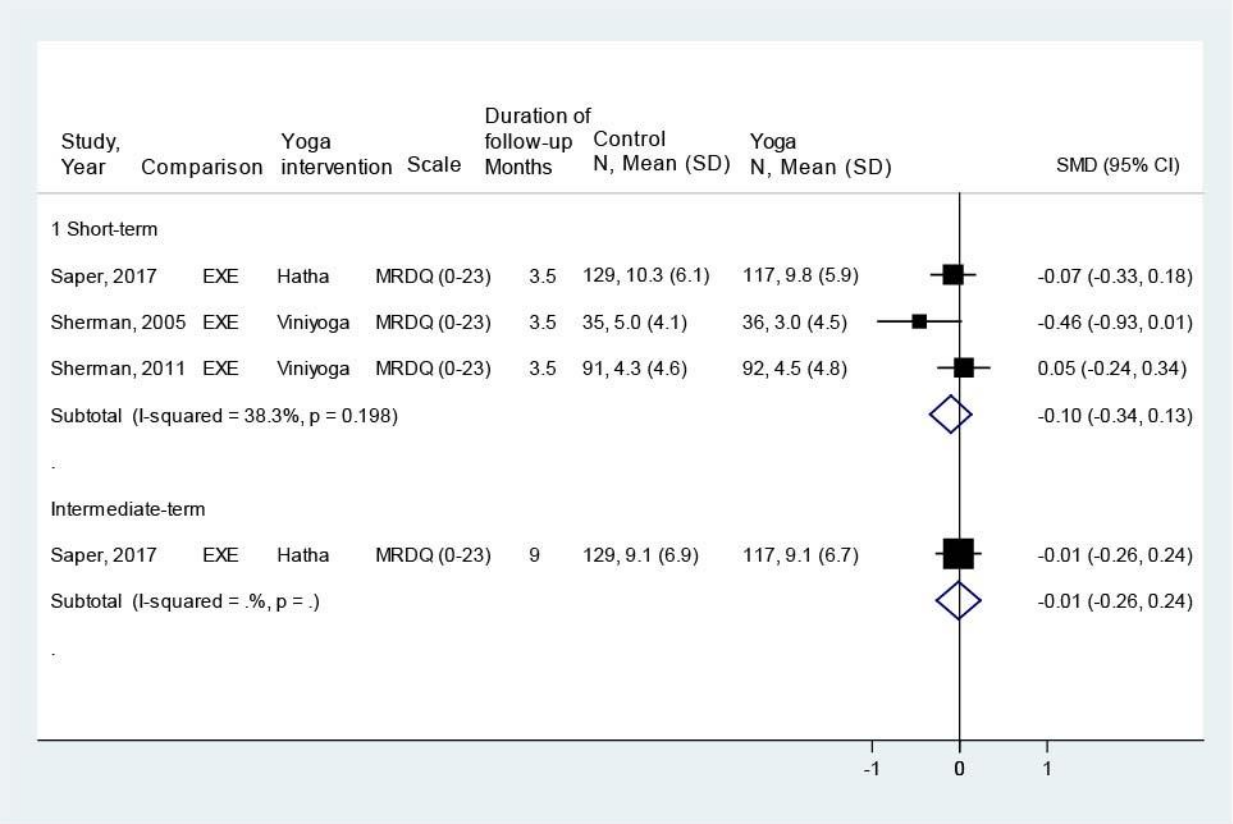
AC/MI = attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; NY/WL = no yoga/waitlist

Figure 11. Yoga versus attention control or wait list: effects on pain



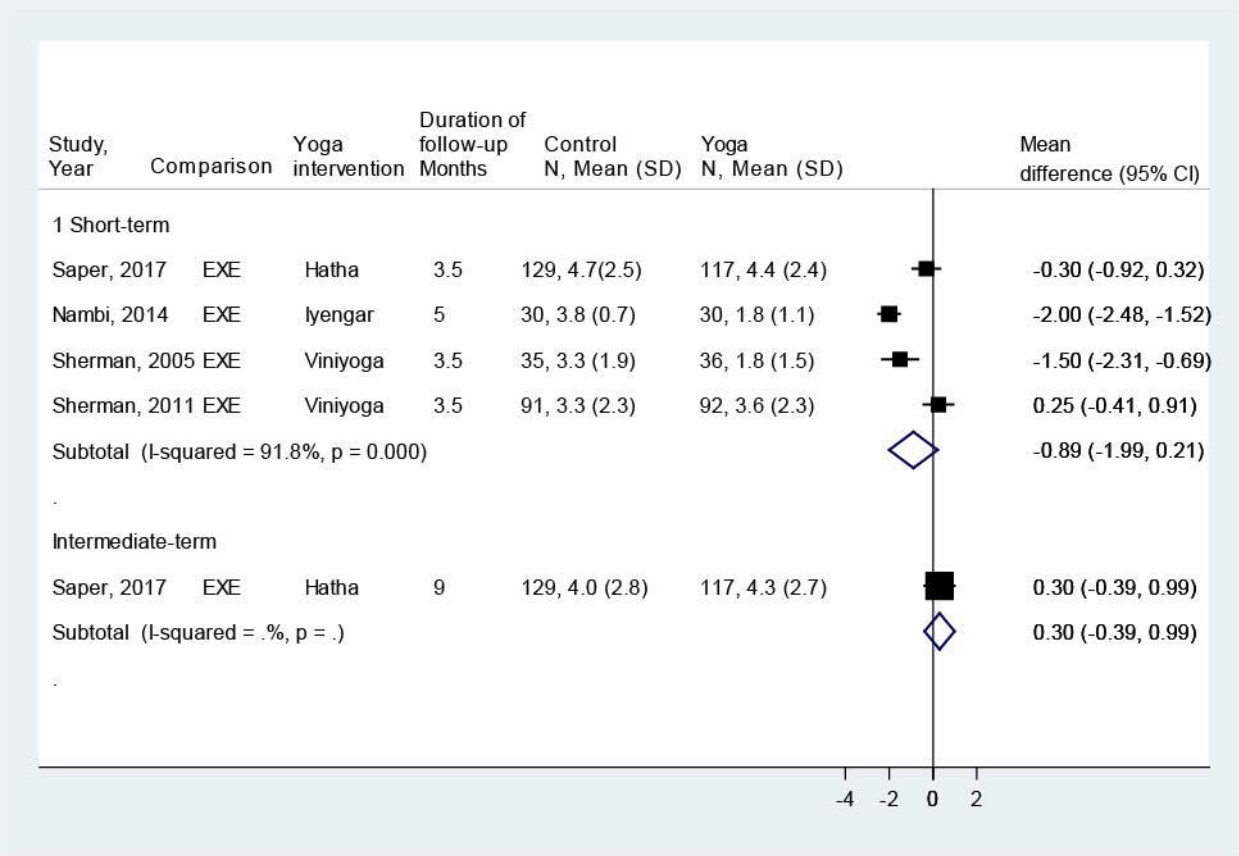
AC/MI = attention control/minimal intervention; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; NY/WL= no yoga/waitlist

Figure 12. Yoga versus exercise: effects on function



CI = confidence interval; EXE=exercise; SD = standard deviation; SMD = standardized mean difference; N = number.

Figure 13. Yoga versus exercise: effects on pain



CI = confidence interval; EXE = exercise; SD = standard deviation; SMD = standardized mean difference; N = number

Low-level Laser Therapy for Low Back Pain

Key Points

- One trial found no differences between low-level laser therapy versus exercise therapy in intermediate-term pain or function (SOE: Low).
- One trial found low-level laser therapy associated with moderately greater effects than sham laser on short-term pain (difference -16.0 on a 0 to 100 scale, 95% CI -28.3 to -3.7) and slightly greater effects on function (difference -8.2 on the 0 to 100 ODI, 95% CI -13.6 to -2.8) (SOE: Low).
- One trial of low-level laser therapy reported no adverse events (SOE: Low).

Detailed Synthesis

Three trials of low-level laser therapy (n=34, 56, and 71) met inclusion criteria (Table 12 and Appendix D).[111,112,136] One trial[112] evaluated neodymium:yttrium-aluminum-garnet (Nd:YAG) laser and two trials[111,136] evaluated gallium-arsenide (GaAs) laser. Two trials compared low-level laser therapy versus sham laser therapy[111,112] and one trial low-level laser therapy versus exercise plus sham laser.[136] One trial was conducted in the United States,[112] one in Iran,[136] and one in Argentina.[111] The duration of laser therapy ranged from 2 to 6 weeks and the number of sessions ranged from 10 to 12. One trial[111] reported intermediate-term outcomes and the other two trials reported short-term outcomes.

Two trials[112,136] were rated fair-quality and one trial[111] poor-quality (Appendix E). The major methodological limitation in the fair-quality trials was unclear allocation concealment methods.[112,136] The poor-quality trial also did not report randomization methods, did not conduct intention-to-treat analysis at intermediate-term followup, and reported high attrition; it was also unclear if timing of followup was the same in all patients.[111]

Table 12. Summary of results for low back pain: physical modalities (low-level laser therapy)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Basford, 1999[112] 2 months Duration of pain: 4.5 vs. 6.5 months *Fair*	A. Nd:YAG laser (542 mW/cm^2, 90 seconds, two sites, applied to eight points along L2 to S3 paraspinal tissues) (n=27) 12 sessions over 4 weeks B. Sham laser (n=29)	A vs. B Age: 48 vs.48 years Female: 40% vs. 55% Baseline maximal pain, last 24 hours (0-100 VAS): 35.2 vs. 37.4 Baseline Oswestry Disability Index: 21 vs. 25	A vs. B <u>2 months</u> Oswestry Disability Index (0-100): 14.7 vs. 22.9, difference -8.2 (95% CI -13.6 to -2.8); p=0.004 Maximal pain in last 24 hours (0-100 VAS): 19.1 vs. 35.1, difference -16.0 (95% CI -28.3 to -3.7); p=0.012	A vs. B <u>2 months</u> Patient perception of benefit (Visual analog scale, lower = less pain): 28.3 vs. 37.8 (95% CI -20.9, 1.9); p=0.101
Djavid, 2007[136] 1.5 months	A. GaAs laser (wavelength 810 nm, 50 mW wave, and 0.2211 cm^2 spot area laser	A vs. B vs. C Age: 40 vs. 38 vs. 36 years Female: 5% vs. 7% vs. 2%	A vs. C <u>1.5 months</u> Pain (0-10 VAS): 4.4 vs. 4.3, difference in change from baseline -0.9 (95% CI -2.5 to	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Duration of pain: 29 months vs. 29 months vs. 25 months *Fair*	applied to 8 points along L2 to S2-S3 paraspinal tissues, dose 27 J/cm^2) (n=16) 12 sessions over 6 weeks B. Low level laser therapy plus exercise (n=19) C. Exercise plus sham laser (strengthening, stretching, mobilizing, coordination) (n=18)	Baseline pain (0-10 VAS): 7.3 vs. 6.3 Baseline Oswestry Disability Index (0-100): 33.0 vs. 31.8	0.7) Oswestry Disability Index (0-100): 20.8 vs. 24.1, difference in change from baseline -4.4 (95% CI -11.4 to 2.5) A vs. B <u>1.5 months</u> Pain (0-10 VAS): 4.4 vs. 2.4, difference in change from baseline -0.9 (95% CI -2.5 to 0.7) Oswestry Disability Index (0-100): 20.8 vs. 16.8 difference in change from baseline -4.4 (95% CI -11.4 to 2.5)	
Soriano, 1998[111] 6 months Duration of pain: greater than 3 months *Poor*	A GaAs laser (wavelength 904 nm, pulse frequency 10,000 Hz, pulse width 200 nsec, peak power 20W, average power 40mW, administered at dose of 4 J/cm^2 per point to pain areas) (n=38) 10 sessions over 5 weeks B. Sham laser (n=33)	A vs. B Age: 63 vs. 64 years Female: 58% vs. 52% Baseline pain (1 to 10): 7.9 vs. 8.1 Baseline function: NR	<u>6 months</u> No pain: 44.7% vs. 15%; p<0.01	Pain recurrence in subgroup of patients with a good or excellent response at end of treatment: 35 % vs. 70%; p=NR

CI =confidence interval; NR = not reported; ODI = Oswestry Disability Index; RDQ, Roland Morris Disability Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Low-level Laser Therapy Compared With Sham Laser

One fair-quality trial found Nd:YAG laser therapy associated with moderately lower pain (difference -16.0 on a 0 to 100 scale, 95% CI -28.3 to -3.7) and slightly better function (difference -8.2 points on the 0 to 100 Oswestry Disability Index, 95% CI -13.6 to -2.8) at short-term followup.(Basford) A poor-quality trial found GaAs laser therapy associated with increased likelihood of having no pain at intermediate-term followup (44.7% vs. 15%, p<0.01), but the analysis was restricted to patients who reported that laser therapy was effective at the end of a two-week course of treatment.[111]

Low-Level Laser Therapy Compared With Pharmacological Therapy

No trial of low-level laser therapy compared with pharmacological therapy met inclusion criteria

Low-Level Laser Therapy Compared With Exercise Therapy

One fair-quality trial found no clear differences between GaAs laser therapy versus exercise plus sham laser in function (difference in change from baseline -4.4 on the 0 to 100 ODI, 95% CI -11.4 to 2.5) or pain (difference in change from baseline -0.9 on a 0 to 10 scale, 95% CI -2.5 to 0.7) at intermediate-term followup.[136] For pain, the difference at followup was similar to the baseline difference (mean 7.3 vs. 6.3), and final scores were very similar (4.4 vs. 4.3)

Harms

No adverse events were reported in any of the three trials of low-level laser therapy.[111,112,136]

Traction for Low Back Pain

Key Points

- Two trials found no differences between traction versus sham traction in short-term pain or function (SOE: Low).
- Harms were not reported in either trial.

Detailed Synthesis

Two trials of traction (n=151 and 60) met inclusion criteria (Table 13 and Appendix D).[107,108] One trial[107] evaluated continuous traction (12 sessions in 5 weeks) and the other[108] evaluated intermittent traction (20 sessions in 6 weeks). The comparator in both trials was sham traction (traction at <10% or 20% of body weight, compared with 35-50% for active traction). Both trials were conducted in the Netherlands and reported only short-term outcomes. The trials were rated fair-quality due to failure to blind care providers (Appendix E).

Table 13. Summary of results for low back pain: physical modalities (traction)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Beurskens, 1997[107] 1.75 and 5 months Duration of pain: 1.5 months *Fair*	A. Continuous traction (n=77) B. Sham traction (20% body weight) (n=74) 12 sessions, 5 weeks	A vs. B Age: 39 vs. 42 years Female: 44% vs. 43% Baseline pain (0-100 VAS): 61 vs. 55 Baseline RDQ (0-24): 2 vs. 12	A vs. B <u>1.75 months</u> Pain at the moment (0-100 VAS): 28.5 vs. 22.8, difference 5.7 (95% CI -4.6 to 15.9) RDQ: 4.4 vs. 4.3, difference 0.1 (95% CI -1.8 to 1.9) <u>5 months</u> Pain at the moment (0-100 VAS): 23.8 vs. 20.1, difference 3.7 (95% CI-8.4 to 15.8) RDQ: 4.7 vs. 4.0, difference 0.7 (95% CI -1.1 to 2.6)	A vs. B <u>1.75 months</u> ADL disability (0 to 100 VAS): 27.1 vs. 29.4, difference -2.4 (95% CI -13.6 to 8.9) Work absence (days): 23.5 vs. 27.8, difference -4.3 (95% CI -14.7 to 6.1) Medical consumption: 34% vs. 25%, difference 9% (95% CI -6% to 24%) <u>5 months</u> ADL disability: 25.7 vs.

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				25.8, difference 0.1 (95% CI -11.5.0 to 11.2) Work absence (days): 35.7 vs. 43.7, difference -8.0 (95% CI -27 to 11) Medical consumption: 45% vs. 42%, difference 3% (95% CI -13% to 19%)
Schimmel, 2009[108] 2 months Duration of pain: 1 year *Fair*	A. Intermittent traction (n=31) B. Sham traction (<10% body weight) (n=29) 20 sessions, 6 weeks	A vs. B Age (mean): 42 vs. 46 years Female: 39% vs. 52% Baseline back pain (0-100 VAS): 61 vs. 53 Baseline ODI: 36 vs. 33	A vs. B <u>2 months</u> Pain (0-100 VAS): 32 vs. 36; p=0.70 ODI (0-100): 25 vs. 23 (SD, p not reported)	A vs. B <u>2 months</u> SF-36, total (0-100): 66 vs. 65 (SD, p not reported)

CI= confidence interval; SF-36 =Short-Form 36 Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Traction Compared With Sham Traction

There were no differences between traction versus sham traction at short-term followup in function (25 vs. 23 on the 0 to 100 ODI in one trial and 4.7 vs. 4.0 on the 0 to 24 RDQ, difference 0.7, 95% CI -1.1 to 2.6, in the other) or pain (32 vs. 36 on a 0 to 100 scale, p=0.70 and 24 vs. 20, difference 3.7 [95% CI -8.4 to 15.8]).[107,108] One trial[108] also found no difference between intermittent traction versus sham on the total SF-36 (66 vs. 65 on a 0 to 100 scale) and one trial[107] found no difference between continuous traction versus sham in global perceived effect, work absence, or medical consumption.

Traction Compared With Pharmacological Therapy or With Exercise

No trial of low-level laser therapy compared with pharmacological therapy or with exercise met inclusion criteria.

Harms

Neither trial reported harms.

Mindfulness-Based Stress Reduction for Low Back Pain

Key Points

- There were no differences between mindfulness-based stress reduction (MBSR) versus usual care or an attention control in short-term function (3 trials, pooled SMD -0.22, 95% CI -0.53 to 0.10, I2=63%) or intermediate-term function (1 trial, SMD -0.20, 95% CI -0.47 to 0.06) (SOE: low).
- MBSR was associated with moderately greater effects than usual care or an attention control on short-term pain (4 trials, pooled difference -1.18, 95% CI -2.14 to -0.22, I2=93%), although the effect was small when a poor-quality trial was excluded (3 trials, pooled difference -0.73, 95% CI -1.18 to -0.28; I2=45%); MBSR was also associated with slightly greater effects on intermediate-term pain (1 trial, difference -0.75, 95% CI -1.17 to -0.33) (SOE: Moderate for short-term, Low for intermediate-term).
- One trial reported temporarily increased pain in 29 percent of patients undergoing MBSR, and two trials reported no harms (SOE: Low).

Detailed Synthesis

Four trials of MBSR for low back pain met inclusion criteria (Table 14 and Appendix D).[84,156-158] In all trials, the MBSR intervention was modeled on the program developed by Kabat-Zinn,[229] with the main intervention consisting of 1.5 to 2 hour weekly group sessions for 8 weeks. Sample sizes ranged from 35 to 282 (total sample=590). Two trials compared MBSR versus usual care[84,156] and two trials compared MBSR versus an attention control (education).[157,158] Three trials[84,157,158] were conducted in the United States. and one trial[156] in Iran. One trial reported outcomes through intermediate-term followup[84] and the others only evaluated short-term outcomes.

Three trials[84,157,158] were rated fair-quality and one trial poor-quality (Appendix E).[156] The major methodological limitation in the fair-quality trials was the inability to effectively blind patients and caregivers to the MBSR intervention. The poor-quality trial reported unclear randomization and allocation concealment methods and high attrition.[156]

Table 14. Summary of results for low back pain: mindfulness-based stress reduction

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Banth, 2015[156] 1 month Duration of pain: ≥6 months *Poor*	A. Mindfulness-based stress reduction (n=NR) 8 1.5-hour sessions over 8 weeks B. Usual care (n=NR) 48 of 88 patients were analyzed, n for each group NR	A vs. B (NR) Age: 40 years Female: 100% McGill Pain questionnaire total score (0-45): 26.08 vs. 26.71 Baseline function: NR	A vs. B <u>1 month</u> McGill Pain questionnaire total score (0-45): 13.58 vs. 23.60	A vs. B <u>1 month</u> SF-12 Mental component (0-100): 31.54 (4.3) vs. 24.29 (5.2) SF-12 Physical component (0-100): 28.08 (4.2) vs. 21.08 (3.3)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Cherkin, 2016[84] 10 months Duration of pain: >3 months *Fair*	A. Mindfulness-based stress reduction (n=113), 8 2-hour sessions over 8 weeks (optional 6 hour retreat) B. Usual care (n=112)	A vs. B 50 vs. 49 years Female: 61% vs. 66% Baseline pain bothersomeness (0-10): 6.1 vs. 6.0 Baseline modified RDQ (0-23): 11.8 vs. 10.9	A vs. B <u>4.5 months</u> Modified RDQ (0-23), mean change from baseline: -4.33 (95% CI -5.16, -3.51) vs. -2.96 (95% CI -3.79, -2.14) Pain bothersomeness (0-10), mean change from baseline: -1.48 (95% CI -1.86, -1.11) vs. -0.84 (95% CI -1.21, -0.46) ≥30% improvement in RMDQ: 60.5% (95% CI 52.0, 70.3) vs. 44.1% (95% CI 35.9, 54.2) ≥30% improvement in pain bothersomeness: 43.6% (95% CI 35.6, 53.3) vs. 26.6% (95% CI 19.8, 35.9) <u>10 months</u> Modified RDQ, mean change from baseline:−5.3 (95% CI −6.16, −4.43) vs. −4.78 (95% CI −5.67, −3.89) vs. −3.43 (95% CI −4.33, −2.52) Pain bothersomeness, mean change from baseline: -1.95 (95% CI -2.32, -1.59) vs. -1.10 (95% CI -1.48, -0.71) ≥30% improvement in RMDQ: 68.6% (95% CI 60.3, 78.1) vs. 48.6% (95% CI 40.3, 58.6) ≥30% improvement in pain bothersomeness: 48.5% (95% CI 40.3, 58.3) vs. 31.0% (95% CI 23.8, 40.3)	A vs. B <u>4.5 months</u> SF-12 MCS, mean change from baseline (0-100): 0.45 (95% CI -0.85, 1.76) vs. 2.13 (95% CI 0.86, 3.40) vs. −1.11 (95% CI −2.39, 0.17) SF-12 PCS, mean change from baseline (0-100): 3.58 (95% CI 2.15, 5.01) vs. 3.27 (95% CI 2.09, 4.44) Used medications for LBP: 43.4% (95% CI 35.9, 52.6) vs. 54.2 (95% CI 46.2, 63.6) <u>10 months</u> SF-12 MCS, mean change from baseline: 2.01 (95% CI 0.74, 3.28) vs. 0.75 (95% CI −0.58, 2.08) SF-12 PCS, mean change from baseline: 3.87 (95% CI 2.55, 5.19) vs. 2.93 (95% CI 1.70, 4.16) Used medications for LBP: 46.8% (95% CI 39.2, 55.9) vs. vs. 52.9% (95% CI 45.1, 62.0)
Morone, 2009[158] 4 months Duration of pain: Mean 9.4 to 11 years *Fair*	A. Mindfulness-based stress reduction (n=16), 8 1.5-hour sessions over 8 weeks B. Attention control (education) (n=19)	A vs. B Age 78 vs. 73 years Female: 69% vs. 58% Baseline McGill Pain Questionnaire Current Pain (0-10): 2.9 vs. 4.4 Baseline RDQ: 8.8 vs. 11.3	A vs. B <u>4 months</u> RDQ: 7.6 (95% CI 6.2 to 8.7) vs. 10.0 (95% CI 8.7 to 11.2) SF-36 Pain Score (10-62): 41.4 (95% CI 39.8 to 43.1) vs. 40.5 (95% CI 38.7 to 42.2) McGill Pain Questionnaire Total Score (0-45): 12.4 (95% CI 10.4 to 14.6) vs. 12.0 (95% CI 10.2 to 13.7) McGill Pain Questionnaire Current Pain (0-10): 2.3 (95% CI 1.6 to 2.8) vs. 3.7 (95% CI 3.1 to 4.3)	NR
Morone, 2016[157] 4.5 months Duration of pain: Mean 11	A. Mindfulness-based stress reduction (n=140), 8 1.5-hour sessions over 8 weeks, with 6	A vs. B Age: 75 vs. 74 years Female: 66% vs. 66% Pain (0-20	A vs. B <u>4.5 months</u> Pain (0-20 NRS): 9.5 vs. 10.6, adjusted difference -1.1 (95% CI -2.2 to -0.01) RDQ: 12.2 vs. 12.6, adjusted	A vs. B <u>4.5 months</u> SF-36 Global Health Composite (9-67): 42.4 vs. 41.2, adjusted difference 0.2 (95% CI -

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
years *Fair*	monthly booster sessions B. Control, (health education) (n=142)	NRS): 11.0 vs. 10.5 RDQ (0-24): 15.6 vs. 15.4	difference -0.4 (95% CI -1.5 to 0.7) RDQ improved ≥2.5 points: 49.2% (58/117) vs. 48.9% (66/135), p=0.97 Pain improved ≥30%: 36.7% (43/117) vs. 26.7% (36/135), p=0.09	1.9 to 2.4) SF-36 Physical Health Composite (20 to 65): 41.2 vs. 41.2, adjusted difference -0.1 (95% CI -1.9 to 1.8)

CI = confidence interval; NR = not reported; MCS = Mental Component Summary; ODI = Oswestry Disability Index; PCS = Physical Component Summary; RDQ = Roland Morris Disability Questionnaire; SF-36 = Short-Form 36 Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

MBSR Compared With Usual Care or an Attention Control

MBSR was associated with no statistically significant differences in short-term function compared to usual care or an attention control (3 trials, pooled SMD-0.22 on the RDQ, 95% CI -0.53 to 0.10, I2=63%) (Figure 14).[84,157,158] The difference on the RDQ was less than 1 point (pooled difference -0.95 points on a 0 to 24 scale, 95% CI -2.07 to 0.17). One trial found no difference between MBSR versus an attention control in intermediate-term function (SMD -0.20, 95% CI -0.47 to 0.06).[84]

MBSR was associated with moderately greater effects than usual care or an attention control on short-term pain (4 trials, pooled difference -1.18, 95% CI -2.14 to -0.22, I2=93%) (Figure 15).[84,156-158] Although statistical heterogeneity was substantial, all estimates favored MBSR, with differences ranging from -0.54 to -2.23 points. Excluding a poor-quality trial,[156] which also reported the largest effect, resulted in an attenuated estimate (3 trials, pooled difference -0.73, 95% CI -1.18 to -0.28, I2=45%). Estimates were similar when analyses were stratified according to whether the trial evaluated usual care or an attention control comparator. One trial found MBSR associated with slightly greater effects than an attention control on intermediate-term pain (difference -0.75, 95% CI -1.17 to -0.33).[84]

Three trials found no clear differences between MBSR versus usual care or an attention control on quality of life measured by the SF-12 or SF-36.[84,156,157] One trial found MBSR associated with less medication use for low back pain at short-term (43% vs. 54%) but not at intermediate-term (47% vs. 53%); MBSR was associated with slightly greater decrease in severity of depression (difference 0.63 points on the PHQ-8 at intermediate-term).[84]

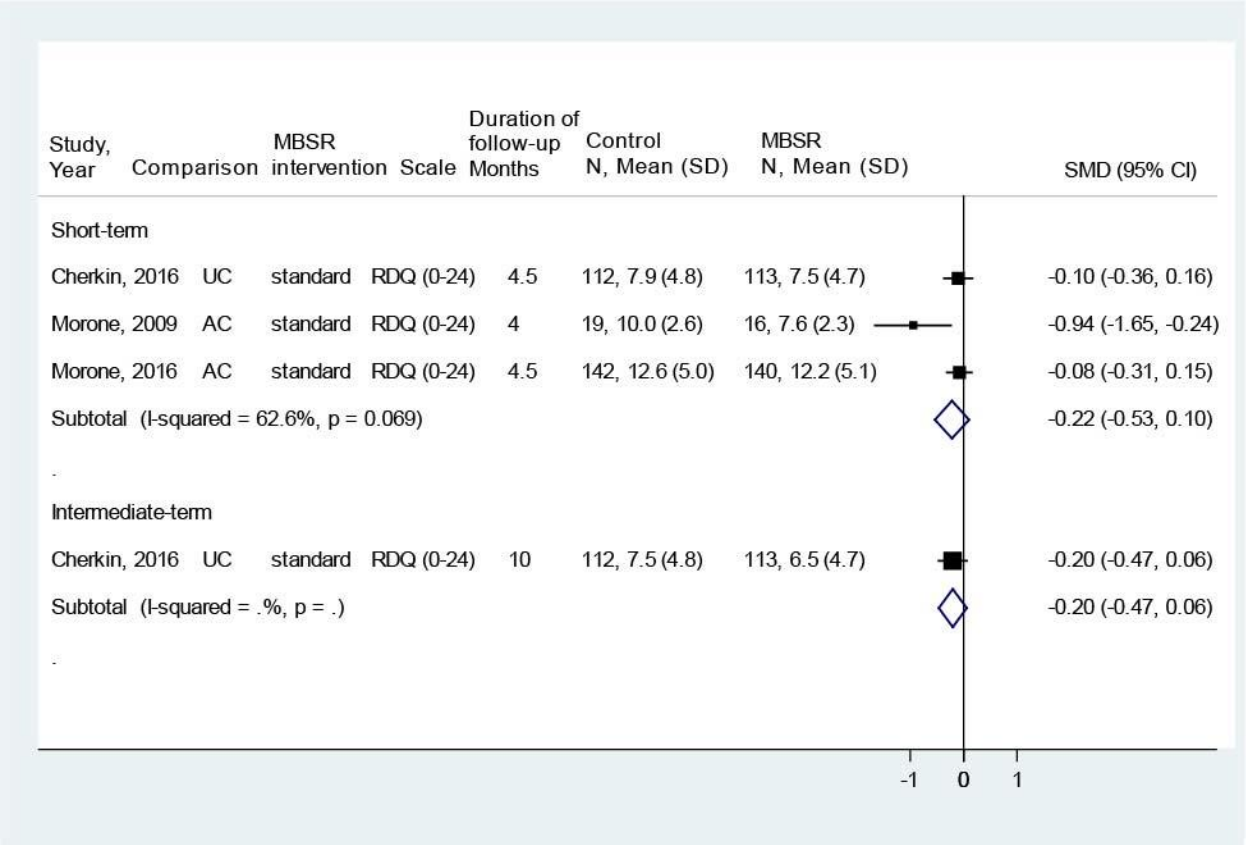
MBSR Compared With Pharmacological Therapy or With Exercise

No trial of MBSR versus pharmacological or versus exercise therapy met inclusion criteria.

Harms

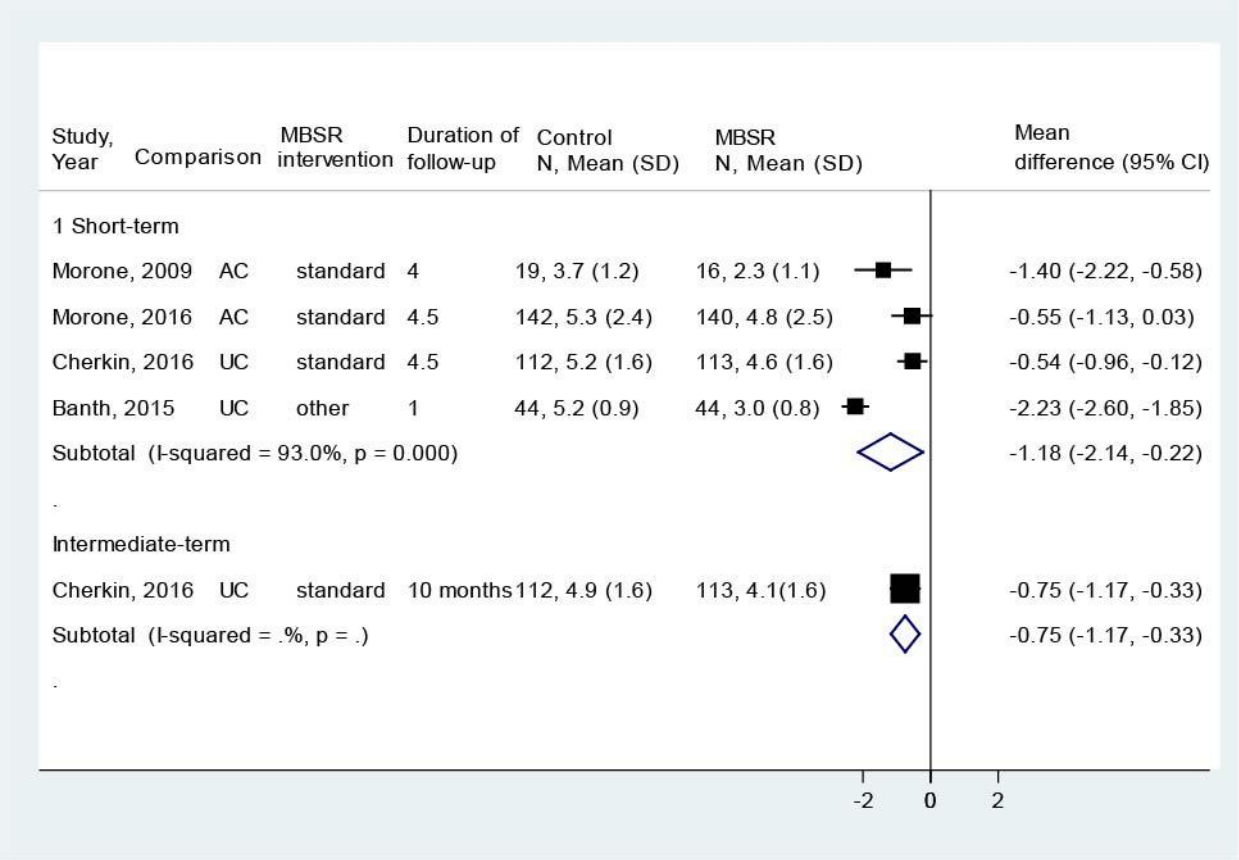
In one trial, 29 percent of MBSR patients reported temporarily increased pain.[84] Two trials[157,158] reported no adverse events and one trial[156] did not report adverse events.

Figure 14. Mindfulness-based stress reduction versus usual care or an attention control: effects on function



AC = attention control; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; UC = usual care

Figure 15. Mindfulness-based stress reduction versus usual care or an attention control: effects on pain



AC = attention control; CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number; UC = usual care

Spinal Manipulation for Low Back Pain

Key Points

- There was no difference between spinal manipulation versus sham manipulation, usual care, an attention control or a placebo intervention in short-term pain (3 trials, pooled difference -0.20 on a 0 to 10 scale, 95% CI -0.66 to 0.26, I2=58%), but manipulation was associated with slightly greater effects than controls on intermediate-term pain (3 trials, pooled difference -0.64, 95% CI -0.92 to -0.36, I2=0%) (SOE: Low for short-term, Moderate for intermediate-term).
- Spinal manipulation was associated with slightly greater effects than sham manipulation, usual care, an attention control, or a placebo intervention in short-term function (3 trials, pooled SMD -0.34, 95% CI -0.63 to -0.05, I2=61%) and intermediate-term function (3 trials, pooled SMD -0.40, 95% CI -0.69 to -0.11, I2=76%) (SOE: Low)
- There were no differences between spinal manipulation versus exercise in short-term pain (3 trials, pooled difference 0.31 on a 0 to 10 scale, 95% CI -0.30 to 0.92; I2=60%) or

intermediate-term pain (4 trials, pooled difference 0.22, 95% CI -0.09 to 0.52, I2=9.4%) (SOE: Low).

- There were no differences between spinal manipulation versus exercise in short-term function (3 trials, pooled SMD 0.01, 95% CI -0.22 to 0.25; I2=62%) or intermediate-term function (4 trials, pooled SMD 0.02, 95% CI -0.13 to 0.18; I2=48%) (SOE: Low).
- No serious adverse events or withdrawals due to adverse events were reported in 7 trials; nonserious adverse events with manipulation (primarily increased pain) were reported in 3 trials (SOE: Low).

Detailed Synthesis

Eight trials of spinal manipulation for low back pain met inclusion criteria (Table 15 and Appendix D).[113,137-140,152-154] All of the trials evaluated standard (high-velocity low-amplitude) manipulation techniques; one trial[154] evaluated flexion-distraction manipulation and one trial[138] evaluated both high-velocity low-amplitude and flexion-distraction manipulation. Sample sizes ranged from 75 to 1,001 (total sample=2,586). The number of manipulation therapy sessions ranged from 4 to 24 and the duration of therapy ranged from 4 to 12 weeks. In one trial, patients were randomized to 12 manipulation sessions over 1 month or to 12 sessions over month plus biweekly maintenance sessions for an additional 10 months.[139] Two trials compared spinal manipulation versus usual care,[138,140] one trial spinal manipulation versus an attention control (minimal massage),[137] one trial spinal manipulation versus sham manipulation,[139] one trial spinal manipulation versus a placebo treatment (sham short-wave diathermy),[113] and four trials spinal manipulation versus exercise.[140,152-154] One trial was conducted in Egypt[139] and the rest in the United States, United Kingdom, or Australia. Six trials reported outcomes through intermediate-term followup,[137,139,140,152-154] and two trials only evaluated short-term outcomes.[113,138]

Two trials[113,139] were rated poor quality and the remainder fair quality (Appendix E). The major methodological limitation in the fair-quality trials was use of an unblinded design. Methodological shortcomings in the poor-quality trials included unclear randomization and allocation concealment methods, failure to report intention-to-treat analysis, and high attrition.

Table 15. Summary of results for low back pain: manual therapies (spinal manipulation)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Bronfort, 2011[152] 9 months Duration of pain: 5 years *Fair*	A. Standard manipulation (n=100), 12-24 sessions over 12 weeks B. Exercise (supervised) (n=100) C. Exercise (home) (n=101)	A vs. B Age: 45.2 vs. 44.5 vs. 45.6 years Female sex: 67% vs. 57% vs. 58% Baseline pain (0-10 NRS): 5.4 vs. 5.1 vs. 5.2 Baseline Modified RDQ (0-23): 8.7 vs. 8.4 vs. 8.7	A vs. B <u>4 months</u> Pain (0-10 NRS): 3.3 vs. 2.9 vs. 3.1, adjusted difference 0.3 (95% CI -0.5 to 1.0) for A vs. B and 0.1 (95% CI -0.6 to 0.9) for A vs. C Modified RDQ (0-23): 4.9 vs. 4.0 vs. 4.2, adjusted difference 0.5 (95% CI -1.0 to 2.1) for A vs. B and 0.7 (95% CI -0.9 to 2.3) for A vs. C <u>9 months</u>	A vs. B <u>4 months</u> SF-36 PCS (norm-based mean=50): 48.6 vs. 50.6 vs. 49.1, adjusted difference -1.8 (95% CI -4.4 to 0.9) for A vs. B and -0.3 (95% CI -3.0 to 2.4) for A vs.C SF-36 MCS (norm-based mean=50): 55.9 vs. 54.8 vs. 55.1, adjusted difference 0.4 (95% CI -2.0 to 2.9) for A vs. B and -0.5 (95% CI -3.0 to 2.1) for A vs.C OTC pain medication use, past week (days): 1.6 vs.

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			Pain (0-10 NRS): 3.3 vs. 2.8 vs. 2.8, adjusted difference 0.3 (95% CI -0.5 to 1.1) for A vs. B and 0.3 (95% CI -0.6 to 1.1) for A vs. C Modified RDQ (0-23): 5.1 vs. 3.8 vs. 4.1, adjusted difference 0.4 (95% CI -1.2 to 2.0) for A vs. B and -0.1 (95% CI -0.7 to 0.5) for A vs. C	1.4 vs. 1.5, adjusted difference 0.4 (95% CI -0.4 to 1.1) for A vs. B and 0.4 (95% CI -0.3 to 1.2) for A vs. C <u>9 months</u> SF-36 PCS (norm-based mean=50): 48.4 vs. 50.4 vs. 49.6, adjusted difference -1.7 (95% CI -4.2 to 0.8) for A vs. B and -1.0 (95% CI -3.5 to 1.5) for A vs.C SF-36 MCS (norm-based mean=50): 55.2 vs. 53.9 (8.6) vs. 56.0, adjusted difference 2.4 (95% CI -0.2 to 5.0) for A vs. B and -2.2 (95% CI -4.9 to 0.5) for A vs. C OTC pain medication use, past week (days): 1.8 vs. 1.8 vs. 1.6, adjusted difference 0.1 (95% CI -0.8 to 0.9) for A vs. B and 0.4 (95% CI -0.4 to 1.3) for A vs. C
Ferreira, 2007[153] 10 months Duration of pain: Not reported *Fair*	A. Standard manipulation and mobilization (n=80), 12 sessions over 8 weeks B. Exercise (motor control) (n=80) C: Exercise (general exercise) (n=80)	A vs. B vs. C Age: 54 vs. 52 vs. 55 years Female: 70 % vs. 66% vs. 70% Baseline pain (0-10 VAS): 6.2 vs. 6.3 vs. 6.5 Baseline RDQ (0-24): 12.4 vs. 14.0 vs. 14.1	A vs. B vs. C <u>4 months</u> Pain (0-10 VAS): 4.3 vs. 4.3 vs. 4.8, difference 0.0 (95% CI -0.9 to 0.8) for A vs. B and -0.5 (95% CI -1.4 to 0.3) for A vs. C RDQ (0-24): 7.7 vs. 8.4 vs. 10.1, difference 0.2 (95% CI -1.5 to 1.9) for A vs. B and -0.9 (95% CI -2.7 to 0.9) for A vs. C <u>10 months</u> Pain (0-10 VAS): 4.9 vs. 4.9 vs. 5.2, difference 0.1 (95% CI -0.8 to 1.0) for A vs. B and -0.2 (95% CI -1.1 to 0.6) for A vs. C RDQ (0-24): 9.2 vs. 8.8 vs. 9.6, difference 1.8 (95% CI 0.0 to 3.6) for A vs. B and 1.2 (95% CI -0.6 to 3.0) for A vs. C	A vs. B vs. C <u>4 months</u> Patient Specific Functional Scale (3-30): 17.3 vs. 16.4 vs. 15.0, difference 0.7 (95% CI -1.3 to 2.7) for A vs. B and 1.7 (95% CI -0.4 to 3.,8) for A vs. C <u>10 months</u> Patient Specific Functional Scale (3-30): 15.2 vs. 15.7 (6.8) vs. 13.9 , difference -0.8 (95% CI -2.9 to 1.2) for A vs. B and 0.3 (95% CI -1.7 to 2.3) for A vs. C

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Gibson, 1985[113] 2 months Duration of pain: 2 to 12 months *Poor*	A. Manipulation (technique unclear) and mobilization (n=41), 4 sessions over 4 weeks B. Placebo (detuned short-wave diathermy) (n=34)	A vs. B 34 vs. 40 years Female: 61% vs. 32% Baseline pain (0-100 VAS): 35 vs. vs. 48	A vs. B <u>1 month</u> Pain (median [range], 0-100 VAS): 28 (0-96) vs. 27(0-80) <u>3 months</u> Pain (median [range], 0-100 VAS): 25 (4-90) vs. 6 (10-96) p<0.01	A vs. B <u>1 month</u> Using analgesics: 25% vs. 50% Using analgesics: 18% vs. 22% <u>3 months</u> Using analgesics: 25% vs. 50% Using analgesics: 18% vs. 22%
Gudavalli, 2006[154] 11 months Duration of pain: >3 months *Fair*	A. Flexion–distraction manipulation (n=123), 8-16 sessions over 4 weeks B. Exercise (n=112)	A vs. B Age: 42 vs. 41 years Female: 34% vs. 41% Baseline pain VAS (0-100: 38.00 vs. 35.70 Baseline RDQ (0-24): 6.64 vs. 6.84	A vs. B <u>2 months</u> Pain (0-100 VAS): 16.52 (SD 2.95) vs.12.04 (SD 2.53) RDQ (0-24): 3.50 (SD 0.50) vs. 3.75 (SD 0.51) <u>5 months</u> Pain (0-100 VAS): 18.26 (SD 2.64) vs. 8.92 (SD 2.89) RDQ (0-24): 3.89 (SD 0.46) vs. 3.42 (SD 0.50) <u>11 months</u> Pain (0-100 VAS): 17.10 (SD 2.55) vs. 12.36 (SD 2.43) RDQ (0-24): 3.90 (SD 0.53) vs. 3.77 (SD 0.44)	
Haas, 2014[137] 10.5 months Duration of pain: 11 to 12 years *Fair*	A. Standard spinal manipulation (n=100), 6 sessions over 6 weeks B. Standard manipulation (n=100), 12 sessions over 6 weeks C. Standard manipulation (n=100), 18 sessions over 6 weeks D: Attention control (minimal massage) (n=100)	A vs. B vs. C vs. D Age: 41 vs. 42 vs. 41 vs. 41 Female: 49% vs. 49% vs. 52% vs. 49% Baseline Pain (0–100 VAS): 51.0 vs. 51.6 vs. 51. vs. 52.2 Baseline Von Korff pain intensity (0–100): 51.0 vs. 51.6 vs. 51.5 vs. 52.2 Baseline Modified Von Korff functional disability (0–100): 44.8 vs.46.1 vs.45.2 vs. 45.2	A vs. B <u>4 months</u> Von Korff pain intensity (0-100): 32.5 vs. 33.7 vs. 32.1 vs. 34.9, adjusted difference -1.7 (95% CI -6.9 to 3.4) for A vs. D, -0.8 (95% CI -6.0 to 4.4) for B vs. D, and -2.4 (95% CI -7.6 to 2.9) for C vs. D Von Korff functional disability (0-100): 25.6 vs. 24.0 vs. 24.1 vs. 27.1, adjusted difference -1.4 (95% CI -7.2 to 4.5) for A vs. D, -3.4 (95% CI -9.3 to 2.4) for B vs. D, and -2.9 (95% CI -8.8 to 2.9) for C vs. D <u>10.5 months</u> Von Korff pain intensity (0-100): 30.7 vs. 31.9 (vs. 28.7 vs. 36.5, adjusted difference -5.4 (95% CI -11.1 to 0.4)	A vs. B <u>4 months</u> Von Korff functional disability improved >=50%: 51.5% vs. 59.8% vs. 54.0% vs. 49.5%, adjusted difference 2.5% (95% CI -11.5 to 16.5%) for A vs. D, 10.4% (95% CI -3.4 to 24.3%) for B vs. D, and 4.8% (95% CI -9.1 to 18.6%) for C vs. D SF-12 PCS (norm-based mean=50): 50.5 vs. 51.4 vs. 50.9 vs. 50.0, adjusted difference 0.0 (95% CI -2.4 to 2.3) for A vs. D, -0.8 (95% CI -3.2 to 1.6) for B vs. C, and -1.3 (95% CI -3.6 to 1.1) for C vs. D SF-12 MCS (norm-based mean=50): 52.8 vs. 50.8 vs. 51.3 vs. 51.8, adjusted difference -2.1 (95% CI -4.2

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			for A vs. D, -4.6 (95% CI -10.3 to 1.2) for B vs. D, and -7.6 (95% CI -13.2 to -2.0) for C vs. D Von Korff functional disability (0-100): 22.6 vs. 22.4 vs. 19.1 vs. 28.0, adjusted difference -5.2 (95% CI -10.9 to 0.5) for A vs. D, -5.9 (95% CI -11.8 to -0.1) for B vs. D, and -8.8 (95% CI -14.4 to -3.3) for C vs. D	to 0.0) for A vs. D, -0.7 (95% CI -2.8 to 1.3) for B vs. D, and -0.1 (95% CI -2.2 to 2.1) for C vs. D EuroQoL (0-100): 77.8 vs. 77.0 vs. 74.5 vs. 73.9, difference -2.9 (95% CI -6.9 to 1.0) for A vs. D, -1.4 (95% CI -5.5 to 2.6) for B vs. D, and -1.5 (95% CI -5.8 to 2.7) for C vs. D <u>10.5 months</u> Von Korff functional disability improved >=50%: 57.6% vs. 57.7% vs. 62.0% vs. 58.9%, adjusted difference -1.1% (95% CI -14.8 to 12.6%) for A vs. D, -1.4% (95% CI -15.4 to 12.6%) for B vs. D, and 2.7% (95% CI -11.0 to 16.5%) for C vs. D SF-12 PCS (norm-based mean=50): 50.8 vs. 52.6 vs. 52.5 vs. 50.7, adjusted difference -0.3 (95% CI -2.1 to 2.7) for A vs. D, -1.4 (95% CI -4.0 to 1.2) for B vs. D, and -2.2 (95% CI -4.5 to 0.2) for C vs. D SF-12 MCS (norm-based mean=50): 50.4 vs. 50.6 vs. 50.4 vs. 51.3, adjusted difference -0.2 (95% CI -2.7 to 2.3) for A vs. D, -1.1 (95% CI -3.7 to 1.6) for B vs. D, and 0.3 (95% CI -2.3 to 2.9) for C vs. D EuroQoL (0-100): 77.1 vs. 77.3 vs. 77.2 vs. 74.8, adjusted difference -1.3 (95% CI -5.4 to 2.7) for A vs. D, -0.9 (95% CI -4.9 to 3.1) for B vs. D, and -3.3 (95% CI -7.2 to 0.5) for C vs. D
Hondras, 2009[138] 4.5 months Duration of pain: Mean 9 to 13 years *Fair*	A. Standard manipulation (n=96), 12 sessions over 6 weeks B. Flexion distraction manipulation (n=95), 12 sessions	A vs. B vs. C Age: 64 vs. 62 vs. 63 years Female: 45% vs. 44% vs. 41% Baseline pain (0-100 VAS): 42.1 (23.6) vs.	<u>1.5 months</u> RDQ (0-24): adjusted difference -1.5 (95% CI -3.1 to 0.1) for A vs. C and -2.2 (95% CI -3.7 to -0.6) for B vs. C Global improvement from baseline (1-10): adjusted difference 1.3 (95% CI 0.2	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	over 6 weeks C: Usual care (n=49)	42.5 (25.2) vs. 42.4 (24.5) Baseline RDQ (0-24), mean (SD): 6.5 vs. 6.6 vs. 5.7	to 2.3) for A vs. C and 1.6 (95% CI 0.5 to 2.7) for B vs. C <u>4.5 months</u> RDQ (0-24): adjusted difference -1.3 (95% CI −2.9 to 0.6) for A vs. C and -1.9 (95% CI -3.6 to -0.2) for B vs. C Global improvement from baseline (1-10): adjusted difference 1.7 (95% CI 0.5 to 2.8) for A vs. C and 1.8 (95% CI 0.6 to 3.0) for B vs. C	
Senna, 2011[139] 9 months Duration of pain: 18-19 months *Poor*	A. Standard manipulation (n=27), 12 sessions over 4 weeks B. Standard manipulation (n=27) , 12 sessions over 4 weeks, then every 2 weeks for 9 months C. Sham manipulation (n=40)	A vs. B Age: 40 vs. 42 vs. 42 years Female: 27% vs. 24% vs. 24% Baseline pain (0-100 VAS): 42 vs. 43 vs. 41 Baseline function (0-100 ODI): 39 vs. 40 vs. 38	A vs. B <u>3 months</u> ODI (0-100): 29.8 vs. 23.1 vs. 33.5; p>0.05 Pain (0-100 VAS): 35.2 vs. 25.9 vs. 35.2; p>0.05 <u>6 months</u> ODI (0-100): 32.2 vs. 22.4 vs. 35.3; p>0.05 Pain (0-100 VAS): 35.5 vs. 25.4 vs. 36.8; p>0.05 <u>9 months</u> ODI (0-100): 34.9 vs. 20.6 vs. 37.4 Pain (0-100 VAS): 38.5 vs. 23.5 vs. 38.3	A vs. B <u>3 months</u> SF-36, total (0-100): 29.2 vs. 32.8 vs. 26.4; p>0.05 <u>6 months</u> SF-36, total (0-100): 27.8 vs. 33.1 vs. 26.1; p>0.05 <u>9 months</u> SF-36, total (0-100): 27.6 vs. 33.70 vs. 25.9; p>0.05
UK BEAM Trial Team, 2004[140] 9 months Duration of pain: >3 months in 59% *Fair*	A: Standard manipulation (n=353), 8 sessions over 12 weeks B: Usual care (n=338) C: Exercise (n=310)	A vs. B vs. C Age: 42 vs. 42 vs. 44 Female: 63% vs. 53% vs. 55% Baseline Von Korff Pain (0-100): 61.4 and 61.6 vs. 60.5 vs. 60.8 Baseline RDQ (0-24): 8.9 and 8.9 vs. 9.0 vs. 9.2	A vs. B <u>9 months</u> RDQ (0-24): 5.15 vs. 6.16, adjusted difference -1.01 (95% CI -1.81 to -0.22) Von Korff Disability (0-100): 29.85 vs. 35.50, adjusted difference -5.65 (95% CI -9.72 to -1.57) Von Korff Pain (0-100): 41.68 vs. 47.56, adjusted difference -5.87 (95% CI -10.17 to -1.58) A vs. C <u>9 months</u> RDQ (0-24): 5.15 (0.29) vs. 5.74 (0.31) Von Korff Disability (0-100): 29.85 (1.50) vs. 29.73 (1.68) Von Korff Pain (0-100): 41.68 (1.58) vs. 41.54 (1.84)	A vs. B <u>9 months</u> SF-36 PCS (0-100): 44.18 vs. 42.50, adjusted difference 1.68 (95% CI 0.18 to 3.19) SF-36 MCS (0-100): 48.09 vs. 46.41, adjusted difference 1.68 (95% CI -0.21 to 3.57) A vs. C <u>9 months</u> SF-36 PCS (0-100): 44.18 (0.55) vs. 44.39 (0.63) SF-36 MCS (0-100): 48.09 (0.69) vs. 46.77 (0.81)

CI = confidence interval; MCS = Mental Component Summary; NR = not reported; ODI = Oswestry Disability Index; OTC = over-the-counter; PCS = Physical Component Score; RDQ = Roland Morris Disability Questionnaire; SF-36 = Short-Form 36 Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Spinal Manipulation Compared With Sham Manipulation, Usual Care, an Attention Control, or a Placebo Intervention

Spinal manipulation was associated with slightly greater effects on function than controls at short-term followup (3 trials, SMD -0.34, 95% CI -0.63 to -0.05, I2=61%)[137-139] and intermediate-term followup (3 trials, SMD -0.40, 95% CI -0.69 to -0.11, I2=76%)[137,139,140] (Figure 16). Based on the original 0 to 100 scales (ODI and VF) used in the trials, the difference was -4.94 (95% CI -9.36 to -0.53) for short-term function and -9.19 (95% CI -12.77 to -5.61) for intermediate-term function. Estimates were similar when a poor-quality trial[139] was excluded. For short-term function, one trial reported similar effects for standard manipulation (difference -1.3 on the RDQ, 95% CI -2.9 to 0.6) and flexion-distraction manipulation (differenced -1.9, 95% CI -3.6 to -0.2); therefore, results for both arms were combined for the pooled analysis.[138]

There was no clear difference between spinal manipulation versus sham manipulation, an attention control, or a placebo intervention in short-term pain (3 trials, pooled difference -0.20 on a 0 to 10 scale, 95% CI -0.66 to 0.26, I2=58%) (Figure 17).[113,137,139] Two of the trials were rated poor quality; the results of the fair-quality trial[137] were consistent with the overall estimate (difference -0.21, 95% CI -0.68 to 0.25). Manipulation was associated with slightly greater effects on intermediate-term pain than sham manipulation, usual care, or an attention control (3 trials, pooled difference -0.64 on a 0 to 10 scale, 95% CI -0.92 to -0.36, I2=0%).[137,139,140] The estimate was similar when a poor-quality trial[139] was excluded (2 trials, difference -0.60, 95% CI -0.98 to -0.22).[137,140]

Two trials found no clear differences between spinal manipulation versus controls on the SF-36 MCS and PCS at short-term.[137,140] One trial[137] found no differences at short-term or intermediate-term followup and the other[140] found manipulation associated with slightly better PCS scores at intermediate-term followup, but the difference was very small (1.68 on a 0 to 100 scale, 95% CI 0.08 to 3.28).

Spinal Manipulation Compared With Pharmacological Therapy

No trial of spinal manipulation versus pharmacological therapy met inclusion criteria.

Spinal Manipulation Compared With Exercise

There were no differences between spinal manipulation versus exercise in function at short-term (3 trials, SMD 0.01, 95% CI -0.22 to 0.25, I2=62%)[152-154] or intermediate-term followup (4 trials, SMD 0.02, 95% CI -0.13 to 0.18, I2=48%)[140,152-154] (Figure 18). Excluding one trial[154] of flexion-distraction manipulation resulted in similar findings.

There were no differences between spinal manipulation versus exercise in short-term pain (3 trials, pooled difference 0.31, 95% CI -0.30 to 0.92, I2=60%)[152-154] or intermediate-term pain (4 trials, pooled difference 0.22, 95% CI -0.09 to 0.52, I2=9.4%) (Figure 19).[140,152-154] Excluding one trial[154] of flexion-distraction manipulation resulted in similar findings.

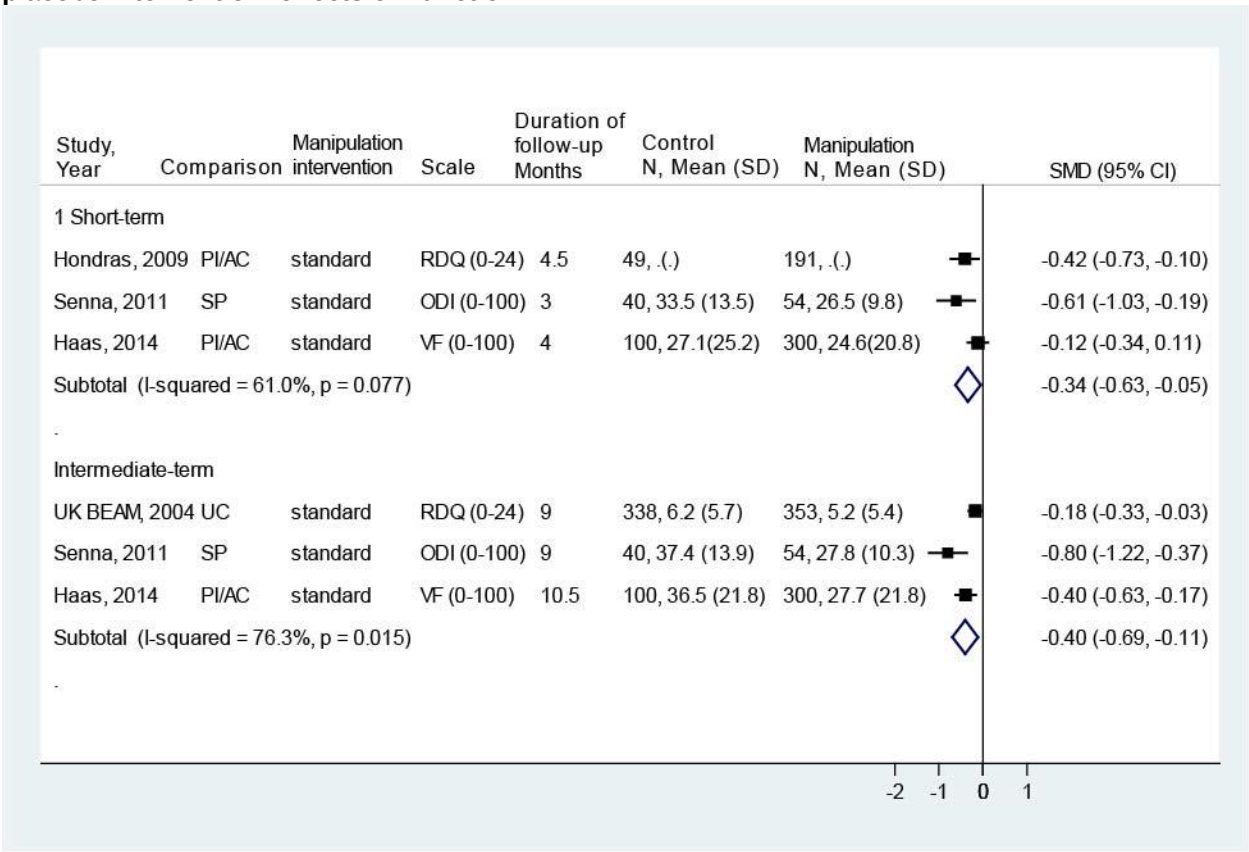
Two trials found no clear differences between spinal manipulation versus controls on the SF-36 MCS and PCS at short-term.[140,152] One trial[137] found no differences at short-term or intermediate-term followup and the other[140] found manipulation associated with slightly better

PCS scores at intermediate-term followup, but the difference was very small (1.68 on a 0 to 100 scale, 95% CI 0.08 to 3.28).

Harms

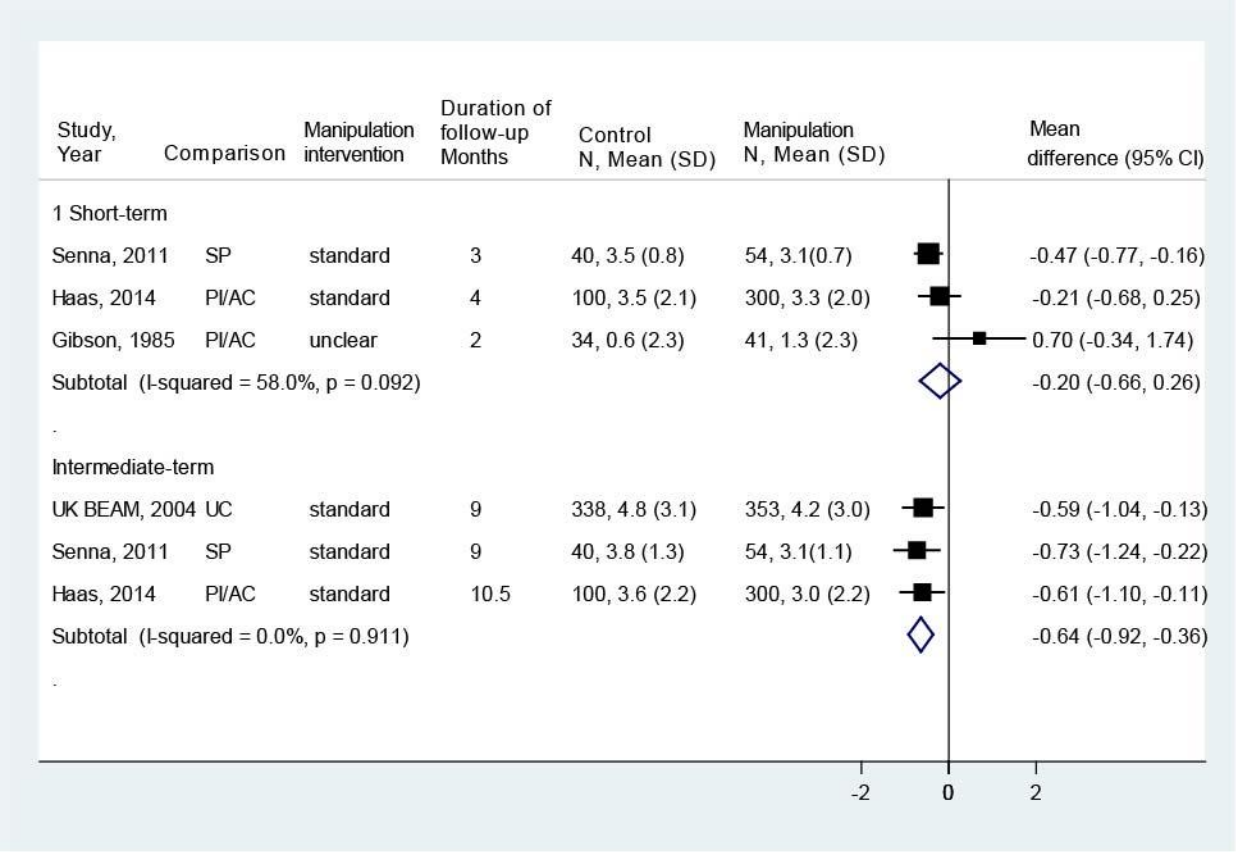
Seven trials reported no serious adverse events or withdrawals due to adverse events.[137-140,152-154] Nonserious adverse events (primarily increased pain) were reported in three trials.[137,139,152]

Figure 16. Spinal manipulation versus sham manipulation, usual care, an attention control, or a placebo intervention: effects on function



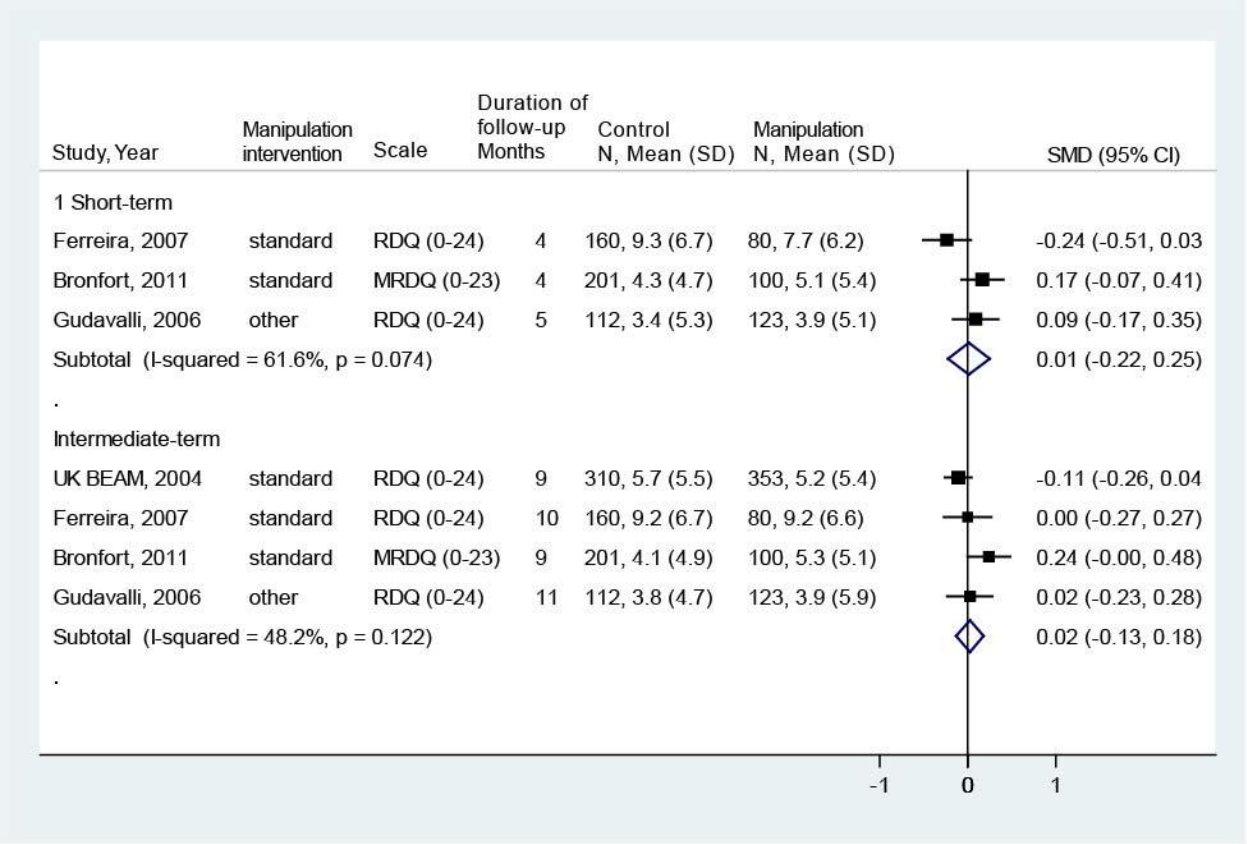
CI = confidence interval; PI/AC = placebo intervention/attention; SD = standard deviation; SMD = standardized mean difference; N = number; SP= sham manipulation, UC = usual care

Figure 17. Spinal manipulation versus sham manipulation, usual care, an attention control, or a placebo intervention: effects on pain



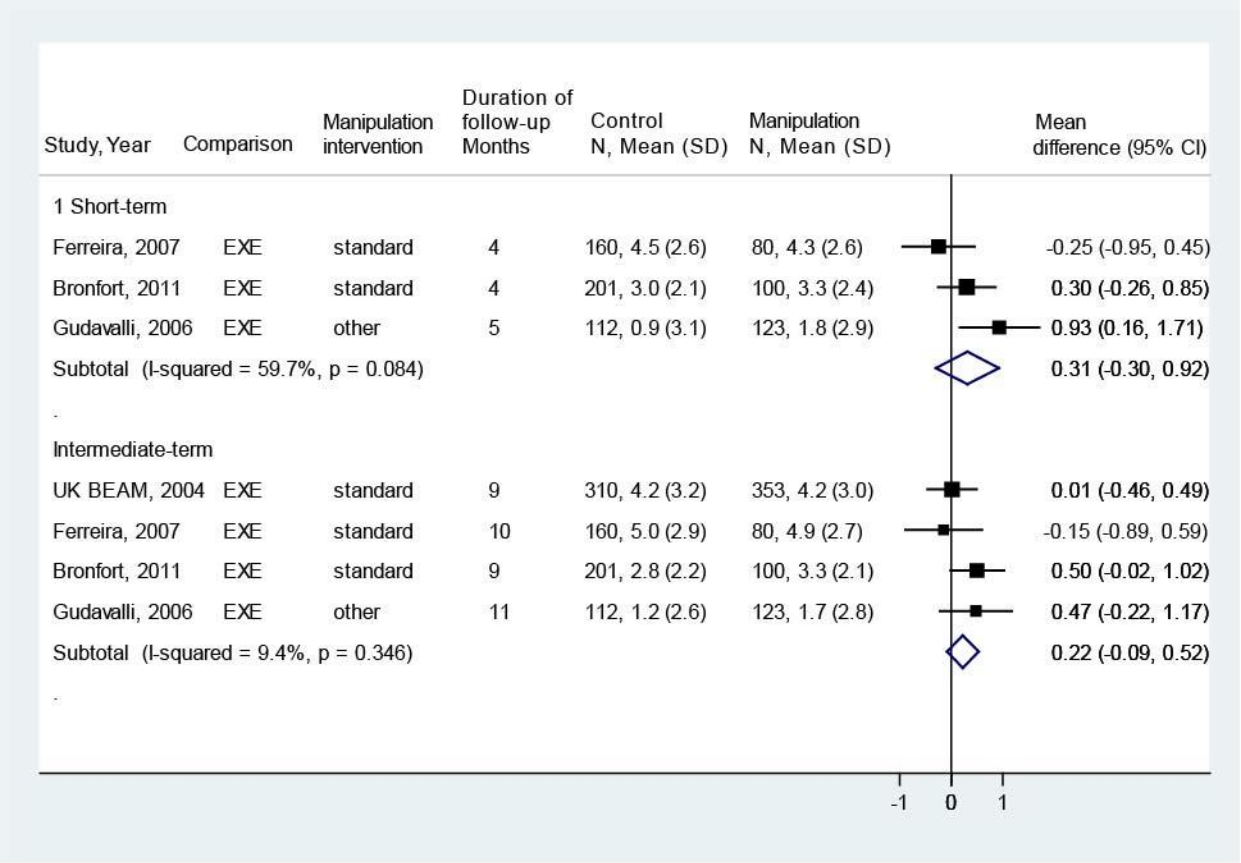
CI = confidence interval; PI/AC = placebo intervention/attention; SD = standard deviation; SMD = standardized mean difference; N = number SP= sham manipulation, UC = usual care

Figure 18. Spinal manipulation versus exercise: effects on function



CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number

Figure 19. Spinal manipulation versus exercise: effects on pain



CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number.
*Comparison: EXE = exercise

Psychological Therapy for Low Back Pain

Key Points

- Psychological therapy was associated with slightly greater effects on function than usual care or an attention control at short-term (3 trials, pooled SMD -0.25, 95% CI -0.38 to -0.12, I2=0%), intermediate-term (3 trials, pooled SMD -0.25, 95% CI -0.37 to -0.13, I2=0%), and long-term followup (2 trials, pooled SMD -0.26, 95% CI -0.39 to -0.12, I2=0%) (SOE: Moderate).
- Psychological therapy was associated with slightly greater effects on pain than usual care or an attention control at short-term (3 trials, pooled difference -0.76 on a 0 to 10 scale, 95% CI -0.99 to -0.53, I2=0%), intermediate-term (3 trials, pooled difference -0.71, 95% CI -0.94 to -0.48, I2=0%), or long-term followup (2 trials, pooled difference -0.53, 95% CI -0.82 to -0.24, I2=3.6%) (SOE: Moderate).
- Evidence from one poor-quality trial was too unreliable to determine effects of psychological therapy versus exercise (SOE: Insufficient).
- One trial reported no serious adverse events and one withdrawal due to adverse events in 468 patients (SOE: Low).

Detailed Synthesis

Five trials of psychological therapies for low back pain met inclusion criteria (Table 16 and Appendix D).[84-88,104] Three trials evaluated group CBT,[84-87] one trial evaluated respondent therapy (progressive muscle relaxation),[88] and one trial evaluated operant therapy.[104] Sample sizes ranged from 49 to 701 (total sample=1311). The number of psychological therapy sessions ranged from 6 to 8 and the duration of therapy ranged from 6 to 8 weeks. In one trial[86,87] the duration of therapy was unclear. Three trials compared psychological therapies versus usual care,[84,85,88] one trial compared psychological therapy versus an attention control (advice),[86,87] and one trial compared psychological therapy versus exercise therapy.[104] All trials were conducted in the United States or the United Kingdom. Three trials reported outcomes through long-term (12 to 34 months) followup,[85-87,104] one trial evaluated outcomes through intermediate-term followup,[84] and one trial only evaluated short-term outcomes.[88]

Three trials[84-87] were rated fair quality and two trials poor quality (Appendix E).[88,104] The major methodological limitation in the fair-quality trials was the inability to effectively blind patients and caregivers to the psychological intervention. Other methodological shortcomings in the poor-quality trials included unclear randomization and allocation concealment methods and high attrition.

Table 16. Summary of results for low back pain: psychological therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Cherkin, 2016[84] 10 months Duration of pain: >1 year in 80% of patients *Fair*	A. Cognitive behavioral therapy (n=116), 8 sessions over 8 weeks B. Usual care (n=112)	A vs. B 49 vs. 49 years Female: 59% vs. 66% Baseline pain bothersomeness (0-10): 6.0 vs. 6.0 Baseline modified RDQ (0-23): 11.5 vs. 10.9	A vs. B <u>4.5 months</u> Pain (0-10): -1.56 (95% CI -2.02, -1.11) vs. -0.84 (95% CI -1.21, -0.46) Modified RDQ (0-23): -4.38 (95% CI -5.3, -3.47) vs. -2.96 (95% CI -3.79, -2.14) <u>10 months</u> Pain (0-10): -1.76 (95% CI -2.14, -1.39) vs. -1.10 (95% CI -1.48, -0.71) Modified RDQ (0-23): −4.78 (95% CI −5.67, −3.89) vs. −3.43 (95% CI −4.33, −2.52) ≥30% improvement in pain: .39.6% (95% CI 31.7, 49.5) vs. 31.0% (95% CI 23.8, 40.3) ≥30% improvement in modified RDQ: 58.8% (95% CI 50.6, 68.4) vs. 48.6% (95% CI 40.3, 58.6)	A vs. B <u>4.5 months</u> PHQ-8 (0–24): -1.80 (95% CI -2.35, -1.26) vs. -0.64 (95% CI -1.23, -0.06) SF-12 Physical component (0-100): 3.78 (95% CI 2.56, 5.00) vs. 3.27 (95% CI 2.09, 4.44) SF-12 Mental component (0-100): 2.13 (95% CI 0.86, 3.40) vs. −1.11 (95% CI −2.39, 0.17) <u>10 months</u> PHQ-8 (0–24): 1.72 (95% CI −2.28, −1.16) vs. −0.88 (95% CI −1.50, −0.27) SF-12 Physical component: 3.79 (95% CI 2.55, 5.03) vs. 2.93 (95% CI 1.70, 4.16) SF-12 Mental component: 1.81 (95% CI 0.59, 3.03) vs. 0.75 (95% CI −0.58, 2.08)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Johnson, 2007[85] 12 months Duration of pain: 6 months *Fair*	A. Cognitive behavioral therapy (n=116), 8 sessions over 6 weeks B. Usual care (n=118)	A vs. B Age: 47 vs. 49 Female: 61% vs. 58% Baseline pain (0-100 VAS): 44.9 vs. 51.6 Baseline RDQ (0-24): 10.6 vs. 10.9	A vs. B <u>6 months</u> Pain (0-100 VAS): 26.1 vs. 35.0, adjusted difference -4.60 (95% CI -11.07 to 1.88) RDQ (0-24): 6.5 vs. 8.0, adjusted difference -1.09 (95% CI -2.28 to 0.09) <u>12 months</u> Pain (0-100 VAS): 27.9 vs. 36.4, adjusted difference -5.49 (95% CI -12.43 to 1.44) RDQ (0-24): 6.7 vs. 8.0, adjusted difference -0.93 (95% CI -2.30 to 0.45)	A vs. B <u>6 months</u> Quality of life (0-1 EQ-5D): 0.75 vs. 0.71, adjusted difference 0.03 (95% CI -0.05 to 0.10) <u>12 months</u> Quality of life (0-1 EQ-5D): 0.75 vs. 0.71, adjusted difference 0.03 (95% CI -0.04 to 0.09)
Lamb 2010[86] and 2012[87] 34 months Duration of pain: 13 years *Fair*	A. Cognitive behavioral treatment (n=468), 8 sessions over unclear number of weeks B. Attention control (n=233)	A vs. B Age: 53 vs. 54 years Female: 59% vs. 61% Baseline pain (0-100 Modified Von Korff): 59 vs. 59 Modified Von Korff disability (0-100): 49 vs. 46 Baseline RDQ (0-24): 9 vs. 9	A vs. B <u>3 months</u> Modified Von Korff pain (0-100): -12.2 (-14.56 to -9.83) vs. -5.4 (-8.40 to -2.49), adjusted difference -6.8 (-10.20 to -3.31) Modified Von Korff disability (0-100): -13.2 (-15.74 to -10.59) vs. -8.9 (-12.27 to -5.56), adjusted difference -4.2 (-8.10 to -0.40) RDQ (0-24): -2.0 (-2.43 to -1.58) vs. -1.1 (-1.54 to -0.35) adjusted difference -1.1 (-1.71 to -0.38) <u>4.5 months</u> Modified Von Korff pain: -13.7 (-16.20 to -11.29) vs. -5.7 (-8.99 to -2.41), adjusted difference -8.0 (-11.80 to -4.28) Modified Von Korff disability: -13.9 (CI -16.25 to -11.55) vs. -5.7 (-9.22 to -2.28), adjusted difference -8.2 (-12.01 to -4.31) RDQ: -2.5 (-3.03 to -1.96) vs. -1.0 (CI -1.67 to -0.40), adjusted difference -1.5 (-2.22 to -0.70) <u>10.5 months</u> Modified Von Korff pain: -13.4 (-15.96 to -10.77) vs. -6.4 (-9.66 to -3.14), adjusted difference -7.0 (-10.81 to -3.12) Modified Von Korff disability: -13.8 (-16.28 to -11.39) vs. -5.4 (-8.90 to -1.99), adjusted difference -8.4 (-12.32 to -4.47)	A vs. B <u>3 months</u> SF-12 PCS (0-100): 3.7 (2.82 to 4.59) vs. 1.5 (0.26 to 2.83), adjusted difference 2.2 (0.74 to 3.57) SF-12 MCS (0-100): 1.3 (0.19 to 2.42) vs. 0 (-1.45 to 1.46), adjusted difference 1.3 (-0.36 to 2.96) <u>4.5 months</u> SF-12 PCS: 3.6 (2.72 to 4.52) vs. 1.8 (0.54 to 3.08), adjusted difference 1.8 (0.37 to 3.25) SF-12 MCS: 2.5 (1.44 to 3.48) vs. -0.09 (-1.61 to 1.43), adjusted difference 2.6 (0.85 to 4.25) <u>10.5 months</u> SF-12 PCS: 4.9 (4.00 to 5.84) vs. 0.8 (-0.52 to 2.11), adjusted difference 4.1 (2.63 to 5.62) SF-12 MSC: 0.9 (-0.10 to 1.90) vs. 0.7 (-0.75 to 2.20), adjusted difference 0.2 (-1.48 to 1.84) <u>34 months</u> EuroQol-5 Dimensions (EQ-5D): 0.07 (0.04 to

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			RDQ: -2.4 (-2.84 to -1.89) vs. -1.1 (-1.72 to -0.39), adjusted difference -1.3 (-2.06 to -0.56) <u>34 months</u> Modified Von Korff pain: -17.4 (-20.35 to -14.44) vs. -12.8 (-17.52 to -7.99), adjusted difference -4.6 (-10.28 to 1.00) Modified Von Korff disability: -16.7 (-19.43 to -13.93) vs. -11.2 (-15.59 vs. -6.86), adjusted difference -5.5 (-10.64 to -0.27) RDQ: -2.9 (-3.42 to -2.38) vs. -1.6 (-2.48 to -0.80), adjusted difference -1.3 (-2.26 to -0.27)	0.10) vs. 0.04 (-0.01 to 0.09), adjusted difference 0.03 (-0.03 to 0.08)
Poole, 2007[88] 4.5 months Duration of pain: 10.6 vs. 9.5 years *Poor*	A. Respondent therapy (progressive muscle relaxation) (n=54), 6 sessions over 6-8 weeks B. Usual care (n=45)	A vs. C Age: 46 vs. 47 Female: 65% vs. 51% Baseline pain (0-100 VAS): 40.7 vs. 40.6 Baseline Oswestry Disability Index (0-100% ODI): 33.2 vs. 36.6	A vs. C <u>4.5 month</u> Pain (0-100 VAS): 41.3 vs. 42.7 ODI (0-100): 31.3 vs. 32.9	A vs. C <u>4.5 month</u> Beck Depression Inventory (0-63): 12.6 (10.9) vs. 12.8 (9.2) SF-36 physical functioning (0-100): 57.3 (31.8) vs. 52.2 (29.5) SF-36 social functioning (0-100): 66.7 (31.6) vs. 61.5 (30.8) SF-36 emotional role limitations (0-100): 63.0 (43.8) vs. 62.0 (44.0) SF-36 pain (0-100): 48.8 (25.9) vs. 44.4 (28.5) SF-36 mental health (0-100): 64.4 (20.7) vs. 67.7 (18.5) SF-36 general health perception (0-100): 52.4 (22.8) vs. 55.0 (24.1)
Turner, 1990[104] 12 months Duration of pain: 12.9 years *Poor*	A. Operant therapy (n=25), 8 sessions over 8 weeks B. Exercise (n=24)	Overall Age: 44 Female: 48% A vs. B vs. C vs. D Baseline pain (0-78 McGill Pain Rating): 21.0 vs. 19.4 vs. 25.5 vs. 21.2 Baseline function (0-100 SIP): 7.9 vs. 8.4 vs. 8.5 vs.	A vs. B <u>6 months</u> McGill Pain Questionnaire Pain Rating Index (0-78): 9.5 (15.7) vs. 15.7 (9.2) Sickness Impact Profile (0-100): 7.6 (9.9) vs. 6.3 (10.1) <u>12 months</u> McGill Pain Questionnaire Pain Rating Index: 16.4 (13.6) vs. 14.9 (7.9) Sickness Impact Profile: 5.3 (6.7) vs. 4.7 (7.9)	A vs. B <u>6 months</u> CES-D Scale (0-60): 11.4 (8.3) vs. 9.3 (8.3) <u>12 months</u> CES-D Scale: 8.3 (7.7) vs. 9.3 (7.7)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
		6.2		

BDI = Beck Depression Inventory; CBT = cognitive behavioral therapy; CES-D = Center for Epidemiologic Studies-Depression; CI = confidence interval; MCS = Mental Component Score; NHP = Nottingham Health Profile; NR = not reported; ODI = Oswestry Disability Index; PCS = Physical Component Score; RDQ = Roland Morris Disability Questionnaire; SF-36, Short-Form 36 Questionnaire; SIP = Sickness Impact Profile; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Psychological Therapy Compared With Usual Care or an Attention Control

Psychological therapy was associated with slightly greater effects on function than usual care or an attention control at short-term (3 trials, pooled SMD -0.25, 95% CI -0.38 to -0.12, I2=0%),[84,86,88] intermediate-term (3 trials, pooled SMD -0.25, 95% CI -0.37 to -0.13, I2=0%),[84-86] and long-term followup (2 trials, pooled SMD -0.26, 95% CI -0.39 to -0.12, I2=0%) (Figure 20).[85,86] Pooled differences on the RDQ or modified RDQ were -1.2 to -1.3 points at all time points. Excluding the trial of progressive relaxation,[88] which found no effect on short-term function (SMD -0.08, 95% CI -0.48 to 0.31), had no effect on the pooled estimate (2 trials, pooled SMD -0.27, 95% CI -0.43 to -0.06).

Psychological therapy was associated with slightly greater effects on pain than usual care or an attention control at short-term (3 trials, pooled difference -0.76 on a 0 to 10 scale, 95% CI -0.99 to -0.53, I2=0%),[84,86,88] intermediate-term (3 trials, pooled difference -0.71, 95% CI -0.94 to -0.48, I2=0%),[84-86] or long-term followup (2 trials, pooled difference -0.53, 95% CI -0.82 to -0.24, I2=3.6%) (Figure 21).[85,87] For short-term pain, two fair-quality trials[84,86,87] evaluated CBT and one poor-quality trial[88] evaluated respondent therapy (progressive relaxation). There was no difference between progressive relaxation versus usual care in short-term pain (mean difference -0.14, 95% CI -1.28 to 1.00). Restricting the analysis to the trials of CBT did not change the pooled estimate (2 trials, pooled difference -0.78, 95% CI -1.06 to -0.49). For intermediate-term and long-term pain, all trials were fair quality and evaluated CBT.

Effects of psychological therapy on short-term or intermediate-term SF-36 PCS or MCS scores were small (differences 0 to 2 points on a 0 to 100 scale) and not statistically significant, except for short-term MCS (2 trials, pooled difference 2.12, 95% CI 0.79 to 3.45).[84,86] One trial found no effect of psychological therapy on work status or health care visits.[87]

Psychological Therapy Compared With Pharmacological Therapy

No trial of psychological versus pharmacological therapy met inclusion criteria.

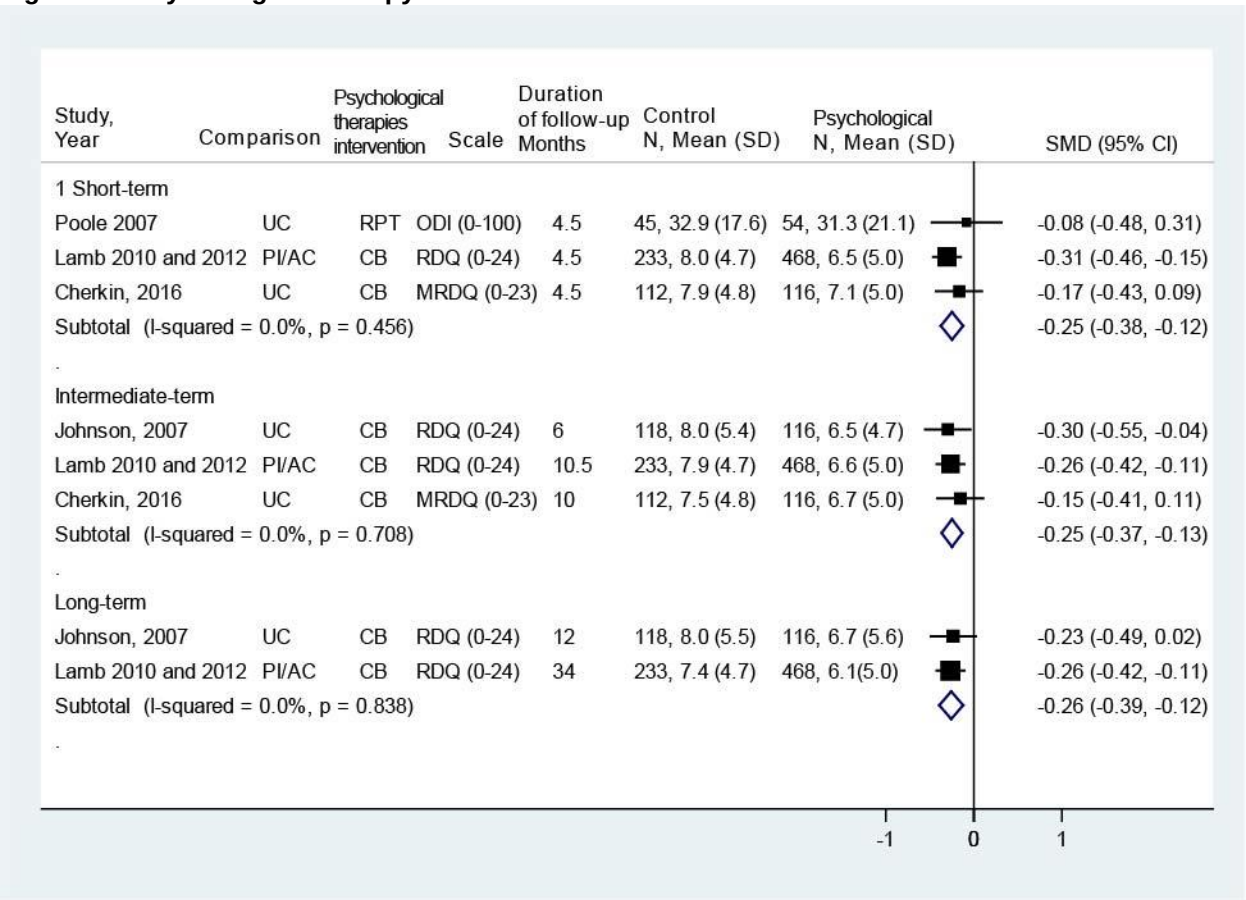
Psychological Therapy Compared With Exercise

One poor-quality trial found no differences between psychological versus exercise therapy in intermediate-term or long-term function.[104] Differences on the McGill Pain Questionnaire were less than 0.5 points on a 0 to 78 scale, and differences on the Sickness Impact Profile were 0.60 to 1.30 points on a 0 to 100 scale.

Harms

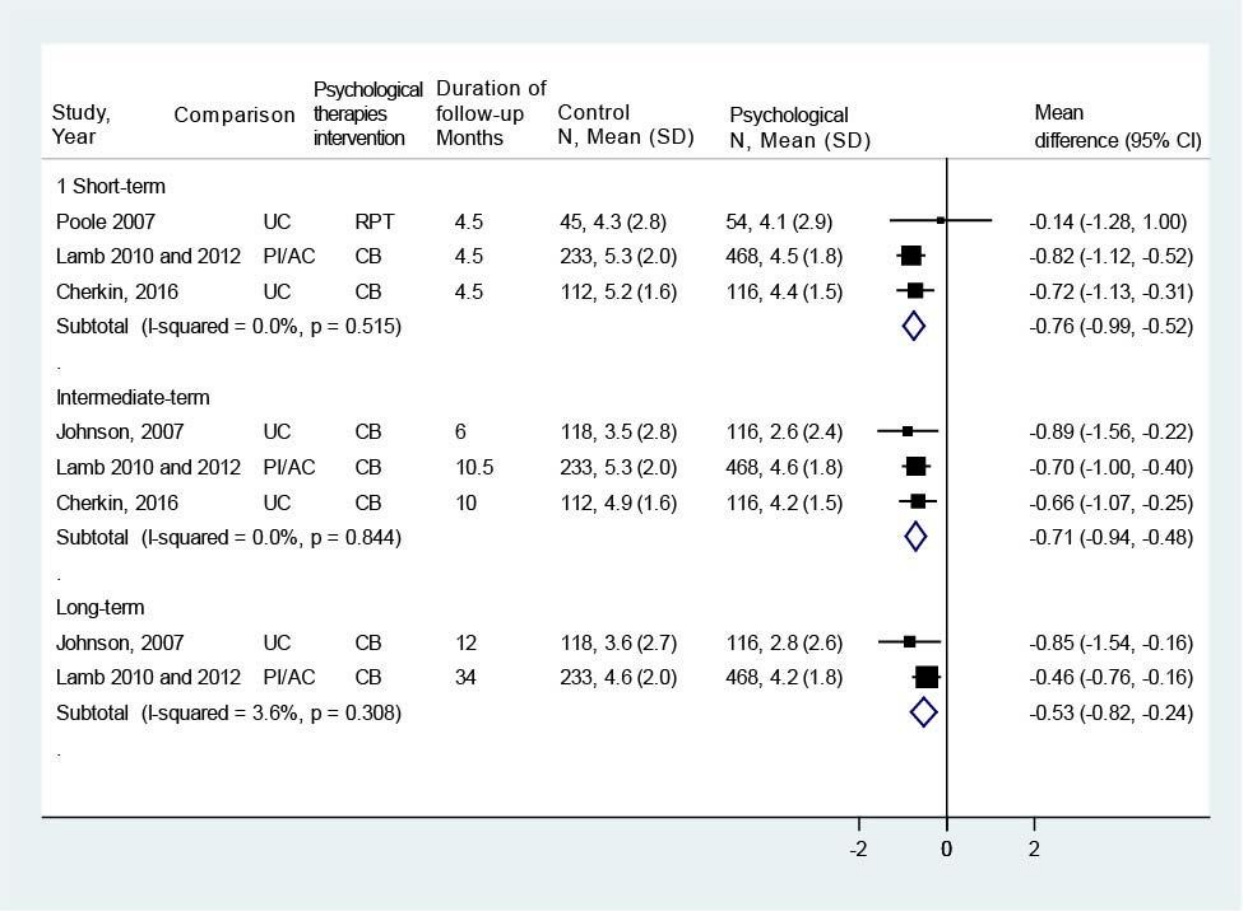
Data on harms were sparse. One trial reported no serious adverse events and one withdrawal due to adverse events among 468 patients randomized to CBT.[86,87]

Figure 20. Psychological therapy versus usual care or an attention control: effects on function



CI = confidence interval; PI/AC = placebo intervention/attention; SD = standard deviation; SMD = standardized mean difference; N = number; UC = usual care

Figure 21. Psychological therapy versus usual care or an attention control: effects on pain



CI = confidence interval; PI/AC = placebo intervention/attention; SD = standard deviation; SMD = standardized mean difference; N = number; UC = usual care

Multidisciplinary Rehabilitation for Low Back Pain

Key Points

- Multidisciplinary rehabilitation was associated with slightly greater effects on function than usual care at short-term (4 trials, pooled SMD -0.31, 95% CI -0.57 to -0.05, I2=70%) and intermediate-term followup (4 trials, pooled SMD -0.37, 95% CI -0.64 to -0.10, I2=50%); there was no difference in long-term function (2 trials, pooled SMD -0.04, 95% CI -0.31 to 0.24, I2=35%) (SOE: Low).
- Multidisciplinary rehabilitation was associated with slightly greater effects on pain than usual care at short-term (4 trials, pooled difference -0.51 on a 0 to 10 scale, 95% CI -0.89 to -0.13, I2=23%) and intermediate-term (4 trials, pooled difference -0.63, 95% CI -1.04 to -0.22, I2=0%) followup; the long-term difference was smaller and not statistically significant (2 trials, pooled difference -0.34, 95% CI -0.86 to 0.18, I2=0%) (SOE: Moderate for short-term and intermediate-term, Low for long-term).

- Multidisciplinary rehabilitation was associated with slightly greater effects than exercise on short-term (6 trials, pooled difference -0.75 on a 0 to 10 scale, 95% CI -1.18 to -0.31, I2=0%) and intermediate-term pain (5 trials [excluding outlier trial], pooled difference -0.55, 95% CI -0.95 to -0.15, I2=0%); there was no effect on long-term pain (2 trials [excluding outlier trial], pooled difference 0.00, 95% CI -0.94 to 0.95) (SOE: Moderate for short-term and intermediate-term, Low for long-term).
- Multidisciplinary rehabilitation was associated with slightly greater effects than exercise on short-term (6 trials, pooled SMD -0.28, 95% CI -0.54 to -0.01, I2=39%) and intermediate-term function (5 trials [excluding outlier trial], pooled SMD -0.22, 95% CI -0.40 to -0.03, I2=0%), there was no effect on long-term function (2 trials [excluding outlier trial], pooled SMD -0.06, 95% CI -0.36 to 0.25, I2=0%) (SOE: Moderate for short-term and intermediate-term, Low for long-term).
- Data on harms were sparse; no serious harms were reported (SOE: Insufficient).

Detailed Synthesis

Sixteen trials (reported in 21 publications) of multidisciplinary rehabilitation for low back pain met inclusion criteria (Table 17 and Appendix D).[29,104,110,151,204-209,216-228] In accordance with our definition for multidisciplinary rehabilitation, the intervention in all trials included a psychological therapy and an exercise therapy component, and therapy was developed by clinicians from at least two disciplines. The intensity of multidisciplinary rehabilitation varied substantially, with treatment ranging from 4 to 150 hours. Five trials evaluated a multidisciplinary rehabilitation intervention that met our criteria for high intensity (>=20 hours/week or >80 hours total).[204,209,217,218,225] The duration of therapy ranged from 4 days to up to 13 weeks. Sample sizes ranged from 20 to 459 (total sample=1,904). Six trials compared multidisciplinary rehabilitation versus usual care,[204-209] nine trials compared multidisciplinary rehabilitation versus exercise therapy,[104,206,217,218,220-225] and one trial compared multidisciplinary rehabilitation versus oral medications.[216] One trial[216] was conducted in Iran and the remainder were conducted in the United States, the United Kingdom, or Australia. Five trials reported outcomes through long-term (12 to 60 months) followup,[104,204,216,217,223] eight trials evaluated outcomes through intermediate-term followup,[104,207-209,218,220,222,225,226] and three trials only evaluated short-term outcomes.[205,221,224]

Ten trials[204,206,207,217,218,221-225] were rated fair quality and six trials poor quality (Appendix Table E).[104,205,208,209,216,220] The major methodological limitation in the fair-quality trials was the inability to effectively blind patients and caregivers to the multidisciplinary rehabilitation. Other methodological shortcomings included unclear randomization and allocation concealment methods and high attrition.

Table 17. Summary of results for low back pain: multidisciplinary rehabilitation

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Abbassi, 2012[208] 10.25 months Duration of pain: ~6 years	A. Multidisciplinary rehabilitation (n=12), 7 sessions over 7 weeks B. Multidisciplinary	A + B + C Overall Age (mean): 45 years Female: 88%	A vs. B vs. C <u>10.25 months</u> Pain (0–10 VAS): 3.7 vs. 2.8 vs. 4.3, p=0.44 RDQ (0–24): 8.8 vs. 8.2 vs. 10.4, p=0.44	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Poor	pain management (spouse-assisted) (n=10). 7 sessions over 7 weeks C: Usual care (n=11)	A vs. B vs. C Baseline pain (0-10 VAS): 4.6 vs. 5 vs. 3.6 Baseline RDQ (0–24): 12.1 vs. 11.2 vs. 8.4		
Bendix, 1995,[217] 1997,[227] 1998[228] 60 months Duration of pain: >6 months *Fair*	A. Multidisciplinary rehabilitation (n=46), 18 sessions over 6 weeks (total ~135 hours) B. Multidisciplinary rehabilitation (n=43), 12 sessions over 6 weeks (total 24 hours) C. Exercise (n=43)	A vs. B vs. C Age: 40 vs. 44 vs. 42 Female: 75% vs. 77% vs. 74% Baseline pain (0-10 NRS): 5.3 vs. 5.9 vs. 5.4 Baseline Low Back Pain Rating Scale (0-30): 15.5 vs. 15.3 vs. 14.4	A vs. B vs. C <u>3.25 months</u> Back pain (0-10 NRS): 2.7 vs. 5.6 vs. 4.4, p<0.001 Low Back Pain Rating Scale (0-30): 8.5 vs. 16.1 vs. 13.5, p=0.002 <u>12 months</u> Back pain (0-10 NRS): 3.3 vs. 6.5 vs. 5.3, p=0.005 Low Back Pain Rating Scale (0-30): 8.9 vs. 16.4 vs. 13.7, p<0.001 <u>24 months</u> Back pain (0-10 NRS): 3 vs. 6 vs. 5, p=0.08 Low Back Pain Rating Scale (0-30): 10 vs. 17 vs. 14, p=0.003 <u>60 months</u> Back pain (0-10 NRS): 4 vs. 6 vs. 5, p=0.3 Low Back Pain Rating Scale (0-30): 8 vs. 16 vs. 14, p=0.02	A vs. B vs. C <u>3.25 months</u> Days of sick leave: 25 vs.122 vs. 13, p=0.005 Health care system contacts: 0.5 vs. 2.8 vs. 1.3, p=0.05 <u>12 months</u> Days of sick leave: 52 vs. 295 vs. 100, p=0.002 Health care system contacts: 4.5 vs. 12.0 vs. 11.8, p=0.002 Days of sick leave: 2.5 vs. 37 vs. 11, p=0.06 <u>24 months</u> Health care system contacts: 5 vs. 21 vs. 14, p=0.03 Overall assessment (1-5): 2 vs. 3 vs. 3, p=0.005 <u>60 months</u> Overall assessment (1-5): 2 vs. 3 vs. 3, p=0.004 Increase in proportion able to work: 30% vs. 23% vs. 0%, p=0.001 Days of sick leave: 13 vs. 11 vs. 88, p=0.2 Health care system contacts: 15 vs. 10 vs. 24, p=0.2 Back surgery: 5% vs. 10% vs. 10%, p=0.7
Bendix,1996,[204] 1998[228] 60 months Duration of pain: >6 months *Fair*	A. Multidisciplinary rehabilitation (n=55), 18 sessions over 6 weeks (total ~135 hours) B. Usual care (n=51)	A vs. B Age 41 vs.40 years Female: 71% vs. 69% Baseline pain (0-10 NRS): 6.1 vs. 6.1 Baseline Low	A vs. B <u>3.25 months</u> Back pain (0-10 NRS): 5.7 vs. 6.9, p=0.05 Low Back Pain Rating Scale (0-30): 12.1 vs. 16.8, p<0.001 <u>24 months</u> Back pain (0-10 NRS): 6 vs.	A vs. B <u>3.25 months</u> Days of sick leave: 10 vs. 122, p=0.02 Contacts to health-care system: 1.6 vs. 5.3, p<0.001 <u>24 months</u>

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
		Back Pain Rating Scale (0-30): 16.9 vs. 15.9	6.5, p=0.5 Low Back Pain Rating Scale (0-30): 16 vs. 15, p=0.9 <u>60 months</u> Back pain (0-10 NRS): 5 vs. 5, p=1.0 Low Back Pain Rating Scale (0-30): 12 vs. 16, p=0.2	Days of sick leave: 15 vs. 123, p<0.001 Health care system contacts: 12 vs. 26, p<0.001 <u>60 months</u> Days of sick leave: 10 vs. 50, p=0.4 Health care system contacts: 16 vs. 48, p=0.1 Back surgery: 7% vs. 12%, p=0.4
Bendix, 2000[218] 10 months Duration of pain: Not reported *Fair*	A. Multidisciplinary rehabilitation (n=59), 18 sessions over 8 weeks (total ~139 hours) B. Exercise (n=68)	A vs. B Age: 40 vs. 43 years Female: 66% vs.65% Baseline pain: NR Baseline function: NR	A vs. B <u>10 months</u> Back pain (0–10): 5.1 vs. 5.7, p=0.33 Low Back Pain Rating Scale (0–30 ADL): 12 vs. 13, p=0.41	A vs. B <u>10 months</u> Overall assessment (1–5): 1.7 vs. 2.7, p=0.03 Work capable: 75% vs. 69%, p=0.64 Health care contacts (number): 2.5 vs. 4, p=0.28
Harkapaa, 1989[205] 1 month Duration of pain: >2 years *Poor*	A. Multidisciplinary rehabilitation (inpatient) (n=156), 3 weeks (number of sessions and total hours unclear) B. Multidisciplinary rehabilitation (outpatient) (n=150), 15 sessions over 8 weeks (total hours unclear) C. Usual care (n=153)	A vs. B vs. C Age: 45 vs. 45 vs. 45 years Female: 37% vs. 39% vs. 35% Baseline Pain Index (0-400): 184.9 vs. 178.6 vs. 175.8 Baseline function, LBP Disability Index (0-45): 16.7 vs. 17.6 vs. 16.7	A vs. B vs. C <u>1 month</u> Pain Index (0-400): 127 vs. 145 vs. 160, p<0.001 for A vs. C and p<0.04 for B vs. C LBP Disability Index (0-45): 13.8 vs. 14.7 vs. 17.3, p<0.004 for A vs. C and p<0.01 for B vs. C	NR
Jousset, 2004[219] 5 months Duration of pain: >4 months *Poor*	A. Multidisciplinary rehabilitation (n=44), 25 sessions over 5 weeks (total 150 hours) B. Exercise (n=42)	A vs. B Age: 41 vs. 40 years Female: 30% vs. 37% Baseline pain (0-10 NRS): 5.0 vs. 4.6 Baseline function Quebec Disability Scale (0-100): 34.6 vs. 31.6	A vs. B <u>5 months</u> Pain (0-10 NRS): 3.1 vs. 4.0, p=0.01 Dallas Pain Questionnaire ADL (0-100): 36.7 vs. 41.5, p=0.36 Quebec Disability Scale (0-100): 22.0 vs. 22.9, p=0.80	A vs. B <u>5 months</u> Hospital Anxiety Depression Scale (0-21): 12.7 vs. 13.4 (6.4), p=0.62 Dallas Pain Questionnaire Social interest (0-100): 19.6 vs. 24.3, p=0.37
Lambeek 2010[207]	A. Multidisciplinary rehabilitation (n=66), 26+	A vs. B Age: 46 vs. 47 years	A vs. B <u>3 months</u> Pain (0-10 VAS): 1.3 vs. 2.3,	A vs. B <u>9 months</u> General practitioner

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
9 months Duration of pain: >4 months *Fair*	sessions over up to 13 weeks (total hours unclear) B. Usual care (n=68)	Female: 44% vs. 40% Baseline pain (0-10 VAS): 5.7 vs. 6.3 Baseline modified RDQ (0-23): 14.7 vs. 15.0	adjusted difference 0.5, 95% CI -0.6 to 1.6 Modified RDQ (0-23): 4.8 vs. 5.0 (0.9), adjusted difference 0.06, 95% CI -2.3 to 2.5 <u>9 months</u> Pain (0-10 VAS): 1.6 vs. 1.9, adjusted difference 0.21, 95% CI -0.8 to 1.2 Modified RDQ (0-23): 7.2 vs. 4.4, adjusted difference -2.9, 95% CI -4.9 to -0.9	visits (# of patients): 13 vs. 29 Medical specialist visits (# of patients): 13 vs. 29 Total costs (pounds): 13165 (SD 13600) vs. 18475 (SD 13616), mean difference -5310 (95% CI -10042 to -391)
Monticone 2013[223] 23 months Duration of pain: 25 vs. 26 months *Fair*	A. Multidisciplinary rehabilitation (n=45), 26 sessions over 5 weeks (total 26 hours) B. Exercise (n=45)	A vs. B Age: 49 vs. 50 years Female: 60% vs. 56% Baseline pain (0-10 VAS): 7.0 vs. 7.0 Baseline RDQ (0-24): 15.3 vs. 15.0	A vs. B <u>11 months</u> Pain (0-10 VAS): 1.4 (1.1) vs. 5.3 (1.2) RDQ (0-24): 1.3 (1.6) vs. 11.0 (2.0) SF-36 physical pain (0-100): 79.0 (14.6) vs. 52.0 (16.2) <u>23 months</u> Pain (0-10 VAS): 1.5 vs. 6.2, difference -4.7, 95% CI -5.1 to -4.3 RDQ (0-24): 1.4 vs. 11.1, difference -9.7, 95% CI -10.4 to -9.0 SF-36 physical pain: 80.4 vs. 61.8, difference 18.6, 95% CI 12.8 to 24.3	A vs. B <u>11 months</u> SF-36 physical functioning (0-100): 85.7 (19.6) vs. 62.1 (19.4) SF-36 general health (0-100): 85.0 (13.8) vs. 56.4 (15.9) SF-36 mental health (0-100): 89.8 (13.0) vs. 54.1 (11.9) <u>23 months</u> SF-36 physical functioning (0-100): 87.6 vs. 65.0, difference 22.6, 95% CI 15.0 to 30.1 SF-36 general health: 86.3 vs. 63.1, difference 23.2, 95% CI 17.3 to 29.1 SF-36 mental health: 91.0 vs. 58.8, difference 32.2, 95% CI 27.4 to 37.0)
Monticone 2014[224] 3 months Duration of pain: 15 vs. 14 months *Fair*	A. Multidisciplinary rehabilitation (n=10), 16 sessions over 8 weeks (total 16 hours) B. Exercise (n=10)	A vs. B Age: 59 vs. 57 years Female: 7% vs. 4% Baseline pain (0-10 NRS): 5 vs. 4 Baseline function (0-100 ODI): 26 vs. 24	A vs. B <u>3 months</u> ODI (0-100): 8 vs. 15, p=0.027 Pain (0-10 NRS): 2 vs. 3, p=1.0 SF-36 bodily pain (0-100): 65 vs. 55, p=0.261	A vs. B <u>3 months</u> SF-36 general health (0-100): 71 vs. 55, p=0.018 SF-36 social function (0-100): 81 vs. 61, p=0.001 SF-36 emotional role (0-100): 77 vs. 57, p=0.007 SF-36 mental health (0-100): 88 vs. 67, p=0.001

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Nicholas, 1991[220] 11 months Duration of pain: 7 years *Poor*	A. Multidisciplinary rehabilitation (cognitive treatment) (n=10) B. Multidisciplinary rehabilitation (behavioral treatment) (n=10) C. Multidisciplinary rehabilitation (cognitive treatment and relaxation treatment) (n=8) D. Multidisciplinary rehabilitation (behavioral treatment and relaxation training) (n=9) E. Exercise + attention control (psychologist-led group discussions) (n=10) F. Exercise (n=11) For all multidisciplinary rehabilitation interventions, 19 sessions over 5 weeks (total 21.5 hours)	Overall Age: 41 years Female: 51% A vs. B vs. C vs. D vs. E vs. F Baseline pain (0-5 categorical scale): 2.78 vs. 2.96 vs. 3.80 vs. 2.27 vs. 2.84 vs. 2.77 Baseline function, (0-100 Sickness Impact Profile): 37.13 vs. 34.24 vs. 33.41 vs. 20.53 vs. 27.12 vs. 28.06	A vs. B vs. C vs. D vs. E vs. F <u>5 months</u> Pain (0-5 categorical scale): 2.18 (0.55) vs. 1.87 (0.73) vs. 3.20 (0.93) vs. 2.22 (0.48) vs. 2.64 (0.90) vs. 3.18 (0.72) Sickness Impact Profile (0-100): 24.42 (11.78) vs. 15.44 (14.12) vs. 25.69 (8.50) vs. 14.86 (9.08) vs. 19.40 (6.89) vs. 29.78 (8.76) <u>11 months</u> Pain (0-5 categorical scale): 2.56 (0.97) vs. 2.66 (1.06) vs. 3.30 (0.83) vs. 1.88 (0.65) vs. 2.70 (0.84) vs. 3.22 (0.69) Sickness Impact Profile (0-100): 23.85 (12.50) vs. 12.80 (8.62) vs. 20.77 (8.29) vs. 12.87 (6.68) vs. 18.94 (12.79) vs. 25.18 (8.08)	A vs. B vs. C vs. D vs. E vs. F <u>5 months</u> Spielberger State Anxiety Inventory (20-80): 57.17 (10.30) vs. 37.57 (12.92) vs. 55.71 (10.47) vs. 36.40 (6.28) vs. 41.13 (11.70) vs. 54.00 (12.03) Beck Depression Inventory (0-63): 18.67 (9.01) vs. 8.14 (5.77) vs. 16.14 (3.80) vs. 9.00 (6.07) vs. 9.88 (5.46) vs. 19.17 (8.78) Medication use (0-5): 1.50 (1.26) vs. 0.57 (0.73) vs. 1.86 (0.64) vs. 1.60 (1.02) vs. 1.50 (0.71) vs. 1.83 (1.07) <u>11 months</u> Spielberger State Anxiety Inventory (20-80): 42.83 (9.42) vs. 37.43 (12.26) vs. 47.17 (17.01) vs. 40.67 (11.81) vs. 46.56 (11.51) vs. 53.40 (18.78) Beck Depression Inventory (0-63): 18.67 (10.04) vs. 8.00 (5.93) vs. 12.83 (6.69) vs. 13.17 (8.51) vs. 10.56 (5.21) vs. 17.60 (6.09) Medication use (0-5): 1.17 (1.37) vs. 0.71 (0.88) vs. 1.67 (1.37) vs. 1.33 (0.75) vs. 1.44 (0.96) vs. 1.60 (1.49)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Nicholas, 1992[221] 5 months Duration of pain: 5.5 years *Fair*	A. Multidisciplinary rehabilitation (n=10), 18 sessions over 5 weeks, (total 31.5 hours) B. Exercise + attention control (psychologist-led group discussions) (n=10)	Overall Age: 44 years Female: 45% A vs. B Baseline pain (0-5 categorical scale): 3.13 vs. 2.84 Baseline function (0-100 Sickness Impact Profile): 30.87 vs. 32.10	A vs. B <u>5 months</u> Pain intensity (0-5 categorical scale): 2.89 (0.64) vs. 2.75 (1.11)	A vs. B <u>5 months</u> Beck Depression Inventory (0-63): 14.44 (5.98) vs. 18.50 (9.26) Using medication: 44% vs. 88%
Roche, 2007,[225] 2011[226] 10.75 months Duration of pain: >4 months *Fair*	A. Multidisciplinary rehabilitation (n=68), 25 sessions over 5 weeks (total 150 hours) B. Exercise therapy (n=64)	A vs. B Age: 41 vs. 39 years Female: 32% vs. 38% Baseline Pain (0-10 VAS): 4.7 vs. 4.5 Baseline function (0-100 Dallas Pain Questionnaire daily activities (0-100): 51.8 vs. 51	A vs. B <u>10.75 months</u> Pain (0-10 VAS): 2.9 vs. 3.5, difference -0.6 (95% CI -1.49 to 0.29) Dallas Pain Questionnaire daily activities (0-100): 31.4 vs. 39.1, difference -7.7 (95% CI -16.15 to 0.75)	A vs. B <u>10.75 months</u> Dallas Pain Questionnaire anxiety/depression (0-100): 21.9 vs. 25.5, difference -3.6 (95% CI -12.56 to 5.36)
Strand, 2001[209] 11 months Duration of pain: 10 vs. 9 years *Fair*	A. Multidisciplinary rehabilitation (n=81), 20 sessions over 4 weeks (total 120 hours) B. Usual Care (n=36)	A vs. B Age: 45 vs. 42 years Female: 59% vs. 64% Baseline pain (0-100 VAS): 48.3 vs. 53.0 Baseline function (0-100 Disability Rating Index): 55.6 vs. 58.3	A vs. B <u>11 months</u> Pain (0-100 VAS): -21.1 (95% CI -31 to -11) vs. -2.3 (95% CI -9.4 to 4.8) vs. -23.1 (95% CI -37 to 9.2) vs. 7.1 (95% CI -7.7 to 22), difference -1.0 (95% CI -11.7 to 9.6) Disability Rating Index (0-100): -27..3 (95% CI -34 to -21) vs. -3.3 (95 % CI -10 to 14) vs. -16.4 (95% CI -26 to -7.3) vs. 0.2 (95% CI -14 to 14), difference -3.8 (95% CI -13.9 to 6.3)	A vs. B <u>11 months</u> Working: 47% vs. 58% difference -11% (95% CI -8 to 30)
Tavafian, 2008[216] 12 months Duration of pain: 9 months *Poor*	A. Multidisciplinary program (n=37), 5 sessions over 0.5 weeks (total hours unclear) B. Medications (acetaminophen, NSAID and chlordiazepoxide)	A vs. B Age: 43 vs. 45 years Female, %: 100 vs. 100 Baseline SF-36 Physical (0-100): 41.2 vs. 42.3 Baseline SF-36	A vs. B <u>3 months</u> SF-36 Physical (0-100): 76.7 vs. 51.2, difference 25.5 (95% CI 14.69 to 36.31) <u>6 months</u> SF-36 PCS (0-100): 66.6 vs. 51.2, difference 15.4 (95% CI 2.35 to 28.45)	A vs. B <u>3 months</u> SF-36 MCS (0-100): 80.4 vs. 57.4, difference 23.0 (95% CI 10.78 to 35.22) <u>6 months</u> SF-36 MCS (0-100): 66.9 vs. 57.9, difference

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	(n=37)	Mental (0-100): 47.5 vs. 47.7	<u>6 months</u> SF-36 PCS (0-100): 64.7 vs. 51.1, difference 13.6 (95%CI -1.48 to 28.68)	9.0 (95% CI -3.88 to 21.88) <u>6 months</u> SF-36 MCS (0-100): 65.1 vs. 60.2, difference 4.9 (95%CI -7.57 to 17.37)
Turner, 1990[104] 12 months Duration of pain: 12.9 years *Poor*	A. Multidisciplinary rehabilitation (n=24), 16 sessions over 2 weeks (total 32 hours) B. Exercise (n=24)	Overall Age: 44 years Female: 48% A vs. B vs. C vs. D Baseline pain (0-78 MPQ): 25.5 vs. 19.4 Baseline function (Sickness Impact Profile): 8.5 vs. 8.4	A vs. B <u>6 months</u> McGill Pain Questionnaire Pain Rating Index (0-78): 25.5 vs. 19.4 Sickness Impact Profile (0-100): 8.5 vs. 8.4 <u>12 months</u> McGill Pain Questionnaire Pain Rating Index (0-78): 13.3 vs. 15.7 Sickness Impact Profile (0-100): 4.5 vs. 6.3	A vs. B <u>6 months</u> Center for Epidemiologic Studies-Depression Scale (0-60): 8.3 vs. 9.3 <u>12 months</u> Center for Epidemiologic Studies-Depression Scale (0-60): 10.0 vs. 9.3
van der Roer, 2008[222] 10 months Duration of pain: ~50 weeks *Fair*	A. Multidisciplinary rehabilitation (n=60), 30 sessions over 10 weeks (total hours unclear) B. Exercise (n=54)	A vs. B Age: 42 vs. 42 years Female: 55% vs. 48% Baseline pain (0-10 NRS): 6.2 vs. 5.9 Baseline function RDQ (0-24): 11.6 vs. 12.1	A vs. B <u>4 months</u> Pain (0-10 NRS): 4.1 vs. 4.8 , adjusted difference -0.97 (95% CI -1.88 to -0.06) RDQ (0-24): 7.4 vs. 7.7, adjusted difference 0.13 (95% CI -2.24 to 2.50) <u>10 months</u> Pain (0-10 VAS): 3.9 vs. 4.6, adjusted difference -1.02 (-2.14 to 0.09) RDQ (0-24): 6.7 vs. 7.1, adjusted difference 0.06 (-2.22 to 2.34)	A vs. B <u>4 months</u> Global Perceived Effect positive (%): 38.2% vs. 39.8%, OR 0.93 (95% CI 0.36 to 2.43) <u>10 months</u> Global Perceived Effect positive (%): 45.0% vs. 32.3%, OR 1.71 (95% CI 0.67 to 4.38)
Von Korff, 2005[206] 22.5 months Duration of pain: >3 months *Fair*	A. Multidisciplinary rehabilitation (n=119), 4 sessions over 5 weeks (total 4 hours) B. Usual care (n=121)	A vs. B Age: 50 vs. 50 years Female: 65% vs. 60% Baseline pain (0-10 NRS): 5.7 vs. 5.8 Modified RDQ (0-23): 12.3 vs. 11.4	A vs. B <u>4.5 months</u> Pain (0-10 NRS): 4.2 (2.0) vs. 4.7 (2.2), p=0.007 Function Modified RDQ (0-23): 9.2 (6.6) vs. 10.1 (6.4), p=0.0003 >1/3 reduction in RDQ: 42.2% vs. 23.7%, adjusted OR 3.5, p=0.0007 <u>10.5 months</u> Pain (0-10 NRS): 4.0 vs. 4.7, p=0.004 Modified RDQ (0-23): 8.4 vs. 9.1, p=0.0063 >1/3 reduction in RDQ: 44.6% vs. 22.7%, adjusted OR 2.1, p=0.03	A vs. B <u>4.5 months</u> SF-36 Social Functioning (0-100): 74.4 vs. 73.6, p=0.26 SF-36 Mental Health (0-100): 70.3 vs. 69.5, p=0.23 <u>10.5 months</u> SF-36 Social Functioning (0-100):74.4 vs. 73.6, p=0.26 SF-36 Mental Health (0-100): 70.3 vs. 69.5, p=0.23 <u>22.5 months</u>

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			22.5 months Pain (0-10 NRS): 4.3 vs. 4.6, p=0.115 Modified RDQ (0-23): 8.1 vs. 9.1, p=0.0078 >1/3 reduction in RDQ: 49.4% vs. 37.0%, adjusted OR 1.8, p=0.08	SF-36 Social Functioning (0-100): 76.7 vs. 76.3, p=0.28 SF-36 Mental Health (0-100): 71.0 vs. 72.4, p=0.98

ADL = activities of daily living; CI = confidence interval; LBO = Low Back Outcome Score; MCS = Mental Component Summary; MPQ = McGill Pain Questionnaire; NHP = Nottingham Health Profile; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; ODI = Oswestry Disability Index; PCS = Physical Component Summary; RDQ = Roland Morris Disability Questionnaire; SF-36 = Short-Form 36Q; STAI-S = Spielberger State Anxiety Inventory; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Multidisciplinary Rehabilitation Compared With Usual Care

Multidisciplinary rehabilitation was associated with slightly greater effects on function than controls at short-term (4 trials, pooled SMD -0.31, 95% CI -0.57 to -0.05, I2=70%),[204-207] and intermediate-term followup (4 trials, pooled SMD -0.37, 95% CI -0.64 to -0.10, I2=50%) (Figure 22).[206-209] There was no difference in long-term function (2 trials, pooled SMD -0.04, 95% CI -0.31 to 0.24, I2=35%).[204,206] In trials that measured function using the RDQ, the difference was 0.70 points at short-term and 1.9 points at intermediate-term. Evaluation of a high-intensity multidisciplinary rehabilitation intervention or exclusion of poor-quality trials had little effect on estimates. At short-term followup, effects on function were somewhat larger with high intensity multidisciplinary rehabilitation interventions (2 trials, pooled SMD -0.51, 95% CI -0.93 to -0.22)[204,205] than with non-high intensity interventions (3 trials, pooled difference -0.20, 95% CI -0.38 to 0.03),[205-207] but the interaction was not statistically significant (p=0.18). At intermediate-term, there were no clear differences between high intensity (1 trial, SMD -0.59, 95% CI -0.99 to -0.19)[209] and non-high intensity (3 trials, pooled difference -0.29, 95% CI -0.68 to 0.06)[206-208] interventions (p=0.47 for interaction).

Multidisciplinary rehabilitation was associated with slightly greater effects than usual care on pain at short-term (4 trials, pooled difference -0.51 on a 0 to 10 scale, 95% CI -0.89 to -0.13, I2=23%)[204-207] and intermediate-term followup (4 trials, pooled difference -0.63, 95% CI -1.04 to -0.22, I2=0%)[206-209] (Figure 23). The long-term difference was smaller and not statistically significant (2 trials, pooled difference -0.34, 95% CI -0.86 to 0.18, I2=0%).[204,206] Excluding poor-quality trials[205,208,209] had little effect on estimates. At short-term followup, effects on pain were somewhat larger with high intensity multidisciplinary rehabilitation interventions (2 trials, pooled difference -0.87, 95% CI -1.56 to -0.33)[204,205] than with non-high intensity interventions (3 trials, pooled difference -0.35, 95% CI -0.70 to 0.14),[205-207] but the interaction between intensity and effects of multidisciplinary rehabilitation was not statistically significant (p=0.45). At intermediate-term, estimates were similar for high intensity (1 trial, difference -0.53, 95% CI -1.36 to 0.30)[209] and non-high intensity (3 trials, pooled difference -0.66, 95% CI -1.19 to -0.12) interventions (p=0.81 for interaction). [206-208]

Data on other outcomes was limited. One trial found no differences between multidisciplinary rehabilitation versus usual care on the SF-36 Social Functioning or Mental Functioning subscales.[206] Three trials reported inconsistent effects on work or disability/sick

leave status.[204,206,209] Two trials found multidisciplinary rehabilitation associated with fewer health system contacts versus usual care.[204,207]

Multidisciplinary Rehabilitation Compared With Pharmacological Therapy

One poor-quality trial (n=74) found multidisciplinary rehabilitation (intensity unclear) associated with greater effects on short-term quality of life than oral medications (acetaminophen, NSAIDs, and chlordiazepoxide).[216] The difference on the SF-36 PCS was 25.5 points (95% CI 14.7 to 36.3) and on the SF-36 MCS was 23.0 points (95% CI 10.8 to 35.2). Effects were smaller at intermediate-term and statistically significant for the SF-36 PCS (difference 15.4, 95% CI 2.35 to 28.45) but not for the SF-36 MCS (difference 9.0, 95% CI -3.88 to 21.9). Effects were not statistically significant at long-term (12 month) followup (differences 13.6 and 4.9 points, respectively).

Multidisciplinary Rehabilitation Compared With Exercise

Multidisciplinary rehabilitation was associated with slightly greater effects on short-term function than exercise (6 trials, pooled SMD -0.28, 95% CI -0.54 to -0.01, I2=39%) (Figure 24).[217,219-222,224] Estimates were similar when a poor-quality trial[220] was excluded and when analyses were restricted to trials of high-intensity multidisciplinary rehabilitation (2 trials, pooled difference -0.14, 95% CI -0.47 to 0.20).[217,219] Multidisciplinary rehabilitation was associated with substantially greater effects than exercise on intermediate-term function (6 trials, pooled SMD -1.01, 95% CI -1.93 to -0.09, I2=96%), but statistical heterogeneity was very large.[104,218,220,222,223,225,226] Excluding an outlier trial (SMD -5.31, 95% CI -6.20 to -4.42)[223] eliminated statistical heterogeneity and resulted in a markedly attenuated (small) effect (5 trials, pooled SMD -0.22, 95% CI -0.40 to -0.03, I2=0%). There was no difference between multidisciplinary rehabilitation versus exercise in long-term function (3 trials, pooled SMD -1.80, 95% CI -4.36 to 0.76, I2=98%).[104,217,223] Excluding an outlier trial[223] resulted in a pooled SMD close to 0 (-0.06, 95% CI -0.36 to 0.25, I2=0%).

Multidisciplinary rehabilitation was associated with slightly greater effects on short-term pain versus exercise (6 trials, pooled difference -0.75 on a 0 to 10 scale, 95% CI -1.18 to -0.31, I2=0%) (Figure 25). Estimates were similar when one poor-quality trial[220] was excluded (5 trials, pooled difference -0.50, 95% CI -1.07 to 0.11) and there were no clear differences when analyses were stratified according to intensity of multidisciplinary rehabilitation. In two trials that evaluated high intensity multidisciplinary rehabilitation, the pooled difference was -0.56 (95% CI -1.53 to 0.36).[217,219] Estimates at intermediate-term (6 trials, pooled difference -1.17 points, 95% CI -2.70 to 0.36, I2=96%)[218,220,222,224-226] and long-term (3 trials, pooled difference -1.63, 95% CI -5.30 to 2.05, I2=99%)[104,217,223] favored multidisciplinary rehabilitation, but effects were not statistically significant. Substantial statistical heterogeneity was present in analyses of intermediate-term and long-term pain, with an outlier trial[223] that reported substantially larger effects than the other trials. For intermediate-term, the outlier trial reported a mean difference of -3.90 points, versus -0.31 to -0.78 points in the other trials. Excluding the outlier trial eliminated statistical heterogeneity and resulted in a small, statistically significance difference in intermediate-term pain that favored multidisciplinary rehabilitation (5 trials, pooled difference -0.55, 95% CI -0.95 to -0.15, I2=0%); there was no difference in long-term pain (2 trials, pooled difference 0.00, 95% CI -0.94 to 0.95, I2=50%). For intermediate-term pain, exclusion of a poor-quality trial[220] (5 trials, pooled difference -1.52, 95% CI -3.34 to 0.39) or restriction of analyses to high intensity multidisciplinary rehabilitation interventions (2 trials, pooled difference -0.60,

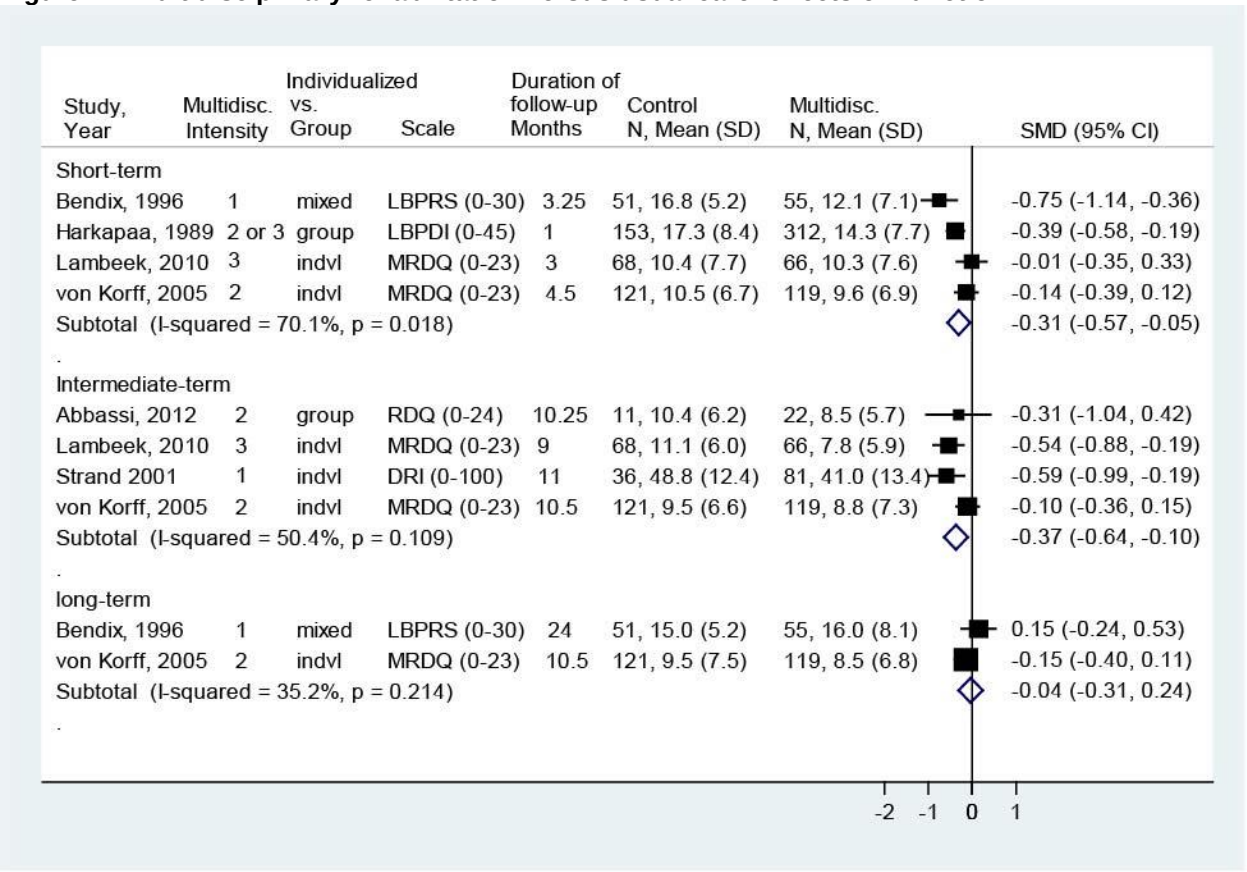
95% CI -1.41 to 0.21)[218,225,226] did not reduce heterogeneity and differences remained not statistically significant.

Data on other outcomes was limited. One trial found multidisciplinary rehabilitation associated with better scores versus exercise on SF-36 subscales at short-term followup (differences 10 to 21 points)[224] Four trials found no clear differences between multidisciplinary rehabilitation versus exercise on severity of depression.[104,219-221] Two trials found no clear effects on work status[217,225,226] and one trial found high intensity multidisciplinary rehabilitation associated with fewer days or sick leave than exercise, but non-high intensity rehabilitation associated with more days of sick leave.[217] Two trials found inconsistent effects on number of health system contacts.[217,218]

Harms

Data on harms were sparse and reported in only two trials. One study reported no clear difference between multidisciplinary rehabilitation versus exercise in risk of transient worsening of pain[224] and one trial reported no harms with either multidisciplinary rehabilitation or medications alone.[216]

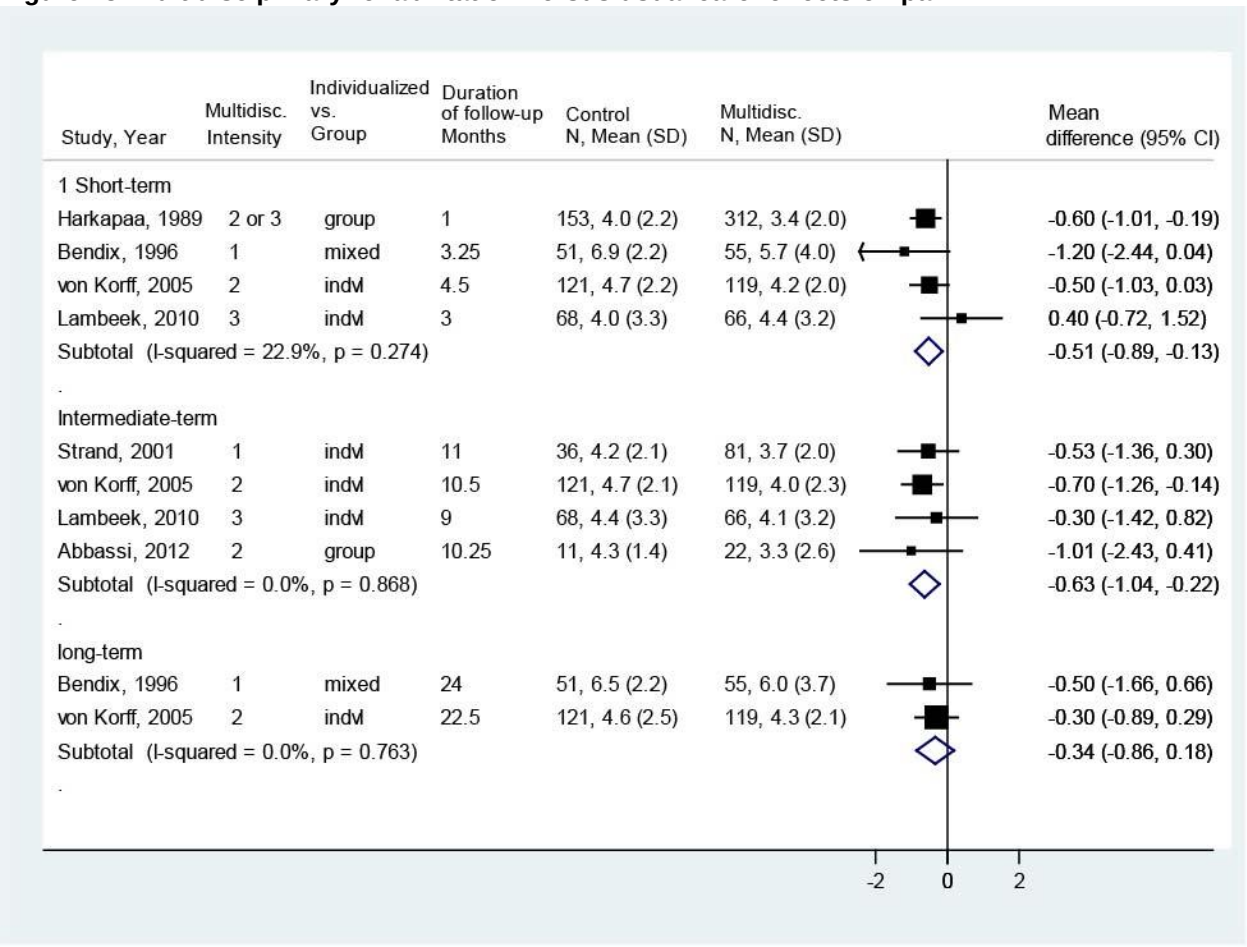
Figure 22. Multidisciplinary rehabilitation versus usual care: effects on function



CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number

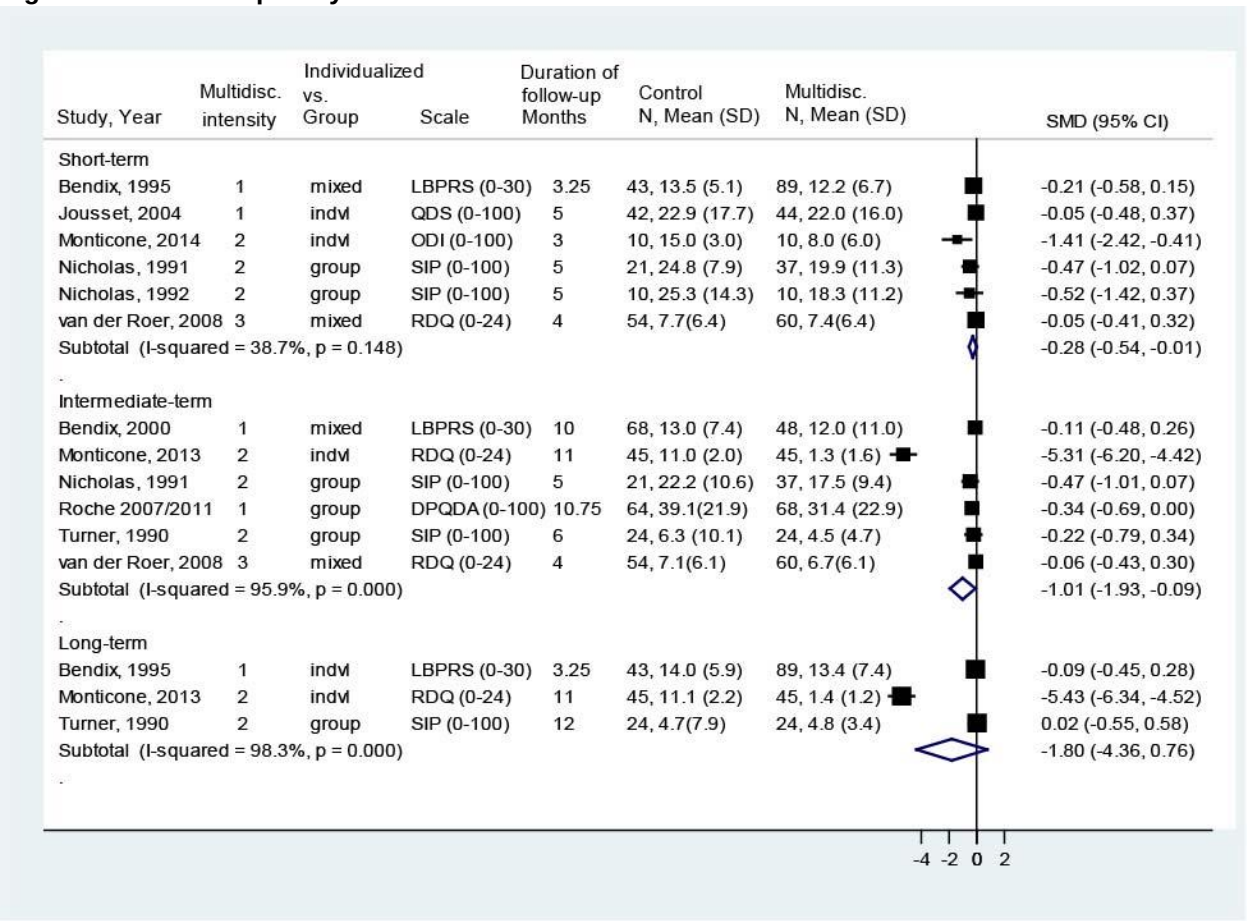
Multidisciplinary rehabilitation intensity: 1= high, 2= not high, 3= unclear or not reported

Figure 23. Multidisciplinary rehabilitation versus usual care: effects on pain



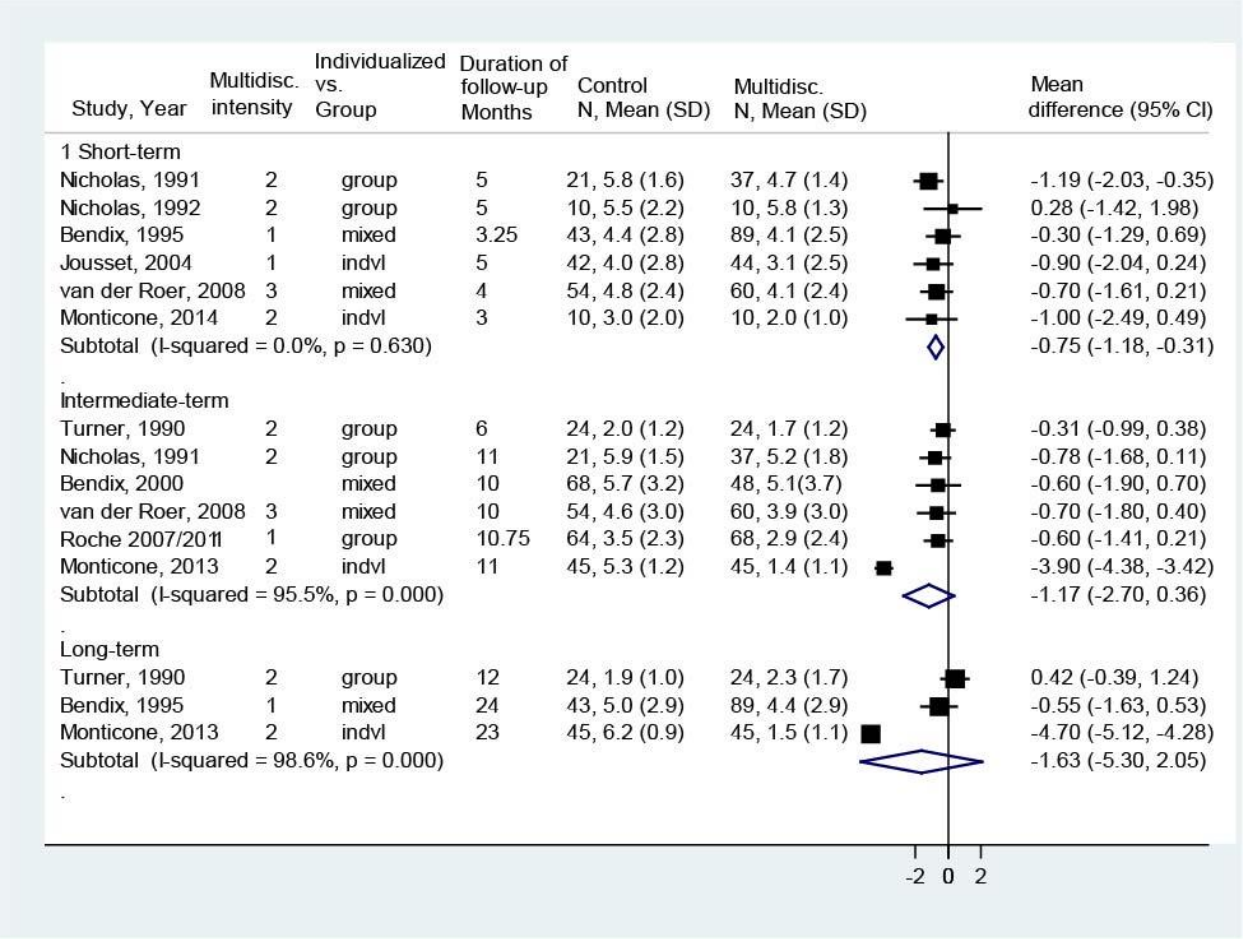
CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number.
*Multidisciplinary rehabilitation intensity: 1= high, 2= not high, 3= unclear or not reported

Figure 24. Multidisciplinary rehabilitation versus exercise: effects on function



CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number.
*Multidisciplinary rehabilitation intensity: 1= high, 2= not high, 3= unclear or not reported

Figure 25. Multidisciplinary rehabilitation versus exercise care: effects on pain



CI = confidence interval; SD = standard deviation; SMD = standardized mean difference; N = number.
*Multidisciplinary rehabilitation intensity: 1= high, 2= not high, 3= unclear or not reported

Key Question 2: Chronic Neck Pain

Exercise for Neck Pain

Key Points

- Across types of exercise, there was no clear improvement in function (3 trials, pooled standardized mean difference (SMD) -0.23, 95% CI -0.71 to 0.15) or pain (3 trials, pooled SMD -0.72, 95% CI -1.49 to 0.06) versus no treatment or advice alone in the short-term (SOE: Low).
- A subgroup of two trials of combination exercises (including 3 of the following 4 exercise categories: muscle performance, mobility, muscle re-education, aerobic) suggests a small benefit in function and pain versus no treatment or advice alone over the short-term and function in the long-term (SOE: Low).

- The effect of exercise versus NSAIDs and muscle relaxants on function and pain was indeterminate at any time period due to insufficient evidence from a single poor-quality trial (SOE: Insufficient).
- Harms were poorly reported in trials of exercise with only two trials describing adverse events. No serious harms were reported in either trial. Minor complaints included muscle pain with exercise, knee pain, and lumbar spine pain (SOE: Low).

Detailed Synthesis

Six trials of exercise therapy for neck pain met inclusion criteria (Table 18 and Appendix D).[31-35,83] Three trials evaluated participants with chronic neck pain associated with office work,[31,34,35] one included patients with chronic neck pain following whiplash,[33] one assessed participants with nonspecific neck pain,[32] and one included patients with cervical arthritis.[83] Across trials, participants were predominately female (>80%) with mean ages ranging from 38 to 52 years.

Three trials evaluated muscle performance exercises (resistive training),[31,34,35] and three combined exercise techniques.[32,33,83] Sample sizes ranged from 40 to 230 (total sample=697). Three trials compared exercise versus an attention control,[31,33,35] one versus no treatment,[34] one versus waitlist[32] and one versus pharmacologic care.[83] Four trials were conducted in Europe,[31,32,34,35] one in Australia[33] and one in Turkey.[83] The duration of exercise therapy ranged from 6 weeks to 12 months, and the number of supervised exercise sessions ranged from 3 to 52. Three trials reported outcomes through long-term followup,[31,33,35] two through intermediate-term followup,[34,83] and one evaluated only short-term outcomes.[32]

Two trials were rated fair quality[33,34] and four poor quality[31,32,35,83] (Appendix E). In the two fair-quality trials, the main methodological limitation was the inability to blind interventions. Limitations in the other trials included inability to blind interventions, unclear randomization and allocation concealment methods, unclear or high loss to followup, and baseline differences between intervention groups.

Table 18. Summary of results for neck pain: exercise therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Andersen, 2008[b31] 6 and 12 months Duration of pain, NR *Poor*	A. Dynamic strengthening exercise (muscle performance exercise) (n=61): for the neck/shoulder muscles, performed in in the workplace; 20 minute sessions, 3 times a week (2 of the 3 weekly sessions were supervised by experienced instructors) B. Lifestyle physical	A + B + C Age: 45 years Female: 78% Office workers: 100% A vs. B vs. C Pain VAS (0-10): 5.0 vs. 5.0 vs. 4.7	A vs. C <u>6 months</u> Pain VAS: 3.4 vs. 4.2, difference -0.80 (-0.87 to -0.73) <u>12 months</u>[c] Pain VAS: 3.8 vs. 4.6, difference -0.80 (-0.87 to -0.73) Days of pain in last 3 months (0-90): 25 vs. 30, p>0.05 B vs. C	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	exercise and activity increase (combination exercise) (n=59): workplace activities such as steppers placed near the copying machines, punch bags in the hall, group sessions of Nordic walking, and strength and aerobic fitness exercise programs C. Control group (n=62): ergonomics, stress management, organization of work, cafeteria food quality Treatment lasted 1 year. All groups were allowed 1 hour per week during working time for activities		<u>6 months</u> Pain VAS: 3.6 vs. 4.2, difference -0.60 (-0.67 to -0.53) <u>12 months</u>[c] Pain VAS: 3.6 vs. 4.6, difference -1.0 (-1.07 to -0.93) Days of pain in last 3 months: 26 vs. 30, p>0.05	
Aslan Telci, 2012[83] 6 months Pain duration: 12 months *Poor*	A. Combination exercises (n=20): consisting of posture, active range of motion, stretching, isometric and dynamic strengthening and endurance exercises, relaxation and proprioception exercises. Clinic followup once a week to maintain motivation and check whether exercises performed correctly for a total of 3 weeks and home exercise for at least another month. B. NSAIDs and muscle relaxants for 15 days (n=20): all patients received verbal advice regarding pain control, posture, and ergonomics.	A vs. B Age: 48 vs. 52 years Female: 85% vs. 75% BMI: 25 vs. 27 Employed: 50% vs. 40% Education year: 12 vs. 11 Pain VAS (0-10): 6.7 vs. 6.4 NDI (0-50): 14.0 vs. 10.7	A vs. B <u>3 month</u> NDI: 9.4 vs. 11.5, difference -2.2 (95% CI -5.8 to 1.5) Pain VAS: 4.1 vs. 5.1, difference -1.0 (95% CI -2.3 to 0.3) <u>6 month</u> NDI: 11.9 vs. 13.7, difference -1.8 (95% CI -5.7 to 2.1) Pain VAS: 4.5 vs. 5.3, difference -0.8 (95% CI -2.3 to 0.7)	A vs. B <u>3 month</u> NHP (0-100): 89.2 vs. 230.0, difference -140.8 (95% CI -214.0 to -67.5) BDI (0-63): 6.8 vs. 10.7, difference -4.0 (95% CI -8.4 to 0.5) <u>6 month</u> NHP: 122.3 vs. 257.6, difference -135.3 (95% CI -209.1 to -61.5) BDI: 8.3 vs. 11.8, difference -3.8 (95% CI -8.5 to 1.0)
Lauche, 2016[32]	A. Combination exercises (n=37): weekly 60-75 minute	A vs. B Age: 47 vs. 49 years Female: 86% vs.	A vs. B <u>3 month</u> NDI: 25.1 vs. 29.4,	A vs. B <u>3 month</u> SF-36 PCS (0-100):

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
3 months Pain duration: NR *Poor*	session for 12 weeks; ergonomic principles, proprioceptive exercises, and isometric and dynamic mobilization, stretching, strengthening neck and core exercises, and relaxation exercises; illustrated written exercises for home use ≥15 minutes/day. B. Wait list (n=39): continuing usual activities/therapies	82% years Pain recently (0-100): 46.2 vs. 51.5 Pain considered tolerable (0-100): 20.5 vs. 20.7	difference -4.3 (95% CI -10.2 to 1.6) Recent pain VAS: 33.1 vs. 44.6, difference -11.5 (95% CI -20.8 to -2.2) Pain with motion VAS: 34.9 vs. 45.5, difference -10.6 (95% CI -18.5 to -2.7)	difference 2.0 (95% CI -1.6 to 5.6) SF-36 MCS (0-100): difference 0.5 (95% CI -3.9 to 4.9)
Stewart, 2007[33] 1.5 and 12 months Pain duration: 9 months *Fair*	A. Combination exercise, plus advice (n=66); aerobic, stretching, functional, speed and endurance, trunk and limb strengthening; 1 hour per session for 12 session over 6 weeks B. Advice alone (n=68): included reassurance of a favorable outcome and encouragement to resume light activity	A vs. B Age: 44 vs. 43 years Female: 73% vs. 62% PSFS (0-10): 3.9 vs. 4.1 NDI (mean, 0-50): 18.2 vs. 19.7 Pain VAS (mean, 0-10): 5.2 vs. 5.3	A vs. B <u>1.5 months</u> PSFS: 6.4 vs. 5.6, difference 0.9 (95% CI, 0.3 to 1.6) NDI: 12.0 vs. 15.7, difference -2.7 (95% CI, -4.5 to -0.9) Pain VAS: 3.2 vs. 4.3, difference -1.1 (95% CI -1.8 to -0.3) <u>12 months</u> PSFS: 6.6 vs. 6.0, difference 0.6 (95% CI, -0.1 to 1.4) NDI: 12.1 vs. 15.5, difference -2.3 (95% CI, -4.9 to 0.3) Pain VAS: 3.5 vs. 3.8, difference -0.2 (95% CI 0.6 to -1.0)	A vs. B <u>1.5 months</u> Bothersomeness (0-10) 3.6 vs. 4.8, p=0.019 SF 36 physical (0-100): 42.1 vs. 38.9, p=0.003 SF 36 mental (0-100): 51.4 vs. 46.4, p=0.005 Global Perceived Effect (-5 to 5) 2.5 vs. 1.5, p=0.006 <u>12 months</u> Bothersomeness 4.1 vs. 4.0, p=0.480 SF 36 physical: 42.3 vs. 38.9, p=0.003 SF 36 mental: 48.4 vs. 46.1, p=0.33 Global Perceived Effect: 2.3 vs. 1.9, p=0.48
Viljanen, 2003[34] 3 and 9 months Pain duration: 11 years *Fair*	A. Dynamic strengthening exercises (muscle performance exercises) (n=135): physical-therapist guided; 3 times per week for 12 weeks, 30 minute sessions B. No intervention (n=130)	A vs. B Age: 45 vs. 44 years Female: 100% vs. 100% Office workers: 100% Computer work >6 hours per day: 33% vs. 35% Depression index (10-40): 16 vs. 16	A vs. B <u>3 months</u> Neck disability scale[e]: 15 vs. 14, adjusted difference -0.1 (95% CI -3.1, 2.9) Pain VAS: 2.9 vs. 2.9, adjusted difference 0.4 (95% CI -0.3, 1.0) <u>9 months</u> Neck disability scale[e]: 19 vs. 17, adjusted difference -	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
		Neck disability scale[e] (0-80): 29 vs. 26 Pain VAS (0-10): 4.8 vs. 4.1	0.1 (95% CI -3.0 to 2.9) Pain VAS: 3.1 vs. 3.2, adjusted difference 0.5 (95% CI -0.1 to 1.0)	
Waling 2002[d35] 6 and 36 months Pain duration: 6.8 years *Poor*	A. Strength training (muscle performance exercise) (n=29): for neck and shoulder muscles, 3 times per week for 10 weeks, 1 hour/session B. Endurance training (muscle performance exercise) (n=28): using arm-cycling and arm exercises, 30 repetition maximum, 3 times per week for 10 weeks, 1 hour/session C. Coordination training (neuromuscular reeducation exercises) (n=25): focus on balance and postural stability 3 times per week for 10 weeks, 1 hour/session D. Reference group (n=21): stress management 1 time per week for 10 weeks, 2 hour/session	A vs. B vs. C vs. D Age: 38 vs. 39 vs. 38 vs. 39 years Female: 100% all groups Office workers: 100% Pain VAS at present (0-10): 2.6 vs. 2.8 vs. 3.3 vs. 3.7	A vs. B vs. C vs. D <u>6 months</u> Frequent pain (% with pain several times per week or more): 76% vs. 91% vs. 78 % vs. 73%, p=0.50 <u>3 years</u> Pain VAS at present: 3.1 vs. 2.2 vs. 2.7 vs. 1.6, p=0.073 Pain VAS in general (0-10): 3.2 vs. 2.9 vs. 2.9 vs. 2.0, p=0.249 Pain VAS at worst (0-10): 6.1 vs. 5.8 vs. 5.7 vs. 5.8, p=0.902 Frequent pain: 47% vs. 50% vs. 58% vs. 39%, p=0.66	NR

BDI, Beck Depression Inventory; BMI, body mass index; CI, confidence interval; HADS, Hospital Anxiety and Depression Scale; MCS: Mental Component Summary; MD, mean difference; NDI, Neck Disability Index; NHP, Nottingham Health Profile; NR, not reported; NSAID = nonsteroidal anti-inflammatory drug; PCS: Physical Component Summary; PSFS, Patient Specific Functional Scale; SF-36, Short-Form 36 questionnaire; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Cluster RCT where clusters were formed from participants working on the same floor

[c] Intervention lasted 12 months and followup is at the end of the intervention

[d] Cluster RCT where clusters were formed from participants selecting a time that best fit their schedule

[e] Neck disability scale was created by investigators from responses to eight questions related to functional limitations due to pain; this scale is not the same as the more common Neck Disability Index (NDI).

Exercise Compared With No Treatment or an Attention Control

Across types of exercise, there was no clear improvement in function versus no treatment or advice alone in the short term (3 trials, pooled standardized mean difference (SMD) -0.23, 95% CI -0.61 to 0.15, I2=72.6%)[32-34] (Figure 26). However, two studies that included combination exercises (3 of the following 4 exercise categories: muscle performance, mobility, muscle re-

education, aerobic) found small improvement in function compared with controls (2 trials, pooled SMD -0.44, 95% CI -0.76 to -0.09, data not shown in figure).[32,33] A fair-quality study reported a continued small benefit with combination exercise in the long term (SMD -0.38, 95% CI -0.74 to -0.03).[33]

Exercise tended toward moderately greater effects on short-term pain compared with no treatment or an attention control (3 trials, pooled SMD -0.72, 95% CI -1.49 to 0.06, I2=63.7%)[32-34] (Figure 27). The effect of exercise on reducing pain was substantially greater in trials assessing combination exercises (2 trials, pooled SMD -1.12, 95% CI -1.82 to -0.43; data not shown in figure).[32,33] There were no differences in pain comparing exercise versus controls in intermediate-term (2 trials, pooled difference -0.25, 95% CI -0.71 to 0.20, I2=0%)[31,34]or long-term (3 trials, pooled difference 0.12, 95% CI -0.52 to 0.76, I2=37.8%).[31,33,35]

Data on effects of exercise on quality of life were limited. One fair-quality trial[33] found significant improvement in SF-36 PCS and MCS in the short term (difference in change score 3.60 on a 0-100 scale, 95% CI 1.23 to 5.97 and 4.00, 95% CI 1.24 to 6.77, respectively) and PCS in the long term (difference in change score 3.80, 95% CI 1.30 to 6.30). A poor-quality trial found no difference in SF-36 PCS or MCS in the short term.[32] No trial evaluated effects of exercise therapies on use of opioid therapies or health care utilization.

There was insufficient evidence to determine effects of duration of exercise therapy or number of sessions on outcomes.

Exercise Compared With Pharmacological Therapy

One poor quality trial (N=40)[83] comparing 1.5 months of home combination exercises (posture, stretching, strengthening and endurance exercises) versus ibuprofen plus thiocolchicoside for 15 days found no between-group difference in function (Neck Disability Index [NDI]) at 3-month (mean difference of -2.2 on 0-50 scale, 95% CI -5.8 to 1.5) or 6-month followup (mean difference of -1.8, 95% CI -5.7 to 2.1). The study reported similar results for pain intensity (mean difference of -1.0 on a 0-10 scale, 95% CI -2.3 to 0.3 at 3 months and mean difference -0.8, 95% CI -2.3 to 0.7 at 6-month followup). The exercise group reported a better quality of life compared with the medication group at 3-month and 6-month followups using the Turkish version of the Nottingham Health Profile (MD -141, scale not stated though usual scale 0-100, 95% CI -214 to -68; MD -135, 95% CI -209 to -62, respectively).[83] The groups scored comparably on the Beck Depression Inventory at both followup periods (Table 18).

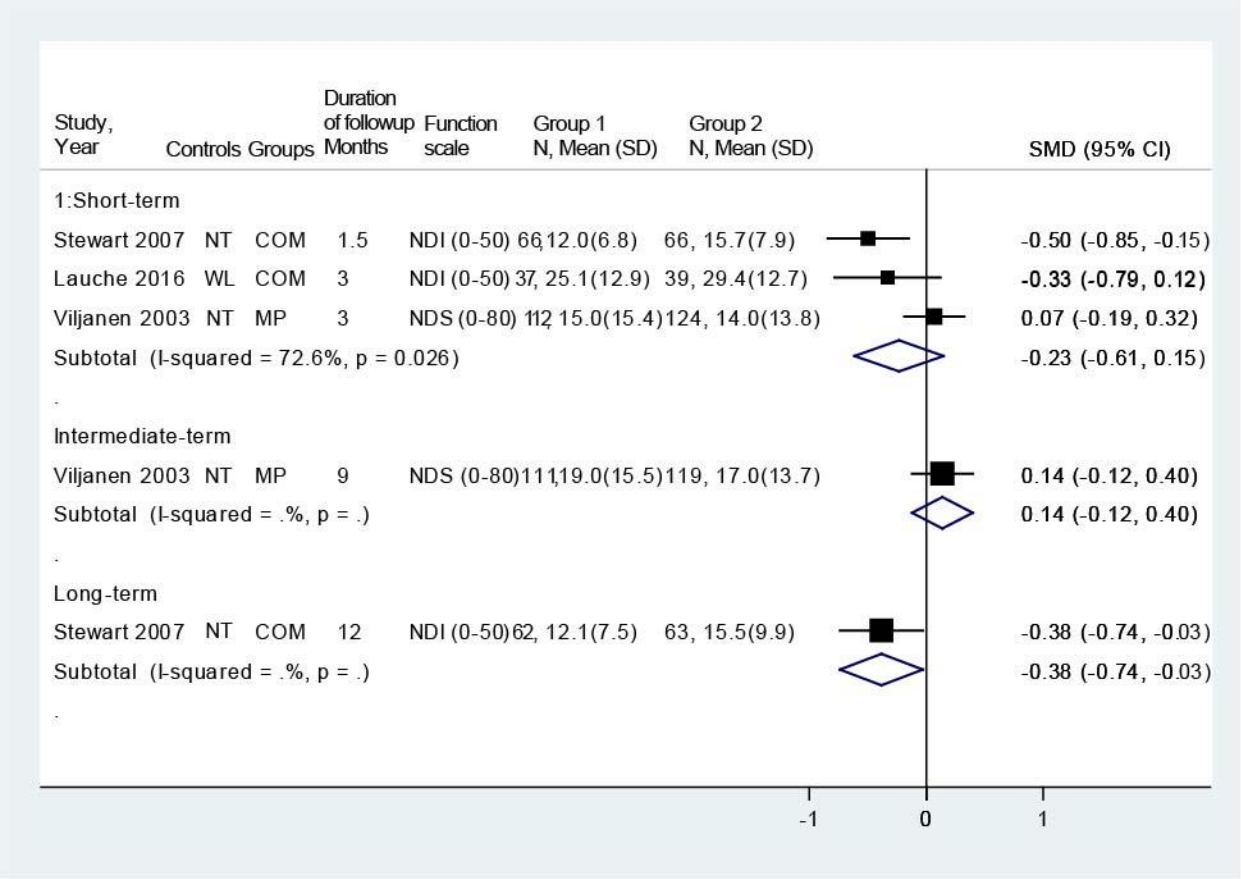
Exercise Compared With Other Non-pharmacological Therapies

Findings for exercise versus other nonpharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

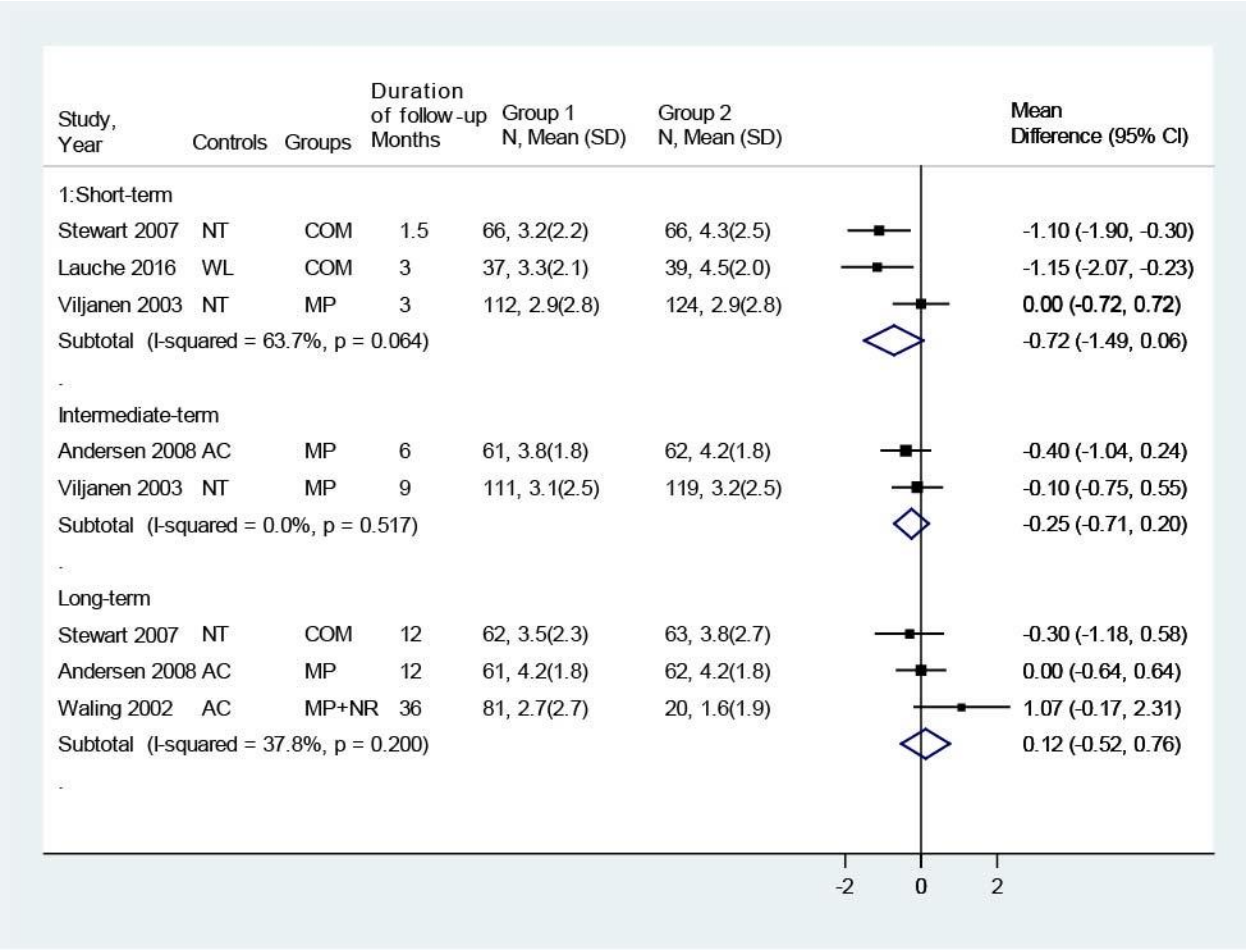
Only two exercise trials reported harms. One reported only mild complaints that included muscle pain with exercise (5%), knee pain (3%), and lumbar spine pain (3%).[33] None required referral to a medical practitioner. In the other, investigators reported no serious harms related to the intervention.[32] One occurrence of minor knee pain was reported in the exercise group.

Figure 26. Exercise versus no treatment or an attention control for neck pain: effects on function



CI = confidence interval; COM = combination exercise therapy; MP = muscle performance exercise; NDI = Neck Disability Index; NDS = neck disability scale; NT = no treatment; SD = standard deviation; SMD = standardized mean difference; WL = waitlist.

Figure 27. Exercise versus no treatment or an attention control for neck pain: effects on pain



AC = attention control; CI = confidence interval; COM = combination exercise therapy; MP = muscle performance exercise; MP+NR = muscle performance plus neuromuscular rehabilitation exercise; NT = no treatment; SD = standard deviation; WL = waitlist

Psychological Therapies for Neck Pain

Key Points

- No difference in function (NDI, 0-80 scale) or pain (VAS, 0-10 scale) in the short-term (adjusted difference 0.1, 95% CI -2.9 to 3.2 and 0.2, 95% CI -0.4 to 0.8, respectively) or intermediate-term (adjusted difference 0.2, 95% CI -2.8 to 3.1 and 0.2, 95% CI -0.3 to 0.8, respectively) from one fair quality study comparing relaxation training and no intervention or exercise (SOE: Low for all). We found no trials with outcomes assessed in the long term.
- We found no evidence comparing relaxation training with pharmacological therapy.
- The only trial of relaxation training did not report harms.

Detailed Synthesis

We found one trial comparing the effects of relaxation training versus no intervention (N=258) or exercise therapy (N=263) in female office workers with chronic neck pain.[34] (Table 19 and Appendix D). Relaxation training and muscle performance exercise therapy were done in 30-minute sessions 3 times per week for 12 weeks, with one week of reinforcement training 6 months after randomization. Patients in the no-treatment group were instructed not to change their usual activities. Adherence to the relaxation schedule during the intervention period was 42 percent of the scheduled sessions. The nature of the intervention and control precluded blinding of participants and people administering the interventions; therefore, this trial was rated as fair quality.

Table 19. Summary of results for neck pain: psychological therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Viljanen, 2003[34] 3 and 9 months Pain duration: 11 years *Fair*	A. Physical therapist guided relaxation training (n=128): progressive relaxation, autogenic training, functional relaxation, and systematic desensitization (goal was to teach correct activation and relaxation of muscles used in daily activities); 3 times per week for 12 weeks, 30 minute sessions B. Physical therapist guided dynamic strengthening exercises of the shoulder and cervical musculature (muscle performance exercises) (n=135): 3 times per week for 12 weeks, 30 minute sessions C. No intervention (n=130)	A vs. B vs. C Age: 43 vs. 45 vs. 44 years Female: 100% Performing physical activity ≥3x/week: 34% vs. 44% vs. 41% Duration of office work: 20 vs. 23 vs. 21 years Sedentary work >6 hours per day: 75% vs. 76% vs. vs. 73% Computer work >6 hours per day: 39% vs. 33% vs. vs. 35% Absent from work due to neck pain: 12% vs. 12% vs. 12% Pain duration: 11 vs. 11 vs. 10 years Neck disability scale[a] (0-80): 29 vs. 29 vs. 26 Pain VAS (0-10): 4.8 vs. 4.8 vs. 4.1 Depression index: 16 vs. 16 vs. 16	A vs. C <u>3 months</u> Neck disability scale[b]: 15 vs. 14, adjusted difference 0.1 (95% CI -2.9 to 3.2) Pain VAS: 3.0 vs. 2.9, adjusted difference 0.2 (95% CI -0.4 to 0.8) <u>9 months</u> Neck disability scale[b]: 19 vs. 17, adjusted difference 0.2 (95% CI -2.8 to 3.1) Pain VAS: 3.3 vs. 3.2, adjusted difference 0.2 (95% CI -0.3 to 0.8) A vs. B <u>3 months</u> Neck disability scale[a]: 15 vs. 15, adjusted difference 0.2 (95% CI -2.8 to 3.2) Pain VAS: 3.0 vs. 2.9, adjusted difference -0.2 (95% CI -0.8 to 0.4) <u>9 months</u> Neck disability scale[a]: 19 vs. 19; adjusted difference 0.2 (95% CI -2.7 to 3.2) Pain VAS: 3.3 vs. 3.1, adjusted difference -0.2 (95% CI -0.8 to 0.3)	NR

CI, confidence interval; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Neck disability scale was created by investigators from responses to eight questions related to functional limitations due to pain. This scale is not the same as the more common Neck Disability Index (NDI).

Relaxation Training Compared With No Treatment

The one fair-quality trial found no between-group differences in the short term (3 months) or intermediate term (9 months) as measured by a neck disability scale (mean difference 0.1 on a 0-80 scale, 95% CI -2.9 to 3.2, and mean difference 0.2, 95% CI-2.8 to 3.1, respectively),[34] (Table 19). The neck disability scale, a non-validated instrument, asked whether the participant had pain or difficulty on eight functional activities, with each activity scored from 0 (no pain or hindrance) to 10 (unbearable pain or maximum hindrance), for a total of 80 points. Likewise, there were no differences in pain intensity between groups at the same time frames, (mean difference 0.2 on a 10-point scale, 95% CI -0.4 to 0.8, and mean difference 0.2, 95% CI -0.3 to 0.8, respectively). There were no trials evaluating relaxation in the long term.

Relaxation Training Compared With Pharmacological Therapy

We did not find any trials meeting our criteria that compared a relaxation training with pharmacological therapy.

Relaxation Training Compared With Exercise Therapy

The one fair-quality trial found no differences between relaxation training and exercise therapy in the short term (3 months) or intermediate term (9 months) as measured by a neck disability scale described above (mean difference 0.2 on a 0-80 scale, 95% CI -2.8 to 3.2, and mean difference 0.2, 95% CI -2.7 to 3.2, respectively),[34] (Table 19). Similarly, there were no differences in pain intensity between groups at the same time frames, (mean difference -0.2 on a 10-point scale, 95% CI -0.8 to 0.4, and mean difference -0.2, 95% CI -0.8 to 0.3, respectively). There were no trials comparing relaxation with exercise therapy in the long term.

Harms

The trial on relaxation therapy did not report harms.[34]

Physical Modalities for Neck Pain

Key Points

- Low-level laser therapy was associated with a moderate improvement in short-term function (2 trials, pooled difference -14.98 , 95% CI -23.88 to -6.07, I^2=39%, 0-100 scale) and pain (3 trials, pooled difference -1.81 on a 0-10 scale, 95% CI -3.35 to -0.27, I^2=75%) compared with sham (SOE: Moderate for function and pain).
- Data from two small, poor-quality trials, one evaluating cervical traction versus attention control (infrared irradiation) and the other electromagnetic fields versus sham, were insufficient to determine effects on function, pain, or harms over the short term (SOE: Insufficient).
- No trials assessed outcomes in the intermediate term or long term, or compared a physical modality to pharmacological therapy or exercise.
- Harms were poorly reported in trials of low-level laser. Adverse effects occurred with similar frequency in the laser and sham groups in the one trail reporting such effects. The most frequently reported adverse effects included mild (78%) or moderately (60%) increased neck pain, increased pain elsewhere (78%), mild headache (60%), and tiredness (24%) (SOE: Low).
- The trials of cervical traction and electromagnetic fields did not report harms.

Detailed Synthesis

A total of five trials (N=53 to 90)[114-118] evaluating physical modalities for the treatment of chronic neck pain met inclusion criteria (Table 20 and Appendixes D and E). Interventions included traction, laser therapy, and electromagnetic field therapy.

One trial (N=79) conducted in Hong Kong compared intermittent cervical traction versus attention control (infrared irradiation)[115] Each treatment was administered for 20 minutes twice weekly for 6 weeks. This trial was considered poor quality due to lack of patient and caregiver blinding, high and unequal attrition (41% in traction group, 58% in control), and dissimilar baseline characteristics between groups.

Three trials (N=53 to 90)[114,117] compared low level laser therapy with sham. The mean duration of pain varied from 4 years in two trials[114,117] to 15 years in a third.[116] Treatment consisted of laser application (wavelength range, 830 to 904 nm) over several myofascial tender points; across the trials, duration ranged from 30 seconds to 3 minutes per tender point and frequency varied from daily to twice weekly over periods of 2 or 7 weeks. One trial was rated good quality[116] and two fair quality[114,117] Common methodological limitations in the two fair-quality trials included inadequate reporting of treatment allocation and no or unclear blinding of the care provider. In addition, baseline characteristics were not similar in one trial, in which the intervention group tended to have more pain and tenderness and longer duration of symptoms.[114]

One trial (N=81) compared the effects of 18, 30-minute sessions (3-5 times per week) of low frequency pulsed electromagnetic fields versus sham.[118] The treatment consisted of an electromagnetic coil against the back of the neck while the participants were lying on a pillow. The investigators covered the set of light emitting diodes that pulse to signal the coil being energized in order to blind the participants to the treatment or sham. This trial was rated as poor quality due to several factors: failure to describe the number randomized in each group; inadequate reporting of treatment compliance and information to calculate participant attrition and intent to treat analysis; care provider not blinded to treatment; and baseline characteristics dissimilar between groups.

Table 20. Summary of results for neck pain: physical modalities

Author, Year, Followup[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Altan, 2005[114] 3 months Pain duration: 4.5 years *Fair*	A. GaAs low level laser treatment (n=26): over the 3 trigger points bilaterally and 1 point in the taut bands in trapezius muscle bilaterally for 2 min over each point once a day for 2 weeks. Laser wavelength of 904 nm. B. Sham laser treatment (n=27)	A vs. B Age: 43 vs. 43 years Female: 87% vs. 48% Pain duration: 4.7 vs. 4.4 years Pain (VAS 0-10): 6.85 vs. 6.24 Pain (5-point scale, 0-5): 2.35 vs. 2.20	A vs. B <u>3 months:</u> Pain (VAS): 3.17 vs. 3.80, difference -0.63 (95% CI -0.95 to -0.31) Pain (5 point scale): 1.09 vs. 1.16, difference -0.07 (95% CI -0.19 to 0.05)	NR
Chiu, 2011[115] 1.5 months	A. Cervical Traction (intermittent) (n=39): ranging from 10-20% of patient body weight, holding time	A vs. B Age: 50.9 vs. 46.8 years	A vs. B <u>1.5 months</u> NPQ Disability[b]:	NR

Author, Year, Followup[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Duration of pain: NR *Poor*	10-25 seconds; resting time 20-50% of holding time; twice/week for 6 weeks; sessions lasting 20 minutes. B. Infrared Irradiation Control (n=40): via infrared lamp positioned so that patients reported minimal warmth over the back of their neck; twice/week for 6 weeks; sessions lasting: 20 minutes.	Female: 65.2% vs. 76.5% NPQ (0-100%): 46.1 vs. 38.5 NPS (0-10): 5.8 vs. 5.2	31.4 vs. 29.6; p > 0.05, 95%CI, 29.66 to 37.50, power=0.15 NPS Pain Severity[b]: 3.5 vs. 2.8; p > 0.05, 95%CI, 3.29 to 4.50, power=0.17	
Chow, 2006[116] 1 month Pain duration: 15 years *Good*	A. Low level laser therapy (n=45): 2x/week for 7 consecutive weeks, maximum half hour per treatment. Up to 50 tender points in the neck were treated for 30 seconds per point. Laser wavelength of 830 nm. B. Sham laser (n=45)	A vs. B Age: 57 vs. 55 years Female: 64% vs. 67% Pain duration: 17 vs. 13 years Pain (VAS 0-10): 5.9 vs. 4.0	A vs. B <u>1 month</u> NPQ (0-100%): -3.5 vs. -0.6, difference -3.0 (95% CI -5.0 to -0.9) NPAD (0-100): -15.2 vs. -3.1, difference -12.1 (95% CI -19.3 to -4.8) Pain, VAS: -2.7 vs. 0.3, difference 3.0 (95% CI -3.8 to -2.1) MPQ VAS (1-5): -2.1 vs. 0.1, difference -2.2 (95% CI -3.5 to -0.9) Improved pain <-3 (%): 40% vs. 7%, RR 6.0 (95% CI 1.9 to 19.0)	A vs. B <u>1 month</u> SF36 PCS: 3.2 vs. -1.3, difference 4.5 (95% CI 0.7 to 8.2) SF 36 MCS: 2.4 vs. 5.4, difference -2.9 (95% CI -7.2 to 1.3), MPQ sensory (0-33): -3.4 vs. -1.9, difference -1.5 (95% CI -4.5 to 1.5) MPQ affective (0-12): -1.3 vs. -0.7, difference -0.6 (95% CI -2.3 to 1.1)
Gur, 2004[117] 2.5 months Pain duration: 43 months *Fair*	A. Active Ga-As low level laser therapy (n=30): daily for 2 weeks, 3 minutes each myofascial tender point. Laser wavelength of 904 nm. B. Sham laser (n=30)	A vs. B Age: 32 vs. 31 years Female: 82% (total pop only) Pain duration: 43 vs. 43 months Employed: 12% vs. 17% Pain at rest (VAS 0-10): 7.43 vs. 6.87 Pain at movement (VAS 0-10): 7.43 vs. 7.19 NPAD (0-100): 65.36 vs. 68.52 NHP (0-100): 78.9 vs. 75.5 BDI (0-63): 21.56	A vs. B <u>2.5 months</u> NPAD: 41.14 vs. 63.29, difference -22.15 (95% CI -36.7 to -7.6) Pain at rest (VAS): 4.18 vs. 6.29, difference -2.11 (95% CI -3.80 to -0.42) Pain at movement (VAS): 5.26 vs. 7.28, difference -2.02 (95% CI -3.31 to -0.73)	A vs. B <u>2.5 months</u> BDI: 14.72 vs. 21.38, difference -6.66 (95% CI -13.24 to -0.08) NHP: 56.41 vs. 72.48, difference -16.1 (95% CI -30.9 to -1.3),

Author, Year, Followup[a] Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
		vs. 20.81		
Trock, 1994[118] 1 month Pain duration: 7.5 years *Poor*	A. Pulsed electromagnetic fields (n=42 treated): extremely low frequency (<2 A, 120 V) applied with stepwise energy characteristics as follows: 5 Hz, 0-15 gauss for 10 minutes; 10 Hz, 15-25 gauss for 10 minutes; and 12 Hz, 15-25 gauss for 10 minutes. Maximum number of pulses/burst was 20. B. Sham (n=39 treated) Treatments were given for 30 minute periods, 3-5 times per week for 18 treatments.	A vs. B Age: 61 vs. 67 years Female: 71% vs. 67% Weight (lb): 161 vs. 162 Duration of symptoms: 7 vs. 8 years Pain (0-10): 7.20 vs. 6.23 ADL difficulty (0-24) 11.9 vs. 11.5	A vs. B <u>1 month:</u> ADL difficulty: 3.78 vs. 2.14, difference 1.6 (95% CI -1.5 to 4.8) Pain: 2.59 vs. 1.47, difference 1.12 (95% CI -0.31 to 2.55)	A vs. B <u>1 month:</u> Patients' assessment of improvement (0-100): 41.2 vs. 40.0, difference 1.2 (95% CI -15.2 to 17.6)

ADLs, activities of daily living; BDI, Beck Depression Inventory; BMI, body mass index; CI, confidence interval; Ga-As, Gallium Arsenide; HADS, Hospital Anxiety and Depression Scale; MD, mean difference; MPQ: McGill Pain Questionnaire; NDI, Neck Disability Index; NHP, Nottingham Health Profile; NPAD, Neck Pain and Disability Scale; NPQ, Northwick Park Questionnaire; NPS, numerical pain scale; NR, not reported; PSFS, Patient Specific Functional Scale; RR, risk ratio; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Results of two-way repeated measures ANOVA

Physical Modalities Compared With Attention Control or Sham

Traction. One poor-quality trial found no short-term differences in function comparing intermittent cervical traction versus attention control (infrared irradiation) using the Northwick Park Questionnaire (NPQ) (mean difference -1.8, 95% CI -10.8 to 7.2, 0-100% scale).[115] Likewise, there was no difference in pain intensity between groups, (mean difference -0.7, 95% CI -2.2 to 0.8, 10 point scale). There were no trials evaluating cervical traction in the intermediate term or long term.

Low-level laser therapy. Laser was associated with moderately greater effects compared with sham on short-term function (2 trials, pooled difference -14.98, 95% CI -23.88 to -6.07, I^2=39%, 0-100 scale) (Figure 28)[116,117] and short-term pain (3 trials, pooled difference -1.81 , 95% CI -3.35 to -0.27, I^2=75%, 0-10 scale) (Figure 29).[114,116,117] Pain improvement of greater than -3.0 on a 10-point VAS scale was substantially more common with laser therapy in the good-quality trial (risk ratio 6.0, 95% CI 1.9 to 19.0).[116] Quality of life improvement also favored low-level laser as measured by the SF-36 PCS (mean difference 4.5, 95% CI 0.7 to 8.2)[116] and the Nottingham Health Profile (mean difference -16.1 on a 0-100 scale, 95% CI -30.9 to -1.3).[117] Measures demonstrating no difference between groups included the SF36 MCS and the McGill Pain Questionnaire component scores,[116] (Table 20). There were no trials evaluating laser therapy in the intermediate term or long term.

Electromagnetic Fields. One poor-quality trial found no between-group differences in short-term difficulty with activities of daily living (ADLs) (mean difference 1.6 , 95% CI -1.5 to 4.8, scale

0-24, non-validated measure).[118] The ADL instrument asked whether the participant had pain or difficulty on eight activities scored from 0 (never) to 3 (always), for a total of 24 points. Likewise, there was no difference in pain intensity between groups, (mean difference 1.1, 95% CI -0.3 to 2.6, 0-10 scale) or in patients' assessment of improvement (mean difference 1.2, 95% CI -15.2 to 17.6, 0-100 scale).[118] There were no trials evaluating electromagnetic fields in the intermediate term or long term.

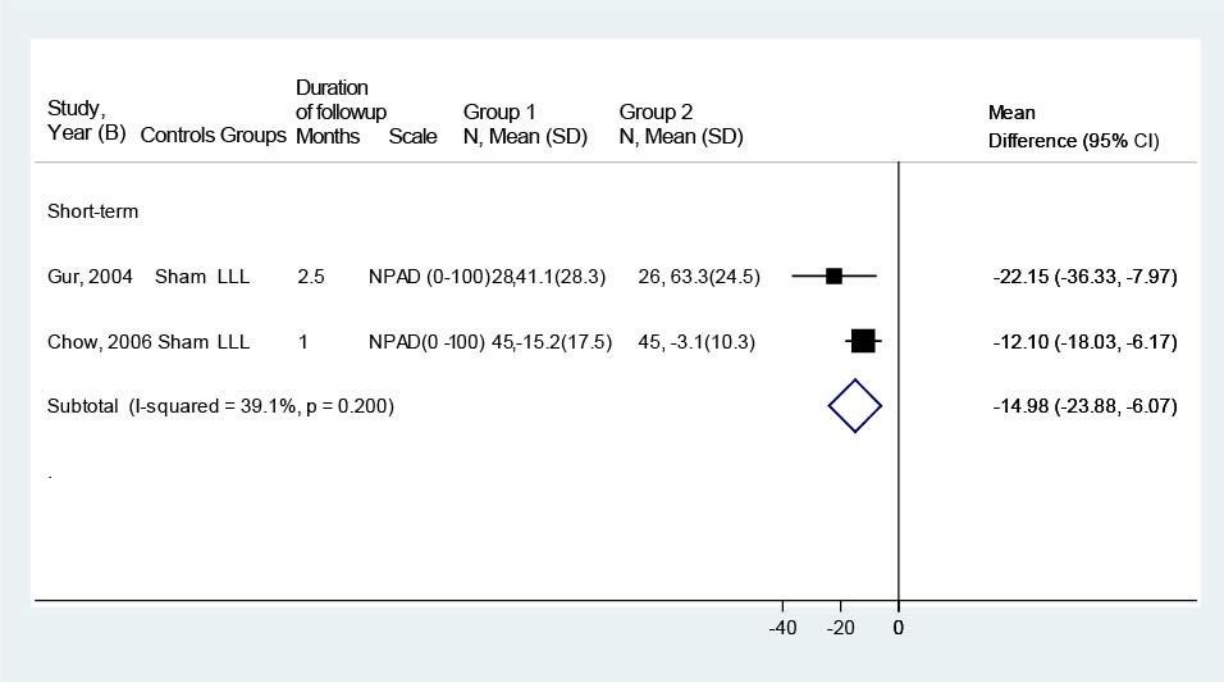
Physical Modalities Compared With Pharmacological Therapy or With Exercise Therapy

We did not find any trials meeting our criteria comparing a physical modality with pharmacological therapy or with exercise.

Harm

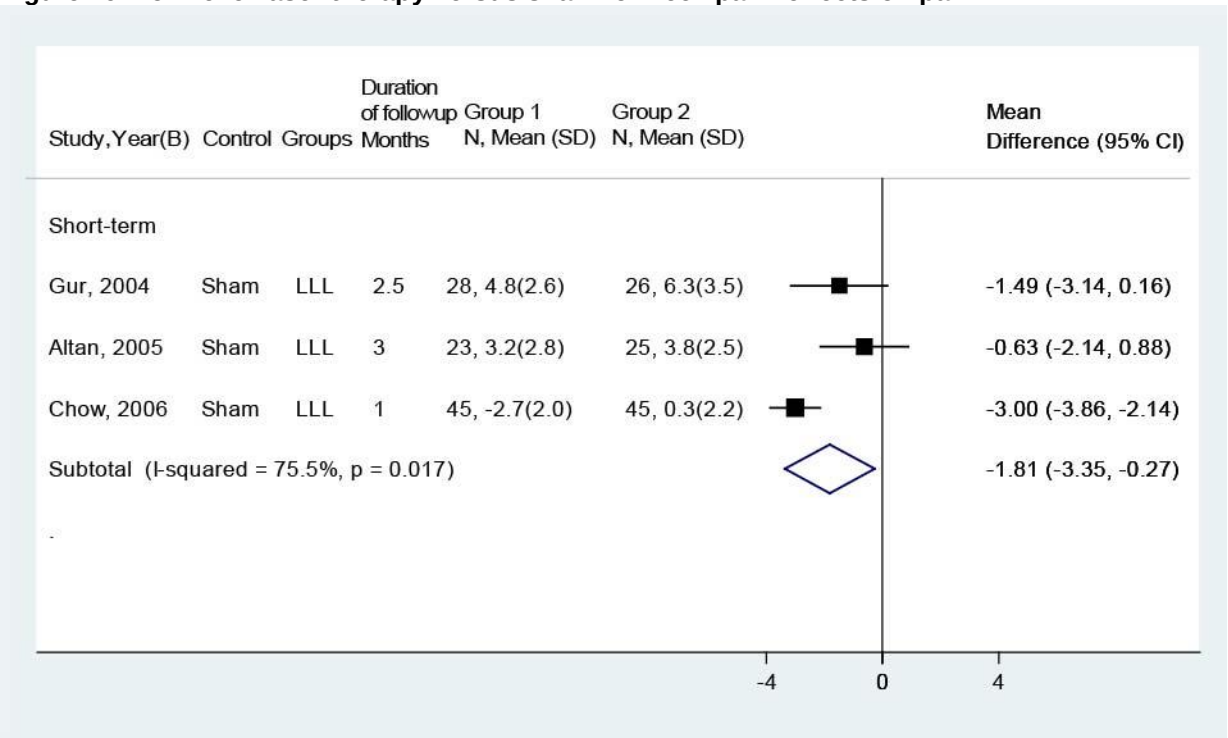
Only one laser trial reported harms.[116] The trial reported a large number of adverse effects with similar frequency in both groups. However, the sham group reported nausea significantly more frequently (42% vs. 20%) while the laser group reported stiffness more frequently (20% vs. 4%). The most frequently reported adverse effects included mild (78%) or moderate (60%) increased neck pain, increased pain elsewhere (78%), mild headache (60%), and tiredness (24%). Harms were not reported by either trial evaluating cervical traction or electromagnetic fields.

Figure 28. Low level laser therapy versus sham for neck pain: effects on function



CI = confidence interval; LLL = low level laser therapy; NPAD = Neck Pain and Disability Scale; SD = standard deviation.

Figure 29. Low level laser therapy versus sham for neck pain: effects on pain



CI = confidence interval; LLL = low level laser therapy; SD = standard deviation.

Manual Therapies for Neck Pain

Key Points

- The effects of massage on function versus self-management attention control were small and not statistically significant in one trial (N=64) in the short term (≥5 point improvement on the NDI, 39% versus 14%, RR 2.7, 95% CI 0.99 to 7.5) and intermediate term (57% versus 31%, RR 1.8, 95% CI 0.97 to 3.5) (SOE: Low for both time periods).
- No clear evidence that massage improved pain in the intermediate term versus exercise ($p>0.05$, data not reported) in one fair-quality trial (SOE: Low).
- Two fair-quality trials reported no serious adverse effects, and more transient nonserious pain or soreness during or after exercise, but not massage (SOE: Low).

Detailed Synthesis

Two trials of classical (N=85)[144] or Swedish massage (N=58)[145] met inclusion criteria (Table 21 and Appendix D). One trial compared massage versus attention control (self-care education),[145] and one trial compared massage versus two types of exercise, (muscle re-education and strength training targeting the neck and shoulder muscles).[144] Muscle reeducation was performed with a newly developed training device strapped to the head and consisted of a plate with 5 exchangeable surfaces that allow for progression of task difficulty. One trial was

conducted in Sweden[144] and one in the U.S.[145] One trial administered 10 massage treatments over 10 weeks,[145] and the other, 22 massage treatments over 11 weeks.[144]

Both trials were rated fair quality (Appendix E). Methodological limitations included the inability to blind interventions in both trials, and 21 percent attrition in the trial comparing massage with exercise.[144]

Table 21. Summary of results for neck pain: manual therapies

Author, Year, Followup, Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Rudolfsson, 2014[144] 6 months Duration of pain: median 84 to 123 months *Fair*	A. Massage, classical (n=36): upper body including the back, neck and shoulders. B. Neck coordination exercise (n=36): performed with a newly developed training device designed to improve the fine movement control of the cervical spine. C. Strength training (n=36): isometric and dynamic exercises targeting the neck and shoulder regions. All 3 interventions consisted of 22 individually supervised single treatment sessions, 30 min each, distributed over 11 weeks	A vs. B vs. C Age: 51 vs. 52 vs. 51 years Female: 100% vs. 100% vs. 100% Weight (kg): 73 vs. 74 vs. 74 Height (cm): 167 vs. 164 vs. 165 Pain duration: 120 vs. 123 vs. 84 months (median) Pain NRS (0-10), 5 vs. 6 vs. 6 (median) NDI: 26 vs. 29 vs. 31 SF-36 PCS (0-100): 43 vs. 39 vs. 39 (median) SF-36 MCS (0-100): 49 vs. 52 vs. 47 (median)	A vs. B: <u>6 months</u> Pain NRS (0-10): 4.0 vs. 3.8, mean difference 0.2 (95% CI -0.82 to 1.22) A vs. C: <u>6 months</u> Pain NRS (0-10): No data given at 6 month, however, authors state no difference among A, B or C.	NR
Sherman, 2009[145] 2.5 and 6.5 months Duration of pain >1 year: 81% *Fair*	A. Massage (n=32): Swedish and clinical techniques and self-care recommendations; 10 massage treatments over a 10-week period B. Self-care book: (n=32) information on potential causes of neck pain, neck-related headaches, whiplash, recommended	A vs. B Age: 47 vs. 46 years Female: 69% vs. 69% White: 87% vs. 81% Married: 78% vs. 59% Smoker: 9% vs. 6% Pain lasted > 1 year: 81% vs. 81% Symptom bothersome (0-10): 4.8 vs. 4.9 NDI (0-50): 14.2 vs. 14.2	A vs. B <u>2.5 months</u> NDI, % ≥5 points: 39% vs. 14%, RR 2.7 (95% CI 0.99 to 7.5) NDI (0-50): mean difference -2.3 (95% CI -4.7 to 0.15) <u>6.5 months</u> NDI, % ≥5 points: 57% vs. 31%, RR 1.8 (95% CI 0.97 to 3.5) NDI: mean difference: -1.9 (95% CI -4.4 to 0.63)	A vs. B <u>2.5 months</u> Bothersome score (0-10): mean difference -1.2 (95% CI -2.5 to 0.1) Bothersome improvement ≥30%: 55% vs. 25%, RR 2.1 (95% CI 1.04 to 4.2) SF-36 PCS: 52.8 vs. 53.3, p=0.982 SF-36 MCS: 45.9 vs. 45.3, p=0.444 <u>6.5 months</u> Bothersome score: mean

Author, Year, Followup, Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	strengthening exercises, body mechanics and posture, conventional treatment, complementary therapies for neck pain, and first aid for intermittent flare-ups.	SF-36 PCS (0-100): 46.0 vs. 44.1 SF-36 MCS (0-100): 51.9 vs. 53.1		difference -0.14 (95% CI -1.5 to 1.2) Bothersome improvement ≥30%: 43% vs. 39%, RR 1.1 (95% CI 0.6 to 2.0) SF-36 PCS and MCS: data not given, no statistical difference Medication use: No change in group A, 14% increase in group B

CI, confidence interval; NDI, Neck Disability Index; NR, not reported; NRS = Numerical Rating Scale; SF-36 MCS, Short-Form 36 Mental Component Summary; SF-36 PCS, Short-Form 36 Physical Component Summary.

Manual Therapies Compared With an Attention Control

The effects of massage on function versus self-management attention control were small and not statistically significant in one trial (N=64) in the short term (≥5 point improvement on the NDI, 39% versus 14%, RR 2.7, 95% CI 0.99 to 7.5) and intermediate term (57% versus 31%, RR 1.8, 95% CI 0.97 to 3.5).[145] A greater proportion of participants in the massage group reported improvement in a symptom bothersomeness scale (≥30%) in the short term (55% versus 25%; RR 2.2, 95% CI 1.04 to 4.2) but not intermediate term (RR 1.1, 95% CI 0.6 to 2.0). There were no differences between groups in SF-36 PCS and MCS. Medication use did not change in the massage group while it increased in the self-management group (14%).

Manual Therapies Compared With Pharmacological Therapy

No trial of manual therapy versus pharmacological therapy met inclusion criteria.

Manual Therapies Compared With Exercise

One fair-quality study reported no difference in intermediate-term pain comparing massage with neck coordination exercises (difference 0.2, 95% CI -0.82 to 1.22, 0-10 scale) or muscle performance exercises (no data given, $p>0.05$).[144]

No trial evaluated effects of manual therapies on use of opioid therapies or health care utilization.

Harms

Neither trial reported serious adverse effects. Nonserious mild adverse effects included discomfort or pain during (n=5) or after massage (n=3) in one trial.[145] There were no serious adverse effects in the massage group in the second trial, though there was transient neck or headache pain in the neuromuscular training exercise group (n=10).

Mind-Body Practices for Neck Pain

Key Points

- Alexander Technique resulted in a small improvement in function in the short term (difference -5.56 on a 0-100% scale, 95% CI -8.33 to -2.78) and intermediate term (difference -3.92, 95% CI -6.87 to -0.97) compared with usual care alone, based on one fair-quality trial (SOE: Low).
- There was no clear evidence that basic body awareness therapy improved function in the short term versus exercise in one fair-quality trial (SOE: Low).
- There is insufficient evidence from one poor quality trial to determine the effects of qigong on intermediate-term or long-term function or pain versus exercise; no data were available for short term outcomes (SOE: Insufficient).
- Both fair-quality trials reported no serious treatment-related adverse events. The trial evaluating Alexander Technique versus usual care found no clear between-group difference for nonserious adverse events, such as pain and incapacity, knee injury, or muscle spasm (RR 2.25, 95% CI 1.00 to 5.04). The other trial reported no differences between basic body awareness and exercise in any nonserious adverse effect (RR 0.65, 95% CI 0.37 to 1.14) (SOE: Low).

Detailed Synthesis

Three trials of mind-body practices met inclusion criteria (Table 22 and Appendix D).[168,175,176] One trial evaluated the Alexander Technique (a method of self-care developed to help people enhance their control of reaction and improve their way of going about everyday activities) plus usual care (N=344),[168] one trial basic body awareness therapy (N=115),[176] and one trial of qigong (N=134).[175] One trial compared mind-body techniques versus usual care[168] and two trials versus individually adjusted cervical and shoulder strengthening and stretching exercises,[175] or group-led exercises for whole body strengthening, aerobic, and coordination exercises.[176] Two trials were conducted in Sweden[175,176] and one in England.[168] The duration of mind-body treatment ranged from 10 to 20 weeks and the number of treatment sessions ranged from 12 to 20. One trial reported outcomes during the intermediate term and long term,[175]one short-term and intermediate-term outcomes,[168] and one short-term outcomes only.[176]

Two of the trials were rated fair quality[168,176] and one trial poor quality[175] (Appendix E). In the two fair-quality trials, the main methodological limitation was the inability to blind interventions. Limitations in the other trial included the inability to blind interventions, high attrition, and unequal loss to followup between groups.

Table 22. Summary of results for neck pain: mind-body practices

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Lansinger, 2007[175] 6 and 12 months	A. Qigong (n=72): 10-12 group sessions of 10-15 people done 1-2	A vs. B Age: 45 vs. 43 Female: 73% vs. 67%	A vs. B <u>6 months</u> NDI, median: 22 vs. 18, p>0.05	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Pain duration: >5 years, 45% *Poor*	times per week over 3 months. Sessions were 1 hour and consisted of information of the philosophy of medical qigong followed by exercises based on the Biyun method B. Exercise (n=67): 10-12 sessions 1-2 times per week over 3 months. Sessions were 1 hour and individualized to target 30%-70% of a person's maximal voluntary capacity, with exercises aiming to maintain/increase circulation, endurance, and strength. All patients: Ergonomic instructions and a pamphlet containing written information on neck pain	Duration of neck pain: 3 mos-1 year: 15% vs. 20% >1 year: 38% vs. 37% >5 years: 22% vs. 24% >10 years: 25% vs. 20% Physical activity: No to light exercise: 67% vs. 65% Med to hard exercise: 33% vs. 35% NDI (0-100%), median: 26 vs. 22 Neck pain VAS (0-10), median: 45 vs. 39	Neck pain VAS (0-10), median: 2.6 vs. 2.3, p>0.05 <u>12 months</u> NDI, median: 22 vs. 18, p>0.05 Neck pain VAS, median: 2.8 vs. 2.1, p>0.05	
MacPherson, 2015[168] 1 and 7 months Duration of pain, 7 years *Fair*	A. Alexander Technique group (n=172): up to 20 one-to-one lessons of 30 minutes' duration (600 minutes total) plus usual care, delivered weekly, with the option of being delivered twice per week initially and every 2 weeks later. B. Usual care (n=172) including general and neck pain–specific treatments routinely provided to primary care patients, such as prescribed	A vs. B Age: 52 vs. 54 years Female: 69% vs. 69% White: 93% vs. 89% Employed: 61% vs. 62% Pain duration (median): 60 vs. 96 months NPQ (0-100%): 39.64 vs. 40.46 SF12v2 physical (0-100): 39.99 vs. 40.98 SF12v2 mental (0-100): 45.07 vs. 46.59	A vs. B <u>1 month</u> NPQ: 35.35 vs. 40.90, mean difference -5.56 (95% CI -8.33 to -2.78) <u>7 months</u> NPQ: 37.07 vs. 40.99, mean difference -3.92 (95% CI -6.87 to -0.97)	A vs. B <u>1 month</u> SF-12v2 physical: data NR, p=NS SF-12v2 mental: data NR, p=NS <u>7 months</u> SF-12v2 physical: 0.68 (95% CI, -1.08 to 2.44), p=0.44 SF-12v2 mental: 1.76 (95% CI, 0.15 to 3.37), p=0.033

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	medications and visits to physical therapists and other health care professionals. Treatment was 12 sessions over 5 months lasting 50 minutes			
Seferiadis,[176] 2015 3 months Pain duration: 9.5 years *Fair*	A. Basic body awareness therapy (n=57): 1.5 hour sessions twice a week for 10 weeks. Sessions consisted of exercises based on activities of daily living, meditation, and Tai chi inspired exercises aiming to improve posture and increase efficient movement patterns B. Exercise (n=56): 1.5 hour sessions twice a week for 10 weeks. Sessions consisted of 45 minutes of muscle strengthening, 15 minutes of stretching, and 20 minutes of progressive muscle relaxation	A vs. B Age: 47 vs. 49 Female: 66% vs. 77% Duration of symptoms (years): 10 vs. 9 WAD classification: 1: 0% vs. 2% 2: 23% vs. 28% 3: 77% vs. 70% NDI (0-50): 20 vs. 18.8 SF-36v2 physical functioning (0-100): 67.5 vs. 69.7 role-physical (0-100): 33.9 vs. 24.5 bodily pain (0-100): 34.3 vs. 35.2 general health (0-100): 54.7 vs. 48.7 vitality (0-100): 39.5 vs. 35.1 social functioning (0-100): 60 vs. 59.4 role-emotional (0-100): 55.4 vs. 51.7 mental health (0-100): 65.9 vs. 62.7	A vs. B <u>3 months</u> NDI: Difference from baseline -2 (95% CI -3.5 to -0.5) vs. -1 (95% CI -2.5 to 0.4), p>0.05	A vs. B <u>3 months</u> SF-36v2 physical functioning: Difference from baseline 7.1 (95% CI 3.7 to 11.4) vs. 0.5 (95% CI -3.2 to 4.1), p>0.05 SF-36 role-physical: Difference from baseline 17.5 (95% CI 5.9 to 29) vs. 19 (95% CI 9.3 to 28.6), p>0.05 SF-36 bodily pain: Difference from baseline 12.2 (95% CI 6.9 to 17.6) vs. 4.9 (95% CI -0.1 to 9.8), p=0.044 SF-36 general health: Difference from baseline 7.5 (95% CI 2.4 to 12.6) vs. 4.5 (95% CI -0.1 to 9), p>0.05 SF-36 vitality: Difference from baseline 7.3 (95% CI 1.0 to13.6) vs. 5.6 (95% CI -0.5 to 11.6), p>0.05 SF-36 social functioning: Difference from baseline 13.3 (95% CI 6.6-19.9 vs. 3.5 (95% CI -3 to 9.9), p=0.037 SF-36 role-emotional: Difference from baseline 9.3 (95% CI-2.3 to 21) vs. 4 (95% CI -8.3 to 16.4), p>0.05 SF-36 mental health: Difference from baseline 2.8 (95% CI -2 to 7.6)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				vs. 1.2 (95% CI -3.6 to 5.9), p>0.05

CI = confidence interval; NDI = Neck Disability Index; NR = not reported; SF-36 =Short-Form 36 Questionnaire; VAS = Visual Analog Scale; WAD = Whiplash Associated Disorders

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Mind-Body Practices Compared With Usual Care

One fair-quality trial found a small improvement in function as measured by the NPQ in favor of the Alexander Technique plus usual care versus usual care alone in the short term (mean difference -5.56 on a 100% scale, 95% CI -8.33 to -2.78) and intermediate term (mean difference -3.92, 95% CI -6.87 to -0.97).[168] There were no significant differences between the intervention group and usual care for the physical component score of the SF-12 v 2 at 1-month or 7-month followups. However, significantly larger improvements in the MCS occurred in the Alexander group versus the usual care group 7 months following treatment (mean difference, 2.12 on a 0-100 scale, 95% CI 0.42 to 3.82).[168]

Mind-Body Practices Compared With Pharmacological Therapy

No trial of mind-body practice versus pharmacological therapy met inclusion criteria.

Mind-Body Practices Compared With Exercise

There were no differences in function as measured by the NDI between basic body awareness therapy (1 fair-quality study, n=113)[176] in the short term (mean change from baseline -2 versus -1, p>0.05) or qigong (poor-quality study, n=139)[175] in the intermediate term or long term (median 22 versus 18, p>0.05, at each time period) versus exercise therapy. The trial assessing qigong found no difference in pain at 6 or 12 months following treatment (median 2.6 versus 2.3 and 2.8 versus 2.3, p>0.05, respectively).[175] Two of the eight sections of the SF-36v2 favored basic body awareness therapy versus exercise in the short term (bodily pain and social functioning) in the fair quality trial.[176] No other section of the SF-36v2 demonstrated a difference between groups.

No trial evaluated effects of mind-body practices on use of opioid therapies or health care utilization.

Harms

One trial of basic body awareness therapy reported no serious adverse effects.[176] One patient in the basic body awareness group and four patients in the exercise group reported that they discontinued treatment due to increased neck symptoms or pain in other joints (p=0.363). The event risk for all nonserious adverse events was 0.27 in the body awareness therapy group and 0.40 in the exercise group (RR 0.65, 95% CI 0.37 to 1.14).

Acupuncture for Neck Pain

Key Points

- Acupuncture was associated with slightly greater effects on short-term and intermediate-term function versus sham acupuncture, a placebo (sham laser), or usual care (short term, 4 trials, pooled SMD -0.32, 95% CI -0.53 to -0.10, I^2=53.1%; intermediate term, 3 trials, pooled SMD -0.19, 95% CI -0.35 to -0.02). One trial reported no difference in function in the long term (SMD -0.23, 95% CI -0.61 to 0.16) (SOE: Low for all time periods).
- There was no difference in pain in studies comparing acupuncture with sham acupuncture, or placebo interventions in in the short term (4 trials, pooled difference -0.2 on a 0-10 scale, 95% CI -0.59 to 0.05, I^2=2%), intermediate term (3 trials, pooled difference 0.45, 95% CI -0.34 to 1.25, I^2=59%) or long term (1 trial, difference -1.8, 95% CI -1.34 to 0.64) (SOE: Low for all time periods).
- There was insufficient evidence from two small poor-quality trials to draw conclusions regarding short-term function or pain for acupuncture versus NSAIDs (SOE: Insufficient).
- No serious adverse events were reported in five trials reporting harms. The most commonly reported nonserious adverse effects in people receiving acupuncture included numbness/discomfort, fainting, and bruising (SOE: Moderate).

Detailed Synthesis

We identified eight trials of acupuncture that met our inclusion criteria[168,184-189,203] (Table 23 and Appendix D). All trials evaluated needle acupuncture to body acupoints; two also evaluated electroacupuncture.[186,189] Control groups included sham acupuncture in four trials,[184-186,188] placebo intervention (sham TENS[187] and sham laser acupuncture[189]) in two trials, usual care in one trial,[168] and pharmacological therapy (Zaltoprofen[203] and Trilisate[184]) in two trials. The duration of acupuncture therapy ranged from 3 weeks to 5 months, and the number of sessions from 5 to 14. Sample sizes ranged from 30 to 355 (total sample=1,091). Across trials, participants were predominately female (from 60% to 90%) with mean ages ranging from 37 to 53 years. One trial was conducted in the United States,[184] one in Turkey,[186] and the rest in Asia[185,189,203] or Europe.[168,187,188] One trial reported outcomes through long-term followup,[188] four trials through intermediate-term followup,[168,187-189] and the remainder only evaluated short-term outcomes.[184-186,203]

Six trials were rated fair quality[168,185-189] and two trials poor quality[184,203] (Appendix E). Common limitations in the fair-quality trials included unclear allocation concealment methods and of care provider blinding; additionally, the poor-quality trials had baseline group dissimilarity (not controlled for) and high attrition.

Table 23. Summary of results for neck pain: acupuncture

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Birch, 1998[184] 3 months Duration of pain, 7.5 years *Poor*	A. Relevant acupuncture, Japanese technique (n=15): using bilateral needles on hands and feet known to be associated with treatment for neck pain and followed by Infrared lamp. B. Irrelevant acupuncture (n=16): using bilateral needles on hands and feet in areas not associated with treatment for neck pain and followed by light. C. NSAIDs only (n=15): 500mg per day of Trilisate 30 minute treatment twice per week for 4 weeks, then once per week for 4 weeks, total 14 treatments	A vs. B vs. C Age: 41 vs. 38 vs. 39 years Female: 86% vs. 77% vs. 86% Pain duration: 82 vs. 92 vs. 91 months Married: 36% vs. 23% vs. 50% Employed: 86% vs. 69% vs. 77% Pain intensity (0-10) 4.8 vs. 4.7 vs. 4.9	A vs. B <u>3 months</u> SF-MPQ[b] (0-33): 9.0 vs. 15.1, p=ns A vs. C <u>3 months</u> SF-MPQ: 9.0 vs. 18.0, p=ns	NR
Cho, 2014[203] 1 month Duration of pain, NR *Poor*	A. Active acupuncture, traditional Chinese (n=15 randomized/15 analyzed), 3x/week for 3 weeks.(length of time for each intervention not reported) B. Zaltoprofen (80mg) alone (n=15 randomized/15 analyzed) 3x/day for 3 weeks.	A vs. B Age: 38 vs. 39 years Female: 60 vs. 80 Pain VAS (0-10): 6.1 vs. 7.1 NDI (0-50): 22.3 vs. 26.3	A vs. B <u>1 month</u> NDI: 17.3 vs. 17.7, difference -0.40 (95% CI -4.55 to 3.75) Pain VAS: 4.5 vs. 3.8, difference 0.7 (95% CI -0.74 to 2.14)	A vs. B <u>1 month</u> BDI (0-63) : 28.5 vs. 27.2, p=ns SF-36 (0-100): 88.6 vs. 84.3, p=ns EQ-5D (scale unclear): 7.3 vs. 6.7, p=ns
Liang, 2011[185] 3 months Duration of pain, NR *Fair*	A. Active acupuncture, traditional Chinese, (n=93) B. Sham acupuncture (n=97) Treatment was 3x/week for 3 weeks (9 treatments total) lasting 20 minutes after needling Both groups received infrared	A vs. B Age: 37 vs. 37 years Female: 72% vs. 73% NPQ (0-100%): 32.7 vs. 33.0 Pain VAS (0-10): 5.3 vs. 5.5	A vs. B <u>3 months</u> NPQ: 19.1 vs. 25.5, difference -6.4 (95% CI -9.9 to -2.9) Pain VAS: 2.9 vs. 3.2, difference -0.3 (95% CI -0.75 to 0.15)	A vs. B <u>3 months</u> SF-36 physical functioning (0-100): 84.3 vs. 85.9, p=0.447 SF-36 mental (0-100): 67.1 vs. 61.6, p=0.001

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
MacPherson, 2015[168] 1 and 7 months Duration of pain, 7 years *Fair*	A. Active acupuncture, traditional Chinese, (n=173): plus usual care 2 weeks later. B. Usual care (n=172): including general and neck pain–specific treatments routinely provided to primary care patients, such as prescribed medications and visits to physical therapists and other health care professionals. Treatment was 12 sessions over 5 months lasting 50 minutes	A vs. B Age: 52 vs. 54 years Female: 69% vs. 69% White: 93% vs. 89% Employed: 61% vs. 62% Pain duration (median): 60 vs. 96 months NPQ (0-100%): 39.64 vs. 40.46	A vs. B <u>1 month</u> NPQ: 35.35 vs. 40.90, difference -5.56 (95% CI -8.33 to -2.78) <u>7 months</u> NPQ: 37.07 vs. 40.99, difference -3.92 (95% CI -6.87 to -0.97)	A vs. B <u>1 month</u> SF-12v2 physical: data NR, p=ns SF-12v2 mental: data NR, p=ns <u>7 months</u> SF-12v2 physical (0-100): mean difference 0.68 (95% CI, 1.08 to 2.44) SF-12v2 mental (0-100): mean difference 1.76 (95% CI, 0.15 to 3.37)
Sahin, 2010[186] 3 months Duration of pain, NR *Fair*	A. Electro-acupuncture (n=15) B. Sham acupuncture (n=16) Treatment was 10 sessions, 3 sessions per week, lasting 30 minutes	A vs. B Age: 39 vs. 35 years Female: 100% vs. 81% University graduate: 54% vs. 94% BMI: 23.9 vs. 24.6 Pain with motion VAS (0-10): 7.38 vs. 6.19 Pain at rest VAS (0-10): 4.00 vs. 5.25	A vs. B <u>3 months</u> Pain with motion VAS: 4.50 vs. 5.38, difference -0.88 (95% CI -2.70 to 0.94) Pain at rest VAS: 4.00 vs. 3.54, difference 0.46 (95% CI -1.88 to 2.80)	NR
Vas, 2006[187] 6 months Duration pain, 3.8 years *Fair*	A. Active acupuncture, traditional Chinese, (n=61) B. Sham TENS (n=62) Treatment was 5 sessions over 3 weeks lasting 30 minutes	A vs. B Age: 46 vs. 47 years Female: 75% vs. 89% Pain duration: 47 vs. 43 months Pain VAS with motion (0-10): 6.9 vs. 7.2 NPQ (0-100): 52.7 vs. 56.5	A vs. B <u>6 months</u> (mean difference from baseline) Pain VAS with motion: 4.1 vs. 2.7, difference 1.4 (95% CI 0.3 to 2.6)	A vs. B <u>6 months</u> SF-36 PCS: (0-100): 9.3 vs. 5.3, p=0.054 SF-36 MCS: (0-100): 8.0 vs. 5.2, p=0.351 Rescue medication (none or occasional): 87% (39/45) vs. 68% (27/40), RR 1.28 (95% CI 1.01 to 1.64)
White, 2004[188] 2, 6, 12 months Duration pain, 6 years *Fair*	A. Active acupuncture, Western technique based on tender local and distal points, (n=70 randomized/54 analyzed) B. Sham electro-acupuncture (n= 65 randomized/53	A vs. B Age: 54 vs. 53 years Female: 66% vs. 63% Pain duration: 4.8 vs. 7.7 years NDI (0-50): 16.8 vs. 17.2 Pain VAS (0-10): 5.0	A vs. B <u>2 months</u> NDI: 11.0 vs. 12.7, difference -1.7 (95% CI -4.3 to 0.9) Pain VAS: 1.7 vs. 2.3, difference -0.6 (95% CI -1.3 to 0.1) <u>6 months</u>	A vs. B <u>2 months</u> SF-36 PCS (0-100): 42.5 vs. 43.8, p=ns SF-36 MCS (0-100): 52.5 vs. 50.3, p=ns

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	analyzed) Treatment was 8 sessions over 4 weeks lasting 20 minutes	vs. 5.4	NDI: 9.9 vs. 10.6, difference -0.7 (95% CI -3.61 to 2.21) Pain VAS: 1.9 vs. 2.1, difference -1.8 (95% CI -1.1 to 0.7) <u>12 months</u> NDI: 8.9 vs. 10.7, difference -1.8 (95% CI -4.84 to 1.24) Pain VAS: 2.1 vs. 2.4, difference -0.3 (95% CI -1.4 to 0.6)	
Zhang, 2013[189] 3 and 6 months Duration of pain, 6.3 years *Fair*	A. Electro-acupuncture, traditional Chinese (n=103 randomized/84 analyzed) B. Sham laser acupuncture (n=103 randomized/76 analyzed) via a mock laser pen 2 minutes, with the pen at a distance of 0.5 to 1 cm from the skin. Treatment 3x/week for 3 weeks, 45 min for electro-acupuncture and 2 min per point for sham laser	A vs. B Age: 46 years (whole population) Female: 70% (whole population) NPQ (0-100%): 40.7 vs. 41.1 Pain with motion (0-10): 5.5 vs. 5.2	A vs. B <u>3 months</u> NPQ: mean 32.9 (95% CI, 30.3 to 35.4) vs. mean 33.3 (95% CI 30.1 to 36.5), p=0.664 Pain with motion VAS: mean 4.7 (95% CI, 4.2 to 5.1) vs. mean 4.5 (95% CI, 4.1 to 5.0), p=0.617 <u>6 months</u> NPQ: mean 33.59 (95% CI, 30.7 to 36.4) vs. mean 34.3 (95% CI 31.1 to 37.6), p=0.808 Pain with motion: mean 4.7 (95% CI, 4.2 to 5.2) vs. mean 4.4 (95% CI, 3.9 to 4.8), p=0.813	A vs. B <u>3 months</u> SF-36 PCS (0-100): mean 52.8 (95% CI, 53.0 to 53.7) vs. mean 53.3 (95% CI, 52.4 to 54.2), p=0.982 SF-36 MCS (0-100): mean 45.9 (95% CI, 46.0 to 46.8) vs. mean 45.3 (95% CI, 44.2 to 46.4), p=0.444 <u>6 months</u> SF-36 PCS: mean 53.0 (95% CI, 52.0 to 53.9) vs. mean 53.2 (95% CI 52.3 to 54.0), p=0.559 SF-36 MCS: mean 45.4 (95% CI, 44.5 to 46.3) vs. mean 44,4 (95% CI, 43.4 to 45.4), p=0.246

BDI = Beck Depression Inventory; CI = confidence interval; EQ-5D = Euroqol 5-D; NDI = Neck Disability Index; NPQ = Northwick Park Neck Pain Questionnaire; NR = not reported; ns = not statistically significant; NSAID = non-steroidal anti-inflammatory drug; SF-36 = Short Form-36 questionnaire; SF-MPQ = McGill Pain Questionnaire Short Form; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Estimated from Figure 1 in Birch et al.[184]

Acupuncture Compared With Sham Acupuncture, Usual Care, or a Placebo Intervention

Acupuncture was associated with slightly greater effects on short-term and intermediate-term function versus sham acupuncture, placebo (sham laser) or usual care (short-term, 4 trials,[168,185,188,189] pooled SMD -0.32, 95% CI -0.53 to -0.10, I2=53.1%; intermediate-term, 3 trials,[168,188,189] pooled SMD -0.19, 95% CI -0.35 to -0.02, I2=0.0%) (Figure 30). Trials measured function using the NDI or the NPQ; across trials the SMD ranged from -0.53 to -0.03 in the short

term and -0.29 to -0.05 in the intermediate term. None of the trials were rated poor quality. One trial reported no difference in function in the long term (SMD -0.23, 95% CI -0.61 to 0.16).[188]

There was no difference between acupuncture versus controls in short-term pain (4 trials, pooled mean difference -0.27, 95% CI -0.59 to 0.05, I2=2%)[185,186,188,189] (Figure 31). Stratified analyses according to the type of control (sham or placebo laser) resulted in similar estimates. Trials reported no differences in pain between acupuncture versus controls in the intermediate term (3 trials, pooled mean difference 0.45, 95% CI -0.34, 1.25, I2=59%)[187-189] or long term (1 trial, mean difference -0.35, 95% CI -1.34 to 0.64).[188]

In general, acupuncture did not improve quality of life compared with sham intervention in the short term or intermediate term as reported in four trials[185,187-189] (Table 23).

No trial evaluated effects of psychological therapies on use of opioid therapies or health care utilization.

Acupuncture Compared With Pharmacological Therapy

Two small poor-quality trials evaluated acupuncture versus NSAIDs. One trial (n=27) compared acupuncture three times per week for 3 weeks versus 80 mg of Zaltoprofen alone three times per day for 3 weeks.[203] The other trial (n=30) compared 14 sessions of acupuncture versus 500 mg of Trilisate per day for 8 weeks.[184] In the short term, one trial reported no difference in NDI (mean difference -0.4, 95% CI -4.6 to 3.8).[203] Both trials reported no difference between groups in pain as measured by the McGill Pain Questionnaire[184] or VAS.[203] One trial found no differences between groups in the Beck Depression Index, the SF-36, or the EQ-5D in the short term[203] (Table 23).

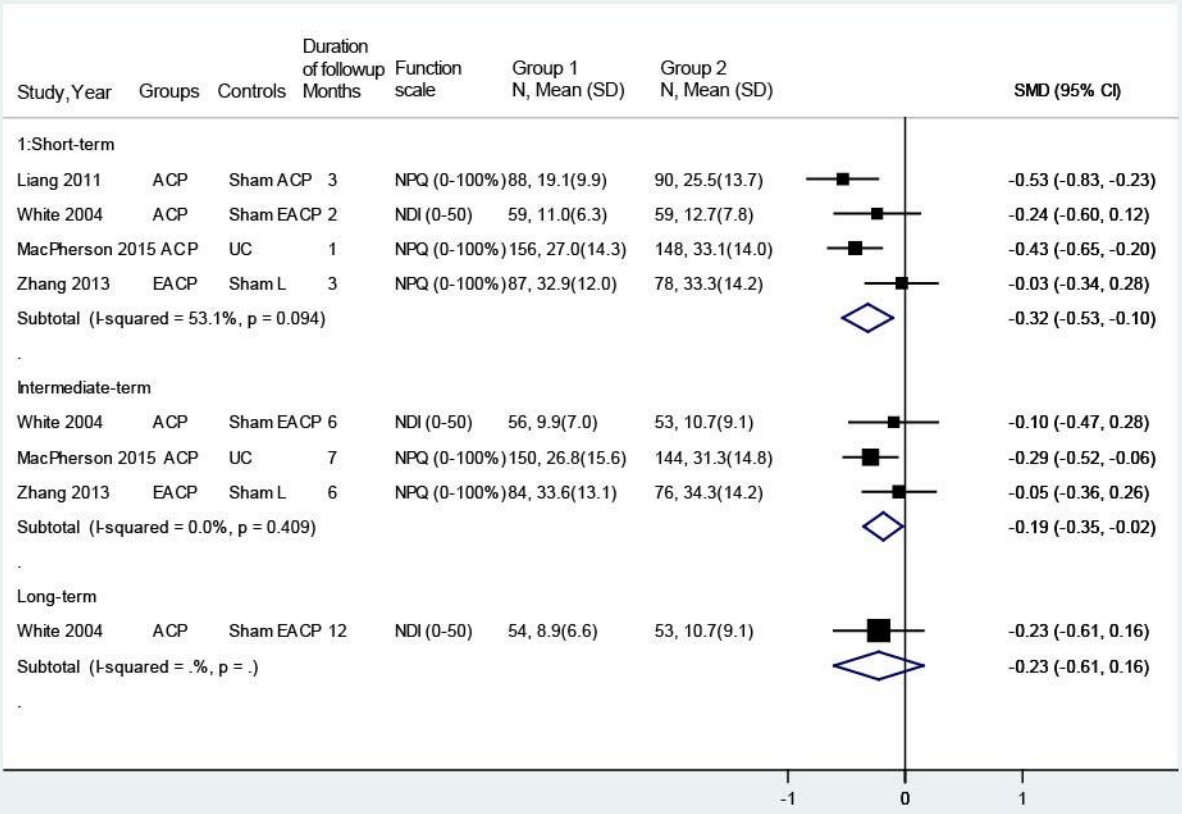
Acupuncture Compared With Exercise Therapy

No trial of acupuncture versus exercise met inclusion criteria.

Harms

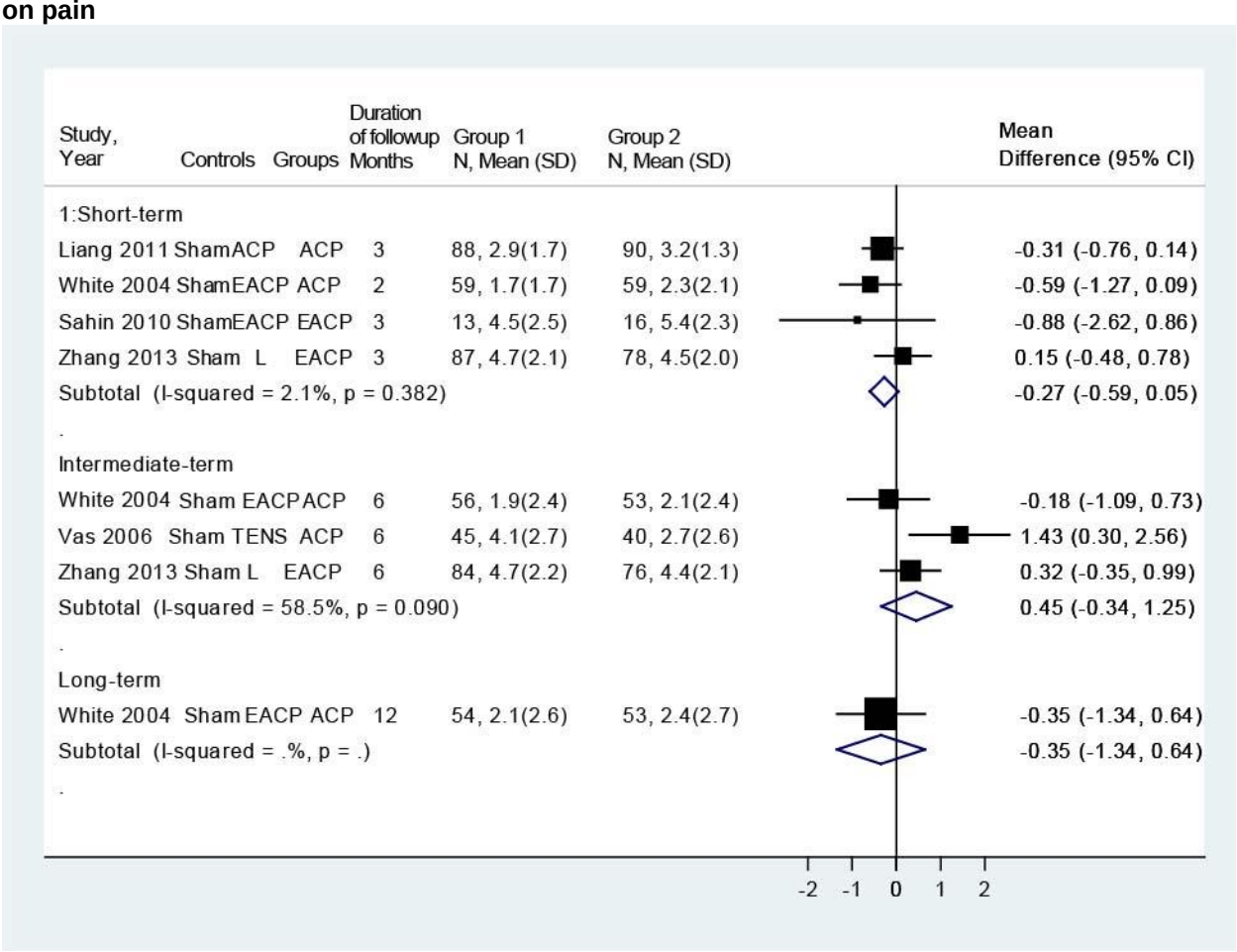
Five of the eight trials assessing acupuncture reported harms.[168,185,187-189] No serious adverse events (defined as involving death, hospitalization, persistent disability, or a life-threatening risk in one trial[168]; undefined in the other four studies) were reported in any trial. The most commonly reported nonserious adverse effects in people receiving acupuncture included numbness/discomfort (2.7%), fainting (1.1%), and bruising (1.1%).

Figure 30. Acupuncture versus sham acupuncture, a placebo intervention, or usual care for neck pain: effects on function



ACP = traditional needle acupuncture; CI = confidence interval; EACP = electroacupuncture; NDI = Neck Disability Index; NPQ = Northwick Park Questionnaire; SD = standard deviation; Sham L = sham laser; SMD = standardized mean difference; UC = usual care

Figure 31. Acupuncture versus sham acupuncture or a placebo intervention for neck pain: effects on pain



ACP = traditional needle acupuncture; CI = confidence interval; EACP = electroacupuncture; SD = standard deviation; Sham L = sham laser; SMD = standardized mean difference; TENS = transcutaneous electrical stimulation; UC = usual care

Key Question 3: Osteoarthritis

Exercise for Osteoarthritis of the Knee

Key Points

- Exercise was associated with slightly greater impact on function than usual care, no treatment, or sham intervention short term (7 trials, pooled standardized mean difference (SMD) -0.25, 95% CI -0.4 to -0.09, I^2 = 0%) , intermediate term (9 trials [excluding outlier trial] pooled SMD SMD -0.78, 95%CI -1.37 to -0.19, I^2 = 91.4%), and long term (2 trials, pooled SMD -0.24, 95% CI -0.37 to -0.11 I^2=0%) (SOE: Moderate for short-term; Low for intermediate and long-term).
- Exercise was associated with a small improvement in pain short term (7 trials, pooled difference -0.44 on a 0 to 10 scale, 95% CI -0.82 to -0.05, I^2= 35%) versus usual care, no treatment, or sham intervention (SOE: Moderate), and with moderately greater effect on

pain in the intermediate term (9 trials, pooled difference -1.61 on a 0 to 10 scale, 95% CI -2.51 to -0.72, I^2=91%) compared with usual care, an attention control, or no treatment (SOE: Low). Long term, there was no clear difference between exercise and improvement in pain but data were limited (2 trials, difference -0.24, 95% CI -0.72 to 0.24) (SOE: Low).

- No trial evaluated exercise versus pharmacological therapy.
- Comparisons involving exercise versus other non-pharmacological therapies are addressed in the sections for the other therapies.
- Harms were not well reported. Across seven trials, one reported minor temporary increase in pain with exercise, four others found no difference in worsening pain versus controls, and one reported no difference in falls or death (SOE: Moderate).

Detailed Synthesis

Twenty-one publications from 18 randomized controlled trials that evaluated exercise interventions for the treatment of knee osteoarthritis (OA) met the inclusion criteria. (Table 24 and Appendix D). Seven trials evaluated muscle performance exercise versus attention control[40,41,43,46,47,55] or no treatment.[38,42,54] In six trials, the interventions consisted of combined exercise approaches compared with usual care,[36,44,45,49,52] an attention control,[53] or no treatment.[39] Muscle performance exercises were a component of 6 trials.[36,39,44,45,49,52,53] One trial had an aerobic exercise arm that consisted of a facility-based, 1-hour walking program three times per week over 3 months, and it used an attention control.[40,46,47] A single trial evaluated a mobility exercise program based on Mechanical Diagnosis and Therapy (MDT) versus a waitlist comparator, where patients were allowed to continue receiving usual care.[50] One trial evaluated gait training (guided strategies to optimize knee movements during treadmill walking with computerized motion analysis with visual feedback) versus usual care.[51] Three trials tested exercise programs as a part of physiotherapy care compared to usual care or sham.[37,48,56] The duration of exercise programs ranged from 2 to 24 weeks; the number of exercise sessions ranged from 4 to 36. No trials comparing exercise with a pharmacological intervention were identified.

Sample sizes ranged from 50 to 786. Across the trials, the majority of patients were female (51% to 88%) with mean ages ranging from 56 to 75 years. Five trials specifically included patients with bilateral knee OA.[38,41-43,55] Five trials were conducted in the United States or Canada,[40,45-47,49-52] five in Europe,[44,48,53,54,56] five in Taiwan,[38,41-43,55] two in Australia or New Zealand,[36,37] and one in Brazil.[39] Most trials had short (5 trials)[36,44,50,51,54] or intermediate followup (10 trials).[38,39,41-43,45,51-53,55] Three trials reported long-term outcomes.[45-47,49,53]

Twelve trials were rated fair quality (one at short-term followup[51]),[36,37,40,41,43-50,54] and eight trials poor quality,[38,39,42,52,53,55,56] including one at intermediate-term followup[51] (Appendix E). In the fair-quality trials, the main methodological limitation was a lack of blinding for the patients or care providers. Additional limitations in the poor-quality trials included unclear randomization and allocation concealment methods, unclear use of intention to treat, unclear baseline differences between intervention groups, and attrition not reported or unacceptable.

Table 24. Summary of results for osteoarthritis of the knee: exercise

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Abbott, 2013[36] 9.75 months Duration of diagnosis: Mean 2.5 to 2.8 years *Fair*	A. Exercise (n=51/29 knee OA): 7 sessions of strengthening, stretching, and neuromuscular control over 9 weeks, with 2 booster sessions at week 16. Individual exercises prescribed as needed. Home exercise prescribed 3 times weekly B. Usual care (n=51/28 knee OA)	A vs. B (total population, includes hip OA) Age: 67 vs. 66 years Female: 52% vs. 58% Percent hip OA: 43% vs. 45% Percent knee OA: 57% vs. 55% Percent both hip OA and knee OA: 20% vs. 26% Baseline WOMAC (0-240): 95.5 vs. 93.8	A vs. B (knee OA only) A vs. C 9.75 months WOMAC mean change from baseline: -12.7 vs. -31.5	None
Bennell, 2005[37] 3 months Duration of pain: 9.6 vs. 8.7 years *Fair*	A.Neuromuscular Re-education (Physiotherapy) (n=73) Knee taping; exercises to retrain the quadriceps, hip, and back muscles; balance exercises; thoracic spine mobilisation; and soft tissue massage. individual sessions lasting 30 to 45 minutes once weekly for four weeks, then fortnightly for eight weeks. Thrice-daily standardized home exercises. B. Control (n=67) Placebo: sham ultrasound and topical non-therapeutic gel. 30 to 45 minutes once weekly for four weeks, then fortnightly for eight weeks.	A vs. B Age: 67 vs. 70 years Female: 68% vs. 66% WOMAC Physical Function (0-68): 27.6 vs. 28.4 WOMAC Pain (0-20): 8.2 vs. 8.0 VAS Pain on movement (0-10): 5.3 vs. 5.2 KPS (0-36): 16.6 vs. 16.4 KPS Frequency (0-30): 23.5 vs. 22.8	A vs. B 3 months Responders, global improvement in pain: 59% vs. 50%, p=0.309 Responders, VAS pain: 58% vs. 42%, p=0.069 WOMAC, Physical Function: 20.0 vs. 21.7, MD -0.9 (95% CI -4.4 to 2.7) WOMAC, Pain: 5.8 vs. 6.0, MD -0.4 (95% CI -1.5 to 0.7) VAS pain on movement: 3.2 vs. 3.5, MD -0.5 (95% CI -1.2 to 0.3) KPS, Severity: 13.5 vs. 14.3, MD -1.0 (95% CI -2.5 to 0.6) KPS, Frequency: 19.4 vs. 20.3, MD -1.7 (95% CI -3.5 to 0.1)	A vs. B 3 months SF-36, Physical Function (0-100): 50.5 vs. 46.2, MD 4.3 (95% CI -1.8 to 10.4) SF-36, Bodily Pain (0-100): 60.4 vs. 61.8, MD 1.8 (95% CI -6.7 to 10.3) SF-36, Role Physical (0-100): 47.0 vs. 46.5, MD 1.6 (95% CI -11.1 to 14.3) AQoL(-0.04 to 1.0): 0.52 vs. 0.48, MD 0.05 (95% CI 0.01 to 0.10) Withdrawals: 18% (13/73) vs. 3% (2/67); RR 6.0 (95% CI 1.4, 25.5) Group A: Minor skin irritation (48%), increased pain with exercises (22%), pain with massage (1%) Group B: Increased pain (2%), itchiness and pain with application of gel (2%) (All were minor and

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				short-lived)
Chen, 2014[38] 6 months Duration of pain: 10-144 months *Poor*	A. Exercise (n=30): 3 sessions per week for 8 weeks. Sessions consisted of a 20 minutes of hot packs and 5 minutes of passive range of motion exercises on a stationary bike, followed by an isokinetic muscle-strengthening exercise program B. Control (n=30): Details NR	A + B Age: 63 Females: 85% A vs. B Lequesne Index (0-26): 7.8 vs. 8.0 Pain VAS (0-10): 5.5 vs. 5.6	A vs. B 6 months Lequesne Index: 5.4 vs. 7.6, (MD -2.2, 95% CI -3.1 to -1.3) Pain VAS: 4.0 vs. 6.5, (MD -2.5, 95% CI -3.3 to -1.7)	A vs. B 6 months Intolerable knee pain: 10% (3/30) vs. 0% (0/30) RR=infinity, p=0.08
Dias, 2003[39] 6 months Duration of pain: NR *Poor*	A. Exercise (n=25): 12 exercise sessions twice a week for the 6 month study period in addition to three supervised walks of 40 minutes each week. Exercise sessions consisted stretching, concentric and eccentric isotonic progressive resistance exercises, and closed kinetic chain weight-bearing exercises B. Control group (n=25): Subjects were instructed to follow the instructions given at an educational session that all participants attended (see information below) All patients: One-hour educational session consisting of a lecture on disease characteristics, joint protection, pain management, and strategies to overcome difficulties in activities of daily life	A vs. B Age, median: 74 vs. 76 Female: 84% vs. 92% Lequesne Index, median (0-24): 12 vs. 12.5 HAQ, median (0-3): 1 vs. 1	A vs. B 6 months Lequesne Index, median: 4.3 vs. 13, p=0.001 HAQ, median: 0.3 vs. 1.1, p=0.006	A vs. B 6 months SF-36 functional capacity, median (0-100): 77.5 vs. 40, p<0.001 SF-36 physical role limitation, median (0-100): 92.5 vs. 75, p=0.001 SF-36 bodily pain, median (0-100): 100 vs. 0, p=0.002 SF-36 general health, median (0-100): 100.5 vs. 51, p=0.021 SF-36 vitality, median (0-100): 93.5 vs. 87, p=0.027 Adverse Events: NR
Ettinger, 1997[40] FAST trial 6 months, 15 months Duration of pain: NR	A. Aerobic Exercise Program (n=144) 3-month facility-based walking program of 3 times per week for 1 hour. Each session consisted of a 10-minute warm-up and cool-down	A vs. B vs. C Age: 69 vs. 68 vs. 69 years Female: 69% vs. 73% vs. 69% African-American: 24% vs. 28% vs. 26%	A vs. C Avg. across all Time-points: FAST Physical Disability Scale Total: 1.72 vs. 1.90 Ambulation subscale: 2.22 vs. 2.64	A vs. B vs. C Adverse Events: Falls- 14% (2/144) vs. 14% (2/146) vs. 0% (0/149) A vs. C: RR=infinity, p=0.15

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fair	phase,including slow walking and flexibility stretches, and a 40-minute period of walking at an intensity equivalent to 50% to 70% of the participants' heart rate reserve. Followed by 15-month home-based walking program. <u>B. Resistance Exercise Program (n=146)</u> 3-month supervised facility-based program, with 3 one-hour sessions per week, and a15-month home-based program. Each session consisted of a 10-minute warm-up and cool-down phase and a 40-minute phase consisting of 2 sets of 12 repetitions of 9 exercises. <u>C. Attention Control (n=149)</u> attended, during the first 3 months, monthly group sessions on education related to arthritis manage-ment, including time for discussions and social gathering. Later, participants were called bimonthly (months 4-6) or monthly (months 7-18) to maintain health updates and provide support	NR	Transfers subscale: 1.75 vs. 1.92 Pain: 2.14 vs. 2.40 B vs. C <u>Avg. across all Time-points:</u> FAST Physical Disability Scale Total: 1.74 vs. 1.90 Ambulation subscale: 2.37 vs. 2.64 Transfers subscale: 1.72 vs. 1.92 Pain: 2.2 vs. 2.40	B vs. C: RR = infinity , p=0.15 Death- 0% (0/144) vs. 0% (0/146) vs. 0.7% (1/149)
Huang, 2003[42] 10 months Duration of pain: range, 0.33(4 months) to 9 years *Poor*	<u>A. Isokinetic Strengthening (n=33)</u> 3 sessions per week for 8 weeks. 60% of average peak torque the initial dose of isokinetic exercise. An increasing dose program was used in the initial first to fifth sessions (1 set to 5 sets), and a dose of 6 sets was applied from sixth to the twenty-fourth sessions. Each set consists of 5 repetitions of	A+B+C+D Age: 62 years Female: 70% A vs. B vs. C vs. D Lequesne Index (0-26): 6.9 vs. 7.1 vs. 6.8 vs. 7.2 VAS pain (0-10): 4.8 vs. 4.6 vs. 4.7 vs. 4.6	A vs. D <u>10 months</u> Lequesne Index: 3.1 vs. 7.6, MD -4.5 (95% CI -5.3 to -3.7), VAS Pain: 2.5 vs. 6.1; p<0.05 B vs. D <u>10 months</u> Lequesne Index: 3.1 vs. 7.6, MD -3.6 (95% CI -4.4 to -2.8) VAS Pain: 2.0 vs. 6.1; p<0.05	A vs. B vs. C vs. D <u>10 months</u> Withdrawals: 3% (1/33) vs. 6% (2/33) vs. 3% (1/33) vs. 18% (6/33) Withdrawals RR (95% CI): A vs. D: 0.17 (0.02, 1.3) B vs. D: 0.33 (0.07,1.53) C vs. D: 0.17 (0.02, 1.3)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	concentric and eccentric contraction in angular velocity 30°/second and 120°/second for extensors, and 5 repetitions of eccentric and concentric contraction in angular velocity 30°/second and 120°/second forflexors. B. Isotonic Strengthening (n=33) The same protocol was used as in the isokinetic exercise. The isotonic muscle strengthening exercise program consisted of 5 repetitions of concentric and eccentric the maximum velocity that the lever arm could achieve. C. Isometric Strengthening (n=33) The same protocol of was used as in the isokinetic exercise. The speed of passive forward or backward motion was set at 30°/second. All intervention groups exercised 3 times weekly for 8 weeks. The patients in all groups also received 20 minutes of hot packs and passive range motion exercise by an electric stationary bike (20 cycles per minute) for 5 min-utes to both knees before muscle strengthening exercise. D. Control (n=33) Description NR		C vs. D 10 months Lequesne Index: 4.8 vs. 7.6, MD -2.8 (95% CI -3.6 to -2.0) VAS Pain: 3.2 vs. 6.1; p<0.05	Stopped therapeutic exercise due to intolerable pain during exercise: 12.1% (4/33) vs. 6.1% (2/33) vs. 6.1% (2/33)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Huang, 2005[43] 10 months Duration of pain: 0.42 (5 months) to 12 years *Fair*	<u>A. Isokinetic Exercise (n=35)</u> 3 times per week for 8 weeks. Began with 60% of the mean peak torque, increasing dose program was used in the first 5 sessions (1 set to 5 sets), and a dose of 6 sets was applied from the sixth to twenty-fourth sessions, with the density rising from 60% to 80% of the mean peak torque as the patient was able. Each set consisted of 5 repetitions of concentric contraction in angular velocities of 30°/second and 120°/second for extensors, and 5 repetitions of eccentric and concentric (Ecc/Con) contractions in angular velocities of 30°/second and 120°/second for flexors. <u>B. Control (n=35)</u> Warm-up exercises only	A+B Age: 65 years Female: 81% A vs. B Lequesne Index(1-26): 7.6 vs. 7.4 VAS pain(0-10): 5.3 vs. 5.4	A vs. B <u>10 months</u> Lequesne Index: 5.8 vs. 8.1, MD -2.3 (95% CI -3.2, -1.4) VAS Pain: 3.9 vs. 6.6, p<0.05	A vs. B <u>10 months</u> Withdrawals 11% (4/35) vs. 11% (4/35) Discontinuation of exercise due to intolerable pain during exercise: 14% (5/35) vs. NA
Huang 2005[41] 10 months Duration of pain: 0.5 (6 mos.) to 11 years *Fair*	<u>A. Isokinetic Exercise (n=30)</u> 3 times per week for 8 weeks. Began with 60% of the average peak torque. Intensity of isokinetic exercise increased from 1 set to 5 sets during the first through fifth sessions and remained at 6 sets for the remaining 6th through 24th sessions. Each set consisted of 5 repetitions of concentric contraction in angular velocities of 30°/s and 120°/s for extensors, and 5 repetitions of eccentric and concentric contractions in angular velocities of 30°/s and 120°/s for flexors. <u>B. Control (n=30)</u>	A+B Age: 62 (range, 42-72) years Female: 81% A vs. B Lequesne Index(1-26): 6.7 vs. 7.0 VAS pain(0-10): 4.9 vs. 4.8	A vs. B <u>10 months</u> Lequesne Index: 5.1 vs. 7.8, MD -2.7 (95% CI -3.8, -1.6) VAS Pain: 3.5 vs. 6.0; p<0.05	A vs. B <u>10 months</u> Withdrawals 13% (4/30) vs. 13% (4/30) Discontinuation of exercise due to intolerable pain during exercise: 17% (5/30) vs. NA

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	Heat for 20 minutes and 5 minutes of passive range of motion on bike only.			
Lund, 2008[44] 3 months Duration of pain: 8.5 vs. 7.8 vs. 4.5 *Fair*	<u>A. Aquatic Exercise (n=27):</u> 2x per week for 8 weeks. warm-up, strengthening and endurance exercise, balance exercise and stretching exercise. Each session lasted 50 min, comprising 10 min warm-up, 20 min resistance exercises, 10 min balance and stabilizing exercises, 5 min lower limb stretches and 5 min cool-down period. Compliance was 92%. <u>B. Land-based Exercise (n=25):</u> 2x per week for 8 weeks. warm-up, strengthening/endurance exercise, balance exercise and stretching exercise. Each session lasted 50 min, comprising 10 min warm-up, 20 min resistance exercises, 10 min balance and stabilizing exercises, 5 min lower limb stretches and 5 min cool-down period. Compliance was 85%. <u>C. Control (n=27):</u> No exercise All 3 groups were asked to continue any other treatment as usual.	A vs. B vs. C Age: 65 vs. 68 vs. 70 years Female: 83% vs. 88% vs. 66% VAS Pain at rest (0-100): 29.8 vs. 23.3 vs. 15.5 VAS Pain during walking (0-100): 59.8 vs. 53.0 vs. 48.5 KOOS symptom (0-100): 50.5 vs. 50.9 vs. 50.1 KOOS pain (0-100): 47.1 vs. 41.0 vs. 37.9 KOOS Activities of Daily Living (0-100): 44.7 vs. 40.6 vs. 39.6 KOOS Sport (0-100): 79.1 vs. 75.6 vs. 70.0 KOOS Quality of Life (0-100): 63.7 vs. 57.0 vs. 60.8	A vs. C <u>3 months</u> KOOS symptom: 64.1 vs. 63.7; MD 0.5 (95% CI -6.6, 7.6) KOOS Activities of Daily Living: 63.0 vs. 61.4; MD 1.6 (95% CI -5.7, 8.9) KOOS sport: 24.2 vs. 23.5; MD 0.7 (95% CI -9.3, 10.7) KOOS quality of life: 42.8 vs. 41.4; MD 1.7 (95% CI -5.4, 8.2) KOOS pain: 60.7 vs. 62.6; MD -1.5 (95% CI -8.7, 5.8) VAS pain at rest: 18.1 vs. 23.8; MD -5.7 (95% CI -13.3, 2.0) VAS pain: 52.9 vs. 58.3; MD -5.4 (95% CI -16.2, 5.4) B vs. C <u>3 months</u> KOOS symptom: 66.1 vs. 63.7; MD 2.4 (95% CI -4.8, 9.5) KOOS Activities of Daily Living: 63.9 vs. 61.4; MD 2.5 (95% CI -5.0, 9.9) KOOS sport: 31.6 vs. 23.5; MD 8.1 (95% CI -2.0, 18.2) KOOS quality of life: 43.1 vs. 41.4; MD 1.7 (95% CI -5.3, 8.7) KOOS pain: 62.0 vs. 62.6; MD -0.3 (95% CI -7.5, 7.0) VAS pain at rest: 15.6 vs. 23.8; MD -8.1 (95% CI -15.8, -0.4) VAS pain walking: 50.1 vs. 58.3; MD -8.2 (95% CI -19.7, 2.7)	A vs. B vs. C <u>3 months</u> Withdrawals: 4% (1/27) vs. 20% (5/25) vs. 7% (2/27) A vs. C: RR 0.5 (95% CI 0.05, 5.2) B vs. C: RR 2.5 (95% CI 0.6, 12.7) Increased pain during and after exercise: 11% (3/27) vs. 32% (8/25) vs. NR Swollen knees: 0% (0/27) vs. 12% (3/25) vs. NR Withdrawals due to adverse events: 0% (0/27) vs. 12% (3/25) vs. NR
Messier, 2004[45] (Same trial as Rejeski 2002) 6 months, 18 months Duration of pain: NR	<u>A. Exercise (n=80):</u> Three 1 hour sessions per week done at the study facility for 4 months. Option to undergo a 2 month transition phase alternating between	A vs. B Age: 69 vs. 69 Female: 74% vs. 68% WOMAC physical function (0-68): 24.0 vs. 26.0	A vs. B <u>6 months</u> WOMAC physical function*: 22.0 vs. 22.0 WOMAC pain: 6.2 vs. 6.2, MD 0.0 (95% CI -0.2 to 0.2)	A vs. B <u>3 months</u> Accident related to treatment: 1% (1/80) vs. 0% (0/78)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fair	facility and home sessions, after which they carried out the program at home. Sessions consisted of 15 minutes of aerobic exercises, 15 minutes of resistance-training, an additional 15 minutes of aerobic exercises, and a 15 minute cool down phase. B. Control (n=78): 1 hour sessions monthly for three months consisting of presentations on osteoarthritis, obesity, and exercise and a question and answer session. Monthly phone contact was maintained for months 4-6 and bimonthly phone contact was maintained for months 7-18. All subjects: Instructed to continue use of all medications and other treatments as prescribed by their personal physicians	WOMAC pain (0-20): 6.6 vs. 7.3	18 months WOMAC physical function: 21.0 vs. 22.6 WOMAC physical function, mean change: 3.1 vs. 3.4 WOMAC pain: 6.2 vs. 6.0, MD 0.2 (95% CI 0.04 to 0.4)	
Penninx, 2001[46] FAST trial (substudy in patients with no baseline ADL disability) (same trial as Pennix 2002 below) 6 months, 15 months Duration of pain: NR *Fair*	A. Aerobic Exercise Program (n=88) 3-month facility-based walking program of 3 times per week for 1 hour. Each session consisted of a 10-minute warm-up and cool-down phase, including slow walking and flexibility stretches, and a 40-minute period of walking at an intensity equivalent to 50% to 70% of the participants' heart rate reserve. Followed by 15-month home-based walking program. B. Resistance Exercise Program (n=82) 3-month supervised facility-based program,	A vs. B vs. C Age: 70 vs. 69 vs. 69 years Female: 66% vs. 72% vs. 66% African-American: 25% vs. 21% vs. 28% Pain intensity (1-6): 2.2 vs. 2.1 vs. 2.1 Disability (scale NR): 1.7 vs. 1.7 vs. 1.6	A vs. B vs. C 15 months ADL Disability (overall): 36.4% vs. 37.8% vs. 52.5% Disability in transferring from a bed to a chair: 29.5% vs. 36.6% vs. 50.0% Disability in bathing: 12.5% vs. 13.4% vs. 27.5% Disability in toileting: 19.4% vs. 13.4% vs. 25.0% Disability in dressing: 5.7% vs. 7.3% vs. 17.5% Disability in eating: 0% vs. 1.2% vs. 5.0%, p=0.02 15 months ADL Disabiltiy (overall) A vs. C: Adj RR 0.53 (95%	A vs. B vs. C 15 months Increased severity of knee OA leading to withdrawal: n=3 (not reported by exercise group)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	with 3 one-hour sessions per week, and a15-month home-based program. Each session consisted of a 10-minute warm-up and cool-down phase and a 40-minute phase consisting of 2 sets of 12 repetitions of 9 exercises. C. Attention Control (n=80) attended, during the first 3 months, monthly group sessions on education related to arthritis manage-ment, including time for discussions and social gathering. Later, participants were called bimonthly (months 4-6) or monthly (months 7-18) to maintain health updates and provide support		CI 0.33, 0.85), B vs. C: Adj RR 0.60 (95% CI 0.38, 0.97), Disability in transferring from a bed to a chair A vs. C: Adj RR 0.46 (95% CI 0.28, 0.76) B vs. C: Adj RR 0.68 (95% CI 0.42, 1.09) Disability in bathing A vs. C: Adj RR 0.31 (95% CI 0.15, 0.68) B vs. C: Adj RR 0.44 (95% CI 0.21, 0.93) Disability in toileting A vs. C: Adj RR 0.58 (95% CI 0.29, 1.15) B vs. C: Adj RR 0.61 (95% CI 0.28, 1.31) Disability in dressing A vs. C: Adj RR 0.20 (95% CI 0.07, 0.64) B vs. C: Adj RR 0.46 (95% CI 0.17, 1.22) Disability in eating: incidence too small to calculate risks.	
Penninx, 2002[47] FAST trial (substudy looking at baseline depressive symptoms) 6 months, 15 months Duration of pain: NR *Fair*	A. Aerobic Exercise Program (n=149) 3-month facility-based walking program of 3 times per week for 1 hour. Sessions consisted of a 10-minute warm-up and cool-down phase, including slow walking and flexibility stretches, and a 40-minute period of walking at an intensity equivalent to 50% to 70% of the participants' heart rate reserve. Followed by 15-month home-based walking program. B. Resistance Exercise Program (n=146) 3-month supervised facility-based program, with 3 one-hour sessions per week, and a15-month home-based program. Each session consisted of a 10-minute warm-up and cool-down	A + B + C Age: 69 years Female: 70% CES-D (cutoff of 5 points): 22% A vs. B vs. C CES-D (scale NR): 2.74 vs. 2.74 vs. 2.74	A vs. C Avg across all time-points CES-D: 2.12 vs. 2.80, p<0.001 B vs. C Avg across all time-points CES-D: 2.59 vs. 2.80, p=0.27	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	phase and a 40-minute phase consisting of 2 sets of 12 repetitions of 9 exercises. C. Attention Control (n=144) attended, during the first 3 months, monthly group sessions on education related to arthritis manage-ment, including time for discussions and social gathering. Later, participants were called bimonthly (months 4-6) or monthly (months 7-18) to maintain health updates and provide support			
Quilty, 2003[48] 2.5 months, 10.5 months Duration of pain: NR *Fair*	A. Combination (Physiotherapy) (n=40) 9 sessions over a 10 week period. Sessions consisted of patellar taping, 7 individualized exercises, posture correction, and footwear advice. All exercises were performed 10 times each, 5 times a day B. Control (n=43): Baseline discussion with the physiotherapist concerning diagnosis, prognosis, footwear, weight reduction, and activity. General exercise was encouraged but no specific quadriceps exercises were advised	A vs. B Age: 69 vs. 67 years WOMAC Function (0-68): 27.4 vs. 27.8 VAS pain (0-100): 51.0 vs. 53.4	A vs. B 2.5 months WOMAC function: 26.5 vs. 27.5; Adjusted MD -0.6 (95% CI -3.7, 2.4) VAS Pain: 42.8 vs. 50.5; Adjusted MD -6.4 (95% CI -15.3, 2.4) 10.5 months WOMAC function: 29.7 vs. 28.3; Adjusted MD 1.7 (95% CI -1.8, 5.2) VAS Pain: 48.1 vs. 54.1; Adjusted MD -4.9 (95% CI -13.6, 3.8)	A vs. B Withdrawals 2% (1/43) vs. 0% (0/44) Adverse Events: None
Rejeski, 2002[49] (Same trial as Messier 2004) 6 months, 18 months Duration of pain: NR *Fair*	A. Exercise (n=80): Three 1 hour sessions per week done at the study facility for 4 months. Option to undergo a 2 month transition phase alternating between facility and home sessions, after which they carried out the program at home. Sessions consisted of 15 minutes of aerobic	A vs. B Age: 68 vs. 69 Female: 74% vs. 67%	NR	A vs. B 6-18 months (average) SF-36 PCS: 37.1 vs. 34.4 SF-36 PCS, adjusted mean: 37.6 vs. 35.3 SF-36 MCS: 52.9 vs. 53.5 SF-36 MCS, adjusted mean: 54.1 vs. 53.7

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	exercises, 15 minutes of resistance-training, an additional 15 minutes of aerobic exercises, and a 15 minute cool down phase. B. Control (n=78): 1 hour sessions monthly for three months consisting of presentations on osteoarthritis, obesity, and exercise and a question and answer session. Monthly phone contact was maintained for months 4-6 and bimonthly phone contact was maintained for months 7-18. All subjects: Instructed to continue use of all medications and other treatments as prescribed by their personal physicians			
Rosedale, 2004[50] 2.5 months Duration of pain: NR *Fair*	A. Exercise (n=120): given end-range exercises in the direction they had responded to, to be performed 10 times every 2 to 3 hours. A nonresponder subgroup was given exercises to strengthen quadriceps and aerobic exercises. All subjects in the exercise group attended 4 to 6 physiotherapy sessions, 2 to 3 assessment sessions lasting up to 1 hour and the rest followup sessions lasting 20 minutes, over a 2 week period. B. Waiting list (n=60): Subjects were followed up in the orthopedic department at the surgeon's discretion and continued receiving their usual care.	A vs. B vs. C Age: 66 vs. 64 Female: 56% vs. 60% Median comorbidities: 3 vs. 3 KOOS function(0-100): 56 vs. 51 KOOS function in sport and recreation(0-100): 22 vs. 20 KOOS pain(0-100): 51 vs. 46 P4 pain scale: 21 vs. 23 KOOS knee symptoms(0-100): 50 vs. 48 KOOS quality of life(0-100): 28 vs. 27	A vs. B 2.5 months KOOS function: 61 vs. 52, (Adj MD 5, 95% CI 1 to 9) KOOS function in sport and recreation: 31 vs. 24, (Adj MD 6, 95% CI 0 to 11) KOOS pain: 56 vs. 46, (Adj MD 7, 95% CI 3 to 11) P4 pain scale: 24 vs. 21, (Adj MD -2, 95% CI -4 to 1) KOOS knee symptoms: 56 vs. 52, (Adj MD 2, 95% CI -2 to 6) KOOS quality of life: 34 vs. 32, (Adj MD 1, 95% CI -3 to 6)	NR
Segal, 2015[51]	A. Gait Training (n=24):	A vs. B	A vs. B, between group	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
3 and 9 months Duration of pain: NR *Fair (3 months)* *Poor (9 months)*	guided strategies to optimize knee movements during treadmill walking; computerized motion analysis with visual biofeedback; individualized home programs from physical therapist; Twice weekly sessions (45 minutes) for 12 weeks (24 total sessions) <u>B. Usual Care (n=18)</u> Usual care for knee osteoarthritis and were not asked to make changes in their lifestyle (e.g., annual visit to their physician, use of pain medications, knee surgery and/or physical therapy); ask to keep a diary	Age: 70 vs. 69 years Female: 76% vs. 53% Race: NR LLFDI basic lower limb function score: 65.8 vs. 63.5 KOOS Pain: 62.7 vs. 59.8 KOOS Symptoms: 60.1 vs. 63.0	difference in change score compared with baseline <u>3 months</u> LLFDI basic lower limb function score: 2.3 (95% CI -1.8 to 6.3) KOOS Pain: 3.7 (95% CI -4.7 to 12.1) KOOS Symptoms: 6.2 (95% CI -2.9 to 15.4) <u>9 months</u> LLFDI basic lower limb function score: 1.0 (95% CI -7.4 to 9.4) KOOS Pain: 7.2 (95% CI -2.0 to 16.5) KOOS Symptoms: 6.0 (95% CI -6.2 to 18.2)	
Sullivan, 1998[52] 10 months Duration of pain: NR *Poor*	<u>A. Exercise (n=52):</u> 3 group sessions of 10-15 subjects per week were done for 8 weeks. Sessions were structured as a hospital-based supervised fitness walking and supportive patient education program. Sessions consisted of stretching and strengthening exercises, expert speakers, group discussions, instructions in safe walking techniques, and up to 30 minutes of walking. At the end of the 8 week treatment period, subjects were encouraged to continue walking and given guidelines for managing individualized programs of fitness walking. <u>B. Usual care (n=50):</u> Subjects continued to receive the standard routine medical care they had been receiving	A vs. B Age: 71 vs. 68 Female: 77% vs. 90% AIMS physical activity subscale(0-10): 6.3 vs. 6.4 AIMS arthritis impact subscale(0-10): 4.6 vs. 4.5 AIMS pain subscale(0-10): 4.9 vs. 5.5 Pain VAS(0-10): 4.1 vs. 6.3 AIMS general health perception subscale(0-10): NR	A vs. B <u>10 months</u> AIMS physical activity subscale: 6.1 vs. 6.2, MD -0.1, (95% CI -1.7 to 1.5) AIMS arthritis impact subscale: 3.3 vs. 3.8, MD -0.5, (95% CI -1.8 to 0.8) AIMS pain subscale: 4.6 vs. 5.5, MD -0.9, (95% CI -2.2 to 0.4) Pain VAS: 5.0 vs. 5.4, MD -0.4, (95% CI -2.0 to 1.2) AIMS general health perception subscale: 3.7 vs. 3.3, MD 0.4 (95% CI -1.0 to 1.8)	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	prior to enrollment in the study. Subjects were interviewed weekly during the 8 week treatment period about their functional and daily activities.			
Thomas, 2002[53] 6 months, 12 months, 18 months, 24 months Duration of pain: NR *Poor*	<u>A. Exercise (n=470):</u> Two year, self-paced program that started with four 30 minute visits in the first two months followed by visits every six months. Designed to maintain and improve strength of muscles around the knee, range of motion at the knee joint, and locomotor function. 121 of the 470 patients also received attention control which consisted of monthly phone calls by a study researcher that sought to monitor symptoms and offer simple advice on knee pain management. 114 of the 470 patients received the attention control and a placebo tablet in addition to the exercise program. The remaining 235 participate in the exercise program only.* <u>B. Control (n=316):</u> 160 subjects received attention control consisted of monthly phone calls by a study researcher that sought to monitor symptoms and offer simple advice on knee pain management. 78 subjects took a placebo tablet. 78 patients had no contact with the researchers between assessment visits.	A vs. B Age: 62 vs. 62 Female: 63% vs. 66% WOMAC pain score(0-20): 7.15 vs. 7.35	A vs. B <u>6 months</u> WOMAC physical function, mean difference (95% CI): NR WOMAC pain, mean difference (95% CI): -0.6 (-1.0 to -0.2) <u>24 months</u> WOMAC physical function, mean difference (95% CI): -2.6 (-4.1 to -1.1) WOMAC pain: -0.82 (-1.3 to -0.3)	A vs. B <u>6 months</u> HADS: NR SF-36: NR <u>24 months</u> HADS: NR (NS) SF-36: NR (NS)
Thorstensson, 2005[54] 5 months Duration of pain:	<u>A. Exercise (n=30):</u> 1 hour group exercise sessions of 2 to 9 participants, twice a week for 6 weeks.	A vs. B Age: 55 vs. 57 Female: 50% vs. 52%	A vs. B <u>5 months</u> KOOS pain, mean change: 3.1 vs. -1.1, p=0.32	A vs. B <u>5 months</u> KOOS QOL, mean change (0-100): 5.1 vs. -2.3, p=0.02

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
NR *Fair*	Sessions consisted of weight-bearing exercises to increase postural control and to increase endurance and strength in the lower extremity. Patients were given daily exercises to perform at home. B. Control group (n=31): Subjects were told not to make any lifestyle changes. Subjects met with the physical therapist at baseline, at 6 weeks, and at 6 months	KOOS Pain (0-100): 60 vs. 64 KOOS ADL(0-100): 69 vs. 71 KOOS Symptoms (0-100): 63 vs. 66 KOOS sports and recreation (0-100): 34 vs. 37	KOOS ADL, mean change: 0.9 vs. -1.9, p=0.61 KOOS symptoms, mean change: 1.0 vs. -3.4, p=0.31 KOOS sports and recreation, mean change: 0.5 vs. -8.3, p=0.32	SF-36 PCS, mean change (0-100): 3.0 vs. -0.7, p=0.09 SF-36 MCS, mean change (0-100): 0.7 vs. -0.7, p=0.40 Adverse Events: A vs. B Increased knee pain: 3% (1/30) vs. 0% (0/31)
Weng,[55] 2009 10 months Duration of pain: 42.5 months *Poor*	A. Isokinetic exercise (n=33): 3 sessions a week for 8 weeks. Sessions consisted of sets of concentric and eccentric contractions at varying angular velocities and start and stop angles. Hot packs for 10 minutes and passive range of motion exercises B. No intervention (n=33): Warm-up cycling for 10 minutes. Hot packs for 10 minutes and passive range of motion exercises	A+B Age: 64 Female: 75% A vs. B Lequesne Index (0-24): 7.3 vs. 7.1 Pain VAS (0-10): 4.7 vs. 4.5	A vs. B 10 months Lequesne Index: 6.3 vs. 7.3 Pain VAS: 3.6 vs. 5.0	A vs. B 10 months Treatment related pain causing withdrawal: 9% (3/33) vs. 0% (0/33) RR=infinity, p=0.08
Williamson, 2007[56] 1.5 months Duration of pain: NR *Poor*	A. Combination (Physiotherapy) (n=60): Groups of 6–10 patients, hourly, once a week for 6 weeks. Exercise circuit of static quadriceps contractions; inner range quadriceps contractions; straight leg raises; sit to stands, stair climbing; calf stretches; theraband resisted knee extensions; wobble board balance training; knee flexion/extension sitting on gym ball and free standing peddle revolutions.	A vs. B Age: 70 vs. 70 years Female: 52% vs. 54% OKS (0-48): 39.3 vs. 40.5 WOMAC (unclear scale): 50.2 vs. 51.1 VAS pain (0-10): 6.8 vs. 6.9	A vs. B 1.5 months OKS: 38.8 vs. 40.8 WOMAC: 49.4 vs. 52.3 VAS Pain: 6.4 vs. 7.2	A vs. B 1.5 months HAD Anxiety (0-21): 7.1 vs. 6.5 HAD Depression (0-21): 6.8 vs. 7.1 Withdrawals: 17% (10/60) vs. 0% (0/61) Adverse Events: None

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	B. Control (n=61): Usual Care (home exercise and advice leaflet)			

ADL = activities of daily living; AIMS = Arthritis Impact Measurement Scale; AQoL = Assessment of Quality of Life; CES-D = Center for Epidemiologic Studies Depression; CI = confidence interval; HADS = Hospital Anxiety and Depression Scale; HAQ = Health Assessment Questionnaire; ITT = intention-to-treat; KOOS = Knee Injury and Osteoarthritis Outcome Score; KPS = Knee Pain Scale; MCS = Mental Component Score; MD = mean difference; NA = not applicable; NR = not reported; NS = not significant; PCS = Physical Component Score; RR = risk ratio; OKS = Oxford Knee Score; QoL = quality of life; SF-36 = Short-Form-36; VAS, Visual Analog Scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Exercise Compared With Usual Care, No Treatment, Sham, or an Attention Control

Functional outcomes. Exercise was associated with slightly greater short-term impact on function than usual care, no treatment, or sham intervention (7 trials, pooled SMD -0.25, 95% CI -0.4 to -0.09, $I^2 = 0\%$)[37,44,48,50,51,54,56] (Figure 32). Estimates were similar following exclusion of poor-quality trials and when analyses were stratified by exercise and control type. In the short term, a small improvement in the KOOS Sport and Recreation scale (0-100) with exercise compared with usual care or no treatment was seen (3 trials, pooled mean difference 5.88, 95% CI 0.80 to 10.96, $I^2 = 0\%$, plot not shown).[44,50,54]

Exercise was also associated with moderately greater effect on function that usual care, no treatment, or attention control at the intermediate-term (10 trials, pooled SMD -1.15, 95% CI -1.85 to -0.46, $I^2 = 93.9\%$)[38,39,41-43,45,48,51,52,55] (Figure 32). Substantial heterogeneity was present with one outlier trial[39] of combination exercise versus no treatment in elderly patients (median age 75 years) which had higher (worse) baseline Lequesne Index scores compared with other studies and a larger change from baseline scores in the intervention group. Removal of this poor quality trial did not improve heterogeneity but did change the pooled estimate (9 trials, pooled SMD -0.78, 95%CI -1.37 to -0.19, $I^2 = 91.4\%$), suggesting smaller effects of exercise on function. Stratification by exercise type and control type may partially explain the heterogeneity. Muscle performance exercise was associated with a moderately greater effect on function compared with attention control or no treatment (5 trials, pooled SMD -1.44, 95% CI -2.08 to -0.17)[38,41-43,55] and when compared with attention control only (3 trials, pooled SMD -1.12, 95% CI -1.83 to -0.47).[41,43,55] No difference was seen across studies of exercise versus usual care (4 trials, pooled SMD 0.02, 95% CI -0.20 to 0.24).[45,48,51,52]

Analyses confined to trials that evaluated function on the 0-24 point Lequesne Index also suggests a moderately greater effect on function compared with attention control or no treatment (6 trials, pooled mean difference -3.42, 95%CI -5.49 to -1.35 $I^2 = 97\%$, plot not shown).[38,39,41-43,55] Again, removal of the poor quality outlier trial[39] did not impact the heterogeneity, but yielded a slightly lower effect estimate (5 trials, pooled mean difference -2.40, 95CI CI -3.32 to -1.44), still consistent with a moderate effect for exercise. Results were similar to this estimate for muscle performance exercise, use of attention control, and when the two fair-quality trials were retained.

One fair-quality trial (n=101 with knee OA)[36] compared combined exercise programs to usual care for intermediate-term function using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). The exercise group had improvement in function from baseline, which was not statistically significant (mean change from baseline -12.7, 95% CI -27.1 to 1.7), while the usual care group had no change in function (mean change from baseline 1.6, 95% CI -10.5 to 13.7). Data were insufficient to determine effect size or include in the meta-analysis.

One trial separately analyzed participants free of disability for activities of daily living at baseline (n=250) and followed them to compare cumulative incidence of disability over 15 months. The aerobic exercise group had decreased risk of disability compared to the attention control group, RR=0.53 (95% CI 0.33, 0.85), as did the muscle performance exercise group compared to the attention control group, RR=0.60 (95% CI 0.38, 0.97).[46]

A small improvement in function long term was seen across two trials of combination exercise compared with usual care, one fair[45] and the other poor quality,[53] (pooled SMD -0.24, 95%CI -0.37 to -0.11 I^2=0%), although separately neither trial reached statistical significance (Figure 32).

Pain outcomes. Exercise was associated with slight improvement in pain short term compared with usual care, no treatment, or sham (7 trials, pooled difference on a 0-10 scale -0.44 , 95% CI -0.82 to -0.05, I^2=35 %), Figure 33. Six were fair-quality trials[37,44,48,50,51,54] and one poor-quality.[56] Across studies comparing exercise with usual care, results were also similar (5 trials, pooled difference -0.53, 95% CI -1.07 to -0.02).[44,48,50,51,56] The estimate was similar following exclusion of the poor quality trials, but results were no longer significant (5 trials, pooled difference -0.40, 95% CI -0.85 to 0.08). One trial found no difference between exercise or sham procedure in the percentage of patients who reported clinically relevant reduction in VAS pain (58% versus 42%, p=0.069) or global improvement (59% versus 50%, p= 0.309).[37]

Exercise was associated with moderately greater effect on pain in the intermediate term compared with usual care, attention control, or no treatment across pain measures (9 trials, pooled difference -1.61, 95% CI -2.51 to -0.72, I2=91% on a 0-10 scale) across four fair-quality trials[41,43,45,48]and five poor-quality trials[38,42,51,52,55] (Figure 33). Results differed somewhat by type of exercise and type of control. Three trials showed no difference between combination exercise and usual care;[45,48,52] however, a substantial effect on pain was seen across trials comparing muscle performance exercise with an attention control (3 trials, pooled difference -2.18, 95% CI -3.15 to -1.24)[41,43,55] and with no treatment (2 poor quality trials, pooled difference -3.01, 95% CI -4.0 to -1.90).[38,42] Substantial improvement in pain was seen across trials of muscle performance exercise versus attention control or no treatment (5 trials, pooled difference on 0-10 scale -2.53, 95% CI -3.23 to -1.80).[38,41-43,55] Results were no longer significant when four poor quality trials[38,42,52,55] were excluded (3 trials, pooled difference on a 0-10 scale -1.69, 95% CI -3.74 to 0.30).[41,43,45]

There was no clear difference between exercise and usual care or attention control on pain long term (pooled difference -0.24 on a 0 to 10 scale, 95% CI -0.72 to 0.24, I2=54.9%), but data are limited to two trials, the largest of which was of poor quality[45,53] (Figure 33).

Most trials evaluated pain using a traditional 0 to 10 VAS. A small improvement in pain short term was observed (3 trials, pooled difference -0.51, 95% CI -1.01 to -0.01, I2=0%) across three trials (2 fair, 1 poor quality)[37,48,56] but was marginally statistically significant. Findings for intermediate-term were similar to the above findings across pain measures, showing a moderate improvement in pain (6 trials, pooled difference -2.29, 95% CI -3.02 to -1.55, I2=78%).[38,41-

[43,52,55] The pooled estimate was slightly larger when four poor-quality trials[38,42,52,55] were excluded, leaving two fair-quality trials (pooled difference -2.62, 95%CI -3.33 to -1.89).[41,43] Stratification by control type among studies reporting VAS pain yielded similar findings to those across multiple measures. Estimates were similar when analyses were stratified according to the type of exercise. No trial employing VAS reported on long-term pain.

Other outcomes. Health-related quality of life outcomes had mixed results (Table 24). Two fair-quality trials found no association between exercise and short-term quality of life on the KOOS 0 to 100 scale (pooled difference 1.76, 95% CI -2.45 to 5.97, I^2=0%, plot not shown).[44,50] A fair-quality trial (n=65) reported no differences in mean change for short-term SF-36 PCS (mean change of 3.0 (95% CI -5.9 to 16.3) versus -0.7 (95% CI -14.8 to 9.8)) and SF-36 MCS (mean change of 0.7 (95% CI -18.1 to 13.2) vs. -0.7 (95% CI -16.8 to 12.8)).[54] One fair-quality trial (n=158) reported similar health-related quality of life scores between a combined exercise group and usual care using averaged intermediate- and long-term scores. The adjusted mean (standard error (SE)) SF-36 PCS were 37.6 (0.9) vs. 35.3 (0.8), respectively, and adjusted mean (SE) SF-36 MCS were 54.1 (0.8) vs. 53.7 (0.8), respectively.[49] A poor-quality trial (n=50) reported intermediate-term SF-36 scores for individual domains. Functional capacity, physical role, bodily pain, general health, and vitality were slightly improved with exercise versus attention control.[39]

A fair-quality trial (n=438) reported no difference in depressive symptoms compared with attention control, 2.59 vs. 2.80 (p=0.27) for muscle performance exercise, while aerobic exercise was associated with fewer depressive symptoms on the CES-D compared to attention control, 2.12 vs. 2.80 (p<0.001).[47]

There was insufficient evidence to determine effects of duration of exercise therapy or number of sessions on outcomes. No trials reported on changes in opioid use as a result of exercise programs.

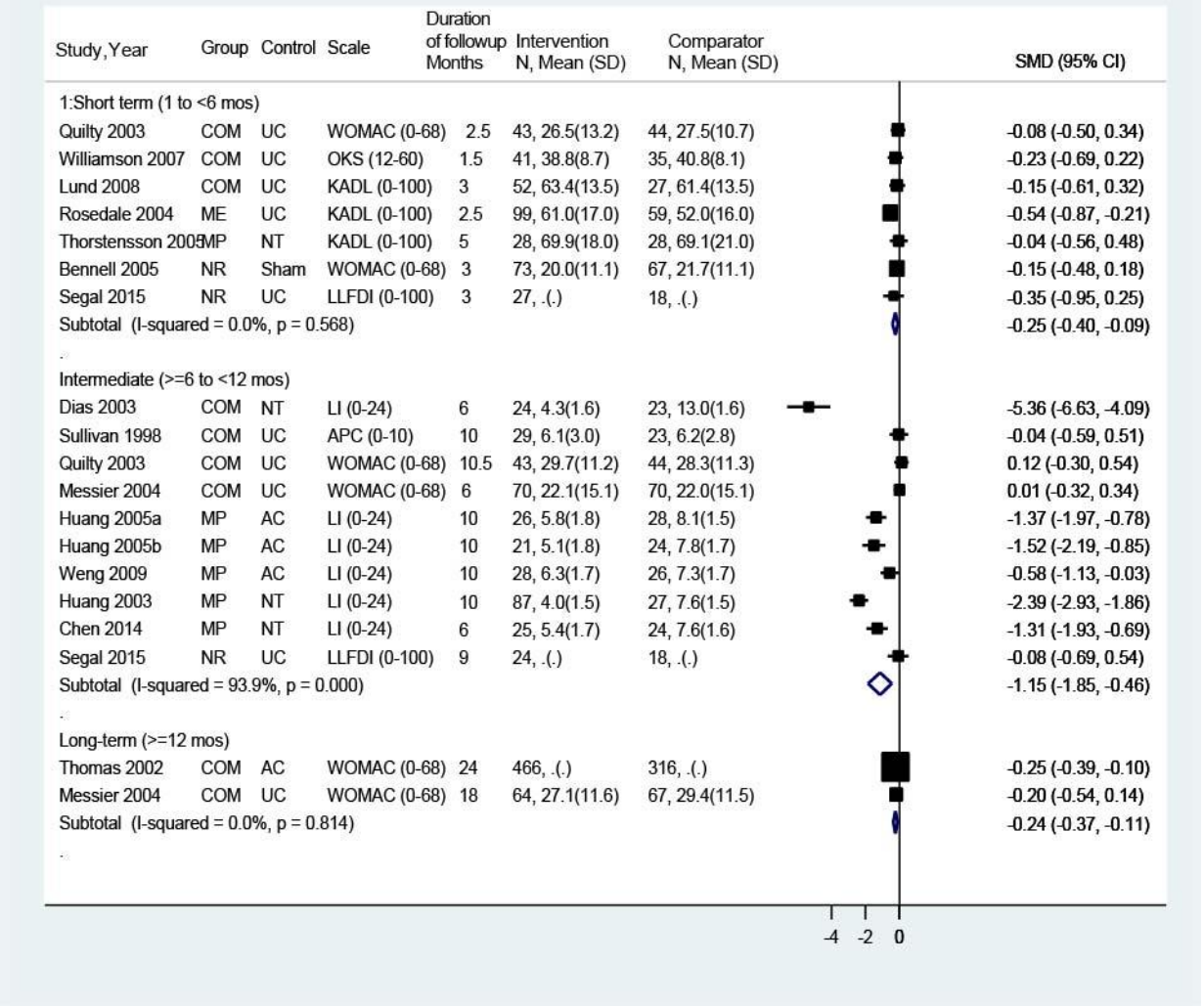
Exercise Compared With Pharmacological Therapy or With Other Non-pharmacological Therapies

No trial of exercise therapy versus pharmacological therapy met inclusion criteria. Findings for exercise versus other nonpharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

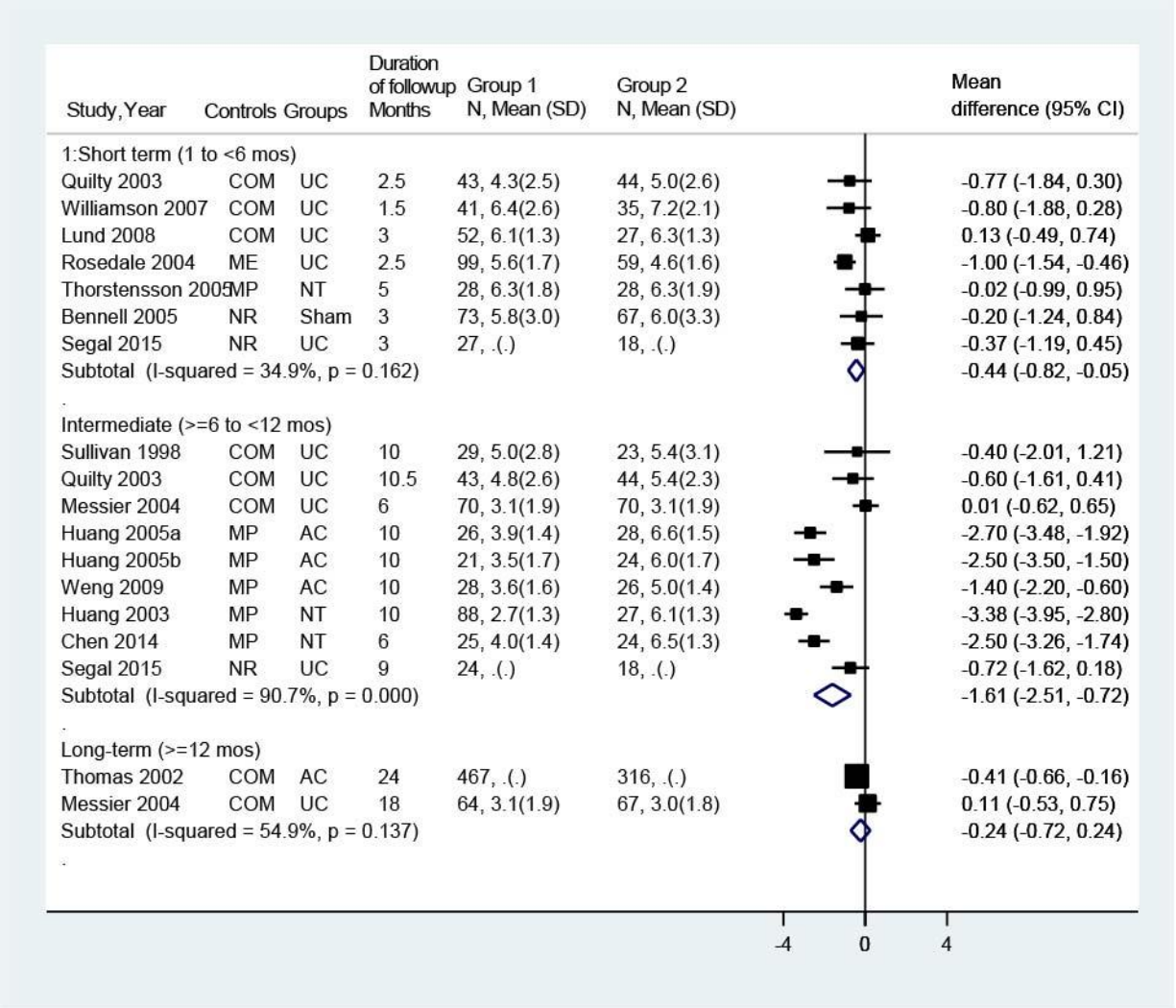
Most trials did not report harms. One trial reported greater temporary, minor increases in pain in the exercise group versus a sham group, RR=14.7 (95% CI 2.0, 107.7); however, the confidence interval is wide.[37] Four studies found no difference in worsening of pain symptoms with exercise versus comparators.[38,42,54,55] One trial found no difference in falls or deaths.[40]

Figure 32. Exercise versus usual care, no treatment, sham, or an attention control for osteoarthritis of the knee: effects on function



AC = attention control; APC = Arthritis Impact Measurement Scale (AIMS) physical activity component; CI = confidence interval; COM = combination exercise therapy; KADL, Knee Injury and Osteoarthritis Outcome Score (KOOS) ADL subscore; LI = Lequesne Index; LLFDI: Late Life Function and Disability Index Basic Lower Limb Function Score; ME = mobility exercise; MP = muscle performance exercise; NR = neuromuscular reeducation exercise; NT = no treatment; OKS = Oxford Knee Score; SD = standard deviation; SMD = standardized mean difference; UC = usual care; WOMAC = Western Ontario and McMaster's Universities Osteoarthritis Index

Figure 33. Exercise versus usual care, no treatment, sham, or an attention control for osteoarthritis of the knee: effects on pain



AC = attention control; CI = confidence interval; COM = combination exercise therapy; ME = mobility exercise; MP = muscle performance exercise; NR = neuromuscular re-education exercise; NT = no treatment; SD = standard deviation; SMD = standardized mean difference; UC = usual care

Psychological Therapy for Osteoarthritis of the Knee

Key Points

- One fair-quality and one poor-quality study of pain coping skills training and cognitive behavioral training versus usual care found no differences in function (WOMAC physical, 0-100) or pain (WOMAC pain, 0-100); treatment effects were averaged over short-term to intermediate-term followup (difference -0.3, 95% CI -8.3 to 7.8 for function and -3.9, 95% CI -1.8 to 4.0 for pain) and intermediate-term to long-term followup (mean 35.2, 95% CI 31.8 to 38.6 vs. mean 37.5, 95% CI 33.9 to 41.2, and mean 34.5, 95% CI 30.8 to 38.2 vs. mean 38.0, 95% CI 34.1 to 41.8), respectively (SOE: Low).
- No serious harms were reported in either trial (SOE: Low).

Detailed Synthesis

Two trials (n=111 for each) comparing psychological therapies conducted in group settings with usual care for OA met the inclusion criteria[89,90] (Table 25 and Appendix D). One trial evaluated CBT led by a trained psychologist and physiotherapist,[89] and the second trial evaluated pain coping skills training delivered by clinical psychologists.[90] Usual care was defined as routine care provided by the patient's primary care doctor and was not well-described in either trial. Across the two trials, respectively, duration of treatment was 6 and 24 weeks and the total number of sessions was six (1 per week, 2 hours duration) and 18 (1 per week for 12 weeks then 1 every other week for 12 weeks, 1 hour duration). One trial was conducted in Finland,[89] and the other in the United States.[90] One trial[89] was rated fair quality and the other poor quality[90] (Appendix E). The primary methodological limitation in the fair quality trial was the inability to effectively blind patients. Additional methodological shortcomings in the poor quality trial included poor treatment compliance and high attrition (32%).

Table 25. Summary of results for osteoarthritis of the knee: psychological therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Helminen, 2015[89] 31.5 to 10.5 months Duration of pain: 7.8 years *Fair*	Cognitive-Behavioral Training plus usual care (n=55): 2-hour groups sessions, weekly for 6 weeks (6 sessions total); included attention diversion methods (relaxation, imagery, distraction), activity-rest cycling and pleasant activity scheduling, cognitive restructuring, and homework assignments B. Usual Care (n=56)	A vs. B Age: 64.5 vs. 63 years Female: 71% vs. 68% BMI: 30 vs. 30 kg/m^2 Bilateral knee OA: 33% vs. 30% Kellgren-Lawrence grade 2: 60% vs. 61% Duration of Chronicity: 6.6 vs. 8.9 years WOMAC Function (0-100): 53.0 vs. 48.4 WOMAC Pain (0-100): 57.6 vs. 56.4 WOMAC Function: 53.0 (48.1−57.9) vs. NRS pain (0-10), average past week: 6.6 vs. 6.4 NRS pain (0-10), worst past week: 8.0 vs. 7.5 NRS pain (0-10), average 3 months: 6.8 vs. 6.6 NRS pain (0-10), worst 3 months: 8.2 vs. 8.0	A vs. B <u>Post-Treatment Average (1.5 to 10.5 months)</u> WOMAC Function: 36.5 vs. 36.7, mean difference−0.3 (95% CI −8.3 to 7.8) WOMAC Pain: 35.6 vs. 39.5, mean difference−3.9 (95% CI −11.8 to 4.0) NRS pain, average past week: 5.0 vs. 4.9, mean difference 0.02 (95% CI −0.89 to 0.93) NRS pain, worst over week: 6.1 vs. 5.9, mean difference 0.1 (95% CI −0.8 to 1.1) NRS pain, average 3 months: 5.2 vs. 5.4 mean difference −0.2 (95% CI −1.0 to 0.6) NRS pain, worst 3 months: 6.4 vs. 6.6, mean difference −0.1 (95% CI −0.9 to 0.7)	A vs. B <u>Post-Treatment Average (1.5 to 10.5 months)</u> WOMAC Stiffness (0-100): 46.2 vs. 49.0 mean difference −2.7 (95% CI −11.4 to 5.9) BDI (0−63): 5.8 vs. 5.9, mean difference −0.1 (95% CI −2.2 to 2.0) BAI (0−63): 8.0 vs. 7.1, mean difference0.9 (95% CI −1.3 to 3.1) HRQoL, 15D: 0.82 vs. 0.85, mean difference −0.03 (95% CI −0.06 to 0.00) RAND-36 Physical Functioning: 48.0 vs. 49.4 mean difference −1.4 (95%CI −10.2 to 7.3) RAND-36 Role-Physical: 44.4 vs. 44.5 mean difference −0.09 (95% CI −14.4 to 14.3) RAND-36 Bodily Pain: 57.3 vs. 57.4, mean difference −0.1 (95% CI −8.0 to 7.7) RAND-36 General Health: 53.1) vs. 58.2, mean difference−5.0 (95% CI −12.3 to 2.3)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				RAND-36 Vitality: 62.7 vs. 67.5, mean difference−4.8 (95% CI −12.6 to 3.1) RAND-36 Social Functioning: 75.0 vs. 82.8, mean difference−7.8 (95% CI −16.4 to 0.81) RAND-36 Role-Emotional: 67.9 vs. 74.7, mean difference−6.7 (95% CI −20.2 to 6.8) RAND-36 Emotional Well-Being: 75.3 vs. 78.5, mean difference−3.2 (95% CI −9.5 to 3.1) RAND-36 Health Change: 46.6 vs. 47.4, mean difference −0.8 (95% CI −9.2 to 7.6)
Somers, 2012[90] 6-12 months Duration of pain: NR *Poor*	A. Pain Coping Skills Training (n=60): 1-hour group sessions, weekly for 12 weeks then every other week for 12 weeks (total of 18 sessions over 24 weeks); consisted of informational lectures, problem solving, skills training, relaxation exercises, homework assignments, and feedback B. Usual Care (n=51)	A vs. B Age: 58 vs. 58 years Female: 67% vs. 78% Caucasian: 62% vs. 61% Mean Duration of Chronicity: NR Kellgren-Lawrence score (0-4): 2.5 vs. 2.3 WOMAC function subscale (0-100): 46.2 vs. 46.1 WOMAC pain subscale (0-100): 42.8 vs. 43.4	A vs. B <u>Post-treatment Average (6-12 months)</u> WOMAC function: 35.2 vs. 37.5, p=NS AIMS physical disability subscale: 1.5 vs. 1.4, p=NS WOMAC pain subscale: 34.5 vs. 38.0, p=NS AIMS pain subscale: 4.4 vs. 4.7, p=NS	A vs. B <u>Post-treatment Average (6-12 months)</u> WOMAC stiffness subscale: 44.5 vs. 46.4, p=NS AIMS psychological subscale: 2.6 vs. 2.5, p=NS

AIMS = Arthritis Impact Measurement Scale; BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; CI = confidence interval; HRQoL = health-related quality of life; NR not reported; NRS = Numerical Rating Scale; NS = not statistically significant; OA = osteoarthritis; WOMAC = Western Ontario and McMaster Universities Osteoarthritis index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Psychological Therapies Compared With Usual Care

Both trials reported outcomes averaged over all post-treatment followup times; the trial of CBT averaged results from 1.5 to 10.5 months post treatment (spanning short to intermediate

term)[89] and the trial of pain coping skills training averaged results from 6 to 12 months post treatment (spanning intermediate to long term).[90]

No significant differences in function or pain were found between the psychological therapy and the usual care groups in either trial. Function was measured using the WOMAC physical function subscale (0-100) in both trials, over the short to intermediate term (mean difference −0.3, 95% CI −8.3 to 7.8)[89] and intermediate to long term (mean 35.2, 95% CI 31.8 to 38.6 vs. mean 37.5, 95% CI 33.9 to 41.2),[90] and using the Arthritis Impact Measurement Scale (AIMS) physical disability subscale in one trial[90] (Table 25). Both trials measured pain using the WOMAC pain subscale (0-100), one trial over short to intermediate term followup (mean difference -3.9, 95% CI -11.8 to 4.0)[89] and the other over intermediate- to long-term followup (mean 34.5, 95% CI 30.8 to 38.2 vs. mean 38.0, 95% CI 34.1 to 41.8).[90] Results were similar for the AIMS pain subscale and the Numerical Rating Scale (NRS) pain scale, reported by one trial each (Table 25). Neither trial reported any differences between groups in any secondary outcome measure.

No trial evaluated effects of psychological therapies on use of opioid therapies or health care utilization.

Psychological Therapies Compared With Pharmacological Therapy or With Exercise Therapy

No trial of psychological therapy versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

In the two trials of psychological interventions,[89,90] no adverse events were observed.

Physical Modalities for Osteoarthritis of the Knee

Key Points

Ultrasound

- One fair-quality trial found continuous and pulsed ultrasound associated with better short-term function (difference of -6.2, 95% CI -8.36 to -4.20, and -5.71, 95% CI -7.72 to -3.70 on a 0-24 scale) and short-term pain intensity (difference -3.3, 95% CI -4.64 to -1.96, and -3.37, 95% CI -4.73 to -2.01 on a 0-10 scale) (SOE: Low).
- One fair-quality trial found no difference between continuous and pulsed ultrasound versus sham in intermediate-term function (difference -2.9, 95% CI -9.19 to 3.39 and 1.6, 95% CI -3.01 to 6.22, on a 0-68 scale) or pain (difference -1.6, 95% CI -3.26 to 0.06 and 0.2, 95% CI -1.34 to 1.74, on a 0-20 scale). There was also no difference between groups for VAS pain during rest or on movement (SOE: Low).
- No adverse events were reported during the two trials (SOE: Low).

Transcutaneous Electrical Nerve Stimulation (TENS)

- One trial found no difference between TENS and placebo TENS in intermediate-term function as measured by the WOMAC function subscale ((proportion of patients who achieved MCID (≥9.1), 38% vs. 39%, RR 1.2 (95% CI 0.6 to 2.2); and difference -1.9 (95% CI -9.7 to 5.9) on a 0-100 scale)) or intermediate-term pain ((proportion of patients who achieved MCID (≥20) in VAS pain, 56% vs. 44%, RR 1.3 (95% CI 0.8 to 2.0); and

mean difference –5.6 (95% CI –14.9 to 3.6) on the 0-100 WOMAC pain subscale)) (SOE: Low for function and pain).
- One trial of TENS reported no difference in the risk of minor adverse events (RR 1.06 (95% CI 0.38 to 2.97) (SOE: Low).

Low Level Laser Therapy
- Evidence was insufficient from one small fair-quality and two poor-quality trials to determine effects or harms of low-level laser therapy in the short or intermediate term; No data were available for the long term (SOE: Insufficient)

Microwave Diathermy
- There was insufficient evidence to determine short-term effects or harms from one small, fair-quality trial (SOE: Insufficient).

Pulsed Short-wave Diathermy
- There was insufficient evidence to determine effects or harms from one poor-quality trial in the short term or from another poor quality trial in the long term (SOE: Insufficient).

Electromagnetic Field
- One fair-quality trial found pulsed electromagnetic fields associated with small effects on function (difference -3.48, 95% CI -4.44 to -2.51 on a 0-85 WOMAC ADL subscale) and pain (difference -0.84, 95% CI -1.10 to -0.58 on a 0-25 WOMAC pain subscale) versus sham short-term but may not be clinically significant (SOE: Low).
- More patients who received real versus sham electromagnetic field therapy reported throbbing or warming sensations or aggravation of pain (29% versus 7%); however, the difference was not significant (RR 1.95, 95% CI 0.81 to 4.71) (SOE: Low).

Superficial Heat
- Evidence was insufficient from one small fair-quality trial to determine effects or harms of trial superficial heat versus placebo in short-term pain (SOE: Insufficient).

Braces
- There is insufficient evidence from one poor-quality study to determine the effects of bracing versus usual care for intermediate-term and long-term function or pain (SOE: Insufficient).
- Harms were not reported.

Detailed Synthesis

A total of 13 trials[119-131] reported the use of a physical modality for the treatment of knee OA (Table 26 and Appendixes D and E). Physical modalities evaluated included ultrasound, TENS, low-level laser therapy, microwave diathermy, pulsed short-wave diathermy, electromagnetic fields, superficial heat, and bracing. All but one intervention (bracing vs. usual care)[121] were compared to a sham procedure.

Table 26. Summary of results for osteoarthritis of the knee: physical modalities

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Al Rashoud, 2014[119] 1.5 and 6 months Duration of pain: 11 years *Fair*	A. Low level laser therapy (n=26), continuous laser (30 mW, 830 nm wavelength) applied to 5 acupuncture points over approximately of 10 sessions B. Placebo laser (n=23), placebo laser applied to 5 acupuncture points over approximately 10 sessions	A vs. B Age: 52 vs. 56 years Female: 62% vs. 65% Baseline pain on movement VAS (0-10): 6.4 vs. 5.9 Baseline Saudi Knee Function Scale (SKFS) (0-112), median: 61.0 vs. 60.0	A vs. B 1.5 months Pain on movement VAS: 3.0 vs. 4.2[b] SKFS, median: 31 vs. 40, median difference -10 (95% CI -23 to -4) p=0.054 6 months Pain on movement VAS: 3.4 vs. 5.2[b] SKFS, median: 31 vs. 51, median difference -21 (95% CI -34 to -7) p=0.006	NR
Battisti, 2004[120] 1 month Duration of pain: 11 years *Poor*	A. Therapeutic Application of Musically Modulated Electromagnetic Field (TAMMEF) (n=30) The anatomical region treated is placed between opposing faces of low frequency electromagnets (3x4 cm). The current from amplifier B feeds a loud speaker that plays music. The music modifies parameters (frequency, intensity, waveform) of the electromagnetic field in time, randomly varying within respective ranges. 15 consecutive daily sessions, 30 minutes each B. Extremely Low Frequency (ELF) (n=30) Similar treatment as Intervention A except the electromagnetic field is stabilized at a frequency of 100Hz in a sinusoidal waveform. 15 consecutive daily	A + B + C Age: 58.9 (7.4) Female: 70% Race: NR Mean Duration of Chronicity: 11 (3.1) A vs. B vs. C Mean Lequesne Pain Score (0-10)[c]: 6.88 vs. 6.28 vs. 6.15 Mean Lequesne Function Score(0-10)[c]: 3.65 vs. 4.28 vs. 3.48	A vs. C 1 month Mean Lequesne Pain Score: 1.4 vs. 6.85 Mean Lequesne Functionality: 6.5 vs. 3.83 B vs. C 1 month Mean Lequesne Pain Score: 1.4 vs. 6.85 Mean Lequesne Functionality: 7.1 vs. 3.83	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	sessions, 30 minutes each C. Simulated (Sham) Frequency Field (n=30) Functionally similar operation to the other groups except a simulated (noneffective) field is used, but the patients remain blinded to its effectiveness. 15 consecutive daily sessions, 30 minutes each			
Brouwer, 2006[121] 6 and 12 months Duration of pain: 6.7 vs. 4.9 years *Poor*	A. Brace (n=60) Patients were fitted with a commercially available knee brace that allowed medial unloading or lateral unloading and also received usual care. Device: Oasys brace, Innovation Sports, Irvine, CA, USA B. Usual Care (n=57) Usual care was identical in both groups and consisted of patient education (adaptation of activities and/or weight loss), and (if needed) physical therapy and analgesic	A vs. B Age[f]: 59.2 Female: 48% vs. 51% Race: NR VAS Pain Severity(0-10): 6.6 vs. 5.5 Hospital for Special Surgery (HSS) Knee Function Score (0-100): 64.9 vs. 69.0 EuroQol-5D Quality of Life (0-1): 0.50 vs. 0.56	A vs. B 6 months VAS Pain Severity: MD -0.58 (95%CI -1.48 to 0.32) HSS Knee Function: MD 3.2 (95%CI -0.58 to 6.98) 12 months (post-treatment) VAS Pain Severity: MD -0.81 (95%CI -1.76 to 0.14) HSS Knee Function: MD 3.0 (95%CI -1.05 to 7.05)	A vs. B 6 months EQ-5D Quality of Life: MD 0.01 (95%CI -0.08 to 0.10) 12 months (post-treatment) EQ-5D Quality of Life: MD 0.01 (95%CI -0.08 to 0.10) Adverse Events: NR
Cakir, 2014[122] 6 months Duration of pain: Mean 4.0 to 5.1 years *Fair*	A. Continuous ultrasound (n=20), Therapeutic ultrasound given 5 times a week for 2 weeks B. Pulsed ultrasound (n=20), Therapeutic pulsed ultrasound given 5	A vs. B vs. C Age: 57 vs. 58 vs. 57 years Female: 70% vs. 80% vs. 85% WOMAC physical mean function (0-68): 55.7 vs. 52.4 vs. 52.5 WOMAC pain (0-	A vs. C 6 months WOMAC physical function: 32.6 vs. 35.5, difference -2.9 (95% CI -9.2 to 3.4) WOMAC pain: 9.5 vs. 11.1, difference -1.6 (95% CI -3.3 to 0.1) Pain at rest VAS: 21.4 vs. 22.3, difference 1.2 (95% CI -9.1 to 11.5)	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	times a week for 2 weeks <u>C. Sham</u> (n=20), Sham ultrasound given 5 times a week for 2 weeks All patients performed home exercise program 3 days a week for 8 weeks	20):15.9 vs. 14.5 vs. 14.9 WOMAC stiffness (0-8): NR Pain at rest VAS (0-10): 57.9 vs. 55.7 vs. 53.6 Pain on movement VAS (0-10): 75.5 vs. 73.0 vs. 72.2 Disease severity VAS (0-10): 73.9 vs. 67.9 vs. 68.4	Pain on movement VAS: 38.7 vs. 38.1, difference 0.6 (95% CI -13.7 to 14.9) Disease severity VAS: 30.0 vs. 29.5, difference 0.5 (95% CI -6.7 to 7.7) B vs. C <u>6 months</u> WOMAC physical function: 37.1 vs. 35.5, difference 1.6 (95%CI -3.01 to 6.22) WOMAC pain: 11.3 vs. 11.1, difference 0.2 (95%CI -1.34 to 1.74) Pain at rest VAS: 20.2 vs. 22.3, difference -2.1 (95%CI -11.2 to 7.0) Pain on movement VAS: 37.5 vs. 38.1, difference -0.6 (95%CI -16.98 to 15.78) Disease severity VAS: 32.5 vs. 29.5, difference 3.0 (95%CI -3.95 to 9.95)	
Fary, 2011[123] 6.5 months Duration of pain: 12 years *Good*	<u>A. Pulsed electrical stimulation</u> (n=34), pulsed electrical stimulator worn 7 hours a day daily for 26 weeks <u>B. Placebo electrical stimulation</u> (n=36), placebo pulsed electrical stimulator worn 7 hours a day daily for 26 weeks	A vs. B Age: 71 vs. 69 years Female: 50% vs. 44% Baseline WOMAC total (0-100): 36 vs. 34 Baseline WOMAC function (0-100): 35 vs. 34 Baseline WOMAC stiffness (0-100): 45 vs. 41 Baseline WOMAC pain (0-100): 35 vs. 36 Baseline pain VAS (0-100): 51 vs. 52	A vs. B <u>6.5 months</u> Proportion of patients who achieved MCID (≥9.1) in WOMAC function: 38% vs. 39%, RR 1.2 (95% CI 0.6 to 2.2) Proportion of patients who achieved MCID (≥20) in pain VAS: 56% vs. 44%, RR 1.3 (95% CI 0.8 to 2.0) Mean change in WOMAC total: 6 vs. 7, MCD -1.3 (-8.8 to 6.3) Mean change in WOMAC function: 5 vs. 7, MCD -1.9 (95% CI -9.7 to 5.9) Mean change in WOMAC stiffness: 9 vs. 5, MCD 3.7 (95% CI -6.0 to 13.5) Mean change in WOMAC pain: 5 vs. 10, MCD -5.6 (95% CI -14.9 to 3.6) Mean change in pain VAS: 20 vs. 19, MCD 0.9 (95% CI -11.7 to 13.4)	A vs. B <u>6.5 months</u> Mean change in SF-36 physical component score (0-100): -1.0 vs. -2.6, MCD 1.7 (95% CI -1.5 to 4.8) Mean change in SF-36 mental component score (0-100): -1.2 vs. -2.4, MCD 1.2 (95% CI -2.9 to 5.4)
Fukuda,[124] 2011 12 months Duration of pain: NR *Poor*	<u>A. Low-dose Pulsed Short Wave (PSW) (n=32)</u> Patients administered a pre-calibrated device to the anterior area of	A vs. B vs. C Age: 62 vs. 63 vs. 57 Female: 100% Race: NR Knee Injury and	A vs. C <u>12-months</u> KOOS Symptoms Subscale: 61.6 vs. 40.7, difference 20.9 (95% 8.92 to 32.88) KOOS Daily Activities	A vs. C <u>12-months</u> KOOS Quality of Life Subscale: 31.8 vs. 33.0

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	the thigh, 5 cm above superior border of the patella. Device was set to output a specific frequency and pulse duration, with a care provider nearby but without direct input from care provider. Three, 19 minute applications per week for three weeks (9 total) Total Energy: 17 kJ Frequency: 27.12 MHz Mean Power Output: 14.5 W Pulse Duration: 400 microseconds Pulse Frequency: 145 Hz <u>B. High-dose PSW (n=31)</u> Treatment characteristics were identical to Group A except length of treatment (and received total energy) were doubled. Three, 38 min applications per week for three weeks (9 total) Total Energy: 33 kJ <u>C. Sham (n=23)</u> Treatment characteristics were identical to Group A except the device was kept in standby mode without any electrical current applied. Three, 19 min applications per week for 3 weeks (9 total)	Osteoarthritis Outcome Score Symptoms Subscale (0-100): 46.5 vs. 47.0 vs. 42.0 KOOS Daily Activities Subscale (0-100): 45.8 vs. 51.7 vs. 45.7 KOOS Recreational Activities Subscale (0-100): 16.6 vs. 15.3 vs. 18.2 KOOS Pain Subscale (0-100): 37.4 vs. 42.5 vs. 38.0 NRS Pain(0-10): 7.1 vs. 6.7 vs. 7.7 KOOS Quality of Life Subscale (0-100): 26.1 vs. 32.4 vs. 27.8	Subscale: 68.9 vs. 41.6, difference 27.30 (95% 13.73 to 40.87) KOOS Recreational Activities Subscale: 24.6 vs. 11.0, difference 13.6 (95% -0.73 to 27.93) KOOS Pain Subscale: 57.5 vs. 33.0, difference 24.5 (95% 12.12 to 36.88) NRS Pain: 5.7 vs. 7.5, difference -1.8 (95% -3.60 to 0.00) B vs. C <u>12-months</u> KOOS Symptoms Subscale: 54.9 vs. 40.7, difference 14.2 (95% 1.21 to 27.19) KOOS Daily Activities Subscale: 51.9 vs. 41.6, difference 10.30 (95% -1.24 to 21.84) KOOS Recreational Activities Subscale: 15.9 vs. 11.0, difference 4.9 (95% -5.32 to 15.12) KOOS Pain Subscale: 57.6 vs. 33.0, difference 24.6 (95% 14.59 to 34.61) NRS Pain: 5.2 vs. 7.5, difference -2.3 (95% -3.68 to -0.92)	<u>B vs. C</u> <u>12-months</u> KOOS Quality of Life Subscale: 41.2 vs. 33.0 A vs. B vs. C <u>Adverse Events:</u> Went on to have a Total Knee Replacement during 12 month followup: 3.1% (1/32) vs. 6.5% (2/31) vs. 4.3% (1/23)
Giombini, 2011[125] 3 months Duration of pain: 3 years *Fair*	<u>A. Microwave diathermy (n=29)</u> hyperthermic treatment 3 times a week for 4 weeks <u>B. Sham diathermy (n=25)</u> sham	A vs. B Age: 67 vs. 67 years Female: 66% vs. 68% Baseline WOMAC total (0-1.20): 103.1 vs. 101.3	A vs. B <u>3 months</u> Mean change in WOMAC total: -46.8 vs. -0.4, difference -46.4 (95% CI -58.3 to -34.5) Mean change in WOMAC pain; -8.6 vs. -0.6,	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	hyperthermic treatment 3 times a week for 4 weeks	Baseline WOMAC pain (0-25): 19.2 vs. 18.5 Baseline WOMAC stiffness (0-10): 9.7 vs. 9.7 Baseline WOMAC activities of daily living (0-85): 74.3 vs. 73.1	difference -8.1 (95% CI -10.7 to -5.3) Mean change in WOMAC activities of daily living: -33 vs. 0.3, difference -33.2 (95% CI -42.0 to -24.6) Mean change in WOMAC stiffness: -5.2 vs. -0.1, difference -5.1, p<0.01	
Hegedus, 2009[126] 2 months Duration of pain NR *Poor*	<u>A. Low Level Laser Therapy (n=18)</u>: 50 mW, continuous wave laser (wavelength 830 nm). Treatment provided over the femoral and tibial condyles. Total dose of 48 J/cm2 per session. Twice a week for four weeks. <u>B. Placebo (n=17)</u>: Placebo probe (0.5 mW power output) used twice a week for four weeks.	Age: 49 Female: 81% A vs. B Pain VAS (0-10): 5.75 vs. 5.62	A vs. B <u>2 months</u> Pain VAS: 1.18 vs. 4.12, difference -2.94 (no estimate of variability provided or calculable)	NR
Laufer, 2005[127] 3 months Duration of pain: NR *Poor*	<u>A. Low Intensity Pulsed Shortwave Diathermy (n=38)</u> Treatment Protocol: Shortwave diathermy was applied to anterior aspect of the affected knee; Three, 20 min sessions per week for 3 weeks (9 total) Pulse Duration: 82 µs Pulse Frequency: 110 Hz Peak Power: 200 W (mean 1.8W)) <u>B. High Intensity Pulsed Shortwave Diathermy (n=32)</u> Treatment protocol identical to Group A except with a higher intensity (pulse duration and frequency) Pulse Duration: 300 µs	A vs. B vs. C Age: 75 vs. 73 vs. 73 Female: 82% vs. 90.6% vs. 67% Race: NR Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Overall: 5.13 vs. 4.60 vs. 5.02 WOMAC Pain: 4.89(3.30) vs. 4.43(3.35) vs. 4.97(3.52); WOMAC Stiffness : 4.87(3.50) vs. 4.25(3.47) vs. 4.92(3.58) WOMAC Activities of Daily Living: 5.16(3.52) vs. 4.69(3.41) vs. 5.05(3.45);	A vs. C <u>3 months</u> WOMAC Overall: 4.82 vs. 4.60, difference 0.22 (95% CI -1.51 to 1.95) WOMAC Pain: 4.48 vs. 4.33, difference 0.15 (95% CI -1.57 to 1.87) WOMAC Stiffness: 4.43 vs. 3.60, difference 0.83 (95% CI -0.98 to 2.64) WOMAC Activities of Daily Living: 4.98 vs. 4.82, difference 0.16 (95% CI -1.51 to 1.83) B vs. C <u>3 months</u> WOMAC Overall: 4.56 vs. 4.60, difference -0.04 (95% CI -1.75 to 1.67) WOMAC Pain: 4.09 vs. 4.33, difference -0.24 (95% CI -2.02 to 1.54) WOMAC Stiffness: 3.81 vs. 3.60, 0.21 (95% CI -1.55 to 1.97) WOMAC Activities of Daily Living: 4.8 vs. 4.82, difference -0.02 (95% CI -	A vs. B vs. C <u>Adverse Events:</u> No adverse reactions to the treatment were reported by the subjects.

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	Pulse Frequency: 300 Hz Peak Power: 200 W (mean 18W) <u>C. Sham Shortwave Diathermy (n=33)</u> Identical treatment except the apparatus was turned on but the power output was not raised.		1.67 to 1.63)	
Mazzuca, 2004[128] 1 month Duration of pain: NR *Fair*	<u>A. Superficial Heat (sleeve) (n=25)</u> Participants word a cotton and lycra sleeve with a heat retaining polyester and aluminum substrate. Patients were asked to wear the sleeve at least 12 hours each day and to continue their usual pain medication(s). <u>B. Placebo Sleeve (n=24)</u> Placebo sleeves and treatment protocol were identical except placebo sleeves did not contain the heat retaining substrate layer.	A + B Age: 62.7 Female: 77% Race: 67% white WOMAC Pain (5-25)[d]: 15.2 vs. 14.7* WOMAC Stiffness (2-10)[e]: 6.5 (1.4) WOMAC Function (17-85)[e]: 51.8 (11.8)	A vs. B <u>1 month</u> WOMAC Pain: 13.7 vs. 13.9	NR
Tascioglu, 2004[129] 6 months Duration of pain: 7 years *Poor*	<u>A. Active laser 3 joule (n=20)</u> continuous laser therapy (50 mW, 830 mm wavelength) applied to 5 painful points 5 days a week for 2 weeks <u>B. Active laser 1.5 joule (n=20)</u> continuous laser therapy (50 mW, 830 mm wavelength) applied to 5 painful points 5 days a week for 2 weeks	A vs. B vs. C Age: 63 vs. 60 vs. 64 years Female: 70% vs. 75% vs. 65% Baseline WOMAC function (0-68):36.6 vs. 38.0 vs. 39.5 Baseline WOMAC stiffness (0-8): 4.1 vs. 4.6 vs. 4.5 Baseline WOMAC pain (0-20): 10.3 vs. 11.6 vs. 9.6 Baseline pain at rest VAS (0-100): 39.1 vs. 41.6 vs. 37.9 Baseline pain at	A vs. C <u>6 months</u> WOMAC function: 34.8 vs. 38.7, difference -3.8 (95% CI -9.8 to 2.1) WOMAC stiffness: 3.9 vs. 4.2, difference -0.3 (95% CI -1.6 to 0.9) WOMAC pain: 10.4 vs. 9.9, difference 0.6 (95% CI -1.5 to 2.7) Pain at rest VAS: 38.7 vs. 38.9, difference -0.3 (95% CI -9.8 to 9.3) Pain at activation VAS: 66.8 vs. 62.0, difference 4.8 (95% CI -4.9 to 14.5) B vs. C <u>6 months</u>	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	C. Placebo laser (n=20), sham laser therapy applied to 5 painful points 5 days a week for 2 weeks	activation VAS (0-100): 68.0 vs. 65.7 vs. 63.9	WOMAC function: 38.5 vs. 38.7 WOMAC stiffness: 4.5 vs. 4.2 WOMAC pain: 11.3 vs. 9.9 Pain at rest VAS: 40.0 vs. 38.9 Pain at activation VAS: 61.8 vs. 62.0	
Thamsborg,[130] 2005 1.5 month Duration of pain, 8 years *Fair*	A. Pulsed Electromagnetic Fields (PEMF) (n=42) Two sets of two adjacent coils were placed on the medial and lateral regions of the study knee, with the interspace between the coils being at the level of the koin line. The coils were placed on an insulating bandage of 3-5 mm thickness. 2-hour daily treatment 5 days per week for 6 weeks 30 total Device: ±50V in 50Hz pulses changing voltage in 3 ms intervals. B. Sham Electromagnetic Field (n=41) Patients in the control group were subjected to a noneffective placebo electromagnetic field. No. of Treatments: daily treatment 5 days per week for 6 weeks (30 total Length of Treatments: 2 hours each	A vs. B Age: 60 vs. 60 Female: 47.6% vs. 61% Race: NR WOMAC Activities of Daily Living (0-85): 43.83 vs. 46.49 WOMAC Stiffness (0-10): 5.74 vs. 5.85 WOMAC Joint Pain(0-25): 13.15 vs. 14.49	A vs. B 1.5 months WOMAC Activities of Daily Living: 37.89 vs. 41.3, difference -3.48 (95%CI -4.44 to -2.51) WOMAC Stiffness: 4.81 vs. 5.15, difference -0.34(95%CI -0.48 to -0.20) WOMAC Joint Pain: 11.40 vs. 12.24, difference -0.84 (95%CI -1.10 to -0.58)	A vs. B Adverse Events: throbbing sensation, warming sensations or aggravation of pain 28.5% (12/42) vs. 14.6% (6/41)
Yildiz,[131] 2015 2 months Duration of pain: Mean	A. Continuous ultrasound (n=30), Therapeutic ultrasound given 5 times a week for 2	A vs. B vs. C Age: 56 vs. 55 vs. 58 years Female: 83% vs. 80% vs. 87%	A vs. C 2 months Lequesne Index: 5.5 vs. 11.7, difference -6.2 (95% CI -8.4 to 4.2)	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
2.8 to 5.1 years *Fair*	weeks <u>B. Pulsed ultrasound</u> (n=30), Therapeutic pulsed ultrasound given 5 times a week for 2 weeks <u>C. Sham</u> (n=30), Sham ultrasound given 5 times a week for 2 weeks All patients performed home exercise program 3 days a week for 8 weeks	Lequesne Index score (0-24): 13.2 vs. 12.9 vs. 12.4 Pain at rest VAS (0-10): NR Pain on movement VAS (0-10): 9.0 vs. 8.6 vs. 8.9	Pain at rest VAS: NR Pain on movement VAS: 3.9 vs. 7.2, difference -3.3 (95% CI -4.6 to -2.0) B vs. C <u>2 months</u> Lequesne Index: 6.0 vs. 11.7, difference -5.7 (95% CI -7.7 to -3.7) Pain at rest VAS: NR Pain on movement VAS: 3.8 vs. 7.2, difference -3.4 (95% CI -4.7 to -2.0)	

EQ-5D, EuroQol Quality of Life Instrument 5-D; HSS, Hospital for Special Surgery; Hz, hertz; KOOS, Knee Injury and Osteoarthritis Outcome Score; MD, difference between means; NR, not reported; NRS; numeric rating scale; RR, risk ratio; SKFS, Saudi Knee Function Score; SF-36.; VAS, visual analog scale; W, watts; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index, μs, microsecond

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Values estimated from graph

[c] The study separated outcome values out into slight, moderate and severe disease patient groups for each treatment arm. These values are combined values for each intervention groups estimated from graphs in the study.

[d] Values estimated from graph

[e] Separate group baseline values not given for stiffness and function subscales

[f] Age only reported for population as a whole

Two fair-quality randomized controlled trials that evaluated ultrasound for knee OA met the inclusion criteria.[122,131] Both trials required grade 2 or 3 radiographic knee OA using the Kellgren–Lawrence criteria for inclusion. Both trials had a continuous and a pulsed ultrasound group, which they compared to sham ultrasound. Both ultrasound groups received 1 MHz treatments five times per week for 2 weeks at an intensity of either 1 or 1.5 W/cm^2. Sham ultrasound used the same protocols, but the power was switched off. All participants were also instructed to perform a home exercise program of mostly muscle performance exercises three times per week. Compliance with the intervention protocol was not reported. One trial reported short-term outcomes,[131] the other intermediate-term outcomes. The methodological shortcomings were unclear adherence to an intention-to-treat analysis,[131] and unclear blinding of the provider or assessor.[122]

We found one good-quality (n=70) trial that compared active TENS with sham TENS for knee OA.[123] Inclusion criteria required a confirmed diagnosis of knee OA using the American College of Rheumatology criteria. The TENS protocol had patients wear a pulsed TENS device 7 hours daily for 26 weeks. The sham TENS groups followed the same protocol as the active treatment, but the device turned off after three minutes. Compliance was unacceptable for time the TENS device was worn.

We identified three small trials (n=30, 49 and 60) that investigated low-level laser therapy versus sham laser for knee OA.[119,126,129] The mean age ranged from 49 to 64 years and most patients were female (62% to 75%). Two studies included patients meeting the American College of Rheumatology criteria for knee OA.[119,129] Two trials also required an average pain intensity of greater than 3 or 4 on a 0-10 VAS,[119] while the other trial had an additional inclusion criteria of radiographic knee OA of Kellgren–Lawrence grade of 2 or 3.[129] Treatment duration ranged from 2 to 4 weeks and the number of total sessions from eight to ten. Low-level laser therapy protocols differed across the trials with doses ranging from 1.2 to 6 Joules per point (range, 5 to 6 points) and length of irradiation from 40 seconds to 2 minutes; all trials used a continuous laser beam. The sham laser comparison groups followed the same respective protocols, but the device was inactive. One trial was rated fair quality[119] and two poor quality.[126,129] In the fair-quality trial, blinding of the care provider was unclear. The two poor quality trials suffered from insufficient descriptions of allocation concealment methods, unclear application of intention to treat, lack of clarity regarding patient blinding, and no reporting of or unacceptable attrition.

One small (n=63), fair-quality trial compared microwave diathermy (three, 30-minute sessions per week for 4 weeks) to sham.[125] The inclusion criteria required radiographic knee OA of a Kellgren and Lawrence grade 2 or 3. The power was set to 50 watts. Sham diathermy followed the same protocol, but the machine was set to off. Compliance with the treatment regimen for each group was unclear. Methodological limitations of this study included no blinding of the care providers.

Two trials (n=86 and 115) examined pulsed short-wave diathermy compared to sham diathermy.[124,127] The mean age ranged from 62 to 75 years, and the proportion of female participants ranged from 67 percent to 100 percent. Both trials included patients meeting radiographic criteria for knee OA. Each trial compared two doses of short-wave diathermy to a sham diathermy group; dosages varied by intensity in one trial (mean power output of either 1.8 or 18 Watts for 20 minutes)[127] or by length of session (19 or 38 minutes at 14.5 Watts) in the other.[124] Both trials applied diathermy three times per week for 3 weeks (total of 9 sessions). Each sham diathermy group followed the same treatment protocol, but the electrical current was not applied. Compliance with the treatment regimens was acceptable for both trials. Both trials were rated poor quality due to unclear concealment of treatment allocation, a lack of care provider blinding, and unacceptable attrition.

Two trials (n=90 for both) compared the application of electromagnetic fields to sham interventions for knee OA.[120,130] The mean age of participants was 59 and 60 years, and the proportion of female participants ranged from 48 percent to 70 percent. The mean duration of chronicity ranged from 9-11 years. The good-quality trial enrolled participants meeting the American College of Rheumatology criteria for knee OA.[130] The inclusion criteria was not clearly presented in the poor-quality trial.[120] The intervention group in the good-quality study received 2-hours of pulsed electromagnetic fields five days a week for 6 weeks.[130] The poor-quality trial had a musically modulated electromagnetic field group that received 15 daily 30-minute sessions. Music from a connected speaker modulated the parameters of the electromagnetic field. The study also had an extremely low frequency electromagnetic field group that had 15 daily 30 minutes sessions, but the electromagnetic field was set at a frequency of 100Hz.[120] The sham group in each trial followed the same respective treatment protocol, but used a non-effective electromagnetic field during the sessions. Compliance to the treatment sessions was acceptable in both trials. One trial was rated fair quality[130] and the other was rated

poor quality.[120] Methodological limitations in both trials included unclear methods for allocation concealment. Additionally, in the poor-quality trial, there were baseline dissimilarities between groups, no blinding of patients, providers, or outcome assessors, and attrition was not reported.[120]

A single trial compared superficial heat with placebo (n=52).[128] Participants were included if they had grade 2 or higher using the Kellgren-Lawrence grading for radiographic knee OA. Superficial heat was provided using a knee sleeve with a heat retaining polyester and aluminum substrate. Participants were instructed to wear the sleeve at least 12 hours per day. The placebo sleeves were identical and participants received the same instructions, but the sleeve did not contain the heat retaining substrate; the extent to which patients could be truly blinded is unclear (sleeve may retain body heat and feel warmer). Compliance with wearing the sleeve was acceptable. This trial was rated fair quality due to unclear concealment of treatment allocation, and a lack of clarity regarding whether it was the provider or outcomes assessor that was blinded.

We identified one trial comparing use of a knee brace to usual care (n=118).[121] Inclusion criteria required unicompartmental knee OA, and either a varus or valgus malalignment. Patients in the intervention group were fitted with a commercially available knee brace that allowed medial unloading or lateral unloading. Usual care consisted of patient education and physical therapy and analgesics as needed. Compliance with continued use of the brace was unacceptable. This trial was rated poor quality due to lack of patient, provider, or assessor blinding, and unacceptable attrition.

Physical Modalities Compared with Sham or Usual Care

Ultrasound. One fair-quality trial reported short-term function using Lequesne Index (0-24) and VAS pain (0-10) during activity.[131] Both the continuous ultrasound group and the pulsed ultrasound group had substantially better short-term function versus sham ultrasound (mean difference of -6.2, 95% CI -8.36 to -4.20, and -5.71, 95% CI -7.72 to -3.70, respectively). Continuous and pulsed ultrasound was also associated with substantially less pain during activity compared to sham ultrasound (mean difference of -3.3, 95% CI -4.64 to -1.96, and -3.37, 95% CI -4.73 to -2.01, respectively, on a 0 to 10 scale).

Intermediate-term results at 6 months from the other fair-quality trial showed no difference on the WOMAC Physical Function subscale (0 to 100) between either the continuous or pulsed ultrasound group versus sham ultrasound (mean difference of -4.5, 95% CI -10.34 to 1.34, and -2.9m 95% CI -9.19 to 3.39, respectively).[122] Results for pain intensity were not consistent with regard to ultrasound method. The continuous ultrasound group had slightly less pain on the WOMAC pain scale compared to sham (mean difference -1.8, 95% CI -3.34 to -0.26), but no statistical difference was seen between pulsed ultrasound and sham (mean difference of -1.6, 95% CI -3.26 to 0.06). There was no difference between either ultrasound group versus sham ultrasound for VAS pain during rest or on movement (Table 26).

TENS. No effect was seen for TENS versus placebo TENS for function or pain over the intermediate term for any outcome measured in one good-quality trial.[123] Function was measured via the WOMAC-function subscale (0 to 100); the difference in mean change scores was -1.9 (95% CI -9.7 to 5.9) and the proportion of patients who achieved MCID (≥9.1) was 38 percent versus 39 percent, RR 1.2 (95% CI 0.6 to 2.2). Pain was measured using a VAS pain scale (difference 0.9 on a scale of 0 to 10, 95% CI –11.7 to 13.4) and the WOMAC pain subscale (difference -5.6 on a 0 to 100 scale, 95% CI –14.9 to 3.6). The proportion of patients who

achieved MCID (≥20) in pain VAS was 56 percent versus 44 percent, RR 1.3 (95% CI 0.8 to 2.0). Health-related quality of life measured with the SF-36 was not different between the two groups for the physical component and mental component score (Table 26).

Low-level laser therapy. One fair-quality trial reported no difference between low-level laser therapy and sham for short-term function based on median Saudi Knee Function Scale scores (range 0-112 with higher scores indicating greater severity), median difference -10 (interquartile range of -23 to -4), p=0.054.[119] There were inconclusive results for intermediate-term function. One fair-quality trial reported the low-level laser therapy group had less functional severity at 6 months compared to sham on the Saudi Knee Function Scale (median difference -21.0, 95% CI -34.0 to -7.0), p=0.006.[119] For the other poor-quality trial, neither the higher dose nor the lower dose low-level laser therapy group differed from sham on the WOMAC physical function (0 to 96) subscale (mean differences -3.82, 95% CI -9.75 to 2.11 and -0.14, 95% CI -6.59 to 6.31, respectively).[129]

Low-level laser therapy was associated with moderately less pain over the short term in one fair-quality and one poor-quality trial (pooled difference -2.03, 95% CI -3.74 to -0.33) (Figure 34).[119,126] There was no difference between low-level laser therapy versus sham for intermediate-term pain (pooled difference -0.93, 95%CI -2.82 to 0.96).[119,129]

Microwave diathermy. Data were insufficient from one small, fair-quality trial evaluating microwave diathermy.[125] The microwave diathermy group showed substantial short-term improvement compared with sham for function (difference -33.2 on a 0-85 scale, 95% CI -42.0 to -24.6, WOMAC ADLs subscale) and pain (difference -8.1 on a 0-25 scale, 95% CI -10.7 to -5.3, WOMAC pain subscale). Substantial imprecision was noted.

Pulsed short-wave diathermy. There was no difference in short-term function or pain for either the low intensity or high intensity group compared to sham diathermy based on the WOMAC in in one poor-quality trial.[125] There was no difference on the WOMAC function subscale (0 to 10) between either the low intensity group versus sham, mean difference of 0.16 (95% CI -1.51 to 1.83), or the high intensity group versus sham, mean difference -0.02 (95% CI -1.67 to 1.63). There was also no difference on the WOMAC-pain subscale (0 to 10) for either the low or high intensity group versus sham, mean differences of 0.15 (95% CI -1.57 to 1.87) and -0.24 (95% CI -2.02 to 1.54), respectively.

The other trial found inconsistent results among the high and low dose groups for long-term function using the Knee Injury and Osteoarthritis Outcome Score (KOOS) (0 to 100).[124] The low dose group had substantially greater improvement on the KOOS-Daily Activities subscale compared to sham (mean difference 27.30, 95% CI 13.73 to 40.87), but there was no difference between the high dose group and sham on the KOOS-Daily Activities subscale (mean difference 10.30, 95% CI -1.24 to 21.84). Neither the low or high dose group differed from sham on the KOOS-recreational activities subscale (Table 26). Regarding pain intensity, the low dose group had moderately better pain NRS (0 to 10) that was not statistically significant (mean difference -1.8, 95% CI -3.60 to 0.00). The high dose group experienced substantially greater pain reduction than the sham group (mean difference -2.3, 95% CI -3.68 to -0.92).

Electromagnetic fields. The fair-quality trial found use of pulsed electromagnetic fields did not appear to provide clinically meaningful short-term improvements in function or pain compared

with sham, although statistical significance was achieved. The pulsed electromagnetic field group had better function on the WOMAC ADL subscale (0 to 85) compared with the sham group, (mean difference -3.48, 95% CI -4.44 to -2.51), and it had lower scores on the WOMAC pain subscale (0 to 25) versus sham (mean difference -0.84, 95% CI -1.10 to -0.58).[130] Based on estimated values from a graph for the poor =-quality trial,[120] each group using electromagnetic fields had better function and substantially less pain in the short term on the Lequesne Index. The musically modulated electromagnetic field group had moderately better Lequesne Function scores (0-10) versus sham (mean of 6.5 vs. 3.8) and substantially lower Lequesne Pain scores (0 to 10) (mean of 1.4 vs. 6.9). The low frequency electromagnetic field group had similar benefits for function (mean of 7.1 vs. 3.83) and pain (mean of 1.4 vs. 6.85, standard deviation and statistical testing not reported), compared with sham.

Superficial heat. Evidence from one small fair-quality trial was insufficient to determine the effects of superficial heat on short-term pain. WOMAC pain subscale scores were similar between the heat and placebo group at 1 month post treatment: 13.7 versus 13.9, respectively.[128]

Brace. There was no difference between bracing and usual care for intermediate-term or long-term function, pain, and quality of life outcomes.[121] Function was measured using the Hospital for Special Surgery (HSS) score (mean difference 3.2, 95% CI -0.58 to 6.98 for intermediate-term function and 3.0, 95% CI -1.05 to 7.05 for long-term function). Pain intensity was assessed using a visual analog scale (VAS). The mean difference was -0.58 (95% CI -1.48 to 0.32) for intermediate-term pain and -0.81 (95% CI -1.76 to 0.14) for long-term pain. Health-related quality of life was measured using the EQ-5D (mean difference 0.01, 95% CI -0.08 to 0.10 for both intermediate-term and long-term health-related quality of life).

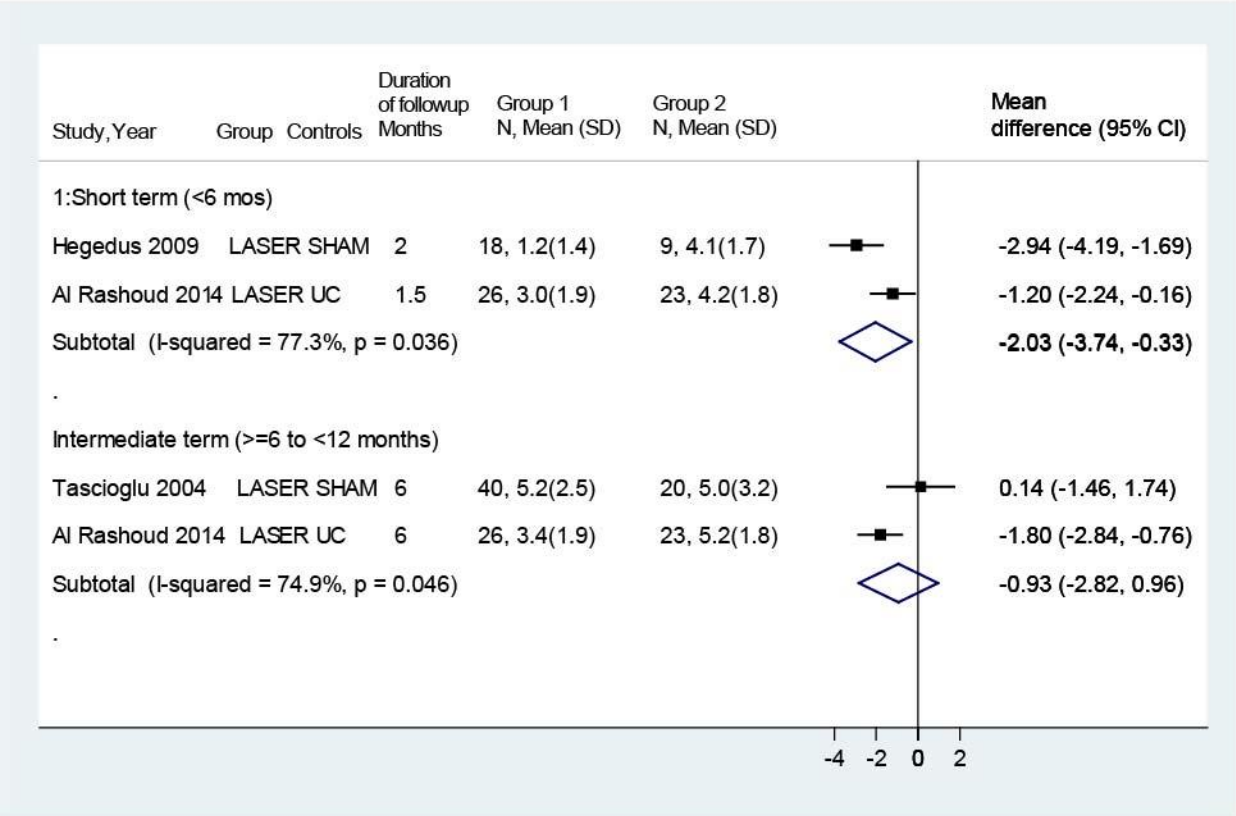
Physical Modalities Compared With Pharmacological Therapy or With Exercise Therapy

No trial of physical modalities versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

In general, harms were poorly reported across the physical modality trials. Six trials (two of low-level laser therapy,[119,129] two of ultrasound therapy,[122,131] one of pulsed short-wave diathermy,[127] and one of superficial heat[128]) reported that no adverse events or side effects occurred in either group. The good-quality trial that evaluated TENS found no difference between active and sham TENS in the risk of localized, mild rashes (18% vs.17%; RR 1.06, 95% CI 0.38 to 2.97).[123] One trial of microwave diathermy reported two cases of symptom aggravation in the intervention group; the events were transient and neither patient withdrew from the trial.[125] More patients who received real versus sham electromagnetic field therapy reported throbbing or warming sensations or aggravation of pain (29% versus 7%); however, the difference was not significant (RR 1.95, 95% CI 0.81 to 4.71) in one fair quality trial.[130]

Figure 34. Low level laser therapy versus usual care or sham for osteoarthritis of the knee: effects on pain



CI = confidence interval; SD = standard deviation; UC = usual care

Manual Therapies for Osteoarthritis of the Knee

Key Points

- There is insufficient evidence from one trial to determine the effects of joint manipulation on intermediate-term function or harms versus usual care or versus exercise due to inadequate data to determine effect sizes or statistical significance (SOE: Insufficient).
- There is insufficient evidence from one trial to determine the effects of massage versus usual care on short-term function, pain or harms, or to evaluate the effect of varying dosages of massage on outcomes (SOE: Insufficient).

Detailed Synthesis

Two trials were identified that met inclusion criteria and evaluated manual therapies for the treatment of knee OA[36,146] (Table 27 and Appendixes D and E); both trials required patients to have radiographically established knee OA meeting the American College of Rheumatology criteria.

Table 27. Summary of results for osteoarthritis of the knee: manual therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Abbott, 2013[36] 9.75 months Duration of diagnosis: 2.6 years *Fair*	<u>A. Manual therapy</u> (n=54/30 knee OA): 7 manual therapy sessions in 9 weeks with 2 additional booster sessions <u>B. Exercise</u> (n=51/29 knee OA), 7 exercise sessions in 9 weeks with 2 additional booster sessions <u>C. Usual care</u> (n=51/28 knee OA)	A vs. B vs. C (total population, includes hip OA) Age: 67 vs. 67 vs. 66 years Female: 49% vs. 52% vs. 58% Percent knee OA: 56% vs. 57% vs. 55% Percent hip OA: 44% vs. 43% vs. 45% Percent both hip OA and knee OA: 22% vs. 20% vs. 26% Baseline WOMAC (0-240): 114.8 vs. 95.5 vs. 93.8	A vs. C (knee OA only) <u>9.75 months</u> WOMAC mean change from baseline: -31.5 vs. 1.6, p NR A vs. B <u>9.75 months</u> WOMAC mean change from baseline: -31.5 vs. -12.7, p NR	NR
Perlman,[146] 2012 4 months Duration of pain: NR *Fair*	<u>A1. Massage Therapy Group 1 (MT) (n=25)</u> Participants received a uniform massage protocol designed to address symptoms of osteoarthritis of the knee with a series of standard Swedish massage strokes, and specified time allocated to various body regions (therapists agreed not to deviate from protocol); one, 30-minute session per week for 8 weeks (8 total sessions) <u>A2. MT Group 2 (n=25)</u> Identical to group A1 except differing 'dosage' of massage; two, 30-min sessions per week for 4 weeks, then once weekly for four weeks (12 total sessions) <u>A3. MT Group 3 (n=25)</u> Identical to group A1 except differing 'dosage' of massage; one, 60-min per week	A1 vs. A2 vs. A3 vs. A4 vs. B Age: 70 vs. 62 vs. 63 vs. 64 vs. 64 Female: 60% vs. 72% vs. 76% vs. 68% vs. 76% Race: 92% vs. 88% vs. 76% vs. 80% vs. 88% white WOMAC Total (0-100): 52.9 vs. 50.2 vs. 53.6 vs. 48.0 vs. 53.2 WOMAC Physical Function (0-100): 52.9 vs. 49.5 vs. 49.8 vs. 48.3 vs. 50.5 WOMAC Pain (0-100): 52.3 vs. 42.4 vs. 52.5 vs. 44.4 vs. 46.3 VAS Pain (0-100): 61.2 vs. 64.0 vs. 66.4 vs. 59.2 vs. 57.6	A1 vs. A2 vs. A3 vs. A4 vs. B <u>4 months:</u> WOMAC Total, mean change from baseline (95% CI): -14.3 (-22.9 to -5.7) vs. -7.0 (-15.6 to 1.6) vs. -14.2 (-23.4 to -5.0) vs. -15.1 (-25.1 to -5.1) vs. -6.0 (-12.6 to 0.5) WOMAC Physical Function, mean change from baseline (95% CI): -15.3 (-24.5 to 26.1) vs. -7.4 (-14.8 to 0) vs. -12.1 (-22.0 to -2.1) vs. -14.4 (-23.4 to -5.4) vs. -4.2 (-11.1 to 2.7) WOMAC Pain, mean change from baseline (95% CI): -12.2 (-22.4 to -2.0) vs. -3.9 (-12.7 to 4.9) vs. -13.7 (-23.4 to -4.0) vs. -14.2 (-24.5 to -3.8) vs. -7.5 (-16.0 to 1.1) VAS Pain, mean change from baseline (95% CI): -14.4 (-25.9, -2.8) vs. -14.0 (-24.7 to -3.3) vs. -18.5 (-29.0 to -8.1) vs. -22.8 (-35.5 to -10.1) vs. -11.5 (-21.0 to -2.0)	A1 vs. A2 vs. A3 vs. A4 vs. B <u>4 months:</u> WOMAC Stiffness, mean change from baseline (95% CI): -15.4 (-26.4 to -4.5) vs. -9.6 (-20.6 to 1.3) vs. -16.9 (-28.5 to -5.2) vs. -16.8 (-29.7 to -3.9) vs. -6.4 (-13.2 to 0.4)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	for 8 weeks (8 total sessions) <u>A4. MT Group 4 (n=25)</u> Identical to group A1 except differing 'dosage' of massage; two, 60-min sessions per week for 4 weeks, then once weekly for four weeks (12 total sessions) <u>B. Usual Care (n=25)</u> Participants continued with their current treatment without the addition of massage therapy.			

CI = confidence interval; NR = not reported; OA = osteoarthritis; VAS = Visual Analog Scale; WOMAC = Western Ontario and McMaster Universities Arthritis Index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

One fair-quality trial (N=58 with knee OA) compared manual therapy with usual care (continued routine care from general practitioner and other providers) and with combination exercise.[36] The manual therapy intervention consisted of nine 50-minute sessions, seven delivered in the first 9 weeks, and two booster sessions at week 16. All participants were prescribed a home exercise program three times per week. Compliance with the intervention was acceptable in all groups, and the methodological shortcomings of this trial were a lack of blinding for the patients and care providers. Only intermediate-term outcomes were reported.

One fair-quality trial (N=125) compared four different dosages of massage therapy with usual care (continued current treatment).[146] The massage protocol consisted of standard Swedish massage strokes applied in each intervention group over 8 weeks. The dosage varied from 240 to 720 minutes based on the frequency (once or twice per week) and duration of massage (30-60 minutes per session). Compliance was acceptable in all groups, and the methodological shortcomings of this trial were a lack of blinding for the patients and care providers in the usual care arm. Only short-term outcomes were reported.

Manual Therapies Compared With Usual Care

Manual therapy. Data were insufficient from one fair-quality trial (n=58 with knee OA)[36] to evaluate effects of joint manipulation versus usual care over the intermediate-term. Although the manual therapy group showed a statistically significant improvement from baseline in function as measured by the WOMAC (mean change -31.5 on a 0-240 scale, 95% CI, -52.7 to -10.3), whereas the usual care group showed no improvement (mean change 1.6, 95% CI, -10.5 to 13.7), insufficient data was provided to calculate an effect estimate (number of patients with knee OA in each group were not provided). Pain outcomes were not reported.

Massage. Data were insufficient from one fair-quality trial (n=125) to evaluate the short-term effects of massage therapy (4 different dosages) compared with usual care.[146] Function was measured using the WOMAC total and physical function subscale scores (both 0 to 100 scales) and pain was measured using the WOMAC pain subscale and the VAS (both 0 to 10). No significant effects were seen in any outcome measure at 4 months post-massage treatment versus usual care (Table 27). Authors report a trend for greater magnitude of change in function and pain with higher massage dosages versus lower massage dosages and versus usual care (statistical tests not provided).

Manual Therapies Compared With Pharmacological Therapy

No trial of manual therapy versus pharmacological therapy met inclusion criteria.

Manual Therapies Compared With Exercise Therapy

The trial evaluating manual therapy also included an exercise group that received aerobic warm-up, muscle strengthening, muscle stretching, and neuromuscular control exercises.[36] Both groups showed improvement from baseline in function (WOMAC) over the intermediate term, but the change was statistically significant in the manual therapy group only: mean change of -31.5 (95% CI, -52.7 to -10.3) versus -12.7 (95% CI, -27.1 to 1.7) for exercise. However, insufficient data was provided to calculate an effect estimate (number of patients with knee OA in each group were not provided). Pain outcomes were not reported.

Harms

No serious treatment-related adverse events occurred in either trial;[36,146] one non-trial related death was reported in the usual care group in the trial evaluating manual therapy.[36]

Mind-Body Therapies for Osteoarthritis of the Knee

Key Points

- Data were insufficient from two small, unblinded trials to determine the effects or harms of Tai chi versus attention control in the short or intermediate terms. No data on long-term outcomes were available (SOE: insufficient).

Detailed Synthesis

Two small trials (n=41 and 40) of Tai chi versus attention control in older adults met the inclusion criteria[169,170] (Table 28 and Appendix D). Tai chi was practiced 40 to 60 minutes two or three times per week for 24 or 36 sessions. Attention control consisted of group education classes with one trial[170] including 20 minutes of stretching for sessions 18 to 24. Blinding was not possible in either trial and was the primary methodological limitation in one fair-quality

trial.[170] Additional methodological concerns in the other poor-quality trial included unclear concealment of treatment allocation and high attrition[169] (Appendix E).

Table 28. Summary of results for osteoarthritis of the knee: mind-body therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Brismee, 2007[169] 1.5 months Duration of pain: NR *Poor*	A. Tai chi (n=18) Subjects in the tai chi group attended group tai chi classes for six weeks followed by six weeks of home video tai chi practice. No. of Treatments: 3/week for 12 weeks (36 total) Length of Treatments: 40 min/session B. Attention Control (n=13) Subjects in the attention control group attended group lectures and discussions covering health-related topics. They did not take part in any further activity past 6 week group period. No. of Treatments: 3/week for 6 weeks (18 total) Length of Treatments: 40 min/session	A vs. B Age: 71 vs. 69 Female: 86.4% vs. 78.9% Race: NR WOMAC Total (26-13)]: 64.6 vs. 59.6 WOMAC Physical Function (17-85): 42.7 vs. 37.6 WOMAC Pain (7-35): 16.5 vs. 16.9 VAS Pain (0-10): 4.7 vs. 4.2 WOMAC Stiffness (2-10): 5.6 vs. 5.1	A vs. B 1.5 months WOMAC Total: 60.28 vs. 57.73m p=ns WOMAC Physical Function: 38.61 vs. 37.58, p=ns WOMAC Pain: 16.39 vs. 16, p=ns VAS Pain: 3.46 vs. 3.19, p=ns WOMAC Stiffness: 5.28 vs. 4.54, p=ns	NR
Wang, 2009[170] 3 and 9 months Duration of pain: 9.7 years *Fair*	A. Tai chi (n=20) Subjects in the tai chi group attended group tai chi classes where they learned 10 forms from the classic Yang style tai chi. They were also instructed to practice tai chi at least 20 minutes per day at home with a tai chi DVD. Home practice continued after group sessions ended until the 48 week followup. B. Attention Control (n=20) Subjects in the attention control group attended group classes where they received nutritional and medical information paired with 20 minutes	A vs. B Age: 63 vs. 68 Female: 80% vs. 70% Race: NR WOMAC Physical Function (0-1,700): 707.6 vs. 827 WOMAC Pain (0-500): 209.3 vs. 220.4 VAS Patient-Assessed Pain (0-10): 4.2 vs. 4.8 VAS Physician-Assessed Pain (0-10): 4.8 vs. 5.8 WOMAC Stiffness (0-200):	A vs. B 3 months (mean change from baseline) WOMAC Physical Function: -440.5 (95% CI -574.4 to -306.6) vs. -257.3 (95% CI -391.2 to -123.4); difference -183.2 (95%CI -372.6 to 6.2) WOMAC Pain: -131.6 (95% CI -177.4 to -85.7) vs. -64.6 (95% CI -110.5 to -18.7); difference -70.0 (95% CI -131.8 to -2.1) VAS Patient Assessed Pain: -2.4 (95% CI -3.5 to -1.2) vs. -1.7 (-2.9 to -0.5); MD -0.7 (-2.3 to 1.0) VAS Physician Assessed Pain: -2.6	A vs. B 3 months (mean change from baseline) SF-36 PCS (0-100): 10.8 (95% CI 7.3 to 14.3) vs. 6.3 (95%CI 2.8 to 9.8); difference 4.5 (95%CI -0.4 to 9.5) SF-36 MCS (0-100): 4.4 (95% CI -0.11 to 8.9) vs. 4.5 (95% CI 0.0 to 9.0); difference -0.1 (95% CI -6.5 to 6.3) CES-D (0-60): -6.4 (95%CI -9.9 to -2.9) vs. -1.1 (95%CI -4.6 to 2.4); difference -5.3 (95% CI -10.2 to -0.4) 9 months SF-36 PCS: 10.4

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	of stretching. Instruction to practice at least 20 minutes of stretching exercises per day at home. In both groups, treatments were 2x/week for 12 weeks (24 total), 60 minute sessions	105.7 vs. 120.7	(95% CI -3.3 to -1.9) vs. -2.1 (95% CI -2.8 to -1.3); difference -0.5 (95% CI -1.6 to 0.5) WOMAC Stiffness: -65.0 (95% CI -86.3 to -43.7) vs. -50.2 (95% CI -71.5 to -28.9); difference -14.8 (95% CI -44.9 to 15.3) <u>9 months</u> WOMAC Physical Function: -405.9 (95% CI -539.8 to -271.9) vs. -300.6 (95% CI -434.5 to -166.6); difference -105.3 (95% CI -294.7 to -84.1) WOMAC Pain: -115.4 (95% CI -161.2 to -69.5) vs. -69.2 (95% CI -115.1 to -23.3); difference -46.2 (95% CI -111.0 to 18.7) VAS Patient Assessed Pain: -1.7 (95% CI -2.8 to -0.5) vs. -1.7 (95% CI -2.9 to -0.5); difference 0.04 (95% CI -1.6 to 1.7) VAS Physician-Assessed Pain: -2.5 (95% CI -3.3 to -1.8) vs. -1.5 (-2.3 to -0.8); difference -1.0 (95% CI -2.1 to 0.02) WOMAC Stiffness: -64.2 (95% CI -85.5 to -42.8) vs. -60.5 (95% CI -81.8 to -39.2); MD -3.7 (95% CI -33.8 to 26.5)	(95%CI 6.9 to 13.9) vs. 4.1 (95% CI 0.6 to 7.6); difference 6.3 (95% CI 1.4 to 11.3) SF-36 MCS: 5.8 (95% CI 1.3 to 10.3) vs. 1.0 (95% CI -3.5 to 5.5); difference 4.8 (95% CI -1.6, 11.1) CES-D: -7.3 (95% CI -10.7 to -3.8) vs. 1.7 (95% CI -1.8 to 5.1); difference -8.9 (95% CI -13.8 to -4.0)

AIMS = Arthritis Impact Measurement Scale; BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; CI = confidence interval; HRQoL = health-related quality of life; NR = not reported; NRS = Numerical Rating Scale; NS = not statistically significant; OA = osteoarthritis; WOMAC, Western Ontario and McMaster Universities Osteoarthritis index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Mind-Body Therapies Compared With Attention Control

There is no clear difference between tai chi and an attention control on functional outcomes across the two trials over the short term on a 0 to 85 point scale WOMAC physical function (difference 1.03, 95% CI -9.87 to 11.93)[169] or WOMAC physical function 0 to 1700 point scale (difference -183.2, 95% CI -372.6 to 6.2)[170] or at intermediate term in one of the trials

(difference -105.3, 95% CI -294.7 to -84.1, 0 to 1700 scale).[170] Results for short-term pain improvement were inconsistent with no difference between groups on WOMAC pain scale (difference 0.39 on a 0-35 point scale, 95% CI -4.21 to 4.99)[169] and the other marginally favoring tai chi on 0 to 500 point WOMAC pain scale (difference -67.0, 95% CI -131.8 to -2.1)[170] but demonstrating no difference between the groups in 0 to 10 VAS pain (difference -0.65, 95% CI -2.31 to 1.02).[170] There were no differences between groups at intermediate term in this latter trial (WOMAC pain 0 to 500 scale, difference -183.2, 95% CI -372.6 to 6.2).[170] One trial noted improvement in health-related quality of life (SF-36) in the intermediate term only and depression (CES-D) and self-efficacy in the short and intermediate terms.

Mind-Body Therapies Compared With Pharmacological Therapy or With Exercise Therapy

No trial of mind-body therapy versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

In the two trials of mind-body interventions, harms were poorly reported. One trial reported no serious adverse events[170] and the other reported sporadic complaints of muscle soreness and foot or knee pain.[169]

Acupuncture for Osteoarthritis of the Knee

Key Points

- There were no clear differences between acupuncture versus control interventions (sham condition, waitlist, or usual care) on function in the short term (5 trials, pooled SMD -0.18, 95% CI -0.55 to 0.20 I^2 =82%) or the intermediate term (3 trials, pooled SMD -0.12, 95% CI -0/30 to 0.07, I^2 = 41%) (SOE: Low for short-term; Moderate for intermediate-term).
- There were no clear differences between acupuncture versus control interventions on pain in the short term (pooled SMD -0.27, 95%CI -0.56 to 0.02, I^2 = 75%) or in the intermediate term (pooled SMD -0.11, 95% CI -0.30 to 0.07, I^2 = 0%) (SOE: Low for short term; Moderate for intermediate term).
- Data from one poor-quality trial were insufficient to determine the effects of acupuncture versus exercise (SOE: Insufficient).
- There is no apparent difference in risk of serious adverse events between any form of acupuncture and the control group. Worsening of symptoms (7% to 14%) and mild bruising, swelling, or pain at the acupuncture site (1% to 18%) were most common; one case of infection at an electroacupuncture site was reported (SOE: Moderate).

Detailed Synthesis

Eight trials of acupuncture for knee OA were identified that met inclusion criteria[56,190-196] (Table 29 and Appendix D). Four trials evaluated traditional acupuncture,[56,191,193,195] three electroacupuncture,[190,192,194] and two laser acupuncture.[191,196] Three studies compared acupuncture with usual care (provision of educational leaflets, instructions to remain on current oral medications, or no changes to their ongoing treatments)[56,190,193] and one study each to no treatment[191] or to waitlist control.[194] Five studies compared acupuncture with sham procedures, which consisted of inactive laser treatment (red light on but no power applied),[191,196] superficial

needling or acupuncture performed at non meridian sites,[194,195] or non-penetrating sham acupuncture.[192] No trials of acupuncture versus pharmacologic therapy or exercise were identified. Sample sizes ranged from 30 to 455 (total sample 1,364). Duration of acupuncture treatment ranged from 2 to 12 weeks, with the number of sessions ranging from 6 to 16. Four studies were conducted in Europe,[56,192,193,195] two in the United States,[190,194] and one study each was conducted in Australia[191] and Turkey.[196] Short-term outcomes were reported by six trials[56,190,192,194-196] and intermediate-term outcomes by three;[191,193,195] no trial reported outcomes over the long term.

Two studies were rated good quality (for the comparison of acupuncture versus sham only).[191,194] Six studies were rated fair quality (to include the comparison of acupuncture with no treatment/waitlist in the two trials described previously)[190-192,194-196] and two were considered poor quality[56,193] (Appendix E). The primary methodological shortcoming in the fair-quality trials was lack of blinding; additionally, the poor-quality trials suffered from unclear allocation concealment methods and high rates of attrition (30% to 35%).

Table 29. Summary of results for osteoarthritis of the knee: acupuncture

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Berman, 1999[190] 1 month Duration of pain: mean 7.2 years *Fair*	A. Acupuncture + usual care (n=36): 20 minute treatments, 2/week for 8 weeks using Traditional Chinese Medicine theory; 9 acupoints points (5 local, 4 distal) with elicitation of de qi; electrical stimulation was used at local points (2.5 to 4 Hz, pulses of 1.0 ms); patients asked not to begin any new physiotherapy or exercise programs B. Usual care alone (n=37): asked to remain on their current level of oral therapy throughout the trial	A vs. B Age: 66 vs. 66 Female: 47% vs. 72% Caucasian: 92% vs. 74% BMI: 32 vs. 32 Duration of symptoms: 7.5 vs. 6.9 years WOMAC total (scale unclear): 48.4 vs. 51.4 WOMAC function (scale unclear): 34.3 vs. 34.4 Lequesne Index (0-24): 11.7 vs. 12.3 WOMAC pain (scale unclear): 9.6 vs. 9.9	A vs. B 1 month WOMAC total: 31.6 vs. 50.4, difference -18.9 (95% CI -26.5 to -11.2) WOMAC function: 23.2 vs. 36.8, difference -13.6 (95% CI -19.4 to -7.8) Lequesne Index: 9.3 vs. 12.4, difference -3.1 (95% CI -4.8 to -1.3) WOMAC pain: 5.6 vs. 9.5, difference -4.0 (95% CI -5.5 to -2.4)	NR
Hinman, 2014[191] 9 months Duration of pain: mean 7.2 years *Good (sham)*	A. Needle acupuncture (n=70): combination of Western and traditional Chinese acupuncture; maximum of 6 points (4 on study limb and 2 distal points) at initial session, in other sessions points were	A vs. B vs. C vs. D Age: 64 vs. 63 vs. 63 vs. 64 years Female: 46% vs. 39% vs. 56% vs. 56% Duration of symptoms ≥ 10 years: 41% vs. 38%	A vs. C 9 months WOMAC function: 22.4 vs. 23.6; adjusted difference -3.7 (95% CI -8.2 to 0.8) Activity restriction, NRS: 3.4 vs. 4.1; adjusted difference -1.1 (95% CI -2.1, -0.2) WOMAC pain: 6.7 vs. 7.4;	A vs. C 9 months AQoL-6D (-0.04 to 1.00): 0.74 vs. 0.77; adjusted difference: -0.01 (95% CI -0.07 to 0.05) SF-12 PCS (0-

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fair (no treatment)	added at therapist's discretion. Needles were left in while patient rested. B. Laser acupuncture (n=71): combination of Western and traditional Chinese acupuncture; delivered to selected points using standard Class 3B laser devices (measured output 10mW and energy output 0.2 J/point) C. No treatment (n=71): did not receive acupuncture; continued in an observational study, unware they were in an acupuncture trial D. Sham laser acupuncture (n=70): same as true laser but no laser was emitted, only red nonlaser light at the probe tip lit up. For all acupuncture and sham groups, sessions were 20 minutes in duration, 1-2 times per week for 12 weeks (8 to 12 sessions total)	vs. 27% vs. 50% Bilateral symptoms: 64% vs. 66% vs. 51% vs. 63% Opioid use: 1% vs. 3% vs. 1% vs. 1% Previous acupuncture for knee pain: 7% vs. 13% vs. 7% vs. 3% WOMAC function (0-68): 31.3 vs. 27.0 vs. 26.1 vs. 27.5 NRS activity restriction (0-10): 5.0 vs. 4.3 vs. 4.1 vs. 4.5 WOMAC pain (0-20): 9.0 vs. 8.3 vs. 7.8 vs. 8.6 NRS average pain overall (0-10): 5.3 vs. 4.9 vs. 5.1 vs. 5.0 NRS pain on walking (0-10): 5.5 vs. 4.8 vs. 4.8 vs. 5.2 NRS pain on standing (0-10): 4.6 vs. 3.8 vs. 4.1 vs. 4.3	adjusted difference -1.4 (95% CI -2.7 to 0.0) Overall Pain, NRS: 4.0 vs. 4.6; adjusted difference -0.7 (95% CI -1.6 to 0.2) Pain on walking, NRS: 4.1 vs. 4.4; adjusted difference -0.6 (95% CI -1.5 to 0.4) Pain on standing, NRS: 3.7 vs. 4.0; adjusted difference -0.5 (95% CI -1.4 to 0.5) B vs. C <u>9 months</u> WOMAC function: 22.6 vs. 23.6; adjusted difference -0.6 (95% CI -1.5 to 0.3) Activity restriction, NRS: 3.7 vs. 4.1; adjusted difference -0.4 (95% CI -1.4, 0.5) WOMAC pain: 7.1 vs. 7.4; adjusted difference -0.4 (95% CI -1.8 to 1.0) Overall Pain, NRS: 4.0 vs. 4.6; adjusted difference -0.6 (95% CI -1.5 to 0.3) Pain on walking, NRS: 4.1 vs. 4.4; adjusted difference -0.3 (95% CI -1.2 to 0.7) Pain on standing, NRS: 3.8 vs. 4.0; adjusted difference -0.2 (95% CI -1.1 to 0.8) B vs. D <u>9 months</u> WOMAC function: 22.6 vs. 21.6; adjusted difference 1.1 (95% CI -4.8 to 7.0) Activity restriction, NRS: 3.7 vs. 3.9; adjusted difference -0.1 (95% CI -1.1 to 1.0) WOMAC pain: 7.1 vs. 6.9; adjusted difference 0.0 (95% CI -1.9 to 1.9) Overall pain, NRS: 4.0 vs. 3.9; adjusted difference 0.0 (95% CI -0.9 to 1.0) Pain on walking, NRS: 4.1 vs. 4.2; adjusted difference 0.0 (95% CI -1.0 to 1.1) Pain on standing, NRS: 3.8 vs. 3.5; adjusted difference 0.5 (95% CI -0.7 to 1.6)	100): 41.7 vs. 38.9; adjusted difference 2.3 (95% CI -1.7 to 6.3) SF-12 MCS (0-100): 51.1 vs. 54.4; adjusted difference -0.9 (95% CI -5.2 to 3.4) Opioid use: 0% (0/70) vs. 1% (1/71) B vs. C <u>9 months</u> AQoL-6D: 0.73 vs. 0.77; adjusted difference: 0.01 (95% CI -0.05 to 0.06) SF-12 PCS: 38.8 vs. 38.9; adjusted difference -0.4 (95% CI -4.4 to 3.6) SF-12 MCS: 52.1 vs. 54.4; adjusted difference -0.9 (95% CI -5.5 to 3.7) Opioid use: 2% (1/71) vs. 1% (1/71) B vs. D <u>9 months</u> AQoL-6D: 0.73 vs. 0.74; adjusted difference 0.01 (95% CI -0.05 to 0.08) SF-12 PCS: 38.8 vs. 38.2; adjusted difference 0.4 (95% CI -3.8 to 4.5) SF-12 MCS: 52.1 vs. 52.8; adjusted difference -0.6 (95% CI -5.4 to 4.2) Opioid use: 2% (1/71) vs. 0% (0/70)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Jubb, 2008[192] 1 month Duration of pain: mean 10 years *Fair*	A. Acupuncture (n=34): manual acupuncture (10 minutes, total of 9 points; depth of 1-1.5 cm; elicitation of de qi) and electro-acupuncture (10 minutes each on anterior and posterior part of the knee (20 minutes total); low frequency, delivered at 6 Hz at a constant current) B. Sham (n=34): sham needles, did not penetrate the skin; electrical stimulation apparatus produced sound signals but no electrical current. Both groups received 30 minute treatments, 2/week for 5 weeks, with 10 sessions in total	A vs. B Age: 64 vs. 66 years Female: 85% vs. 76% Caucasian: 74% vs. 85% Duration of symptoms: 10 vs. 9.6 years WOMAC function (0-1700): 1028 vs. 979 WOMAC pain (0-500): 294 vs. 261 Total body pain, VAS (0-100): 49 vs. 49 Night pain knee, VAS (0-100): 61 vs. 52 Overall pain knee, VAS (0-100): 63 vs. 53 Weight-bearing pain knee, VAS (0-100): 71 vs. 60 EuroQoL VAS (0-100): 63 vs. 54	A vs. B 1 month WOMAC function: change from baseline, 137 (95% CI 20 to 255) vs. 134 (95% CI 9 to 258); difference, 4 (95% CI -163 to 171) WOMAC pain: change from baseline, 59 (95% CI 16 to 102) vs. 13 (95% CI -22 to 50); difference, 46 (95% CI -9 to 100) Weight-bearing knee pain (VAS), change from baseline, 19 (95% CI 9 to 30) vs. 8 (95% CI -1 to 16); difference, 11 (95% CI -2 to 25) Overall knee pain (VAS), change from baseline, 14 (95% CI 5 to 24) vs. 2 (95% CI -6 to 10); difference, 12 (95% CI -1 to 24) Nighttime knee pain (VAS), change from baseline, 10 (95% CI -1 to 22) vs. 5 (95% CI -3 to 14); difference, 5 (95% CI -9 to 19) General body pain (VAS), change from baseline, 5 (95% CI -5 to 15) vs. -8 (95% CI -1 to 18); difference: 13 (95% CI 0 to 27) EuroQoL-VAS: mean 63 vs. 52, p=0.98	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Lansdown, 2009[193] 9.5 months Duration of pain NR *Poor*	A. Acupuncture + usual care (n=15): once per week for up to 10 weeks, with maximum of 10 sessions, which varied in length and content (mean number of acupoints was 12, range 4-24; de qi was usually elicited; variety of stimulation methods used including tonification and reduction; retention time for needles ranged from 10-30 minutes); auxiliary treatment included moxibustion (3/14, 21%) and acupressure massage (3/14, 21%); life style advice 11/14 (79%) B. Usual care (n=15): any appointments, medications prescribed or over the counter) and interventions sought by participants from any health practitioner	A vs. B Age: 63 vs. 64 years Female: 60% vs. 60% Caucasian: 100% vs. 100% Duration of symptoms: NR WOMAC total (0-96): 31 vs. 37.5 WOMAC function (0-68): 20.5 vs. 26.3 OKS (12-60): 30.9 vs. 30.6 WOMAC pain (0-20): 7.3 vs. 7.4	A vs. B 9.5 months WOMAC total: 24.8 vs. 25.6 (17.6), adjusted difference -2.9 (95% CI 9.5 to -15.4) WOMAC function: 17.4 vs. 17.6, adjusted difference -1.36 (95% CI 8.7, -11.4) WOMAC pain: 4.7 vs. 5.3 (3.9), adjusted difference -1.4 (95% CI 0.8 to -3.6) OKS: 24.5 vs. 28.1; difference -3.6 (95% CI -9.8 to 2.6)	A vs. B 9.5 months (SF-36 scales are 0-100 for all) SF-36 physical functioning: 54.2 vs. 55.6, difference -1.4 (95% CI -21.8 to 19.0) SF-36 social functioning: 81.3 vs. 76.6, difference 4.7 (95% CI -10.6 to 20.0) SF-36 role physical: 71.4 vs. 57.8, difference 13.6 (95% CI -6.3 to 33.5) SF-36 role mental: 79.2 vs. 67.7, difference 11.5 (95% CI -5.8 to 28.8) SF-36 mental health: 73.1 vs. 65.0, difference 8.1 (95% CI -5.4 to 21.6) SF-36 vitality: 58.2 vs. 46.9, difference 11.3 (95% CI -0.22 to 22.8) SF-36 pain: 65.2 vs. 65.9, difference -0.7 (95% CI -15.6 to 14.2) SF-36 general health: 67.7 vs. 62.4, difference 5.3 (95% CI -4.8 to 15.4), EQ5D: 0.66 vs. 0.63, difference 0.03 (95% CI -0.13 to 0.19)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Suarez-Almazo, 2010[194] 1.5 months Duration of pain: mean 8 years *Good (sham)* *Fair (waitlist)*	A. Electro-acupuncture (n=153): Traditional Chinese Medicine points; TENS equipment emitted a dense disperse wave (50Hz, dispersed at 15 Hz, 20 cycles/minute); voltage increased from 5V to 60V until maximal tolerance achieved. Patients rested for 20 minutes with needles retaining and with continuing TENS. B. Sham (n= 302) 40Hz adjustable wave; voltage increased until the patient could feel it and then immediately turned off. Patients rested for 20 minutes with the needles retained, but without TENS stimulation; non-relevant acupoints used and depth of needle placement was shallow C. Waitlist (n=72)	A vs. B vs. C Age: 65 vs. 65 vs. 64 Female: 66% vs. 65% vs. 58% Caucasian: 70% vs. 68% vs. 65% Mean duration of chronicity: 9.2 vs. 8.6 vs. 11.5 years WOMAC function (0-100): 42.9 vs. 44.6 vs. 40.1 WOMAC pain (0-100): 44.5 vs. 45.0 vs. 44.1 VAS pain (0-100): 58.3 vs. 57.4 vs. 54.6 J-MAP (1-7): 4.4 vs. 4.4 vs. 4.3	A vs. B 1.5 months WOMAC function: 31.2 (vs. 32.1; difference -0.9 (95% CI -4.4 to 2.6) WOMAC pain: 30.8 vs. 31.0; difference -0.2 (95% CI -3.8 to 3.4) VAS pain: 36.2 vs. 36.7; difference -0.5 (95% CI -6.1 to 5.1) J-MAP: 3.3 vs. 3.4; difference -0.1 (95% CI -0.39 to 0.19) A vs. C 1.5 months WOMAC function: 31.2 vs. 41.7; difference -10.5 (95% CI -15.6 to -5.5) WOMAC pain: 30.8 vs. 42.4; difference -11.6 (95% CI -16.5 to -6.7) VAS pain: 36.2 vs. 53.2; difference -17.0 (95% CI -24.7 to -9.3) J-MAP: 3.3 vs. 4.2; difference -0.9 (95% CI -1.3 to -0.5)	A vs. B 1.5 months SF-12 PCS (0-100): 39.5 vs. 38.7; difference 0.8 (95% CI -1.1 to 2.7) SF-12 MCS (0-100): 54.1 vs. 53.2; difference 0.9 (95% CI -0.8 to 2.6) A vs. C 1.5 months SF-12 PCS: 39.5 vs. 35.8; difference 3.7 (95% CI 1.0 to 6.4) SF-12 MCS: 54.1 vs. 51.6; difference 2.5 (95% CI 0.04, 5.0)
Williamson, 2007[56] 1.5 months Duration of symptoms: NR *Poor*	A. Acupuncture (n=60): conducted by a physiotherapist in a group setting (6-10 patients); needles inserted into 7 acupoints until de qi was achieved and left in place for 20 minutes; treatments were once per week for 6 weeks, with 6 sessions in total B. Combination Exercise (Physiotherapy) (n=60): supervised group (6-10 people) exercise comprised of	A vs. B vs. C Age: 72 vs. 70 vs. 70 years Female: 55% vs. 52% vs. 54% BMI: 30.9 vs. 32.8 vs. 32.7 WOMAC total (scale unclear): 50.9 vs. 50.2 vs. 51.1 OKS (12-60): 40.2 vs. 39.3 vs. 40.5 Pain VAS (0-10): 7.3 vs. 6.8 vs. 6.9 HAD Anxiety (0-21): 7.3 vs. 7.5 vs. 6.7	A vs. B 1.5 months WOMAC: 48.4 vs. 49.4, difference -1.0 (95% CI -6.7 to 4.7) OKS: 38.1 vs. 38.8, difference -0.7 (95% CI -3.5 to 2.1) Pain VAS: 6.6 vs. 6.4, difference 0.22 (95% CI -0.67 to 1.11) A vs. C 1.5 months WOMAC: 48.4 vs. 52.3, difference -3.9 (95% CI -9.5 to 1.6) OKS: 38.1 vs. 40.8,	A vs. B 1.5 months HAD Anxiety: 6.9 vs. 7.1, difference -0.20 (95% CI -1.89 to 1.49) HAD Depression: 6.7 vs. 6.8, difference -0.03 (95% CI -1.30 to 1.24) A vs. C 1.5 months HAD Anxiety: 6.9 vs. 6.5, difference 0.34 (95% CI -1.11 to 1.8) HAD Depression:

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	strengthening, aerobic, stretching, and balance training; 60 minutes, once per week for 6 weeks; C. Usual care (n=61): exercise and advice leaflet; told they were enrolled in the "home exercise group"	HAD Depression (0-21): 7.1 vs. 7.1 vs. 7.4	difference -2.6 (95% CI -5.4 to 0.1) Pain VAS: 6.6 vs. 7.2, difference -0.66 (95% CI -1.45 to 0.12)	6.7 vs. 7.1, difference, -0.41 (95% CI -1.63 to 0.8)
Witt, 2005[195] 4 and 10 months Duration of pain: mean 9.4 years *Fair*	A. Acupuncture (n=150): semi-standardized; patients received at least 6 local and at least 2 distant Traditional Acupuncture points; elicitation of de qi; needles stimulated manually at least once during each session B. Minimal acupuncture (n=76): superficial insertion of at non-acupuncture sites away from knee; manual stimulation of the needles and provocation of de qi were avoided Both groups underwent 12 sessions of 30 minutes duration, administered over 8 weeks	A vs. B Age: 65 vs. 63 years Female: 70% vs. 65% Duration of symptoms: 9.1 vs. 9.9 years Bilateral OA: 74% vs. 77% Previous acupuncture: 9% vs. 7% WOMAC total (scale unclear): 50.8 vs. 52.5 PDI (Disability) (0-70): 27.9 vs. 27.8 VAS pain (0-100): 64.9 vs. 68.5	A vs. B 4 months WOMAC total: 30.4 vs. 36.3; difference -5.8 (95% CI -12.0 to 0.3) WOMAC physical function: 30.4 vs. 36.5; difference -6.2 (95% CI -12.4 to 0.1) PDI: 18.6 vs. 22.8; difference -4.2 (95% CI -8.3 to -0.0) WOMAC pain: 28.9 vs. 33.8; difference -4.8 (95% CI -11.2 to 1.6) 10 months WOMAC Total: 32.7 vs. 38.4; difference -5.7 (95% CI -12.1 to 0.7) WOMAC physical function: 33.0 vs. 38.9; difference -5.9 (95% CI -12.5 to 0.7) PDI: 20.0 vs. 23.6; difference -3.6 (95% CI -7.7 to 0.5) WOMAC pain: 30.0 vs. 33.5; difference -3.5 (95% CI -10.0 to 3.0)	A vs. B 4 months SF-36 Physical: 35.1 vs. 33.0; difference 2.1 (95% CI -0·5 to 4.8) SF-36 Mental: 52.6 vs. 51.7; difference 0.9 (95% CI 2.3 to 4.2) ADS (Depression): 48.2 vs. 48.7; difference -0·5 (95% CI -3.6 to 2.5) 10 months SF-36 Physical: 35.0 vs. 32.8; difference 2.2 (95% CI -0.6 to 5,1) SF-36 Mental: 52.9 vs. 51.1; difference 1.9 (95% CI -1.3 to 5.1) ADS: 48.6 vs. 49.8; difference -1.2 (95% CI -4.3 to 1.8)
Yurtkuran, 2007[196] 3 months Duration of pain: mean 5.4 years *Fair*	A. Laser acupuncture (n=28): applied to the medial side of the knee to the acupuncture point on the sural nerve; infrared 27 GaAs diode laser instrument (output 4 mW, 10 mW/cm2 power density, 120-sec treatment time and 0.48 J dose per session); irradiation was pulsed (duration of 1 pulse was 200	A vs. B Age: 52 vs. 53 years Female: 96% vs. 96% Duration of symptoms: 5.2 vs. 5.6 months WOMAC total: 66.5 vs. 51.3 WOMAC physical function: 47.5 vs. 35.3 WOMAC pain: 13.7 vs. 11.6	A vs. B 2.5 months WOMAC total: 62.4 vs. 50.6, difference 11.8 (95% CI -1.0 to 24.6) WOMAC physical function: 44.2 vs. 35.3, difference 11.9 (95% CI 2.9 to 20.9) WOMAC pain: 13.5 vs. 11.5, difference 2.0 (95% CI -1.3 to 5.3) VAS pain on movement: 5.6 vs. 4.8, difference 0.8 (95% CI -0.9 to 2.5)	A vs. B 2.5 months NHP (0-38): 7.6 vs. 6.4. difference 1.2 (95% CI -2.1 to 4.4)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	nanosecond), and only one point was treated with contact application technique. B. Sham laser acupuncture (n=27): performed in the same location and under the same conditions as the true laser acupuncture; patients could see a red light but the machine was turned off Both groups: 20 minutes sessions, 5 days per week for 2 weeks (total duration of therapy was 10 days, 10 sessions total); in addition, all patients received a home-based, standardized exercise program	VAS pain on movement (0-10): 6.5 vs. 6.1		

AIMS = Arthritis Impact Measurement Scale; BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; CI = confidence interval; HRQoL = health-related quality of life; NR = not reported; NRS = Numerical Rating Scale; NS = not statistically significant; OA = osteoarthritis; WOMAC = Western Ontario and McMaster Universities Osteoarthritis index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Acupuncture Compared With Usual Care, Waitlist or Sham

Functional outcomes. There were no clear differences between acupuncture versus control interventions on WOMAC function in the short-term (5 trials, pooled SMD -0.18, 95% CI -0.55 to 0.20, I^2 = 82%)[190,192,194-196] (Figure 35). All trials were considered fair quality. No differences were found whether standard needle acupuncture (2 trials, pooled SMD -0.63, 95% CI -1.65 to 0.30)[190,195] or electroacupuncture were used (2 trials, pooled SMD -0.11, 95% CI -0.34 to 2.74),[192,194] compared with control interventions. When stratified by control type no differences were found between any form of acupuncture and sham treatments (4 trials, pooled SMD -0.02, 95% CI -0.28 to 0.39);[192,194-196] however, when acupuncture was compared with waitlist and usual care, estimates suggest moderate improvement in function (2 trials, pooled SMD -0.74, 95% CI 1.40 to -0.24).[190,194] One small trial (N = 55)[196] that applied low-level laser to acupuncture points (no needles were used) reported a difference in WOMAC function score that favored the sham control (difference 11.9, 95%CI 2.9 to 20.9, scale not provided).

Similarly, there were no differences in short-term function between acupuncture and sham, waitlist, and usual care across trials based on total WOMAC score (4 trials, pooled SMD -0.30,

95% CI -0.81 to 0.21, I2 = 85%, plot not shown).[56,190,195,196] Stratification by acupuncture type, control type, and exclusion of one poor quality trial yielded similar estimates. No differences between acupuncture and any of the control conditions across other measures of function were reported in five trials,[56,191-193,195] (Table 29). One trial (N = 73)[190] reported that acupuncture improved function compared with usual care based on the WOMAC function subscale, WOMAC total score (scales not provided) and the Lequesne Index (0 to 24 scale). The trial of low-level laser application to acupuncture points reported a difference in WOMAC Total score that again favored the sham control.[196]

There were no clear differences between acupuncture versus control interventions on the WOMAC function scale in the intermediate term (3 trials, pooled SMD -0.12, 95% CI -0.30 to 0.07, I2 = 0%),[191,193,195] (Figure 35). Estimates were similar when stratified by study quality, acupuncture type, and control type; however, sensitivity analyses were limited by the small number of trials. No differences in WOMAC Total score were found for standard needle acupuncture versus usual care or sham (2 trials, pooled SMD -0.23, 95% CI -0.49 to 0-.03 I2 = 0%, plot not shown).[193,195] In the two trials reporting on intermediate-term function there were no differences between standard needle acupuncture and sham acupuncture on the 0 to 70 scale Pain Disability Index (mean difference -3.5, 95% CI -7.7 to 0 .5) in one fair-quality trial[195] or between acupuncture and usual care in one small poor-quality trial on a 12 to 60 scale Oxford Knee Score (mean difference 3.6, 95% CI -9.8 to 2.6).[193]

No trials reported data on long-term function.

Pain outcomes. There were no clear differences between acupuncture versus control interventions pain in the short term (6 trials, pooled SMD -0.27, 95%CI -0.56 to 0.02, I^2=75%)[56,190,192,194-196] (Figure 36). All but one trial used the WOMAC pain score. Estimates were similar after exclusion of one poor-quality trial and for stratification by acupuncture type and for analyses of VAS or NRS instead of WOMAC pain score if more than one pain measure was reported. There were no differences between acupuncture and sham control (4 trials, pooled SMD -0.06, 95%CI -0.24 to 0.14);[192,194-196] however, when acupuncture was compared with waitlist and usual care, estimates suggest slight effects on pain (2 trials, pooled SMD -0.68, 95% CI -1.28 to -0.15).[190,194]

There were no differences between acupuncture versus control interventions for pain in the intermediate term (3 trials, pooled SMD -0.11, 95% CI -0.30 to 0.07, I^2=0%)[191,193,195] (Figure 36). Stratification based on acupuncture type, type of control intervention, and study quality yielded similar results.

No trial reported data on long-term pain.

Other outcomes. Data on the effects of acupuncture on quality of life were limited. The effects were small and not statistically significant for individual trials on a 0 to 100 scale SF-12 or SF-36 PCS or MCS. In the short term a small effect favoring acupuncture versus control conditions was seen for 0 to 100 scale SF-12 or SF-36 PCS (2 trials, pooled difference 1.6, 95%CI 0.08 to 3.11 I^2 = 0%), but no difference in the MCS Score was seen (2 trials, pooled difference 1.14, 95%CI -0.27 to 2.56, I2 = 0%).[194,195] Individual trials reported no differences between acupuncture and control interventions on other quality of life measures or on measures of anxiety or depression. Similarly for the intermediate term a small effect favoring acupuncture over controls interventions was seen for PCS (2 trials, pooled difference 1.94, 95% CI 0.03 to 3.86, I^2 = 0%) but not for the MCS Score (2 trials, pooled difference -0.25.95% CI -4.05 to 3.54).[191,195]

There were no differences between acupuncture and control interventions on other quality of life measures or on measures of anxiety or depression (Table 29).

A small (1%) change in opioid use at 9 months post intervention was seen with needle acupuncture (decrease from 1% to 0%) and laser acupuncture (decrease from 3% to 2%), while use remained the same in the no treatment group.[191] A small change in the sham control group (decrease from 1% to 0%) was also observed (Table 29).

Acupuncture Compared With Pharmacological Therapy

No trial of acupuncture versus pharmacological therapy met inclusion criteria.

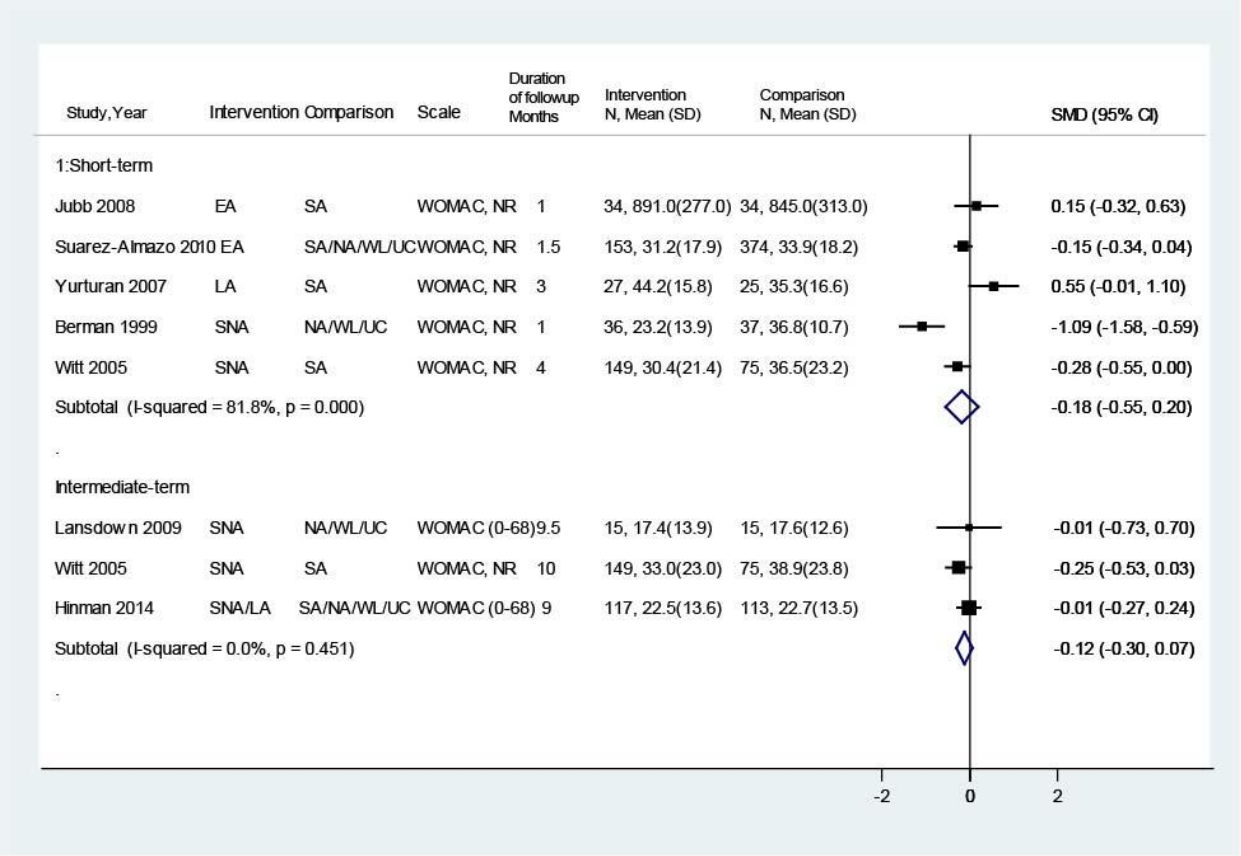
Acupuncture Compared With Exercise Therapy

Data were insufficient from one poor-quality trial (n=120)[56] to evaluate the effects of weekly acupuncture versus 60 minutes of combination exercise (strengthening, aerobics, stretching, and balance training) for 6 weeks for knee OA (Table 29 and Appendix D). Methodological limitations included lack of patient or care provider blinding, unclear adherence, unacceptable attrition and differential loss to followup (Appendix E). There were no differences between groups with regard to function on the Oxford Knee Score questionnaire (difference -0.7, 95% CI -3.5 to 2.1 on 12-60 scale) or WOMAC score (difference -1.0, 95% CI -6.7 to 4.7; scale not provided by author). Similarly there was no difference between treatments for VAS pain on a 0 to 10 scale (difference 0.22, 95% CI -0.67 to 1.11) or for anxiety or depression based on the Hospital Anxiety and Depression Scale.

Harms

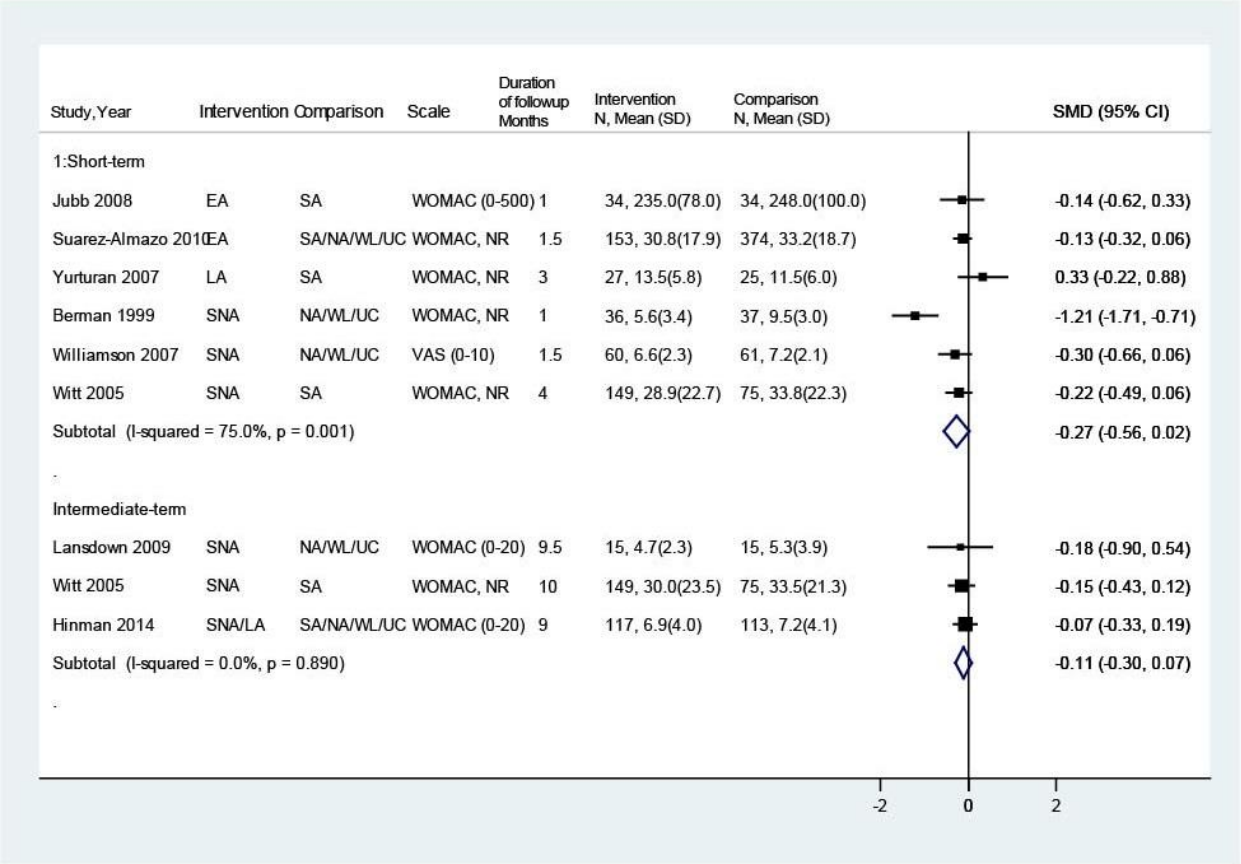
All trials reported adverse events. One trial reported similar rates of serious adverse events in patients who received real versus sham acupuncture (2.1% vs. 2.7%, respectively; RR 0.75, 95% CI 0.13 to 4.39), to include hospitalizations and one case of death from myocardial infarction in the control group; none were considered to be related to the study condition or treatment.[195] All other events reported were classified as mild and there was no apparent difference in risk of adverse events between any form of acupuncture and the control groups. The most common adverse events reported were worsening of symptoms (7% to 14%) in three trials[191,193,194] and mild bruising, swelling, or pain at the acupuncture site (1% to 18%) in five trials.[56,191,193-195] One trial reported one case of an infection at the electroacupuncture site (n=455 for real and sham acupuncture groups).[194]. In only one trial did an adverse event (not treatment related) lead to withdrawal: one patient (3%) in the acupuncture group had a flare-up of synovitis (non-septic).[192]

Figure 35. Acupuncture versus usual care, waitlist, sham, or a placebo intervention in osteoarthritis of the knee: effects on function



EA = electroacupucnture; LA = laser acupuncture; NR = not reported; SA = sham acupuncture; SNA = standard needle acupuncture; SD = standard deviation; SMD = standardized mean difference; NR = not reported; UC = usual care; WL = waitlist; WOMAC = Western Ontario and McMaster's Universities Osteoarthritis Index

Figure 36. Acupuncture versus usual care, waitlist, sham, or a placebo intervention for osteoarthritis of the knee: effects on pain



EA = electroacupucnture; LA = laser acupuncture; NR = not reported; SA = sham acupuncture; SNA = standard needle acupuncture; SD = standard deviation; SMD = standardized mean difference; NR = not reported; UC = usual care; WL = waitlist; WOMAC = Western Ontario and McMaster's Universities Osteoarthritis Index

Exercise for Osteoarthritis of the Hip

Key Points

- Exercise was associated with a small improvement in function versus usual care in the short term (3 trials, pooled standardized mean difference (SMD) -0.33, 95% CI, -0.53 to -0.12, I^2=0.0%), intermediate term (2 trials, pooled SMD -0.28, 95% CI -0.50 to -0.05, I^2=0.0%), and long term (1 trial, SMD -0.37, 95% CI -0.74 to -0.01) (SOE: Low for short and intermediate-term, Insufficient for long-term).
- Exercise tended toward slightly greater improvement in short-term pain compared with usual care (3 trials, pooled SMD -0.34, 95% CI, -0.63 to -0.04, I^2=48.2%) but the results were no longer significant at intermediate term (2 trials, pooled SMD -0.14, 95% CI -0.37 to 0.08, I^2=0%) or long term (1 trial, SMD -0.25, 95% CI -0.62 to 0 .11) (SOE: Low for short and intermediate-term, insufficient for long-term).

- Evidence for harms was insufficient in trials of exercise with only two trials describing adverse events. However, no serious harms were reported in either trial. (SOE: Insufficient).

Detailed Synthesis

Four trials of exercise therapy for hip OA met the inclusion criteria; three were conducted in Europe[57-59] and the other in New Zealand[36] (Table 30 and Appendix D). Three trials evaluated participants with chronic hip pain diagnosed as OA using American College of Radiology criteria[36,57,59] and one assessed participants with hip OA diagnosed clinically who were on a waitlist for hip replacement.[58] Sample sizes ranged from 45 to 203 (total number randomized=477). Across trials, participants were predominately female (>50%) with mean ages ranging from 64 to 69 years. Three trials were conducted in Europe[57-59] and the other in New Zealand.[36]

All trials compared exercise with usual care, defined as care routinely provided by the patient's primary care physician, which could include physical therapy referral. Two trials also provided education about hip OA to all participants.[57,59] The exercise interventions included 8 to 12 supervised sessions of 30 to 60 minutes duration once per week over 8 to 12 weeks; the interventions were comprised of strengthening and stretching exercises (all studies), as well as neuromuscular control exercises in one trial[36] and endurance exercise in another.[59] All trials reported compliance rates with the scheduled exercise sessions between 76 percent and 88 percent. However, in one study,[36] although 88 percent of patients completed more than 80 percent of the scheduled sessions, only 44 percent of participants returned logbooks to demonstrate compliance with the recommended home exercises.

Three trials were rated fair quality[36,57,59] and one was rated poor quality[58] (Appendix E). In all trials, the nature of the intervention and control precluded blinding of participants and researchers; patient-reported outcomes were therefore not blinded. Additionally, in the poor-quality study,[58] concealed allocation was unclear, outcomes were poorly reported, as were attrition rates, which were substantial for pain (68%) and function (73%) outcomes.

Table 30. Summary of results for osteoarthritis of the hip: exercise

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Abbott, 2013[36] 9.75 months Duration of pain: 9 months *Fair*	A. Exercise therapy (n=51/22 hip OA): 7 sessions of strengthening, stretching, and neuromuscular control over 9 weeks, with 2 booster sessions at week 16. Individual exercises prescribed as needed. Home exercise prescribed 3 times weekly B. Usual care (n=51/23 hip OA): Routine care	A vs. B (total population, includes knee OA) Age: 67 vs. 66 Females: 49% vs. 63% % hip OA: 43.1% vs. 45.1% WOMAC (0-240): 95.5 vs. 93.8	A vs. B (hip OA only) <u>9.75 months</u> WOMAC mean change from baseline: -12.4 vs. 6.6	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	provided by patient's own GP and other healthcare providers			
Juhakoski, 2011[57] 3, 9, and 21 months Duration of pain: Mean 8.3 to 8.5 years *Fair*	A. Exercise + usual care (n=57): 12 strengthening and stretching exercise sessions of 45 minutes once per week, with 4 booster sessions 1 year later B. Usual care (n=56): normal routine care offered by patient's own GP. All patients attended an hour-long session on basic principles of non-operative treatment of hip osteoarthritis	A vs. B Age: 67 vs. 66 years Female: 68% vs. 72% Duration of pain: 8.3 to 8.5 years WOMAC function (0-100): 24.7 vs. 28.9 WOMAC pain (0-100): 21.5 vs. 29.1	A vs. B <u>3 months</u> WOMAC function: 22.6 vs. 30.1, (MD -7.5, 95% CI -13.9 to -1.0) WOMAC pain: 23.4 vs. 28.9 (MD -5.5, 95% CI -13.0 to 2.0) <u>9 months</u> WOMAC function: 24.6 vs. 27.6 (MD -3.0, 95% CI -9.2 to 3.2) WOMAC pain: 22.9 vs. 25.0 (MD -2.1, 95% CI -9.2 to 5.0) <u>21 months</u> WOMAC function: 24.4 vs. 30.0 (MD -5.6, 95% CI -12.9 to 1.7) WOMAC pain: 24.1 vs. 27.9 (MD -3.8, 95% CI -12.0 to 4.4)	A vs. B <u>3 months</u> Weak opioid[b] use (p=0.73): Not using: 82.5% vs. 87.7% 1-6 times/week: 10.5% vs. 8.8% Daily: 7.0% vs. 3.5% <u>9 months</u> Mean doctor visits for hip OA: 0.5 vs. 0.8, p=0.07 Mean physiotherapy visits for hip OA: 1.3 vs. 2.0, p=0.05 Weak opioid[b] use (p=0.12): Not using: 81.0% vs. 93.1% 1-6 times/week: 10.4% vs. 1.7% Daily: 8.6% vs. 5.2% <u>21 months</u> Mean doctor visits (between 9 and 21 month followup) for hip OA: 0.5 vs. 1.1, p=0.05 Mean physiotherapy visits (between 9 and 21 month followup) for hip OA: 0.4 vs. 1.3, p<0.001 Weak opioid[b] use (p=0.70): Not using: 80.7% vs. 85.2% 1-6 times/week: 12.3% vs. 7.4% Daily: 7.0% vs. 7.4%
Tak,[58] 2005[c] 6 months, 3 years Mean duration of pain: NR *Poor*	A. Exercise (n=45): Eight weekly group sessions of strength training, information on a home exercise program, ergonomic advice, and dietary advice B. Usual care (n=49): Subject-initiated	A vs. B Age: 68 vs. 69 Female: 64% vs. 71% HHS (0-100): 71.1 vs. 71.0 GARS (18-72): 22.8 vs. 25.3 SIP-136 physical (0-100): 7.2 vs. 7.6	A vs. B <u>3 months</u> HHS: 75.4 vs. 71.1, (MD 4.3, 95% CI -2.2 to 10.8) GARS: 23.7 vs. 26.3, (MD -2.6, 95% CI -6.0 to 0.8) SIP-136 physical: 5.1 vs. 8.4, (MD -3.3, 95% CI -5.3 to -1.3) Pain VAS: 3.5 vs. 5.1,	A vs. B <u>3 months</u> QoL VAS (0-10): 5.0 vs. 4.2, (MD 1.4, 95% CI -0.2 to 3.0) HRQoL (7-39): 28.6 vs. 27.3, (MD 0.9, 95% CI -0.4 to 2.2)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	contact with GP. Reference group (n=NR) consisting of weekly stress management sessions for 10 weeks	Pain VAS (0-10): 3.8 vs. 4.2 HHS pain subscale (0-44): 27.9 vs. 28.8	(MD -1.6, 95% CI -2.6 to -0.6) HHS pain subscale: 29.6 vs. 26.9, (MD -0.9, 95% CI -4.7 to 2.9)	
Teirlinck, 2016[59] 3 and 9 months Duration of pain: Median 1 year *Fair*	A. Exercise therapy (n=101): 12 sessions over 3 months consisting of strengthening, stretching, and aerobic exercise B. Usual care (n=102): Routine care provided by patient's own GP	A vs. B Age: 64 vs. 67 Females: 62% vs. 55% Pain duration median (IQR): 365 (810) vs. 365 (819) days HOOS function (0-100): 35.4 vs. 32.2 HOOS pain (0-100): 37.6 vs. 38.9 ICOAP constant pain (0-20): 5.4 vs. 5.8 ICOAP intermittent pain (0-24): 8.0 vs. 8.4 ICOAP total pain (0-100): 30.4 vs. 32.2	A vs. B <u>3 months</u> HOOS function: 30.8 vs. 35.3, (Adj MD -2.4, 95% CI -6.7 to 1.9) HOOS pain: 34.4 vs. 37.2, (Adj MD -2.2, 95% CI -6.2 to 1.7) ICOAP constant pain: 4.0 vs. 5.3, (Adj MD -0.9, 95% CI -1.9 to 0.1) ICOAP intermittent pain: 7.0 vs. 7.9 , (Adj MD -0.6, 95% CI -1.7 to 0.6) ICOAP total pain: 24.9 vs. 29.8, (Adj MD -3.3, 95% CI -8.0 to 1.4) <u>9 months</u> HOOS function: 26.8 vs. 34.2, (Adj MD -3.0, 95% CI -6.7 to 0.2) HOOS pain: 31.6 vs. 34.6, (Adj MD -1.6, 95% CI -6.2 to 3.0) ICOAP constant pain: 3.6 vs. 4.7, (Adj MD -0.7, 95% CI -1.7 to 0.4) ICOAP intermittent pain: 6.1 vs. 7.2, (Adj MD -0.6, 95% CI -1.8 to 0.6) ICOAP total pain: 22.2 vs. 27.0, (Adj MD -2.8, 95% CI -7.6 to 2.0)	A vs. B <u>3 months</u> EuroQol 5D-3L (-0.329-1.0): 0.77 vs. 0.76, (Adj MD -0.01, 95% CI -0.06 to 0.04) <u>9 months</u> EuroQol 5D-3L: 0.78 vs. 0.78, (Adj MD -0.01, 95% CI -0.06 to 0.04) Total hip replacements: 6 vs. 9

CI = confidence interval; GARS = gait abnormality rating scale; GP = general practitioner; HHS = Harris Hip Score; HOOS = hip disability and osteoarthritis outcome score; HRQoL = Health Related Quality of Life; ICOA =: intermittent and constant pain score; MD = mean difference; NR = not reported; OA = osteoarthritis; QoL = Quality of Life; SIP-136 = Sickness Impact Profile-136; VAS = Visual Analog Scale; WOMAC= Western Ontario and McMaster Universities Osteoarthritis Index
a Unless otherwise noted, followup time is calculated from the end of the treatment period
b Authors defined weak opioids as tramadol or codeine
c Cluster RCT where clusters were formed from participants selecting a time that best fit their schedule.

Exercise Compared With Usual Care

Exercise was associated with a slightly greater effect on function versus usual care in the short term (3 trials, pooled standardized mean difference (SMD) -0.33, 95% CI -0.53 to -0.12, I2=0.0%),[57-59] intermediate term (2 trials, pooled SMD -0.28, 95% CI -0.50 to -0.05, I2=0.0%)[57,59] and long term (1 trial, SMD -0.37, 95% CI -0.74 to -0.01)[57] (Figure 37). The

intermediate-term findings are consistent with the additional trial not included in the meta-analysis (authors did not provide sufficient data),[36] although the small improvement in function in this trial did not reach statistical significance in those with hip OA. The small number of trials precluded meaningful sensitivity analysis.

Exercise tended toward slightly greater improvement on short-term pain compared with usual care (3 trials, pooled SMD -0.34, 95% CI -0.63 to -0.04, I^2=48%)[57-59] (Figure 38), but not at intermediate term (2 trials, pooled SMD -0.14, 95% CI -0.37 to 0.08, I^2=0%).[57,59] There was moderate heterogeneity between studies and the short-term improvement in pain was observed in only one poor quality study,[58] whereas the two fair-quality studies did not demonstrate any significant differences in short-term pain relief.[57,59] There were no identifiable differences in methodology between the studies to explain these inconsistent findings, although the poor-quality study only reported pain outcomes for 68 percent of participants, which may have biased results. There was no difference between exercise versus usual care in the long term based on a single study (SMD -0.25, 95% CI -0.62 to 0.11).[57] The small number of trials precluded meaningful sensitivity analysis.

Data on effects of exercise on quality of life were limited and were reported in only two trials.[58,59] One fair-quality trial[59] found no differences in health-related quality of life between groups in the short term and intermediate term and one poor-quality study[58] found no differences between groups in the short term. One fair-quality study found no differences between groups in terms of opioid use at any time point (proportion of patients using tramadol or codeine daily: 7.0% vs. 3.5% at 3 months, 8.6% vs. 5.2% at 9 months, and 7.0% vs. 7.4% at 21 months, p=0.73), but did report slightly fewer followup physical therapy visits in the exercise group in the intermediate and long terms[57] (Table 30).

There was insufficient evidence to determine effects of duration of exercise therapy or number of sessions on outcomes.

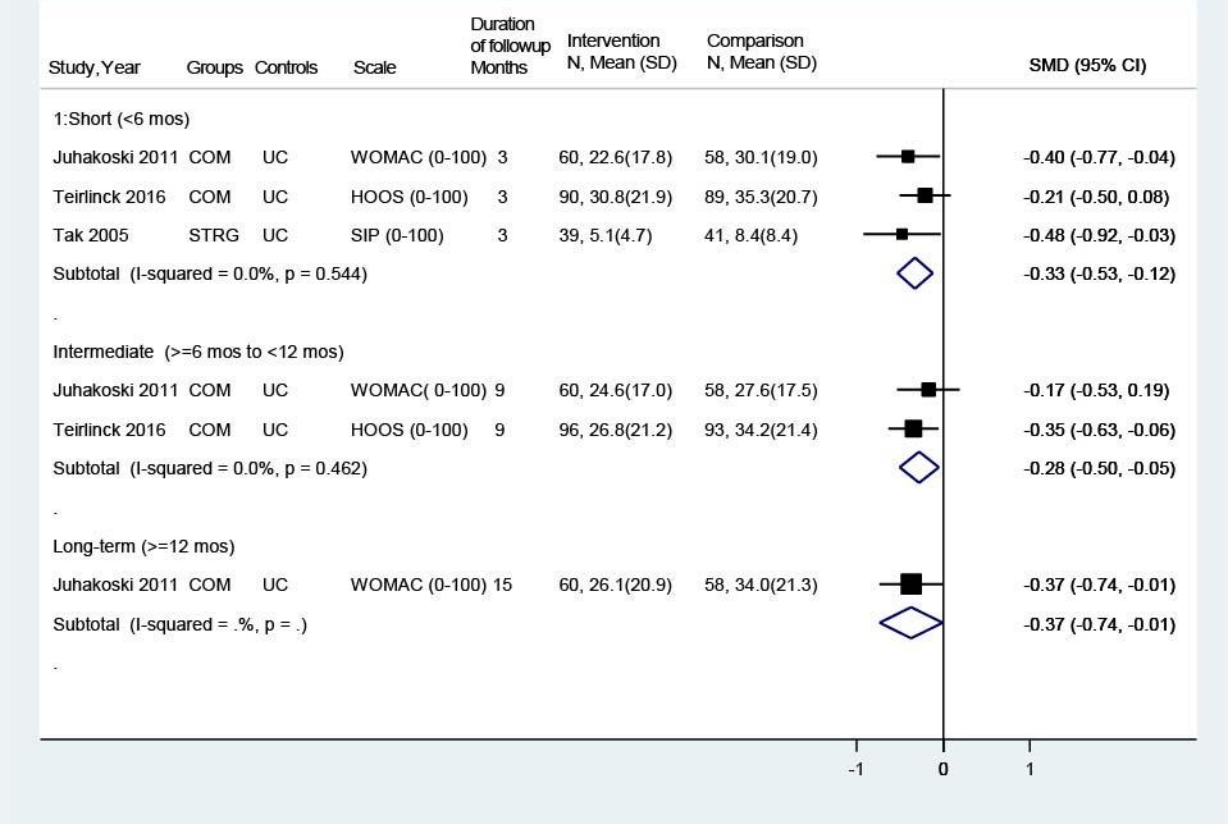
Exercise Compared With Pharmacological Therapy or With Other Nonpharmacological Therapies

No trial of exercise versus pharmacological therapy met inclusion criteria. Findings for exercise versus other non-pharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

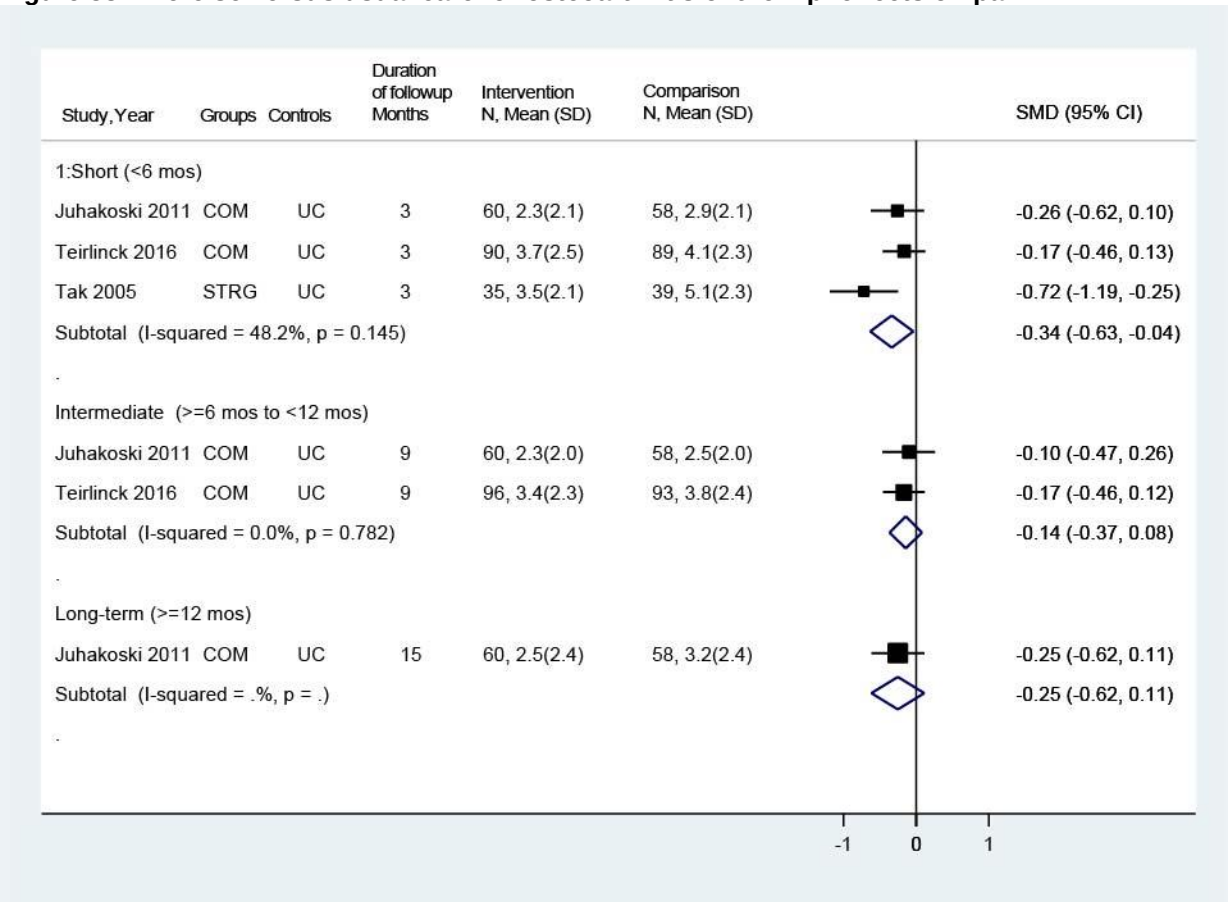
Only two exercise trials reported on harms, and neither reported any adverse events in either the exercise group or usual care groups.[36,58]

Figure 37. Exercise versus usual care for osteoarthritis of the hip: effects on function



CI = confidence interval; COM = combination exercise therapy; HOOS = Hip disability and Osteoarthritis Outcomes Score; SD = standard deviation; SIP = Sickness Impact Profile physical function score; SMD = standardized mean difference; STRG = strength training exercise; UC = usual care; WOMAC = Western Ontario and McMaster's Universities Osteoarthritis Index

Figure 38. Exercise versus usual care for osteoarthritis of the hip: effects on pain



CI = confidence interval; COM = combination exercise therapy; SD = standard deviation; SMD = standardized mean difference; STRG = strength training exercise; UC = usual care.

Manual Therapies for Osteoarthritis of the Hip

Key Points

- There are insufficient data to determine the effects or harms of manual therapy compared with usual care at intermediate term. No effect size could be calculated (SOE: Insufficient).
- Manual therapy was associated with a small short-term effects (mean difference 11.1, 95% CI 4.0 to 18.6, 0-100 scale Harris Hip Score) and small intermediate-term effects (mean difference 9.7, 95% CI, 1.5 to 17.9) on function versus exercise (SOE: Low).
- Manual therapy was associated with a small effect on pain in the short term [mean differences of -0.72 (95% CI -1.38 to -0.05) for pain at rest and -1.21 (95% CI -2.29 to -0.25) for pain walking] versus exercise. (SOE: Low) The impact on pain is not clear at intermediate term; there was no difference in pain at rest (adjusted difference -7.0, 95% CI -20.3 to 5.9 , 0-100 scale) but there was small improvement in pain while walking, adjusted difference -12.7, 95% CI -24.0 to -1.9) (SOE: Insufficient).

- No trials evaluated manual therapies versus pharmacological therapy.
- One trial reported that no treatment related-serious adverse events were detected and in the other, no difference in study withdrawal due to symptom aggravation was seen between manual therapy and exercise (RR 1.42, 95% CI 0.25 to 8.16) (SOE: Low).

Detailed Synthesis

We identified two trials (n=69 and 109) of manual therapy for hip OA that met inclusion criteria (Table 31 and Appendix D); one was conducted in New Zealand[36] and the other in the Netherlands.[155] Mean patient age ranged from 66 to 72 years and females comprised 49 percent to 72 percent of the populations. Both trials required a diagnosis of hip OA meeting the American College of Rheumatology (ACR) criteria for inclusion. The duration of manual therapy ranged from 5 to 16 weeks with a total of nine sessions in both groups; in one trial this included seven sessions over the first 9 weeks and two booster sessions at week 16.[36] One trial compared manual therapy to usual care (continued routine care from a general practitioner and other providers)[36] and both trials compared manual therapy to combination exercise programs.[36,155] The number of exercise sessions matched the manual therapy group of that respective study. All participants were prescribed a home exercise program three times per week. One trial reported short-term outcomes[155] and both reported intermediate-term outcomes.

Both trials were rated fair quality (Appendix E). Compliance with the intervention was acceptable in all groups, and the methodological shortcomings of these trials included a lack of blinding for the patients and care providers.

Table 31. Summary of results for osteoarthritis of the hip: manual therapy

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Abbott, 2013[36] 9.75 months Duration of diagnosis: 2.6 years *Fair*	A. Manual therapy (n=54/24 hip OA): 7 manual therapy sessions in 9 weeks with 2 additional booster sessions B. Exercise (n=51/22 hip OA), 7 exercise sessions in 9 weeks with 2 additional booster sessions C. Usual care (n=51/23 hip OA)	A vs. B vs. C (total population, includes knee OA) Age: 67 vs. 67 vs. 66 years Female: 49% vs. 52% vs. 58% Percent knee OA: 56% vs. 57% vs. 55% Percent hip OA: 44% vs. 43% vs. 45% Percent both hip OA and knee OA: 22% vs. 20% vs. 26% Baseline WOMAC (0-240): 114.8 vs. 95.5 vs. 93.8	A vs. B (hip OA only) 9.75 months WOMAC, mean change from baseline: -22.9 vs. -12.4, p NR A vs. C (hip OA only) 9.75 months WOMAC, mean change from baseline: -22.9 vs. 6.6, p NR	None

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Hoeksma, 2004[155] 3 and 6 months Duration of symptoms: mean NR *Fair*	A. Manual therapy (n=56): Sessions consisted of stretching followed by traction manipulation in each limited position (high velocity thrust technique). B. Exercise therapy (n=53): Sessions implemented exercises for muscle functions, muscle length, joint mobility, pain relief, and walking ability and were tailored to the specific needs of the patient. Instructions for home exercises were given. Both groups received 2 sessions per week for 5 weeks (9 sessions in total).	Age: 72 vs. 71 years Females: 68% vs. 72% Symptom duration of 1 month to 5 years: 76% vs.. 81% Severe OA on radiography: 45% vs. 38% HHS (0-100): 54 vs. 53 Pain at rest VAS (0-100): 22.5 vs. 23.0 Pain walking VAS (0-100): 34.0 vs. 28.8	A vs. B <u>3 months</u> HHS: 68.4 vs. 56.0, adjusted difference 11.1, 95% CI 4.0 to 18.6 Pain at rest VAS: 19.1 vs. 26.9, adjusted difference -7.2, 95% CI -13.8 to -0.5 Pain walking VAS: 16.4 vs. 23.7, adjusted difference -12.1, 95% CI -22.9 to -2.5 <u>6 months</u> HHS: 70.2 vs. 59.7, adjusted difference 9.7, 95% CI 1.5 to 17.9 Pain at rest VAS: 14.0 vs. 21.6, adjusted difference -7.0, 95% CI -20.3 to 5.9 Pain walking VAS: 17.0 vs. 24.3, adjusted difference -12.7, 95% CI -24.0 to -1.9	A vs. B <u>3 months</u> SF-36 physical function: 45.3 vs. 46.6, adjusted difference -2.1, 95% CI -11.7 to 7.7 SF-36 role physical function: 25.4 vs. 29.8, adjusted difference -23.5 to 10.2 SF-36 bodily pain: 47.4 vs. 46.1, adjusted difference -3.2, 95% CI -13.1 to 6.8 <u>6 months</u> SF-36 physical function: 50.4 vs. 45.3, adjusted difference 3.1, 95% CI -4.1 to 10.5 SF-36 role physical function: 36.7 vs. 32.4, adjusted difference 2.2, 95% CI -16.8 to 21.1 SF-36 bodily pain: 51.4 vs. 49.9, adjusted difference -1.5, 95% CI -11.1 to 7.7

CI = confidence interval; HHS = Harris Hip Score; NR =not reported; OA = osteoarthritis; SF-36 = Short Form 36 Questionnaire; VAS = Visual Analog Scale; WOMAC =Western Ontario and McMaster Universities Arthritis Index

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Manual Therapies Compared With Usual Care

A single fair-quality trial (n=69 with hip OA)[36] found that manual therapy resulted in an improvement in function at intermediate term using the total WOMAC score (0 to 240) in the manual therapy group (mean change from baseline -22.9, 95% CI, -43.3 to -2.6), while the usual care group showed little change from baseline (mean change -7.9 (95% CI, -30.9 to 15.3). Lack of information on the number of patients precluded calculation of effect size, and results of statistical testing between groups was not presented.

Manual Therapies Compared With Pharmacological Therapy

No trial of manual therapy versus pharmacological therapy met inclusion criteria

Manual Therapies Compared With Exercise

One trial found that manual therapy resulted in slightly better short-term function compared with exercise (adjusted mean difference on the 0-100 scale Harris Hip Score [HHS] of 11.1, 95% CI 4.0 to 18.6). Regarding intermediate-term function, manual therapy conferred a slight benefit in both trials. The adjusted mean difference on the HHS was 9.7 (95% CI, 1.5 to 17.9) in one trial.[155] The other trial compared function using the total WOMAC score (0 to 240); the manual therapy group experienced a statistically significant improvement from baseline (mean change of -22.9, 95% CI, -43.3 to -2.6), while the exercise group did not (mean change -12.4, 95% CI, -27.1 to 2.3).[36]

Only one of the trials reported pain outcomes. Manual therapy was associated with slightly better short-term pain at rest and during walking compared to exercise (adjusted mean differences on a VAS (0 to 10) of -0.72, 95% CI -1.38 to -0.05, and -1.21, 95% CI -2.29 to -0.25, respectively).[155] Intermediate-term pain results were inconsistent. A moderate effect on VAS pain during walking was seen following manual therapy compared to exercise (adjusted mean difference -1.27, 95% CI -2.40 to -0.19), but there was no difference for pain at rest (adjusted mean difference -0.70, 95% CI -2.03 to 0.59).[155]

There was no difference in one trial[155] between manual therapy and exercise for short-term or intermediate-term quality of life measured with the SF-36 physical function, role physical, or bodily pain subscales (Table 31).

Harms

No trial-related serious adverse events were detected in one trial,[36] and there was no difference in symptom aggravation leading to withdrawal (5% vs. 4%; RR 1.42, 95% CI 0.25 to 8.16) in the other trial.[155]

Exercise for Osteoarthritis of the Hand

Key Points

- Data from one poor-quality trial were insufficient to determine the effects or harms (though no serious harms were reported) of exercise versus usual care in the short term (SOE: insufficient).

Detailed Synthesis

One Norwegian trial (n=130) that evaluated the effects of strengthening and range of motion exercise (3 times weekly for 3 months plus 4 group sessions) versus usual care (treatment recommended by the patient's general practitioner) met inclusion criteria[60] (Table 32 and Appendix D). This trial was rated poor quality due to lack of patient blinding, baseline differences in mental health conditions, and large differential attrition between groups (exercise 29% vs. usual care 7%) (Appendix E). Only short-term data was reported.

Table 32. Summary of results for osteoarthritis of the hand: exercise

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Osteras, 2014[60] 3 months Duration of pain: NR *Poor*	A. Exercise (n=46): ROM/strength exercises, 4 group sessions supplemented by instructions for home exercise 3 times per week for 12 weeks B. Usual care (n=64): Subjects received no particular attention, referral, or treatment from the study.	A vs. B Age: 67 vs. 65 years Females: 89% vs. 91% Fulfilment of ACR criteria for hand OA 91% vs. 91% Self-reported hip OA: 39% vs. 46% Self-reported knee OA: 40% vs. 51% Other rheumatic disease: 13% vs. 15% Severe mental distress: 17% vs. 39% FIHOA (0-30): 10.8 vs. 9.8 PSFS (0-10): 3.5 vs. 3.9 Hand pain NRS (0-10): 4.2 vs. 3.9	A vs. B <u>3 months</u> FIHOA: 10.9 vs. 10.5; adjusted difference -0.5 (95% CI -1.9 to 0.8) Hand pain NRS: 4.3 vs. 4.3 ; adjusted difference -0.2 (95% CI -0.8 to 0.3) OARSI OMERACT no. of responders: 30% vs. 28% (NS)	A vs. B <u>3 months</u> PSFS: 4.3 vs. 4.4 ; adjusted difference 0.1 (95% CI -0.7 to 1.0) Patient global assessment of disease activity: 4.2 vs. 4.1; adjusted difference 0.1 (95% CI -0.5, 0.7) Patient global assessment of disease activity affecting ADL: 3.8 vs. 3.8 ; adjusted difference -0.2 (95% CI -0.8 to 0.4)

ACR = American College of Radiology; ADL = activity of daily living; CI = confidence interval; FIHOA = Functional Index for Hand OsteoArthritis; NR = not reported; NRS = numeric rating scale; NS = not statistically significant; OA = osteoarthritis; OARSI OMERACT = Osteoarthritis Research Society International Outcome Measures in Rheumatology; PSFS = patient-specific function scale; ROM = range of motion.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Exercise Compared With Usual Care

Data were insufficient from one poor-quality trial. No differences between exercise and usual care were observed for function according to the Functional Index for Hand OsteoArthritis (adjusted mean difference -0.5 on a 0-30 scale, 95% CI -1.9 to 0.8), or for pain (adjusted MD -0.2 on a 0 to 10 VAS pain scale, 95% CI -0.8 to 0.3) at 3 months.[60] Similarly, there were no differences between groups in the proportion of OARSI OMERACT responders (30% versus 28%). There were also no differences between groups in any secondary outcome measure, including the patient-specific function scale, hand stiffness, or patient global assessment of disease activity.

The effects of exercise on use of opioid therapies or health care utilization were not reported. There was insufficient evidence to determine effects of duration of exercise therapy or number of sessions on outcomes.

Exercise Compared With Pharmacological Therapy or Other Nonpharmacological Therapies

No trial of exercise versus pharmacological therapy met inclusion criteria. Findings for exercise versus other non-pharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

In this trial,[60] no serious adverse events were reported; 8/130 (6%) patients reported increased pain (3 in hand, 5 in neck/shoulders) but adverse events were not reported by group.

Physical Modalities for Osteoarthritis of the Hand

Key Points

- One good-quality study of low-level laser treatment versus sham demonstrated no improvement in terms of function (difference 0.2, 95% CI -0.2 to 0.6) or pain (difference 0.1, 95% CI -0.3 to 0.5) in the short term (SOE: Low).
- Data were insufficient from one fair-quality trial to determine effects or harms of heat therapy using paraffin compared to no treatment on function or pain (SOE: Insufficient).
- No serious harms were reported in the trial of low level laser therapy (SOE: Low).

Detailed Synthesis

We identified two trials of physical modality use for hand OA (Table 33 and Appendixes D and E). One good-quality double-blind Canadian trial (N=88)[132] compared three, 20-minute sessions of low-level laser treatment to a sham laser probe over a 6-week period. Identical treatment procedures were used in each group. All participants attended three sham laser treatment sessions prior to randomization to ensure ability to comply with the treatment protocol.

One fair-quality trial (n=56) conducted in Turkey compared 15 minutes of paraffin wrapping 5 days per week for 3 weeks with a no treatment control group.[133] Both groups received information about joint protection strategies. Methodological limitations included lack of patient blinding, unclear compliance with treatment, and poorly reported analyses.

Table 33. Summary of results for osteoarthritis of the hand: physical modalities

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Brosseau,[132] 2005 4.5 months Duration of pain: NR *Good*	A. Low-level laser therapy (n=42): 3 J/cm^2 applied for 1 second each to the skin overlying the radial, medial and ulnar nerves (total of 15 points irradiated); 3 sessions lasting 20 minutes per week for 6 weeks B. Sham low-level	A vs. B Age: 64 vs. 65 years Female: 74% vs. 83% Medication use: 60% vs. 61% Diagnosis of OA: 7.5 vs. 8.5 years Pain intensity VAS (0-100): 56.9 vs. 49.4 AUSCAN function (0-4)[b]: 2.2 vs. 2.1 AUSCAN pain (0-4)[b]:	A vs. B 4.5 months AUSCAN function: 1.9 vs. 1.7, difference 0.2 (95% CI -0.2 to 0.6) AUSCAN pain: 1.9 vs. 1.8, difference 0.1 (95% CI -0.3 to 0.5) Pain VAS: NR	A vs. B 4.5 months Patient global assessment: Fully improved: 0% vs. 3% Partially improved: 40% vs. 33.3% No improvement: 60% vs. 52%

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	laser therapy (n=46): same procedure as the active treatment but a sham laser probe was used.	2.4 vs. 2.1		
Dilek, 2013[133] 2.25 months Duration of pain: 5.5 years *Fair*	A. Dip-wrap paraffin bath therapy (n=24): patients dip both hands into 50°C parrafin bath 10 times, paraffin left on for 15 minutes, treatment administered 5 days per week for 3 weeks B. Control group (n=22): Details NR; assumed to be no treatment Only paracetamol intake was permitted during the study	A vs. B Age: 59 vs. 60 years Female: 83% vs. 91% AUSCAN function (0-36)[c]: 16.2 vs. 17.1 AUSCAN pain (0-20)[c] : 10.7 vs. 9.8 Pain at rest, median (VAS 0-10): 5.0 vs. 4.0 Pain during ADL, median (VAS 0-10): 7.0 vs. 8.0	A vs. B 2.25 months AUSCAN function: 13.8 vs. 17.8, difference -4.0 (95% CI -8.6 to 0.6) DFI: data NR, p=0.05 AUSCAN pain: 6.5 vs. 9.5, difference -3 (95% CI -5.5 to -0.5) Pain VAS at rest, median: 0.0 vs. 5.0, p<0.001 Pain VAS during ADL, median: 5.0 vs. 7.0, p=0.05	NR

ADL = activity of daily living; AUSCAN = Australian Canadian Osteoarthritis Hand Index; CI =confidence interval; DFI = Dreiser Functional Index; NR = not reported; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Data for the AUSCAN was presented as an average of all responses, on a 5-point Likert scale (0-4), for both the physical function (9 items) and pain (5 items) subscale.

[c] Data for the AUSCAN was presented as a sum of the values across all items within the physical function (9 items) and pain (5 items) subscales; a 5-point Likert scale (0-4) was used to rate each item resulting in score ranges of 0-36 and 0-20, respectively.

Physical Modalities Compared With Sham or No Treatment

Low-level laser therapy. In the one good-quality trial of low-level laser treatment versus sham (n =88),[132] there were no differences in short-term function (mean difference 0.2 on a 0-4 Australian Canadian Osteoarthritis Hand Index [AUSCAN] functional subscale, 95% CI -0.2 to 0.6) or pain (mean difference 0.1 on a 0-4 AUSCAN pain subscale, 95% CI -0.3 to 0.5) at 4.5 months. Likewise, no difference was seen between groups in improvement based on patient global assessment.

Paraffin treatment. One fair-quality trial (N =56)[133] of paraffin heat treatment demonstrated no difference compared with no treatment on the AUSCAN function scale (0-36) (mean difference -4.0, 95% CI -8.6 to 0.6 at short-term (2.25 months) followup). Regarding pain, no clear difference was identified between the groups over the short term as there was inconsistency across measures used and analyses for outcomes were poorly reported; findings were considered insufficient.[133] While heat treatment was slightly favored based on the AUSCAN pain subscale (mean difference -3 on a 0-20 scale, 95% CI -5.5 to -0.5), it was not statistically significant in the

author's intention-to-treat (ITT) analysis (p=0.07). VAS pain at rest suggested more improvement with heat therapy versus control in the ITT analysis (median 0 vs. 5.0 on a 0-10 scale, p<0.001); however, there was no clear difference between groups on VAS pain during ADL (median 5.0 vs. 7.0, p=0.09 for per protocol analysis, p=0.05 for ITT).

No trial evaluated effects of physical modalities on use of opioid therapies or health care utilization.

Physical Modalities Compared With Pharmacological Therapy or With Exercise Therapy

No trial of a physical modality versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

Only the low-level laser therapy trial reported adverse events; no serious harms were reported.[132] One patient (2%) who received low-level laser treatment experienced erythema at the site.

Multidisciplinary Rehabilitation for Osteoarthritis of the Hand

Key Points

- One fair-quality trial of multidisciplinary rehabilitation versus waitlist control demonstrated no short-term differences between groups in function (adjusted difference 0.49, 95% CI, -0.09 to 0.37 on 0-36 scale), pain (adjusted difference 0.40, 95% CI, -0.5 to 1.3 on a 0-20 scale) or with regard to the proportion of OARSI OMERACT responders (OR 0.82, 95% CI, 0.42 to 1.61) (SOE: Low for all outcomes).
- Data on harms were insufficient, although no serious adverse events were reported in the one trial of multidisciplinary rehabilitation versus waitlist control (SOE: Insufficient).

Detailed Synthesis

One fair-quality trial (n=151) compared four, 2.5- to 3-hour group-based occupational therapy sessions of consisting of self-management techniques, ergonomic principles, daily home exercises, and splint (optional) with a waitlist control[210] (Table 34 and Appendix D). Waitlist control consisted of one 30-minute explanation of OA followed by a 3-month waiting period. Effect estimates were adjusted for baseline function or pain, body mass index (BMI), gender, and presence of erosive arthritis. Methodological limitations included lack of patient blinding and unreported compliance to treatment (Appendix E).

Table 34. Summary of results for osteoarthritis of the hand: multidisciplinary rehabilitation

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Stukstette[210] 2013 3 months Duration of	A. Multidisciplinary treatment program (n=75): 4 group based therapy sessions of 2.5-3	A vs. B Age: 60 vs. 58 Female: 18% vs. 16% Mean duration of	A vs. B <u>3 months</u> AUSCAN function: 18.6 vs. 18.8, adjusted mean difference 0.49 (95% CI -	A vs. B <u>3 months</u> Patient global assessment: 60.4 vs. 66.0, adjusted mean

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
pain: 4 years *Fair*	hours duration (time period NR), supervised by a specialized nurse and occupational therapist B. Waiting list (n=72) All patients: 30 minute explanation of written information about OA	diagnosis: 4 vs. 4 years Proportion taking opioids: 3% vs. 4% AUSCAN function (0-36): 21.0 vs. 21.8 AUSCAN pain (0-20):10.4 vs. 10.2	0.09 to 0.37) AUSCAN pain: 9.4 vs. 9.0, adjusted mean difference 0.40 (95% CI, -0.5 to 1.3) OARSI OMERACT responders: 33% vs. 37%, OR 0.82 (95% CI 0.42 to 1.61)	difference -5.2 (95% CI -11.4, 1.0) SF-36 PCS: 39.8 vs. 39.9, adjusted mean difference -0.14 (95% CI -1.62 to 1.35) SF-36 MCS: 50.3 vs. 51.6, adjusted mean difference 0.27 (95% CI -2.13 to 2.67)

AUSCAN = Australian Canadian Osteoarthritis Hand Index; NR = not reported; OARSI-OMERACT = Osteoarthritis Research Society International Outcome Measures in Rheumatology; SF-36 MCS = Short-Form 36 Mental Component Summary score; SF-36 PCS = Short-Form 36 Physical Component Summary score

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Multidisciplinary Rehabilitation Compared With Waitlist

No short-term (3 months) differences in function on the AUSCAN functional subscale (adjusted MD 0.49, 95% CI -0.09 to 0.37 on 0-36 scale) or on the AUSCAN pain subscale (adjusted MD 0.40, 95% CI, -0.5 to 1.3, scale 0-20) were reported.[210]

There was no difference in the proportion of OARSI OMERACT responders (OR 0.82, 95% CO 0.42 to 1.61) between groups or on any secondary outcome measure, including activities of daily living (Canadian Occupational Measurement Scales), health-related quality of life (SF-36), arthritis self-efficacy, pain coping, muscle strength, or joint mobility.[210]

The effect of multidisciplinary rehabilitation on use of opioid therapies or health care utilization was not evaluated in any of the included studies.

Multidisciplinary Rehabilitation Compared With Pharmacological Therapy or With Exercise Therapy

No trial of a multidisciplinary rehabilitation program versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

No serious adverse events were reported. One patient reported a swollen hand and increased pain after the second treatment session.[210]

Key Question 4: Fibromyalgia

Exercise for Fibromyalgia

Key Points

- Exercise was associated with slightly greater effects on function compared with attention control, no treatmen,t or usual care in the short term (7 trials, pooled difference -7.61 on

a 0 to 100 scale, 95% CI, -12.78 to -2.43, I2= 59.9%) (SOE: Low) and intermediate term (8 trials, pooled difference, -6.04, 95% CI –9.05 to -3.03, I2= 0%) (SOE: Moderate). There were no clear effects long term (3 trials, pooled difference, -4.33, 95% CI -10.18 to 1.52, I2= 0%) (SOE: Low).

- Exercise had a slightly greater effect on VAS pain (0 to 10 scale) compared with usual care, attention control, or no treatment short erm (6 trials (excluding outlier trial) pooled difference -0.89, 95% CI -1.32 to -0.46 I2 = 0%), but there were no clear effects at intermediate term (6 trials, pooled difference -0.31, 95% CI -0.79 to 0.17, I^2= 5.4%) or long term (4 trials, pooled difference -0.18, 95% CI -0.77 to 0.42, I2= 0%) (SOE: Moderate for all time frames).
- Data on harms were insufficient. Most trials of exercise did not report on adverse events at all. One trial reported one non-study-related adverse event. Two trials reported no adverse events (SOE: Insufficient).

Detailed Synthesis

Twenty trials (across 22 publications) of exercise therapy for fibromyalgia met inclusion criteria[61-82] (Table 35 and Appendix D). The exercise interventions varied across the trials and included combinations of different exercise types (11 trials),[62,63,65,68,70,74,76-81] aerobic exercise (9 trials),[64,67,69,71-73,75,77,82] muscle performance exercise/strength training (1 trial),[71] and Pilates (1 trial).[61] The duration of exercise therapy ranged from 1 to 8 months across the trials and the number of exercise sessions ranged from four to six (at a frequency of 1 to 5 times per week). Many trials also included instruction for home exercise practice. Exercise was compared to usual care in eight trials,[64,65,75-77,80-82] no treatment in six trials[68-71,74,78,79] attention control in five trials,[61,63,67,72,73] and waitlist in one trial.[62] Usual care generally included medical treatment for fibromyalgia and continued normal daily activities (which often specifically excluded the exercise intervention being evaluated). Attention control conditions consisted of fibromyalgia education sessions, social support, instructions in coping strategies, relaxation and stretching exercises, and physical activity planning

Sample sizes ranged from 32 to 166 across the trials (total number randomized=1,276). Patient mean age ranged from 44 to 57 years, and the majority were female (89% to 100%). Twelve trials were conducted in Europe,[64,68,70,73-82] five in North America,[63,65-67,69,72] two in Brazil,[62,71] and one in Turkey.[61]

Eleven trials were rated fair quality[61,62,64-67,71,73,74,77,80,82] and nine poor quality[63,68-70,72,75,76,78,79,81] (Appendix E). Methodological limitations in the fair-quality trials were primarily related to unclear allocation concealment methods and lack of blinding (the nature of interventions precluded blinding of participants and researchers). Additionally, poor-quality trials also suffered from unclear randomization methods and high rates of attrition and/or differential attrition.

Table 35. Summary of results for fibromyalgia: exercise therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Altan,[61] 2009 3 months Pain duration	A. Pilates (n=25): 1 hour session 3 times per week for 3 months: Pilates postural education,	A vs. B Age: 48 vs. 50 years Female: 100% vs. 100%	A vs. B <u>3 months:</u> FIQ: 69.3 vs. 77.6, difference -8.3 (95% CI -21.8 to 5.2)	A vs. B <u>3 months:</u> NHP (0-100): 224.2 vs. 246.3, difference -22.1 (95% CI -96.0 to 51.8)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
NR *Fair*	search for neutral position, sitting, antalgic, stretching and, proproceptivity improvement exercises, and breathing education B. Attention control (n=25): Instructions in home exercise relaxation/stretching program of 1 hour sessions 3 times per week for 3 months All patients: Education session about available diagnosis and treatment of FM	FIQ (0-100): 80.8 vs. 80.1 Pain VAS (0-10): 6.1 vs. 6.3	Pain VAS: 5.2 vs. 6.5, difference -1.3 (95% CI -2.6 to 0.03)	
Baptista, 2012[62] 4 months Pain duration NR *Fair*	A. Belly dance (n=40): One hour belly dance classes twice a week for 16 weeks (combination exercise) B. Waiting list control (n=40): dance offered at end of the study	A vs. B: Age: 50 vs. 49 years Female: 100% vs. 100% Race: NR FIQ (0-10): 5.9 vs. 6.3 Pain VAS (0-10): 7.7 vs. 7.5	A vs. B <u>4 months</u> FIQ: 4.3 vs. 5.9; difference -1.6 (95% CI -2.45 to -0.75) Pain VAS: 4.7 vs. 7.3; difference -2.6 (95% CI -3.61 to -1.59)	A vs. B <u>4 months</u> BDI (0-63): 23.1 vs. 23.5; difference -0.40 (95% CI -7.09 to 6.29) STAI part 1: 49.4 vs. 51.8; difference -2.40 (95% CI -6.87 to 2.07) STAI part 2: 49.8 vs. 54.1; difference -4.3 (95% CI -8.72 to 0.12) SF-36 function (0-100): 56.3 vs. 39.1; difference 17.2 (95% CI 7.55 to 26.85) SF-36 limitation due to physical aspects (0-100): 36.5 vs. 13.8; difference 22.7 (95% CI 9.06 to 36.34) SF-36 pain (0-100): 46.0 vs. 29.1; difference 16.9 (95% CI 7.62 to 26.18) SF-36 mental (0-100): 52.3 vs. 46.2; difference 6.1 (95% CI -3.89 to 16.09)
Buckelew, 1998[63] 3 and 24 months Duration of symptoms, 11 years	A. Combination exercise (n=30): included active range of motion exercises, strengthening exercises, low to moderate intensity aerobic exercise,	A vs. B Age: 46 vs. 44 years Female: 93% vs. 90% Duration of symptoms: 12 vs. 10 years Duration of diagnosis: 3.0 vs.	A vs. B <u>3 months</u>: AIMS physical activity subscale: median 4.0 vs. 6.0; median change from baseline 0 vs. 0 Pain VAS: median 5.4 vs. 5.8, median	A vs. B <u>3 months</u>: SCL-90-R Global Severity Index (0-90): median 65.5 vs. 65.0, median change from baseline -3 vs. 0 CES-D (0-60): median 13.5 vs. 13.0, median change from baseline -2.5

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Poor	proper posture and body mechanics instruction, and instructions on use of heat, cold, and massage; one 90 minute session per week for 1.5 months and instructions to train 2 additional times independently per week then 24 months of monthly one-hour groups. B. Attention control (n=30): one 90-180 minute education session weekly for 1.5 months	2.5 years AIMS physical activity subscale (0-10): median 4.0 vs. 6.0 Pain VAS (0-10): median 6.3 vs. 5.9	change from baseline -0.8 vs. -0.5 <u>24 months</u> AIMS physical activity subscale: median 4.0 vs. 6.0, median change from baseline 0 vs. 0 Pain VAS: median 5.5 vs. 5.4, median change from baseline -1.2 vs. -0.6	vs. 3 Sleep scale (0-12), median 8.0 vs. 5.0, median change from baseline 0 vs. 0 <u>24 months</u> SCL-90-R Global Severity Index: median 65.5 vs. 67.0, median change from baseline -2.5 vs. -1 CES-D: median 11.5 vs. 12.0, median change from baseline -3.5 vs. -2 Sleep scale: median 7.5 vs. 6.0, median change from baseline 0 vs. 0
Clarke-Jenssen,[64] 2014 3 and 12 months Symptom Duration, 14 years *Fair*	A. Aerobic exercise (n=44): conducted on land and in warm water provided in a warm climate; also stretching, relaxation, and education; provided in groups 5 days per week for 4 weeks B. Aerobic exercise (n=44): on land and in warm water provided in a cold climate; also stretching, relaxation, education, provided in groups 5 days per week for 4 weeks C. Usual Care (n=44): no intervention	A vs. B vs. C: Age: 46 vs. 46 vs. 45 years Female: 88% vs. 93% vs. 96% Symptom duration: 17 vs. 13 vs.12 years Pain VAS (mean, 0-10): 6.6 vs. 6.9 vs. 6.6	A vs. C, between-group difference in change from baseline: <u>3 months</u> FIQ: data NR, p=ns Pain VAS: -1.2 (95% CI -2.2 to -0.1) <u>12 months</u> FIQ data NR, p=ns Pain VAS: 0.1 (95% CI -0.9 to 1.1) B vs. C, between-group difference in change from baseline: <u>3 months</u> FIQ: data NR, p=ns Pain VAS: -0.9 (95% CI -1.9 to 0.2) <u>12 months</u> FIQ: data NR, p=ns Pain VAS: 0 (95% CI -1 to 1)	A vs. C, between-group difference in change from baseline: <u>3 months</u> HADS: data NR, p=ns SF-36 Physical: data NR, p=ns SF-36 Mental: data NR, p=ns <u>12 months</u> HADS: data NR, p=ns SF-36 Physical: data NR, p=ns SF-36 Mental: data NR, p=ns B vs. C, between-group difference in change from baseline: <u>3 months</u> HADS: data NR, p=ns SF-36 Physical: data NR, p=ns SF-36 Mental: data NR, p=ns <u>12 months</u> HADS: data NR, p=ns SF-36 Physical: data NR, p=ns SF-36 Mental: data NR, p=ns
Da Costa, 2005[65] 3 and 9 months	A. Combination Exercise (n=39): aerobic exercise, stretching, and	A vs. B Age, years: 49 vs. 52 Female: 100% vs.	A vs. B, mean change from baseline <u>3 months:</u> FIQ: -7.8 (95% CI -	A vs. B, mean change from baseline <u>3 months:</u> SCL 90-R GSI (30-81): -

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Symptom duration, years: 11 years *Fair*	strength exercises; 4 visits (initial 90 minutes, others 30 minutes) over 12 weeks with exercise physiologist; individualized home-based program. B. Usual care (n=41): subjects asked to record exercise activity weekly during the 12-week intervention phase and monthly thereafter.	100% Symptom duration: 10.5 vs. 11.2 years FIQ (0-100): 55.1 vs. 48.6 Upper body pain VAS (0-100): 49.5 vs. 47.4 Lower body pain VAS (0-100): 47.0 vs. 47.0	13.9 to -1.7) vs. -0.04 (95% CI -5.2 to 5.1), p = 0.05 Pain VAS, upper body: -10.6 (95% CI -17.8 to -3.4) vs. -1.9 (95% CI -6.9 to 3.2), p = 0.048 Pain VAS, lower body: -8.21 (95% CI -15.7 to -0.74) vs. -2.0 (95% CI -9.4 to 5.4), p=0.24 <u>9 months</u>: FIQ: -10.1 (95% CI -16.1 to -4.0) vs. -0.024 (95% CI -4.4 to 3.9), p = 0.009 Pain VAS, upper body: -7.9 (95% CI -14.3 to -1.4) vs. 2.4 (95% CI 3.7 to 8.5), p = 0.02 Pain VAS, lower body: -5.6 (95% CI -13.3 to 2.2) vs. -0.29 (95% CI -8.6 to 8.0), p = 0.35	0.02 (95% CI -0.3 to -0.04) vs. -0.07 (95% CI -0.2 to 0.05), p=0.26 <u>9 months</u>: SCL 90-R GSI (30-81): -0.16 (95% CI -0.28 to 0.35) vs. -0.09 (95% CI -0.21 to 0.03), p=0.39
Fontaine, 2011[67] 6 and 12 months Mean duration of fibromyalgia 7.4 years *Fair*	A. Aerobic Exercise (n=30): Lifestyle Physical Activity; 6, 60-minute group sessions over 3 months with the goal to increase moderate-intensity physical exercise by accumulating short bursts of physical activity throughout the day to 30 minutes 5-7 days per week. B. Attention control (n=23): FM education, monthly sessions for 3 months. Included education about FM and social support.	A vs. B Age: 46 vs. 49 years Female: 94% vs. 100% Race, white: 78% vs. 82% Years since diagnosis: 5.9 vs. 9.6 FIQ (scale NR): 67.5 vs. 69.7 Pain VAS (0-100): 54.6 vs. 58.9	A vs. B <u>6 months:</u> FIQ: 65.3 vs. 63.9, difference 1.4 (95% CI -10.0 to 12.8) Pain VAS: 54.9 vs. 49.4, difference 5.5 (95% CI -7.8 to 18.8) <u>12 months:</u> FIQ: 64.4 vs. 65.1, difference -0.7 (95% CI -13.6 to 12.2) Pain VAS: 51.6 vs. 50.9, difference 0.7 (95% CI -12.9 to 14.3)	A vs. B <u>6 months:</u> CES-D (scale NR): 18.1 vs. 19.9, difference -1.8 (95% CI -7.5 to 3.9) <u>12 months:</u> CES-D: 19.8 vs. 20.6, difference -0.8 (95% CI -7.1 to 5.5)
Giannotti, 2014[68] 1 and 6 months Pain duration NR *Poor*	A. Combination exercise (n=21): stretching, strengthening, active and passive mobilization, spine flexibility, and aerobic training plus education 2 days a week (60 minutes	A vs. B Age: 53 vs. 51 years Female: 95% vs. 92% FIQ (0-100): 62.7 vs. 59.1 Pain VAS (0-10): 6.1 vs. 6.1	A vs. B <u>1 month:</u> FIQ: 55.5 vs. 50.9, difference 4.6 (95% CI -6.38 to 15.58) Pain VAS: 5.3 vs. 5.5, difference -0.20 (95% CI -1.87 to 1.47) <u>6 months</u>	A vs. B <u>1 month</u> Sleep VAS (0-10): 4.6 vs. 5.0, difference -0.40 (95% CI -2.51 to 1.71) <u>6 months</u> Sleep VAS (0-10): 6.3 vs. 6.1, difference 0.20 (95% CI -2.15 to 2.55)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	per session) for 10 weeks; instructions to perform at home the exercise program at least 3 times per week. B. No intervention (n=20)		FIQ: 48.8 vs. 56.9, difference -8.1 (95% CI -20.33 to 4.13) Pain VAS: 5.8 vs. 5.4, difference 0.4 (95% CI -1.4 to 2.2)	
Gowans, 2001[69] 6 months Duration of symptoms, 9 years *Poor*	A. Aerobic exercise (n=30): 3 pool and walking exercise classes (plus stretching) per week for 6 months B. Control group (n=27): continued ad libitum activity	A vs. B Age: 45 vs. 50 years Female: 89% vs. 87% FIQ (0-80): 57.7 vs. 56.6	A vs. B <u>6 months:</u> FIQ: 48.6 vs. 54.9, p**<0.05; difference -6.3 (95% CI -14.8 to 2.2)	A vs. B <u>6 months:</u> BDI (0-63): 16.9 vs. 21.3, p**<0.05 difference -4.4 (95% CI -10.4 to 1.6), p = 0.15 STAI (20-80): 41.3 vs. 51.7, p**<0.05; difference -10.4 (95% CI -18.2 to -2.6), p=0.01
Gusi, 2006[70] 3 months Duration of symptoms, 22 years *Poor*	A. Combination exercise (n=18): 1-hour pool exercise (warmup, aerobic exercise, mobility and lower-limb strength exercises, cool down) 3 times per week for 12 weeks (subjects instructed to avoid physical exercise for the next 12 weeks) B. Control (n=17): Normal daily activities, which did not include any exercise related to those in the therapy.	A vs. B Age, years: 51 vs. 51 Female: 100% vs. 100% Pain VAS (0-100): 63.1 vs. 63.9	A vs. B Change from baseline <u>3 months</u> Pain VAS: -1.6 (95% CI -12.7 to 0.9) vs. 0.9 (95% CI -7.3 to 9.2), p=0.69	A vs. B Change from baseline <u>3 months</u> EQ-5D (0-1): 0.14 (95% CI -0.03 to 0.32) vs. -0.02 (-0.17 to 0.13), p=0.14 EQ-5D Pain/discomfort (1-3): -0.1 (95% CI -0.4 to 0.3) vs. 0 ((95% CI -0.3 to 0.3), p=0.79 EQ-5D Anxiety/depression (1-3): -0.5 ((95% CI -0.8 to -0.1) vs. 0 (95% CI -0.2 to 0.2), p=0.01
Kayo,[71] 2012 3 months Duration of symptoms, 5 years *Fair*	A. Aerobic exercise (n=30): Walking program, 60 minutes 3 times per week for 16 weeks, supervised by physical therapist. B. Muscle strengthening exercise (n=30): 60 minutes 3 times per week for 16 weeks, supervised by physical therapist.	A vs. B: Age: 48 vs. 47 vs. 46 years Symptom duration: 4.0 vs. 4.7 vs. 5.4 FIQ total (0-100): 63.1 vs. 67.3 vs. 63.8 Pain VAS (0-10): 8.6 vs. 8.7 vs. 8.4	A vs. C <u>3 months</u> FIQ: 38.5 vs. 57.7; overall group X time interaction p=ns Pain VAS: 4.8 vs. 6.7; overall group X time interaction p=ns B vs. C <u>3 months</u> FIQ: 50.5 vs. 57.7; overall group X time interaction p=ns Pain VAS: 5.9 vs. 6.7; overall group X time	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	C. No treatment (n=30)		interaction p=ns	
King, 2002[72] 3 months Duration of symptoms, 8.5 years *Poor*	A. Aerobic exercise (n=30): aerobic land and water activities; three, 10-40 minute supervised exercise sessions per week for 3 months B. Control (n=18): instructions on stretches and coping strategies and contacted 1-2 times during the 3 month treatment period to answer any questions	A vs. B Age: 45 vs. 47 years Female: 100% vs. 100% Duration of symptoms: 7.8 vs. 9.6 years FIQ (0-80): 52.4 vs. 55.2	A vs. B <u>3 months</u> FIQ: 47.5 vs. 51.5, difference -4.0 (95% CI -12.2 to 4.2)	NR
Mannerkorpi, 2009[73] 6-7 months Pain duration NR *Fair*	A. Aerobic exercise (n=81): One 45 minute pool aerobic exercise session per week for 20 weeks, stretching exercise also, plus six 1 hour weekly sessions of strategies to cope with FM symptoms, plan for physical activity for the following week and short relaxation exercise B. Education control (n=85): six 1 hour weekly sessions of strategies to cope with FM symptoms, plan for physical activity for the following week and short relaxation exercise	A vs. B Age: 45 vs. 47 years Female: 100% vs. 100% FIQ (0-100): 61.6 vs. 66.6 FIQ pain subscale (0-100): 67.7 vs. 70.4	A vs. B <u>6-7 months</u> FIQ: mean change from baseline: -3.9 vs. -4.5, p=0.04 FIQ pain: mean change from baseline: -6.5 vs. -2.5, p=0.018	A vs. B <u>6-7 months</u> HADS depression scale (0-21): mean change from baseline -0.4 vs. 0.0, p=0.99 HADS anxiety scale (0-21): mean change from baseline -0.7 vs. 0.4, p=0.15 SF-36 PCS (0-100): mean change from baseline 2.9 vs. 1.3, p=0.13 SF-36 MCS (0-100): mean change from baseline 0.5 vs. 1.3, p=0.15 SF-36 physical functioning (0-100): mean change from baseline 2.2 vs. 1.3, p=0.70 SF-36 role-physical (0-100): mean change from baseline 12.1 vs. 9.3, p = 0.72 SF-36 bodily pain (0-100): mean change from baseline 5.0 vs. 3.6, p = 0.24
Paolucci, 2015[74] Duration of symptoms: NR 3 months *Fair*	A. Combination exercise (n=19): Low-impact aerobic training, agility training balance and postural exercises, hip flexor strengthening, static stretching, diaphragmatic breathing, and	A vs. B Age: 50 vs. 48 years Female: 100% vs. 100% FIQ total (0-100): 64.8 vs. 63.9	A vs. B <u>3 months:</u> FIQ total: 53.8 vs. 64.3, difference -10.50 (95% CI -17.77, -3.23)	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	relaxation; 10, 60-minute sessions, twice a week for 5 weeks B. Control (n=18): No rehabilitation interventions, continued normal activities			
Sanudo, 2010[77] 6 months Pain duration NR *Fair*	A. Combination exercise (n=21): supervised aerobic, muscle strengthening, and flexibility exercises; twice-weekly sessions for 24 weeks B. Aerobic exercise (n=22): warm-up, aerobic exercise, cooldown; two, 45-60 minute sessions/week for 6 months C. Usual care control (n=21): medical treatment for FM and continued normal daily activities, which did not include aerobic exercise.	A vs. B vs. C Age: 56 vs. 56 vs. 57 years FIQ (0-100): 62.2 vs. 60.9 vs. 60.5	A vs. C <u>6 months</u> FIQ: mean change from baseline -8.8 vs. NR; p < 0.01 B vs. C <u>6 months</u> FIQ: mean change from baseline -8.8 vs. NR; p < 0.05	A vs. C <u>6 months</u> BDI (0-63): mean change from baseline -6.4 vs. NR; p < 0.01 SF-36 total (0-100): mean change from baseline 8.4 vs. NR; p < 0.01 B vs. C <u>6 months</u> BDI: -8.5 vs. NR; p < 0.01 SF-36 total: 8.9 vs. NR; p < 0.05
Sanudo,[76] 2012 6, 18 and 30 months Pain duration NR *Poor*	A. Combination exercise (n=21): Twice-weekly 45- to 60-minute sessions of exercise (warmup, aerobic exercise, muscle strengthening exercise, flexibility exercises) for 6 months. B. Usual care (n=20): alternated between 6 months of training and 6 months with no exercise intervention (asked not to participate in any structured	A vs. B Female: 100% vs. 100% FIQ (0-80): 58.6 vs. 55.6	A vs. B <u>6 months:</u> FIQ: 48.5 vs. 55.4, p<0.0005; difference -6.9 (95% CI -14.35 to 0.55), p =0.07 <u>18 months:</u> FIQ: 45.6 vs. 51.3, p=NR; difference -5.7 (95% CI -14.6 to 3.2), p=0.20 <u>30 months</u> FIQ: 38.5 vs. 49.5, p ns; difference -11.0 (95% CI -19.93 to -2.07), p =0.02	A vs. B <u>6 months:</u> SF-36 (0-100): 49.5 vs. 37.9, p=0.13; difference 4.68 (95% CI .096 to 21.104), p = 0.02 BDI (0-63): 14.7 vs. 16.6, p=0.18; difference -1.9 (95% CI -6.5 to 2.7), p = 0.41 <u>18 months:</u> SF-36: 51.8 vs. 41.3, p= NR; difference 10.5 (95% CI 0.5 to 20.5), p=0.04 BDI: 14.3 vs. 14.2, p=NR; difference 0.10 (95% CI -5.4 to 5.6), p=0.97 <u>30 months</u> SF-36: 60.5 vs. 42.0, p=ns

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	exercise program) for 30 months.			BDI: 9.7 vs. 17.9, p=ns
Sanudo, 2015[75] 6 months Pain duration NR *Poor*	A. Aerobic exercise (n=16): consisted of warmup, steady state exercise at 60-65% of predicted maximum heart rate, interval training at 75-80% of predicted maximum heart rate, and cool-down; 2, 45-60 minute sessions per week for 6 months B. Usual care (n=16): normal activities, which did not include structured exercise.	A vs. B Age: 55 vs. 58 years Female: 100% vs. 100% Pain VAS (0-10): 7.4 vs. 7.2	A vs. B <u>6 months:</u> Pain VAS: 6.7 vs. 7.0, difference -0.3 (95% CI -6.3 to 5.7),	A vs. B <u>6 months</u> Anxiety VAS (0-10): 5.7 vs. 7.5, difference -1.8 (95% CI -10.8 to 7.2) Depression VAS (0-10): 5.6 vs. 6.7 (2.2), difference -1.1 (95% CI -10.1 to 7.9) Sleep disturbance VAS (0-10): 7.2 vs. 8.6 (1.9), difference -1.4 (95% CI -8.9 to 6.1)
Tomas-Carus, 2008,[78] 2009[79] 8 months Symptom duration 20 years *Poor*	A. Combination exercise (n=17): Pool exercise in 1 hour sessions 3 times per week for 8 months (warmup, aerobic exercise, mobility and lower limb strength exercises using water resistance and upper limb strength exercises without water resistance, cooldown) B. Control (n=16): normal activities for 8 months, which did not include exercise similar to that in group A.	A vs. B Age: 51 vs. 51 years Female: 100% vs. 100% FIQ Total (0-10): 6.1 vs. 6.3 FIQ Physical Function (0-10): 3.0 vs. 3.7 FIQ Pain (0-10): 5.6 vs. 6.4	A vs. B <u>8 months</u> FIQ Total: 5.2 vs. 6.5, difference -1.3 (95% CI -0.23 to -0.3) FIQ Physical Function: 2.4 vs. 3.7, difference -1.3 (95% CI -2.7 to 0.09) FIQ Pain: 5.3 vs. 6.6, difference -1.3 (95% CI -2.5 to -0.09)	A vs. B <u>8 months</u> FIQ Anxiety (0-10): 4.7 vs. 6.6, difference -1.9 (95% CI -3.7 to -0.1) FIQ Depression (0-10): 4.0 vs. 6.1, difference -2.1 (95% CI -4.1 to -0.1) STAI State Anxiety (20-80): 37.5 vs. 44.4, difference -6.9 (95% CI -13.2 to -0.6) SF-36 physical function (0-100): 54.1 vs. 36.6, difference 17.5 (95% CI 3.4 to 31.6) SF-36 bodily pain (0-100): 51.7 vs. 27.1, difference 24.6 (95% CI 11.6 to 37.6) SF-36 Mental Health (0-100): 67.3 vs. 49, difference 18.3 (95% CI 2.5 to 34.0)
van Eijk-Hustings,[80] 2013 18 months Pain duration NR *Fair*	A. Aerobic exercise (n=47): two group sessions per week for 12 weeks (warmup, aerobic exercise, resistance training to strengthen muscles, cooldown). Subjects were asked to practice exercises at home with videodisc once a	A vs. B Age: 44 vs. 43 years Female: 100% vs. 98% FIQ total (0-100): 60.0 vs. 55.4 FIQ physical function (0-10): 3.6 vs. 3.4 FIQ Pain (0-10): 6.2 vs. 5.5	A vs. B <u>18 months:</u> FIQ total: 52.0 vs. 56.2, ES = 0.22 (95% CI -0.20 to 0.61) FIQ physical function: 3.6 vs. 3.9, ES = 0.11 (95% CI -0.29 to 0.52) FIQ pain: 5.2 vs. 5.3, ES = 0.05 (95% CI -0.36 to 0.44)	A vs. B <u>18 months:</u> FIQ Depression (0-10): 5.0 vs. 4.2, ES = 0.09 (95% CI -0.31 to 0.49) FIQ Anxiety (0-10): 5.0 vs. 4.8, ES = -0.06 (95% CI -0.46 to 0.34) EQ-5D (-0.59 to 1): 0.54 vs. 0.51, ES = 0.10 (95% CI -0.31 to 0.50) GP consultations[b]: 1.0 vs. 0.7, ES = -0.10 (95% CI -

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	week. B. Usual care (n=48): individualized FM education and lifestyle advice within 1-2 consultations, plus care as usual			0.48 to 0.32) Medical specialist consultations[b]: -0.4 vs. 0.2, ES = -0.29 (95% CI -0.58 to 0.22) Physiotherapist consultations[b]: 0.4 vs. 2.8, ES = -0.29 (-0.58 to 0.22) Other paramedical professional consultations[b]: 2.1 vs. 0.2, ES = -0.68 (95% CI -1.00 to -0.18)
van Santen, 2002[81] 6 months Duration of symptoms: 12 years *Poor*	A. Combination exercise (n=58): group sessions (60 minutes) twice a week for 24 weeks (aerobic exercises, stretching, general flexibility and balance exercises, and isometric muscle strengthening); encouraged to attend a third, unsupervised, 60 minute session weekly and to use sauna or swimming pool after all sessions. B. Usual care (n=29): analgesics NSAIDs, or tricyclic antidepressants, if appropriate; GPs informed that aerobic exercises and relaxation should not be prescribed or encouraged	A vs. B Age: 46 vs. 43 years Female: 100% vs. 100% Duration of symptoms: 9.7 vs. 15.4 years SIP physical score (mean, 0-100): 11.3 vs. 9.8 SIP total score (mean, 0-100): 14.4 vs. 11.4 AIMS (mean, 0-10): 1.9 vs. 5.4 Pain VAS (mean, 0-100): 66.8 vs. 62.4	A vs. B, mean change from baseline <u>6 months:</u> SIP physical score: -1.7 (95% CI -3.7 to 0.3) vs. -0.6 (95% CI -2.9 to 1.7), p=ns SIP total score: -1.9 (95% CI -3.9 to 0.1) vs. -1.4 (95% CI -3.4 to 0.6) p=ns AIMS: 0.1 (95% CI -0.6 to 0.8) vs. 0.8 (95% CI -1.8 to -0.2), p=ns Pain VAS: -5.5 (95% CI -10.9 to -0.1) vs. 1.3 (95% CI -4.5 to 7.1), p=ns	A vs. B, mean change from baseline <u>6 months:</u> SCL-90-R Global Severity Index (scale unclear): -6.8 (95% CI -20.1 to 6.5) vs. -8.1 (95% CI -19.8 to 3.6), p=ns SIP psychosocial score (0-100): -3.2 (95% CI -6.2 to 0.2) vs. -3.5 (95% CI -7.0 to 0.0), p=ns Patient global assessment (1-5): 0.5 (95% CI 0.2 to 0.8) vs. 0.5 (95% CI 0.2 to 0.8), p=ns
Wigers, 1996[82] 48 months Duration of symptoms: 10 years *Fair*	A. Aerobic exercise (n=20): sessions consisted of training to music (further details not given) and aerobic games; 45 minute group sessions 3 times a week for 14 weeks B. Treatment as usual (n=20)	A vs. B Age: 43 vs. 46 years Female: 90% vs. 95% Duration of symptoms: 9 vs. 11 years Pain VAS (0-100): 72 vs. 65	A vs. B <u>48 months:</u> Pain VAS: 68 vs. 69, difference -1.0 (95% CI -16.3 to 14.4)	A vs. B <u>48 months</u> Depression VAS (0-100): 32 vs. 30, difference 2.0 (95% CI -18.8 to 22.8) Global subjective improvement: 75% vs. 12%, RR 5.9 (95% CI 1.5 to 22.2)

AIMS = Arthritis Impact Measurement Scale; BDI = Beck Depression Inventory; CI = confidence interval; EQ5D = EuroQoL 5 Dimensions; ES = effect size; FIQ = Fibromyalgia Impact Questionnaire; FM = fibromyalgia; GP = general practitioner; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; SCL-90-R = Symptom Checklist-90-Revised; SF-36 = Short-Form 36 Questionnaire; SIP = Sickness Impact Profile; STAI = State-Trait Anxiety Inventory; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

[b] Total number of consultations over a period of 2 months prior to measurement.

Exercise Compared With Usual Care, Waitlist, an Attention Control or No Treatment

Functional outcomes. Exercise was associated with a slightly greater effect on short-term function than usual care, an attention control, or no treatment based on Fibromyalgia Impact Questionnaire (FIQ) total scores (7 trials, pooled difference -7.61 on a 0 to 100 scale, 95% CI, -12.78 to -2.43, I^2= 60%)[61,62,65,68,71,72,74] (Figure 39). The estimate across fair-quality trials (i.e., not including the poor-quality trials) was somewhat higher (5 trials, pooled difference -9.91, 95% CI -15.75 to -4.07).[61,62,65,71,74]

Exercise was associated with slightly greater effects on intermediate-term function than controls for FIQ total score (8 trials, pooled difference on 0-100 scale, -6.04, 95% CI -9.05 to -3.03.10, I^2= 0%)[65,67-69,73,76-78] (Figure 39). Estimates were slightly smaller across the fair-quality trials only (4 trials, pooled difference -4.04, 95%CI -7.90 to -0.03).[65,67,73,77] Stratification by exercise type yielded similar results for combination exercise (7 trials, pooled difference -5.75, 95% CI -9.29 to -2.54),[65,67,68,73,76-78] but there was no clear difference between aerobic exercise and no treatment or usual care (2 trials, pooled difference -8.13, 95% CI -16.24 to 0.28).[69,77] Estimates were consistent with a slightly greater effect of exercise on function when compared with usual care (3 trials, pooled difference -6.13, 95% CI -11.71 to -1.06)[65,76,77] or no treatment (3 poor quality trials, pooled difference -9.97, 95% CI -16.24 to -3.45),[68,69,78] but there was no clear difference in trials using attention controls (2 fair quality trials, pooled difference -3.25, 95% CI -99.32 to 5.20).[67,73]

Exercise had a slightly greater effect on long-term function than controls but results were no longer statistically significant based on the FIQ total score (3 trials, pooled difference on 0 to 100 scale, -4.33, 95% CI -10.18 to 1.52, I^2= 0%)[67,76,80] (Figure 39). There were no clear differences in estimates when analyses were stratified according to the type of exercise (2 trials of combination exercise pooled difference -4.45, 95%IC -14.39 to 6.24),[67,76] usual care (2 trials, pooled difference -5.34, 95% CI -13.4 to 2.32),[76,80] or for the exclusion of one poor-quality trial (pooled difference -3.11, 95% CI -11.26 to 5.86).[67,80] Findings are based on a small number of trials.

Pain outcomes. Exercise had a moderately greater effect on VAS pain (0 to 10 scale) pain in the short term compared with usual care, attention control, or no treatment short term (7 trials, pooled difference -1.07, 95% CI -1.73 to -0.41, I^2=58.7%)[61-63,65,68,70,71] (Figure 40). Substantial heterogeneity was noted with one outlier trial of belly dance (combination exercise) versus waitlist control reporting substantially higher estimates.[62] Excluding the outlier trial reduced heterogeneity and led to an effect size consistent with a small effect (6 trials, -0.89, 95% CI -1.32 to -0.46, I^2=0%). Estimates were similar when stratified by exercise type and control type. Across the fair-quality trials, the estimate was somewhat larger (4 trials, pooled difference -1.44, 95% CI -2.4 to -0.49, including the outlier).[61,62,65,67,71]

There was no effect of exercise on VAS pain at intermediate term (6 trials, pooled difference -0.31, 0-10 scale, 95% CI -0.79 to 0.17, I^2= 5.4%)[65,67,68,75,78,81] (Figure 40). Removal of poor

quality trials[68,75,78] and stratification by exercise and control types yielded similar estimates (differences ranged from -0.10 to -0.71) with no clear difference identified.

There was no effect of exercise on pain long term (4 trials, pooled difference -0.18, 95% CI -0.77 to 0.42, I^2=0%)[63,67,80,82] (Figure 40). Similar estimates were obtained and no clear differences were seen following exclusion of one poor quality-trial or for the comparisons of aerobic exercise with usual care or combination exercise with attention control; pooled differences ranged from -0.5 to -0.26.

Other outcomes. Data on the effects of exercise on anxiety, depression, and quality of life were often poorly reported (Table 35). Exercise had no clear effect in the short term on measures of mental health, depression, anxiety, psychological distress, or sleep disturbance VAS across five trials,[61-65] with only one small poor-quality trial reporting a positive effect favoring exercise on the EQ-5D anxiety/depression scale (p=0.01).[70] Similarly, exercise had no clear effect on quality of life with only one fair-quality trial reporting improved SF-36 scores.[62]

Across three trials (two of poor quality),[69,76,77] exercise had no clear effect on depression measured by the Beck Depression Inventory at intermediate term (3 trials, pooled mean difference -5.05 on a 0-63 scale, 95% CI -10.13 to 0.020, I2= 52.5%, plot not shown) compared with no treatment or usual care; individually, one trial reached statistical significance.[77] Across various measures, exercise had no clear effect on depression in five trials,[63,64,67,73,75] anxiety in two trials,[73,75] psychological problems in two trials,[63,65] or sleep in 3 trials.[63,68,75] One small poor-quality trial reported that exercise had a greater effect on anxiety (FIQ anxiety and State-Trait Anxiety Inventory [STAI] anxiety scales) and depression (FIQ depression) compared with usual care.[78] An additional poor-quality trial reported a slightly greater effect with exercise on the STAI compared with usual care.[69] Exercise was associated with a positive impact on quality of life based on various components of SF36 in three small trials,[76,77,79] but not in a forth larger fair-quality trial[73] (Table 35).

Long term, exercise had no clear effect on measures of depression, anxiety, or psychological problems in all but one poor-quality trial that reported a small effect of exercise versus usual care on depression based on the Beck Depression Inventory.[76] This same trial also reported improvement in SF-36 total scores, whereas one larger fair-quality trial did not.[64] Additionally, there were no differences between groups in health care utilization in the 2 months prior to the final measurement at 18 months[80] (Table 35).

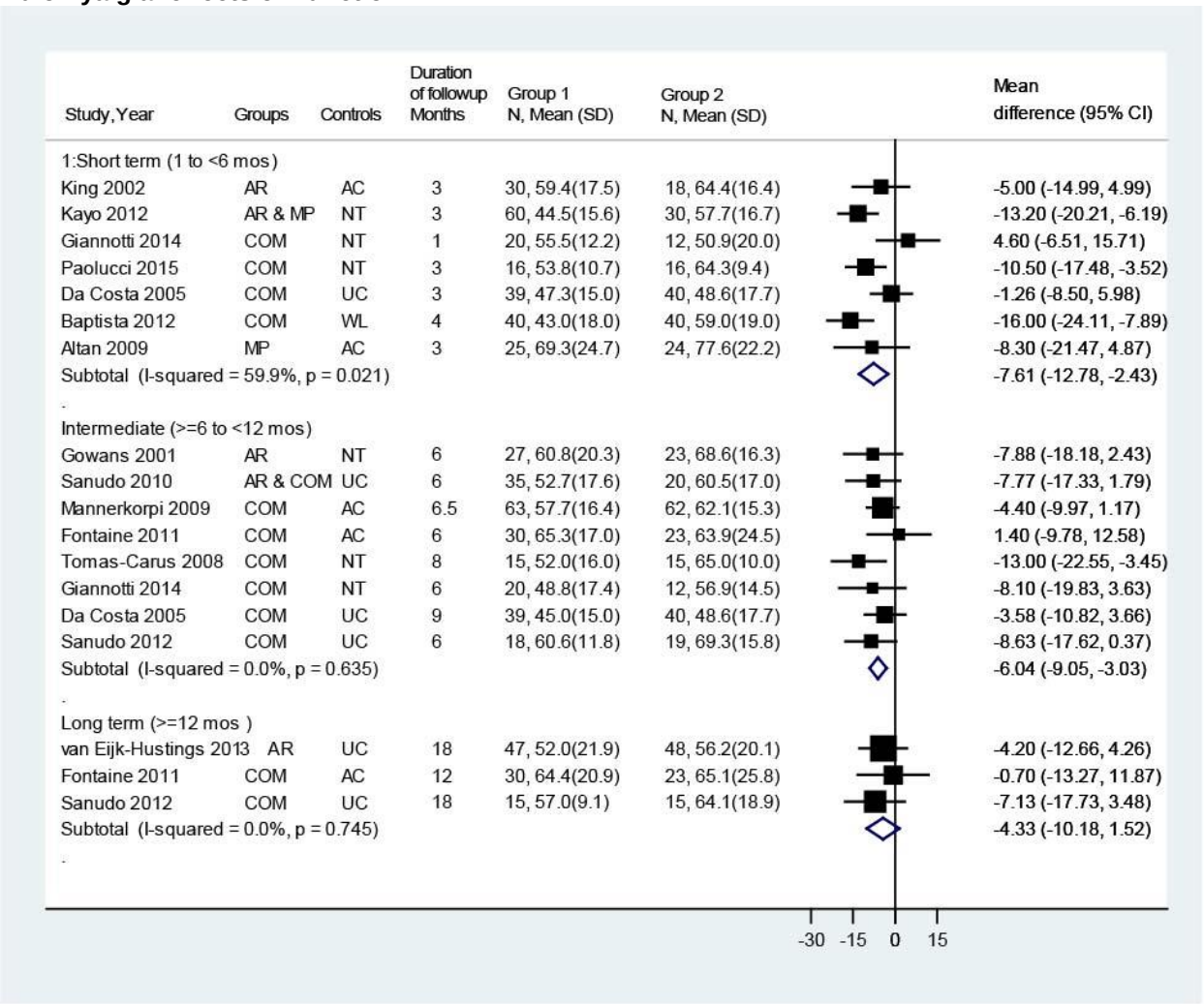
Exercise Compared With Pharmacological Therapy or With Other Nonpharmacological Therapies

No trial of exercise versus pharmacological therapy met inclusion criteria. Findings for exercise versus other nonpharmacological therapies are addressed in the sections for other nonpharmacological therapies.

Harms

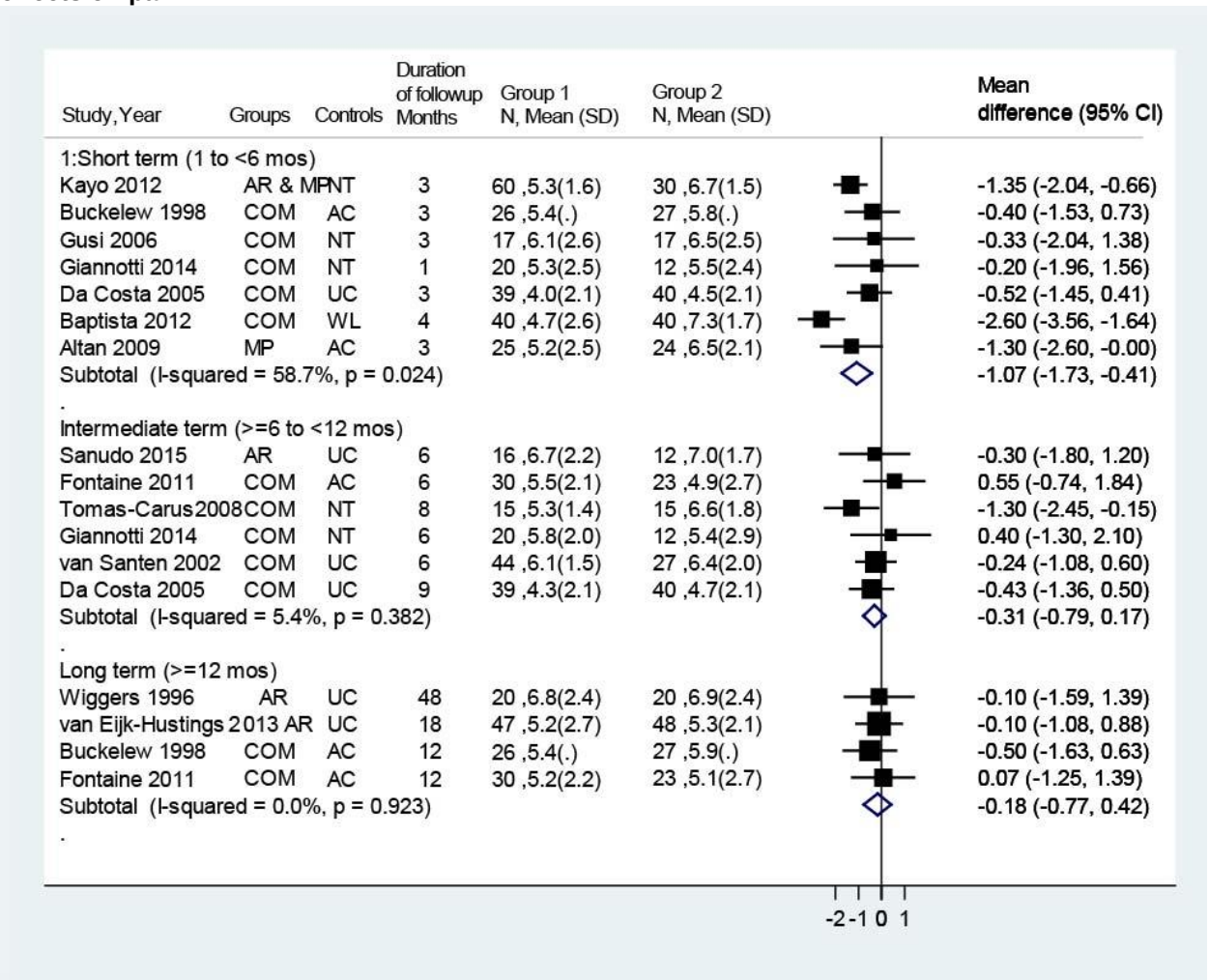
Most trials of exercise did not report on adverse events at all. One trial reported one non-study-related adverse event.[70] Two trials reported no adverse events.[71,74]

Figure 39. Exercise versus usual care, no treatment, waitlist, or an attention control for fibromyalgia: effects on function



AC = attention control; AR = aerobic exercise; AR & COM = aerobic exercise in one arm and combination exercise in another arm; AR & MP = aerobic exercise in one arm and muscle performance exercise in another arm; CI = confidence interval; COM = combination exercise therapy; MP = muscle performance exercise; MP+NR = muscle performance plus neuromuscular rehabilitation exercise; NT = no treatment; SD = standard deviation; UC = usual care; WL = waitlist.

Figure 40. Exercise versus usual care, no treatment, waitlist, or attention control for fibromyalgia: effects on pain



AC = attention control; AR = aerobic exercise; AR & MP = aerobic exercise in one arm and muscle performance exercise in another arm; CI = confidence interval; COM = combination exercise therapy; MP = muscle performance exercise; MP+NR = muscle performance plus neuromuscular rehabilitation exercise; NT = no treatment; SD = standard deviation; UC = usual care; WL = waitlist.

Psychological Therapies for Fibromyalgia

Key Points

- CBT was associated with a slightly greater effect on FIQ Total Score than usual care or waitlist in the short term (2 trials, pooled mean difference -10.67, 95%CI -17 to -4.30 I^2 = 0%, 0-100 scale). The pooled estimate at intermediate term was not statistically significant due to heterogeneity. However, individual trials showed a greater effect than usual care, and a third trial using the 0 to 10 FIQ Physical Impairment Scale showed a greater effect of CBT than attention control (mean difference -1.8, 95% CI -2.9 to -0.70) (SOE: Low for short and intermediate-term).

- Data were insufficient to determine the effects of psychological therapies on function for the following: EMG biofeedback and guided imagery compared with attention controls in the short term (2 poor-quality trials); EMG biofeedback compared with usual care in the intermediate term (1 poor-quality trial); and CBT and EMG biofeedback compared with usual care or attention controls over the long term (3 poor quality trials) (SOE: Insufficient for all time points).
- Psychological therapies were associated with a slightly greater improvement in pain compared with usual care, waitlist, or an attention control in the short term (4 trials, pooled mean difference -0.74, 95%CI -1.20 to -0.28, $I^2 = 0\%$) and intermediate term (3 trials, pooled mean difference -0.67, 95% CI -1.21 to -0.31, $I^2 = 36.7\%$); evidence from two poor-quality trials was insufficient to determine effects on pain long term. (SOE: Low for short-term and intermediate-term, Insufficient for long-term).
- Data were insufficient across two trials (one fair-quality and one poor-quality) to determine the effects of psychological therapies (CBT, EEG biofeedback) versus pharmacological treatments (amitriptyline, escitalopram, pregabalin) on function or pain over the short term (SOE: Insufficient).
- CBT was associated with a small benefit for function (difference -4.0 on the 0-100 FIQ, 95% CI -7.7 to -0.27), but not for pain (difference 0.2 on a 0-100 VAS, 95% CI -4.0 to 4.4), compared with pregabalin over the intermediate term in one fair-quality trial (SOE: Low). Long-term data was not reported.
- There was insufficient evidence from one small poor-quality trial to determine the effects of psychological therapies (biofeedback) versus exercise on function in the short term, across two poor-quality trials (CBT, biofeedback) at intermediate term and across three poor-quality trials (CBT, biofeedback, relaxation training) at long term (SOE: Insufficient for all time points).
- There was insufficient evidence from one small poor-quality trial to determine the effects of psychological therapies (biofeedback) versus exercise on pain in the short term across two poor-quality trials (CBT, biofeedback) at intermediate term and across four poor-quality trials (CBT, biofeedback, relaxation training at long-term (SOE: Insufficient).
- Data on harms were insufficient. Adverse events were poorly reported across the five poor -quality trials but were overall minor and occurred at similar frequencies between groups. In one trial, however, fewer patients who received stress management (4.8%) compared with usual care (50%) withdrew from the trial, citing increased depression and worsening of symptoms, respectively (SOE: Insufficient).

Detailed Synthesis

A total of 14 trials of psychological therapy for fibromyalgia met our inclusion criteria: nine trials (across 10 publications) featured a CBT component,[82,91-95,97,98,101,106] three trials included biofeedback (electromyography or electroencephalography),[63,81,102] and one trial each included relaxation training[105] or guided imagery[96] (Table 36 and Appendix D). The various psychological interventions were compared with usual care, waitlist control or attention control groups (10 trials),[63,81,82,91-98] pharmacological therapy (3 trials),[91,101,102] or exercise therapy (5 trials).[63,81,82,105,106]

The majority of subjects in all the trials were female (range 90 to 100%) and mean ages ranged from 32 to 52 years. Sample sizes ranged between 32 and 169 subjects (total

sample=1,167). Therapy duration and frequency in CBT trials ranged from 6 weekly sessions to 20 sessions over 14 weeks. CBT was delivered in groups in eight studies[91,93-95,97,98,101,106] and by telephone[92] in another. All CBT studies except two were of CBT as traditionally delivered for the treatment of pain problems. The two exceptions were a study of Acceptance and Commitment Therapy (ACT), and a Stress Management Therapy (SMT) intervention that spent equal time between presentations on stress mechanisms and training on pain coping and relaxation strategies; however, the interventions were similar enough to standard CBT to be included in this analysis. Session lengths ranged from 30 minutes up to 120 minutes. In the five trials of biofeedback and associated interventions, therapy duration ranged from 4 to 16 weeks and was delivered individually in the three biofeedback trials and in groups for the remaining two trials. The frequency ranged from one to five times per week with sessions as short as 25 minutes and as long as 3 hours. Short-term outcomes (<6 months) were reported by three CBT studies[92,94,97,101] and three biofeedback trials.[63,96,102] Intermediate outcomes (6 to <12 months) were reported by four CBT trials[91,93,95,106] and one trial of biofeedback.[81] Long-term outcomes (≥12 months) were reported by four CBT trials[82,95,98,106] and two biofeedback trials.[63,105] Studies were conducted in Brazil, the Netherlands, Norway, Spain, Sweden, Turkey, and the United States.

Of the nine CBT trials, three were considered fair-quality,[91,94,97,101] while the remaining six were rated poor-quality[82,92,93,95,98,106] (Appendix E). Among the remaining trials of biofeedback, relaxation, and guided imagery interventions, all were rated poor quality.[63,81,96,102,105] Methodological shortcomings included lack of blinding in fair-quality and poor-quality trials, and unclear allocation concealment methods, poor compliance, and high attrition in the poor-quality trials. The nature of the intervention types precluded blinding of participants in all trials.

Table 36. Summary of results for fibromyalgia: psychological therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Alda,[91] 2011 6 months Years since diagnosis: 12.9 vs. 11.2 vs.11.7 *Fair*	A. CBT (n=57): 10-12 week program; 10 weekly 90-minute group sessions.of cognitive restructuring and training in cognitive and behavioral coping strategies. B. Recommended pharmacological treatment (n=56): pregabalin (300-600 mg/day); duloxetine (60-120 mg/day) for patients with major depressive disorder. C. Usual care (n=56): standard care offered by general practitioners at subjects' health centers who received a guide for the treatment of FM in	A vs. B vs. C Age: 46 vs. 47 years vs. 47 years Females: 95% vs. 93% vs. 96% Race NR FIQ (mean, 0-100): 65.9 vs. 66.4 vs. 64.5 Pain VAS (mean, 0-100): 64.2 vs. 68.1 vs. 64.7	A vs. B <u>6 months</u>: FIQ: 48.8 vs. 52.8; MD -4.0 (95% CI -7.730 to -0.270) Pain VAS: 40.7 vs. 40.5; MD 0.2 (95% CI -3.996 to 4.396) A vs. C <u>6 months:</u> FIQ: 48.8 vs. 53.3, MD -4.5 (95% CI -7.91 to -1.09) Pain VAS: 40.70 vs. 44.3, MD -3.6 (95% CI -7.617 to 0.417)	A vs. B <u>6 months</u>: HAM-D (0-50): 7.9 vs. 8.2; MD -0.3 (95% CI -1.226 to 0.626) HAM-A (0-50): 7.3 vs. 7.4; MD -0.1 (95% CI -1.247 to 1.047) A vs. C <u>6 months</u>: HAM-D: 7.9 vs. 8.6, MD -0.7 (95% CI -1.719 to 0.319) HAM-A: 7.3 vs. 7.6, MD -0.3 (95% CI -1.361 to 0.761)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	primary care.			
Ang, 2010[92] 1.5 months Duration of fibromyalgia, years: 11.8 vs. 12.3 *Poor*	A. CBT (n=17): 6 weekly 30-40 minute sessions of telephone-delivered CBT (activity pacing, pleasant activity scheduling, relaxation, automatic thoughts and pain, cognitive restructuring, and stress management) B. Usual care (n=15): customary care from subject's treating physician	A vs. B Age: 51 vs. 47 years Female: 100% vs. 100% White: 81% vs. 80% FIQ total (mean, 0-100): 62.2 vs. 67.8 FIQ Physical Impairment (PI) (0-10): 5.6 vs. 5.4 FIQ Pain (0-10): 7.6 vs. 7.8	A vs. B <u>1.5 months:</u> Proportion of patients with clinically meaningful improvement from baseline FIQ total (14%): 33% vs. 15%, RR 2.2 (95% CI 0.5 to 9.3) mean change from baseline: FIQ PI: -0.6 vs. 0.5, adjusted p=0.13; FIQ Pain: -0.6 (1.6) vs. -0.3 (1.7), adjusted p=0.60;	A vs. B <u>1.5 months:</u> PHQ-8 (0-24): mean change from baseline -0.9 (5.2) vs. 0.0 (4.1), adjusted p=0.80; overall effect size = 0.60
Buckelew 1998 3, 12, and 24 months Duration of symptoms, years: 11.6 vs. 10.0 vs. 11.6 *Poor*	A. Electromyographic biofeedback and relaxation training (n=29): 1 session for 1.5-3 hours per week for 6 weeks and instructions to train 2 additional times independently per week. Subjects were taught cognitive and muscular relaxation strategies. 6-week individual training was followed by 2-year group maintenance phase of one-hour groups once per month. B. Attention control (n=30): 1 session for 1.5-3 hours per week for 6 weeks. Subjects received educational information on diagnosis and treatment of FM and general health topics information. This was followed by one hour groups once per month for 2 years. C. Combination Exercise (n=30): 1 session for 1.5 hours per week for 6 weeks and instructions to train 2 additional times independently per week.	A vs. B vs. C Age: 44 vs. 44 vs. 46 years Female: 97% vs. 90% vs. 93% Race NR AIMS physical activity subscale (median, 0-10): 6.0 vs. 6.0 vs. 4.0 Pain VAS (median, 0-10): 5.8 vs. 5.9 vs. 6.3	A vs. B <u>3-months:</u> AIMS physical activity subscale, median (median change from baseline): 6.0 (0) vs. 6.0 (0), ns Pain VAS, median (median change from baseline): 5.2 (-0.2) vs. 5.8 (-0.5), ns <u>24-months:</u> AIMS physical activity subscale, median (median change from baseline): 6.0 (0) vs. 6.0 (0), ns Pain VAS, median (median change from baseline): 5.2 (-1.1) vs. 5.4 (-0.6), ns A vs. C <u>3 months:</u> AIMS physical activity subscale, median (median change from baseline): 6.0 (0) vs. 4.0 (0), p≤0.05 Pain VAS, median (median change from baseline): 5.2	A vs. B <u>3-months:</u> SCL-90-R Global Severity Index, median (median change from baseline): 65.0 (-2) vs. 65.0 (0), ns CES-D, median (median change from baseline): 10.0 (-2) vs. 13.0 (3), ns Sleep scale, median (median change from baseline): 7.0 (0) vs. 5.0 (0), ns <u>24-months:</u> SCL-90-R Global Severity Index, median (median change from baseline): 64.0 (-1) vs. 67.0 (-1), ns CES-D, median (median change from baseline): 10.0 (-2) vs. 12.0 (-2), ns Sleep scale, median (median change from baseline): 6.0 (-2) vs. 6.0 (0), ns A vs. C <u>3 months:</u> SCL-90-R Global Severity Index, median (median

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	Sessions consisted of active range of motion exercises, strengthening exercises, low to moderate intensity aerobic exercise, proper posture and body mechanics instruction, and instructions on the use of heat, cold, and massage. 6-week individual training was followed by 2-year group maintenance phase of one-hour groups once per month.		(-0.2) vs. 5.4 (-0.8), ns <u>24 months:</u> AIMS physical activity subscale, median (median change from baseline): 6.0 (0) vs. 4.0 (0), p≤0.05 Pain VAS, median (median change from baseline): 5.2 (-1.1) vs. 5.5 (-1.2), ns	change from baseline): 65.0 (-2) vs. 65.5 (-3), ns CES-D, median (median change from baseline): 10.0 (-2) vs. 13.5 (-2.5), ns Sleep scale, median (median change from baseline): 7.0 (0) vs. 8.0 (0), ns <u>24 months:</u> SCL-90-R Global Severity Index, median (median change from baseline): 64.0 (-1) vs.. 65.5 (-2.5), ns CES-D, median (median change from baseline): 10.0 (-2) vs. 11.5 (-3.5), ns Sleep scale, median (median change from baseline): 6.0 (-2) vs. 7.5 (0), ns
Castel,[93] 2012 3 and 6 months A vs. B Pain duration, years: 13.6 vs. 11.6 *Poor*	A. CBT plus usual pharmacological care (n=34): CBT conducted in groups (except for one individual session); 14 weekly 2 hour sessions. CBT included education about FM and pain, autogenic training, cognitive restructuring, CBT for insomnia, assertiveness training, activity pacing, pleasant activity scheduling, goal setting, and relapse prevention. B. Usual care (n=30): usual pharmacological care, including analgesics, antidepressants, anticonvulsants, and myorelaxants	A vs. B Age: 50 vs. 49 years Female: 94% vs. 100% White: 100% vs. 100% FIQ (scale NR): 62.7 vs. 66.1 Pain NRS (0-10): 6.1 vs. 6.9	A vs. B <u>3 months:</u> Proportion of patients with MCSD (14% improvement from baseline): FIQ: 55.9% vs. 20%; OR 5.1 (95% CI 1.7 to 15.6); RR 2.8 (95% CI 1.3 to 6.1) Pain: 14.6% vs. 10%; RR 1.5 (95% CI 0.4 to 5.7) FIQ: 52.8 vs. 66.3; MD -13.5 (95% CI -15.5 to -11.5) Pain NRS: 5.9 vs. 6.8; MD -0.9 (95% CI -1.1 to -0.7) <u>6 months:</u> Proportion of patients with MCSD: FIQ: 58.8% vs. 20%; OR 5.7 (95% CI 1.9 to 17.8); RR	A vs. B <u>3 months:</u> HADS (scale NR): 15.4 (1.3) vs. 22.3 (1.4); MD -6.9 (95% CI -7.685 to -6.115) MOS Sleep quantity (scale NR): 6.9 (0.2) vs. 5.5 (0.3); MD 1.4 (95% CI 1.254 to 1.546), p <0.0001 MOS Sleep index problems (scale NR): 40.1 (1.6) vs. 28.8 (1.7); MD 11.3 (95% CI 10.340 to 12.260) <u>6 months:</u> HADS: 15.7 (1.3) vs. 23.7 (1.4); MD -8.0 (95% CI -8.785 to -7.215) MOS Sleep quantity: 6.7 (0.2) vs. 5.6 (0.3); MD 1.1 (95% CI 0.954 to 1.25) MOS Sleep index problems: 39.9 (1.5)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			2.9 (95% CI 1.4 to 6.3) Pain: 17.6% vs. 13.3%; RR 1.3 (95% CI 0.4 to 4.2) FIQ: 50.5 vs. 68.5; MD -18.0 (95% CI -20.095 to -15.905) Pain NRS: 5.7 vs. 6.8; MD -1.1 (95% CI -1.333 to -0.867)	vs. 28.0 (1.6); MD 11.9 (95% CI 10.998 to 12.802)
Falcão, 2008[101] 3 months Disease duration, years: 3.5 vs. 3.7 *Fair*	A. CBT plus Amitriptyline (n=30): amitriptyline 12.5/mg per day during first week, then increase dose to 25 mg/day. Those with intolerance or side effects to amitriptyline were given cyclobenzaprine 5 mg/day in the first week and then 10 mg/day. Routine medical visits once a week for 10 weeks for brief discussions with the doctors. Immediately after each visit, they had a group CBT session, consisting of progressive relaxation training with electromyographic biofeedback, cognitive restructuring, and stress management. B. Amitriptyline only (control) (n=30): amitriptyline 12.5/mg per day during first week, then increase dose to 25 mg/day. Those with intolerance or side effects to amitriptyline were given cyclobenzaprine 5 mg/day in the first week and then 10 mg/day. Routine medical visits once a week for 10 weeks for brief discussions with the doctors.	A vs. B Age: 45 vs. 46 years Female: 100% vs. 100% Caucasian: 80% vs. 77% FIQ (0-100): 64.9 vs. 69.6 Pain VAS (0-10): 6.9 vs. 7.0	A vs. B <u>3 months</u>: FIQ: 38.7 vs. 42.8; MD -4.1 (95% CI -18.765 to 10.565) Pain VAS: 4.4 vs. 5.1; MD -0.7 (95% CI -2.841 to 1.441)	A vs. B <u>3 months</u>: BDI (0-63): 10.6 vs. 15.6; MD -5.0 (95% CI -11.122 to 1.122) STAI-State scale (20-80): 45.8 (2.5) vs. 46.8 (2.3); MD -1.0 (95% CI -2.351 to 0.351) SF-36 Physical Capacity (0-100): 59.6 vs. 54.0; MD 5.6 (95% CI -11.905 to 23.105) SF-36 Pain (0-100): 48.4 vs. 45.5; MD 2.9 (95% CI -10.783 to 16.583) SF-36 Mental Health (0-100): 69.9 vs. 56.2; MD 13.7 (95%CI 0.070 to 27.330)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Jensen 2012,[94] Wicksell 2013[97] 3-4 months Time since FM onset, years: 10.5 vs. 11.8 *Fair*	A. Acceptance and Commitment Therapy (ACT) (n=25): 12 weekly 90-minute group sessions: exposure to personally important situations and activities previously avoided due to pain and distress, training to distance self from pain and distress. B. Waiting list control (n=18)	A vs. B Age: 45 vs. 47 years Female: 100% vs. 100% FIQ (0-100): 49.3 vs. 48.7 PDI (scale NR): 40.0 vs. 39.0 Pain VAS (0-100): 61 vs. 65.0 Pain NRS (0-10): 4.2 vs. 4.3	A vs. B <u>3-4 months</u> FIQ: 37.4 vs. 45.7, Cohen's d=0.66 (95% CI -0.06 to 1.37); MD -8.3 (95% CI -17.056 to 0.456) PDI: 28.1 vs. 38.1, Cohen's d=0.73 (95% CI -0.00 to 1.44); MD -10.0 (95% CI -19.740 to -0.260) Pain VAS: means NR but group X time interaction p=0.26 Pain NRS: 3.9 vs. 4.8, Cohen's d= 0.82 (95% CI 0.08 to 1.54); MD -0.90 (95% CI -1.674 to -0.126)	A vs. B <u>3-4 months</u> BDI (0-63): 10.7 vs. 16.4, Cohen's d=0.64 (95% CI -0.08 to 1.35); MD -5.7 (95% CI -12.044 to 0.644) STAI-State: 39.8 vs. 45.4; Cohen's d=0.55 (95% CI -0.17 to 1.26); MD -5.6 (95% CI -12.751 to 1.551) SF-36 Mental: 46.0 vs. 34.7, Cohen's d=1.06 (95% CI 0.28 to 1.82); MD 11.3 (95% CI 3.761 to 18.839) SF-36 Physical (0-100): 28.4 vs. 31.1, Cohen's d=0.28 (95% CI -0.45 to 1.00); MD -2.7 (95% CI -9.401 to 4.001),
Kayiran 2010 4 to 5 months Duration of symptoms: 5 years *Poor*	A. EEG Biofeedback (Neurofeedback) (n=20): 5 sessions based on sensorimotor rhythm training protocol per week for 4 weeks. Each session consisted of 10 sensorimotor rhythm training periods lasting for 3 minutes for a total of 30 minutes B. Escitalopram (n=20): 10 mg/day for 8 weeks (control group)	A vs. B Age: 32 vs. 32 years Female: 100% vs. 100% FIQ (mean, 0-100): 70 vs. 74* Pain VAS (mean, 0-10): 8.9 vs. 9.1	A vs. B <u>4-5 months</u>: FIQ: 19 vs. 48*, p NR Pain VAS: 2.6 vs. 5.3; MD -2.7 (95% CI -3.7 to -1.7)	A vs. B <u>4-5 months</u>: HAM-D (0-50): 6.3 vs. 13.4; MD -7.1 (95% CI -9.1 to -5.1) BDI (0-63) : 4.7 vs. 12.3; MD -7.6, 95% CI -9.7 to -5.5) HAM-A (0-56): 7.1 vs. 15.2; MD -8.1 (95% CI -11.0 to -5.2) BAI (0-63): 7.2 vs.. 16.7; MD -9.5 (95% CI -13.9 to -5.1) SF-36*: Physical functioning (0-100): 77 vs. 65, p<0.05 Bodily pain: 70 vs. 45, p<0.05 Role-physical (0-100): 90 vs. 43, p<0.05 Role-emotional (0-100): 95 vs. 51, p<0.05 Social functioning (0-100): 76 vs. 65, p<0.05 General mental health (0-100): 74

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				vs. 59, p<0.05 General health (0-100): 72 vs. 28, p<0.05 Vitality (0-100): 70 vs. 50, p<0.05
Larsson 2015 13 to 18 months Duration of symptoms: 10 years *Poor*	A. Relaxation therapy (n=63): Two group sessions of 5-8 subjects per week for 15 weeks. The intervention was preceded by an individual meeting covering instructions and allowing for adjustments to the intervention. The sessions lasted 25 minutes and consisted of autogenic training guided by physiotherapist and were followed by stretching. B. Resistance exercise (Strength) (n=67): Two group sessions of 5-7 subjects per week for 15 weeks. The intervention was preceded by an individual meeting going over instructions on the intervention, testing, and modifications of specific exercises. Sessions were based on a resistance exercise program aiming to improve muscle strength, focusing on large muscle groups in the lower extremity.	A vs. B Age: 52 vs. 51 Female: 100% vs. 100% FIQ (0-100): 61.1 vs. 60.5 Pain VAS (0-100): 52.4 vs. 49.3 PDI (0-70): 35.0 vs. 35.3	A vs. B <u>13-18 months</u> FIQ: 55.4 vs. 57.1, (MD -1.7, 95% CI -9.3 to 5.9) Pain VAS: 52.1 vs. 49.2, (MD 2.9, 95% CI -5.5 to 11.3) PDI: 33.7 vs. 33.0, (MD 0.7, 95% CI -4.0 to 5.4)	A vs. B <u>13-18 months</u> SF-36 PCS (0-100): 32.0 vs. 32.2, (MD -0.2, 95% CI -3.8 to 3.4) SF-36 MCS (0-100): 40.0 vs. 39.2, (MD 0.8, 95% CI -4.6 to 6.2) Patient global impression of change (mean, 1-7): Values NR but difference was NS
Redondo, 2004[106] 6 and 12 months Pain duration NR *Poor*	A. Cognitive Behavior Therapy (n=21): 1, 2.5 hour session per week for 8 weeks. Sessions included information about chronic pain and FM, relaxation techniques, and pain coping strategies training. B. Combination Exercise (n=19): 5, 45-minute sessions per week for 8	A vs. B Age NR Female: 100% vs. 100% FIQ total (mean, 0-80): 52.0 vs. 52.0 FIQ pain (mean, 0-10): 7.3 vs. 6.8 FIQ depression (mean, 0-10): 5.2 vs. 5.3 FIQ anxiety (mean, 0-10): 6.4 vs. 6.3	A vs. B <u>6 months:</u> FIQ total: 47.4 vs. 48.0, (MD -0.6, 95% CI -12.6 to 11.4) FIQ pain: 5.9 vs. 6.9, (MD -1.0, 95% CI -2.8 to 0.8) <u>12 months:</u> FIQ: 47.8 vs. 47.7; (MD 0.1, 95% CI -10.5 to 10.7)	A vs. B <u>6 months:</u> FIQ depression (0-10): 5.2 vs. 5.3, (MD -0.1, 95% CI -2.6 to 2.4) FIQ anxiety (0-10): 6.0 vs. 5.8, (MD 0.2, 95% CI -2.2 to 2.6) BAI: 25.2 vs. 22.1, (MD 3.1, 95% CI -5.1 to 11.3) BDI (0-63): 17.1 vs. 15.0, (MD 2.1, 95%

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	weeks. Each week included 1 session of aquatic exercises, 2 sessions of flexibility and endurance exercises, and 2 sessions of cardiovascular exercises. All subjects: Offered ibuprofen or diclofenac, 25 mg of amitriptyline a day, and acetaminophen.		FIQ pain: 6.3 vs. 6.6; (MD -0.3, 95% CI -2.0 to 1.3)	CI -6.6 to 10.8) SF-36 physical functioning (0-100): 52.2 vs. 43.9, (MD 8.3, 95% CI -6.4 to 23.0) SF-36 physical role (0-100): 22.4 vs. 18.3, (MD 4.1, 95% CI -21.2 to 29.4) SF-36 bodily pain (0-100): 31.4 vs. 32.9, (MD -1.5, 95% CI -16.1 to 13.1) SF-36 social functioning (0-100): 66.4 vs. 66.9, (MD -0.5, 95% CI -21.6 to 20.6) SF-36 emotional role (0-100): 68.4 vs. 66.0, (MD 2.4, 95% CI -28.2 to 33.0) SF-36 mental health (0-100): 48.9 vs. 51.8, (MD -2.9, 95% CI -19.3 to 13.5) <u>12 months:</u> FIQ depression: 5.4 vs. 4.9; (MD 0.5, 95% CI -2.0 to 3.0) FIQ anxiety: 6.0 vs. 5.8; (MD 0.2, 95% CI -2.1 to 2.5) BAI: 20.0 vs. 20.0; (MD 0.0, 95% CI -7.4 to 7.4) BDI: 13.0 vs. 13.6; (MD -0.6, 95% CI -7.9 to 6.7) SF-36 physical functioning: 38.9 vs. 41.6; (MD -2.7, 95% CI -19.5 to 14.1) SF-36 physical role: 26.1 vs. 31.0; (MD -4.9, 95% CI -27.9 to 18.1) SF-36 bodily pain: 33.8 vs. 34.3; (MD -0.5, 95% CI -20.9 to 19.9) SF-36 social functioning: 60.7 vs. 57.2; (MD 3.5, 95% CI -17.2 to 24.2)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				SF-36 emotional role: 66.7 vs. 58.7; (MD 8.0, 95% CI -19.2 to 35.2) SF-36 mental health: 56.5 vs. 53.8; (MD 2.7, 95% CI -19.1 to 24.5)
Thieme, 2006[95] 6 and 12 months Duration of symptoms, years: 9.1 vs. 8.7 *Poor*	A. CBT (n=42): 2-hour group sessions weekly for 15 weeks. Sessions focused on changing patients' thinking and problem-solving, stress and pain coping strategies, and relaxation exercises performed during and between sessions. B. Attention control (n=40): 2-hour group sessions weekly for 15 weeks: general discussions about medical and psychosocial problems of fibromyalgia.	A vs. B Age: 49 vs. 47 years Female: 100% vs. 100% FIQ physical impairment (mean, 0-10): 4.4 vs. 4.2 WHYMPI pain intensity (mean, 0-6): 4.2 vs. 3.8	A vs. B <u>6 months</u> FIQ physical impairment: 3.0 vs. 4.8; MD -1.8 (95% CI -2.899 to -0.701) WHYMPI pain intensity: 3.7 vs. 4.1; MD -0.4 (95% CI -0.841 to 0.041) <u>12 months</u> FIQ physical impairment: 3.4 vs. 5.2; MD -1.8 (95% CI -2.855 to -0.745) WHYMPI pain intensity: 3.2 vs. 4.1; MD -0.9 (95% CI -1.537 to -0.263)	A vs. B <u>6 months</u> WHYMPI affective distress: 2.6 vs. 4.0; MD -1.4 (95% CI -1.952 to -0.848) <u>12 months</u> WHYMPI affective distress: 2.6 vs. 4.2; MD -1.6 (95% CI -2.172 to -1.028)
Van Santen 2002 Post 6-month intervention Duration of symptoms, years: 10.1 vs. 15.4 vs. 15.4 *Poor*	A. Electromyographic biofeedback (n=56): Progressive muscle relaxation and frontalis EMG biofeedback; 30-minute individual sessions 2 times per week for 8 weeks; subjects encouraged to practice at home twice daily for the 8 weeks then for 16 more weeks. Subjects randomized to education aimed at compliance with biofeedback training (6 90-minute sessions over 24 weeks). B. Usual care (n=29): General physicians informed not to prescribe or encourage aerobic exercises and relaxation. Intervention duration: 6 months	A vs. B Age: 44 vs. 43 vs. 46 years Female: 100% vs. 100% vs. 100% Race NR SIP Physical score (0-100): 11.4 vs. 9.8 vs.11.3 Pain VAS (0-100): 59.1 vs. 62.4 vs. 66.8 AIMS (0-10): 3.1 vs. 5.4 vs. 1.9 SIP Total score (0-100): 14.0 vs. 11.4 vs. 14.4 SIP Psychosocial score (0-100): 15.8 vs. 18.1 vs. 16.3	A vs. B <u>6-months:</u> SIP physical score, mean change: -1.6 (95% CI -3.4 to 0.2) vs. -0.6 (95% CI -2.9 to 1.7) SIP total score, mean change: -2.3 (95% CI -4.3 to -0.3) vs. -1.4 (95% CI -3.4 to 0.6) AIMS, mean change: 0.4 (95% CI -0.1 to 0.9) vs. 0.8 (95% CI -1.8 to -0.2) SIP total score, mean change: -2.3 (95% CI -4.3 to -0.3) vs. -1.4 (95% CI -3.4 to 0.6) Pain VAS, mean change: -0.6 (95% CI -6.5 to 5.3) vs. 1.3 (95% CI -4.5 to 7.1)	A vs. B <u>6-months:</u> SIP psychosocial score, mean change: -3.7 (95% CI -4.9 to -2.5) vs. -3.5 (95% CI -7.0 to 0.0) Patient global assessment of well-being, mean change: 0.3 (95% CI 0.0 to 0.6) vs. 0.5 (95% CI 0.2 to 0.8) A vs. C <u>6-months:</u> SIP psychosocial score, mean change: -3.7 (95% CI -4.9 to -2.5) vs. -3.2 (95% CI -6.2 to 0.2) Patient global assessment of well-being, mean change: 0.3 (95% CI

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	C. Combination Exercise (n=58): 60-minute group sessions of twice a week for 24 weeks; aerobic exercises, postural strengthening, general flexibility and balance exercises, and isometric muscle strengthening; subjects encouraged to attend third, unsupervised, 60-minute session and to use sauna or swimming pool after sessions.		A vs. C <u>6-months:</u> SIP physical score, mean change: -1.6 (95% CI -3.4 to 0.2) vs. -1.7 (95% CI -3.7 to 0.3), ns SIP total score, mean change: -2.3 (95% CI -4.3 to -0.3) vs.. -1.9 (95% CI -3.9 to 0.1) AIMS, mean change: 0.4 (95% CI -0.1 to 0.9) vs. 0.1 (95% CI -0.6 to 0.8) Pain VAS, mean change: -0.6 (95% CI -6.5 to 5.3) vs.. -5.5 (95% CI -10.9 to -0.1), ns	0.0 to 0.6) vs. 0.5 (95% CI 0.2 to 0.8)
Verkaik, 2014[96] 1.5 months Duration of symptoms, NR *Poor*	A. Guided imagery (n=33): Two 1.5 hour group sessions of 6-12 subjects. The first sessions consisted of group discussion, the theoretical background of guided imagery, and instructions to practice at least one exercise daily for 4 weeks. Each exercise was a CD and contained relaxation techniques, music, positive imagery, and pain management techniques. The second group session took place after the 4 weeks and consisted of a group discussion. B. Attention control (n=37): Two 1.5 hour group sessions of 6-12 subjects held 4 weeks apart. Group sessions were a group discussion and did not contain any information or training on guided imagery.	A vs. B Age: 47 vs. 48 Female: 100% vs. 97% FIQ(0-100): 53.7 vs. 56.4 Pain VAS (0-10): 5.9 vs. 5.8	A vs. B <u>1.5 months</u> FIQ: 54.2 vs. 53.0, MD 1.2, 95% CI -0.2 to 2.6) Pain VAS: NR	NR
Wigers, 1996[82] 48 months	A. Stress management (n=20): 90 minute group sessions of 10 patients	A vs. B Age: 44 vs. 46 vs. 43 years	A vs. B <u>48 months</u> Pain VAS: 70 vs.	A vs. B <u>48 months</u> Depression VAS (0-

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fibromyalgia duration A vs. B vs. C Mean: 11 vs. 9 years *Poor*	done 2 times a week for 6 weeks followed by 1 session per week for the next 8 weeks. Sessions consisted of equal portions of presentations stress mechanisms and strategies for improving quality of life, group discussions on patients' experiences of stress and coping with pain, and relaxation training aimed at helping cope with stress and pain. B. Usual care (n=20): Subjects continued treatments they had been using at baseline. C. Aerobic exercise (n=20): 45 minute group sessions of 10 patients done 3 times a week for 14 weeks. The exercise program involved the whole body and aimed to minimize eccentric muscle strain. Sessions consisted of training to music (further details not given) and aerobic games.	Female: 90% vs. 95% vs.. 90% Pain VAS (0-100): 72 vs. 65 vs. 72	69, (MD 1, 95% CI -12.6 to 14.6) A vs. C <u>48 months</u> Pain VAS: 70 vs. 68, (MD 2, 95% CI -11.6 to 15.6)	100): 40 vs. 30, (MD 10, 95% CI -8.9 to 28.9) Global subjective improvement: 47% (6/13) vs. 12% (2/16), (RR 3.7, 95% CI 0.9 to 15.3) A vs. C <u>48 months</u> Depression VAS: 40 vs. 32, (MD 8, 95% CI -11.9 to 27.9) Global subjective improvement: 47% (6/13) vs. 75% (11/15), (RR 0.6, 95% CI 0.3 to 1.2)
Williams, 2002[98] 12 months Fibromyalgia duration, 8.6 years *Poor*	A. Group CBT plus Usual Care (n=76): 6 1-hour group sessions over 4-week period: progressive muscle relaxation, imagery, activity pacing, pleasant activity scheduling, communication skills and assertiveness training, cognitive restructuring, stress management and problem-solving. B. Usual Care (n=69): Standard pharmacological management (typically low-dose tricyclic antidepressant medication, analgesics, and/or antidepressants)	A + B Age, mean, years: 47.7 Females: 90% Race: White non-Hispanic 88%, black non-Hispanic 9%, Hispanic 2%, Asian American 1% MPQ-Sensory (scale NR): 14.8 MPQ-Affective pain score (scale NR): 4.6	A vs. B <u>12 months</u> M (SD): NR Proportion of subjects who improved more than 12 points from baseline on MPQ-Sensory scale: 3.9% vs. 7.2%RR 0.54 (95% CI 0.14 to 2.2)	A vs. B <u>12 months</u> M (SD) NR Proportion of subjects who improved more than 6.5 points from baseline on SF-36 PCS Score: 25% vs. 11.6%, OR=2.9; RR 2.2 (95% CI 0.98 to 4.99) Proportion of subjects who improved more than 5 points from baseline on MPQ-Affective scale: 9.2% vs. 8.7%,RR 1.1 (95% CI 0.37 to 3.0)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	plus suggestions to engage in aerobic fitness.			

AIMS, Arthritis Impact Measurement Scales; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CES-D, Center for Epidemiologic Studies Depression Scale; CI, confidence interval; FIQ, Fibromyalgia Impact Questionnaire; HAM-D, Hamilton Rating Scale for Depression;
HAM-A, Hamilton Anxiety Rating Scale; HADS, Hospital Anxiety and Depression Scale; MCSD, Minimal Clinically Significant Difference; MPQ, McGill Pain Questionnaire; PDI, Pain Disability Index; PHQ, Patient Health Questionnaire; PI, Physical Impairment; SCL-90-R, Symptoms Checklist 90-Revised; SIP, Sickness Impact Profile; SF-36, Short-Form 36 questionnaire; SF-36 PCS, Short-Form 36 Physical Component Summary Score; SF-36 MCS, Short-Form 36 Mental Component Summary Score; STAI, State-Trait Anxiety Inventory; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Psychological Therapies Compared With Usual Care or Sham

Ten trials compared psychological interventions versus usual care, waitlist, or attention control.[63,81,82,91-98]

Functional outcomes. Across types of psychological therapies, results for function were not consistent short term (Table 36). One trial was fair quality.[94,97] CBT was associated with a slightly greater effect on function based on FIQ Total Score than usual care or waitlist in the short term (2 trials, pooled mean difference -10.67, 95%CI -17 to -4.30, I^2=0%, 0-100 scale)[93,94,97] (Figure 41). One poor-quality trial of CBT reported significantly more patients in the CBT group attaining a clinically important improvement (≥14% on the FIQ total, 0-100 scale) from baseline on FIQ compared with usual care (RR 2.8, 95% CI 1.3 to 6.1),[93] while another smaller poor-quality trial did not (RR 2.2, 95% CI 0.5 to 9.3).[92] One poor-quality trial of guided imagery reported no difference in function versus attention control (FIQ total, 0 to 100 scale difference 1.2, 95% CI -0.2 to 2.6)[96] nor did one poor-quality trial of electromyographic biofeedback and relaxation training versus attention control (median change from baseline 6.0 for both groups, AIMS physical activity subscale, 0-10).[63]

Individual trials showed CBT had a greater effect on function than usual care based on FIQ total score; however, the pooled estimate at intermediate term was not statistically significant due to heterogeneity (2 trials, pooled mean difference -10.36, 95% CI -23.52 to 2.80, I^2 = 84.5%, 0-100 scale)[91,93] (Figure 41). One of the trials reported substantially more CBT patients achieving a clinically important difference (RR 2.9, 95% CI 1.9 to 17.8) than usual care.[93] Findings from an additional trial suggest greater effect of CBT on function based on a 0 to 10 FIQ Physical Impairment Scale (mean difference -1.8, 95%CI -2.9 to -0.70).[95] There was no clear difference between biofeedback and usual care on function in one poor-quality trial (mean changes -1.6, 95% CI -3.4 to 0.2 versus -0.6, 95% CI -2.9 to 1.7, respectively). [81]

Data from three poor-quality trials were insufficient to determine the long-term effects of psychological therapies on function. One trial of CBT versus attention control reported improvement on a 0 to 10 FIQ Physical Impairment Scale (mean difference -1.8, 95% CI -2.85 to -0.745).[95] By contrast, another trial found no difference between CBT and usual care in the proportion of participants achieving a clinically meaningful change in a 12-point MPQ Sensory

Scale (RR 0.54 (95% CI 0.14 to 2.2).[98] A trial of biofeedback versus usual care reported median change in the AIMS Physical Activity subscale of 6.0 in both groups.[63]

Pain outcomes. Psychological therapies were associated with a slightly greater improvement in pain compared with usual care, waitlist, or attention control (4 trials, pooled mean difference -0.74, 95% CI -1.20 to -0.28, I^2=0%)[63,92-94,97] (Figure 42). Estimates were similar when studies were stratified by type of psychological therapy. Estimates were similar when results were stratified by type of control (usual care), but results were no longer statistically significant. Only one trial was considered fair. Psychological therapies were also associated with slightly greater improvement in pain compared with usual care at intermediate term (3 trials, pooled mean difference -0.67, 95% CI -1.21 to -0.31, I^2=36.7%)[81,91,93] (Figure 42). Estimates were similar when the two CBT trials were pooled, but no longer significant. Long term, there was no clear difference between psychological therapies (biofeedback or CBT) and attention control or usual care (2 trials, pooled mean difference 0.04, 95% CI -0.77 to 0.84, I^2=0%);[63,82] however, evidence across the two poor quality trials was considered insufficient (Figure 42).

Other outcomes. Results were mixed across studies for effects of CBT on secondary outcomes. In comparisons of CBT to usual care at short-term followup, one study[92] found no significant differences on a measure of depression, whereas another study[93] found a significant benefit for CBT on the Hospital Anxiety and Depression Scale (HADS) and on measures of sleep. In a study that compared ACT to a waiting list control at short-term followup,[94,97] there were no significant differences on measures of depression or anxiety or on the SF-36 PCS, but there was a significant benefit for ACT on the SF-36 MCS. Results were also mixed for comparisons of CBT with usual care at intermediate-term followups; one study[91] found no significant differences on measures of anxiety or depression, but another study[93] found a significant benefit for CBT on the HADS measure of anxiety and depression and on measures of sleep. The one study[98] that compared CBT to usual care at long-term followup found no difference in the proportion of those who improved more than 6.5 points from baseline on the SF-36 PCS. In an additional trial evaluating stress management training,[82] there was no significant difference compared with usual care on a measure of depression at long-term followup.

Data comparing CBT to attention control conditions on secondary outcomes were very limited. In the one study[95] that met inclusion criteria, there was a significant benefit for CBT on a measure of affective distress at intermediate-term and long-term followups.

Two studies compared EMG biofeedback to attention control conditions; neither found differences on secondary outcomes. One study[63] found no difference at short-term or long-term followups on measures of psychological distress, depression, or sleep. In a study[81] that compared EMG biofeedback to usual care immediately after the 6-month intervention, there were no differences on the AIMS, Sickness Impact Profile (SIP) total score, SIP Psychosocial scale, or a patient global assessment of well-being.

Psychological Therapies Compared With Pharmacological Therapy

Two fair-quality[91,101] and one poor-quality trial[102] compared a psychological therapy with pharmacological treatment. Two trials reported functional outcomes over the short-term with differing results. No clear effect was seen for CBT (plus amitriptyline) compared with amitriptyline alone at 3 months in one fair quality trial (mean difference -4.1, 95% CI -18.8 to 10.6, on the FIQ total score [0 to 100] scale).[101] One poor-quality trial, comparing EEG

biofeedback with escitalopram, reported better mean FIQ total scores in the biofeedback group at 4 to 5 months followup (19 versus 48, 0 to 100 scale), but did not provide enough data to calculate an effect estimate.[102] Intermediate-term function was reported by one fair-quality trial, which found a small benefit for CBT compared with pregabalin (plus duloxetine as needed) on the FIQ at 6 months (difference -4.0 on a 0-100 scale, 95% CI -7.7 to -0.27).[91]

The pattern for pain outcomes was similar over the short term. No differences were seen between groups in the trial of CBT versus amitriptyline (difference -0.7 on a 0-10 VAS, 95% CI -2.8 to 1.4),[101] whereas a moderate effect was seen for EEG biofeedback compared with escitalopram (difference -2.7 on a 0-10 VAS, 95% CI -3.7 to -1.7) in the poor-quality trial.[102] At intermediate-term, VAS pain scores were similar between the CBT and pregabalin groups in the third trial (difference 0.2 on a 0-100 scale, -4.0 to 4.4).[91]

Regarding secondary outcomes, EEG biofeedback was associated with significantly better outcomes on various measures of anxiety, depression, and quality of life compared with escitalopram over short-term followup in the poor-quality trial,[102] whereas the two fair-quality trials evaluating CBT (versus amitriptyline and versus pregabalin)[91,101] found no differences between groups over the short or intermediate term, with the exception of SF-36 Mental Health scores at short-term followup in one trial (difference 13.7 on a 0-100 scale, 95% CI 0.07 to 27.3).[101]

Psychological Therapies Compared With Exercise

Five poor-quality trials compared psychological interventions with some form of exercise; two trials evaluated compared CBT,[82,106] two trials evaluated biofeedback,[63,81] and one evaluated relaxation training[105] (Table 36).

Data were insufficient from one poor-quality trial to determine the effects of biofeedback versus combination exercise on function. The trial reported improved function based on the AIMS physical activity subscale (median change from baseline): 6.0 versus 4.0, $p<0.05$.[63] Intermediate term, data from two poor-quality trials was also considered insufficient to determine effects of psychological therapies on function: no clear differences in function were seen for CBT (mean difference -0.6, 95% CI -12.6 to 11.4 on 0-100 FIQ total score)[106] or biofeedback (mean change -1.6, 95% CI -3.4 to 0.2 vs. -0.6, 95% CI -2.9 to 1.7 on 0-100 SIP Physical score)[81] versus combination exercise. Similarly, no clear differences between psychological therapies and exercise were seen across three trials at longer term and evidence was considered insufficient. Results from two trials were not statistically significant: for one trial of CBT versus combination exercise, (mean difference 0.1, 95% CI -10.5 to 10.7 on 0-100 FIQ total scale)[106] or one trial of relaxation training versus strength training (mean difference -1.7, 95% CI -9.3 to 5.9, on 0-100 FIQ Total Score).[105] The third trial of biofeedback versus combination exercise reported improvement in function, but limited data were provided (median change from baseline, 6.0 versus 4.0, $p<0.05$).[63]

Data were insufficient from one poor-quality trial to determine the effects of biofeedback versus combination exercise pain (median change from baseline, 5.2 vs. 5.4 on 0-10 VAS).[63] Across two poor-quality trials at intermediate-term, no clear differences were seen for CBT (mean difference -1.0, 95% CI -2.8 to 0.8)[106] or biofeedback (mean change: -0.6 (95% CI -6.5 to 5.3) vs. -5.5 (95% CI -10.9 to -0.1), p = ns)[81] compared with combination exercise; evidence was considered insufficient. There were no clear differences between any of the psychological therapies and exercise for pain on a 0 to 10 scale across four trials long-term; CBT versus combination exercise (mean difference 0.3, 95% CI -2.0 to 1.3)[106] or aerobic exercise (difference 2, 95% CI -11.6 to 15.6),[82] for biofeedback versus combination exercise (median change: 5.2 vs.

5.5, p = ns)[63] or for relaxation training versus strength training (difference 2.9, 95% CI -5.5 to 11.3).[105]

There were generally no significant differences on measures of mental health, depression or anxiety, or on SF-36 scales, at any time frame across five poor-quality trials.[63,81,82,105,106] Some trials did not provide data for determination of effect sizes between treatment groups or report results of significance tests (Table 36).

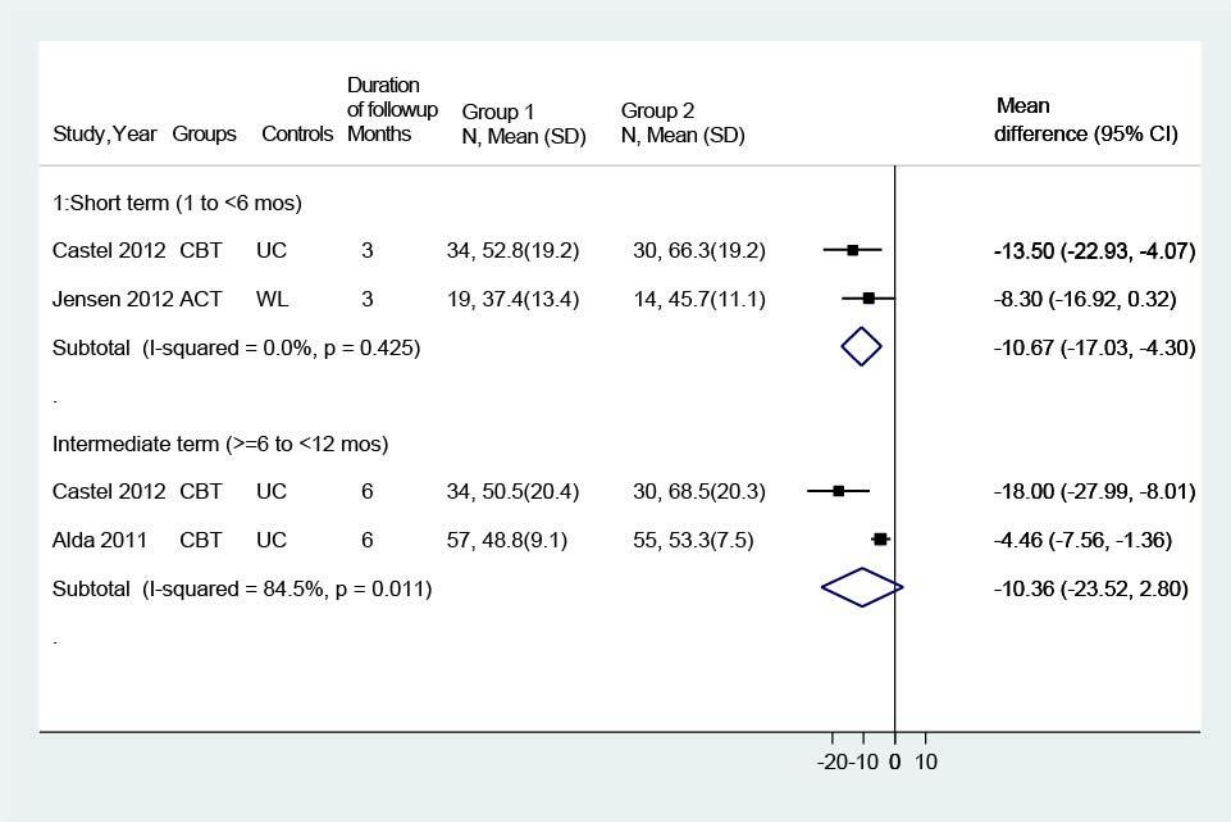
Harms

Only five trials (1 fair quality and 4 poor quality) reported harms, which were poorly described in general. Two trials compared CBT with usual care; one trial reported no withdrawals due to adverse events in the CBT group compared with two (3.6%) in the control group (not further described),[91] and the other trial reported two withdrawals, one in each group, because the nociceptive flexion reflex test being used was too painful.[92] One trial comparing CBT with attention control reported that 4.8 percent (due to depression) versus 50 percent (due to worsening of symptoms) of patients, respectively, withdrew from the study.[95] One trial compared stress management to usual care and reported one withdrawal due to cancer (unrelated to the treatment) in the intervention group compared to no withdrawals or adverse events in the control.[82]

One of the above trials also compared CBT to pharmacological therapy (pregabalin) and reported no withdrawals due to adverse events in the CBT group compared with three (5.5%) in the control group, two due to digestive problems, and one due to dizziness).[91]

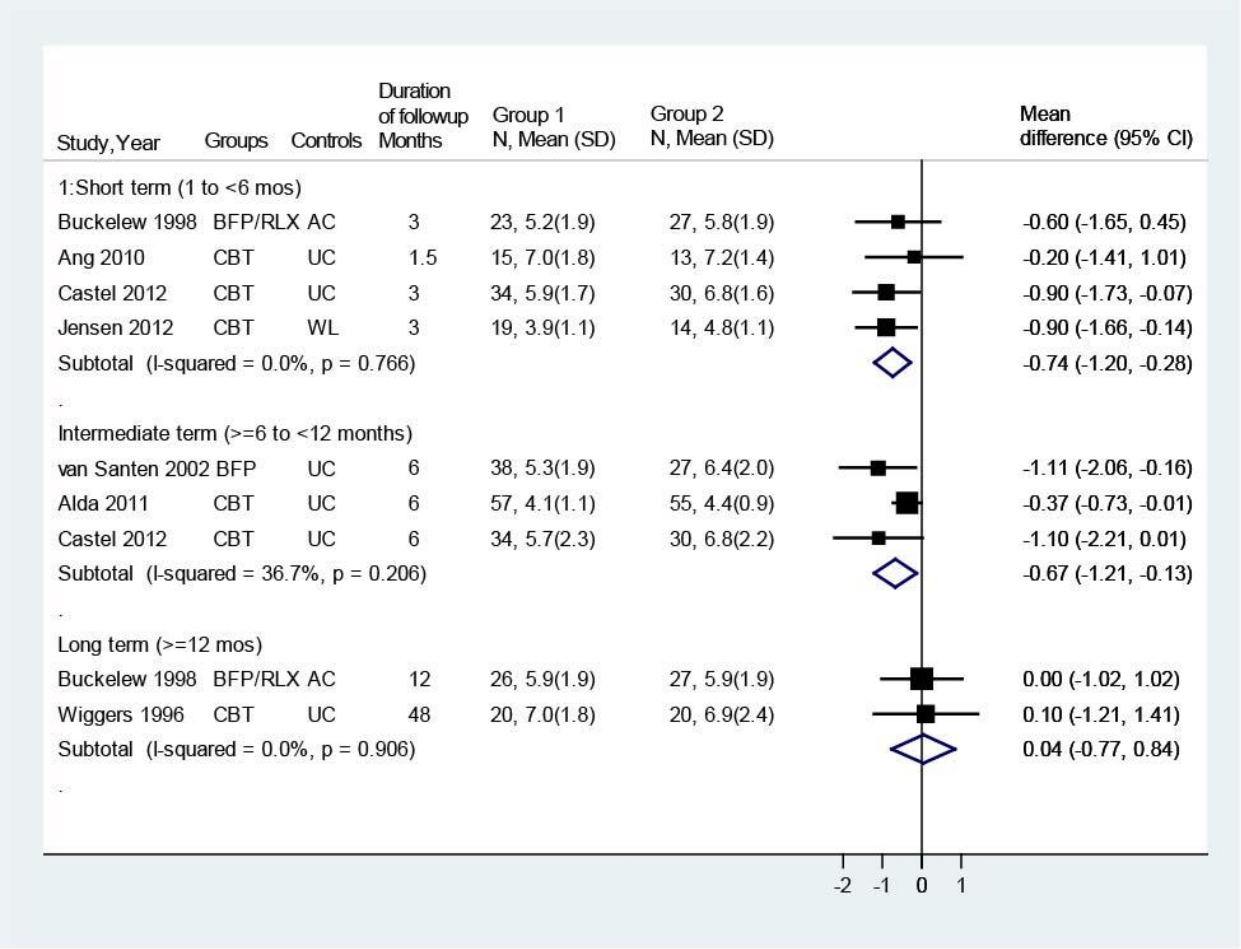
Two trials compared psychological therapies with exercise, one of which reported no adverse effects with relaxation therapy but five (7.5%) reports of adverse effects following strengthening exercises (due to increased pain), three of which withdrew,[105] and the other reported one withdrawal due to cancer (unrelated to the treatment) in the intervention group compared with three withdrawals in the exercise group (1 death, 1 gastritis, 1 ischialgia).[82]

Figure 41. Psychological therapies versus usual care or waitlist for fibromyalgia: effects on function



ACT = Acceptance and Commitment Therapy; CBT = cognitive behavioral therapy; CI = confidence interval; SD = standard deviation; UC = usual care; WL = waitlist

Figure 42. Psychological therapies versus usual care, waitlist, or attention control for fibromyalgia: effects on pain



AC = attention control; BFP = Biofeedback; BFP/RLX = Biofeedback with a Relaxation component; CBT = cognitive behavioral therapy; CI = confidence interval; SD = standard deviation; UC = usual care; WL = waitlist

Physical Modalities for Fibromyalgia

Key Points

- One fair-quality trial showed no differences between magnetic mattress pads compared with sham or usual care in intermediate-term function (difference on the 0 to 80 scale FIQ -5.0, 95% CI -14.1 to 4.1 vs. sham and -5.5, 95% CI -14.4 to 3.4 vs. usual care) or pain (difference -0.6, 95% CI -1.9 to 0.7 and -1.0, 95% CI -2.2 to 0.2, respectively on a 0 to 10 NRS) (SOE: Low).
- There were no differences in adverse events between the functional and sham magnetic mattress pad groups (data not reported); none of the events were deemed to be related to the treatments (SOE: Low).

Detailed Synthesis

One trial (n=119),[134] conducted in the United States, evaluated the efficacy of two different magnetic mattress pads (one with a low, uniform magnetic field of negative polarity and the other a low, static magnetic field that varied spatially and in polarity) for the treatment of fibromyalgia (Table 37 and Appendix D). Comparisons included sham (mattress pads with demagnetized magnets) and usual care (management by primary care provider). All pads were used for 6 months and outcomes were measured immediately post-treatment. This trial was rated fair quality due to deviations from the randomization protocol and unacceptable attrition rate (21%) (Appendix E).

Table 37. Summary of results for fibromyalgia: physical modalities

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Alfano, 2001[134] 6 months Duration of pain: >3 months (mean NR) *Fair*	A. Magnetic mattress pad designed to expose body to a uniform magnetic field of negative polarity (n=37) B. Magnetic mattress pad exposing body to magnetic field that varied spatially and in polarity (n=33) C. Sham magnetic field (n=32): combined group of 2 sham magnetic mattress pads; identical in appearance to real magnetic pads but contained demagnetized magnets. D. Usual care (n=17): maintain current treatment under PCP, refrain from new treatments Treatment period was 6 months for all groups.	A vs. B vs. C vs. D Age: 44 vs. 47 vs. 46 vs. 45 years Female: 92% vs. 87% vs. 96% vs. 100% FIQ (0-80): 51.6 vs. 55.5 vs. 51.5 vs. 53.9 Pain intensity FIQ NRS (0-10): 7.1 vs. 7.0 vs. 6.7 vs. 7.0	A + B vs. C <u>Post 6-month intervention</u> FIQ: 42.9 vs. 47.9, difference -5.0 (95% CI -14.1 to 4.1) Pain intensity NRS: 5.6 vs. 6.2, difference -0.6 (95% CI -1.9 to 0.7) A + B vs. D <u>Post 6-month intervention</u> FIQ: 42.9 vs. 48.4, difference -5.5 (95% CI -14.4 to 3.4) Pain intensity NRS: 5.6 vs. 6.6, difference -1.0 (95% CI -2.2 to 0.2) A vs. C <u>Post 6-month intervention</u> FIQ: 38.3 vs. 47.9, difference -9.6 (95% CI -20.0 to 0.8) Pain intensity NRS: 4.8 vs. 6.2, difference -1.4 (95% CI -2.8 to 0.05) B vs. C <u>Post 6-month intervention</u> FIQ: 47.4 vs. 47.9, difference -0.5 (95% CI -11.2 to 10.2) Pain intensity NRS: 6.3 vs. 6.2, difference 0.1 (95% CI -1.4 to 1.6) A vs. D <u>Post 6-month intervention</u> FIQ: 38.3 vs. 48.4, difference -10.1 (95% CI -21.9 to 1.7) Pain intensity NRS: 4.8 vs. 6.6, difference -1.8 (95% CI -3.4 to -0.2) B vs. D <u>Post 6-month intervention</u> FIQ: 47.4 vs. 48.4, difference -1.0 (95% CI -13.0 to 11.0),	NR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			Pain intensity NRS: 6.3 vs. 6.6, difference -0.3 (95% CI -2.0 to 1.4)	

CI = confidence interval; FIQ = Fibromyalgia Impact Questionnaire; NRS = numerical rating scale; PCP = primary care physician

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

Physical Modalities Compared With Usual Care or Sham

The magnetic mattress pads offered no intermediate-term benefit for either function or pain compared with both sham and usual care.[134] The mean difference between groups on the 0 to 80 scale FIQ at 6 months was -5.0 (95% CI -14.1 to 4.1) (versus sham) and -5.5 (95% CI -14.4 to 3.4) (usual care). Regarding pain, the between-group differences were -0.6 (95% CI -1.9 to 0.7) and -1.0 (95% CI -2.2 to 0.2), respectively, on a 0 to 10 NRS. When the intervention groups were considered separately, only the magnetic mattress pad designed to expose the body to a uniform magnetic field of negative polarity resulted in lower FIQ and NRS pain scores compared with controls; however, the mean differences between groups were not statistically significant.

Physical Modalities Compared With Pharmacological Therapy or Exercise

No trial of physical modality versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

There were no differences in adverse events between the magnetic mattress pad and sham pad groups.[134] Type of adverse events was not reported, but none of the events were judged to be due to magnetic treatments.

Manual Therapies for Fibromyalgia

Key Points

- Myofascial release therapy was associated with a slightly greater effect on intermediate-term function as measured by the FIQ (mean 58.6 ± 16.3 vs. 64.1 ± 18.1 on a 100 point scale, p=0.048 for group by repeated measures ANOVA), but not long-term function (mean 62.8 ± 20.1 vs. 65.0 ± 19.8 on the FIQ, 0-100 scale, p=0.329), compared with sham in one fair-quality trial (SOE: Low). Short-term function was not reported.
- There is insufficient evidence to determine the effects of myofascial release therapy on short-term pain (1 poor-quality trial) and intermediate-term pain (1 fair-quality and 1

poor-quality trial) compared with sham; there were inconsistencies in effect estimates between the intermediate-term trials (SOE: Insufficient).

- Myofascial release therapy was associated with slightly greater effects on long-term pain based on the sensory (18.2 ± 8.3 vs. 21.2 ± 7.9 on a 0-33 scale, p=0.038 for group by repeated measures ANOVA) and evaluative (23.2 ± 7.6 vs. 26.7 ± 6.9 on a 0-42 scale, p=0.036) domains of the McGill Pain Questionnaire (MPQ) in one fair-quality trial; there were no differences for the affective domain of the MPQ or for VAS pain (SOE: Low).
- Data were insufficient for harms; however, no adverse effect occurred in one fair-quality trial (SOE: Insufficient)

Detailed Synthesis

Two trials (n=64, 94)[147,148] evaluating myofascial release therapy versus sham therapy for fibromyalgia met inclusion criteria (Table 38 and Appendix D). Mean patient ages were 48 and 55 years. Baseline pain history characteristics were poorly described in both trials. The duration of myofascial release therapy was 20 weeks in both trials; sessions ranged in length from 60 to 90 minutes and were conducted twice or once a week, respectively. The sham conditions included short-wave and ultrasound electrotherapy or sham (disconnected) magnotherapy. Both trials reported intermediate-term outcomes; short-term and long-term outcomes were also reported by one trial each. One trial was rated fair quality and the other poor quality (Appendix E). Unclear allocation concealment methods and lack of blinding were the major methodological shortcoming in both trials. Additionally, the poor-quality trial did not describe the randomization process employed.

Table 38. Summary of results for fibromyalgia: manual therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Castro-Sanchez, 2011a[147] 6 and 12 months Duration of pain, NR *Fair*	A. Myofascial Release (n=47): myofascial release (across 10 pain regions) administered by a physiotherapist; 60 minutes sessions twice weekly for 20 weeks B. Sham short-wave and ultrasound electrotherapy (n=47): both applied to the cervical, dorsal and lumbar regions using disconnected equipment; 30 minute sessions (10 minutes each region), twice weekly for 20 weeks	A vs. B Age: 55 vs. 54 years Female: NR Race: NR Mean duration of pain: NR FIQ total (0-100): 65.0 vs. 63.9 Pain (FIQ, 0-10): 9.2 vs. 8.9 Pain (VAS, 0-10): 9.1 vs. 8.9 MPQ sensory dimension (0-33): 19.3 vs. 19.9 MPQ affective dimension (0-12): 5.6 vs. 4.9 MPQ evaluative (sensory + affective) dimension (0-45): 24.9 vs. 25.3	A vs. B <u>6 months</u> FIQ Total: 58.6 vs. 64.1, p=0.048 FIQ pain: 8.5 vs. 8.0, p=0.042 VAS pain: 8.25 vs. 8.94, p=0.043 MPQ sensory: 17.3 vs. 20.7, p=0.042 MPQ affective: 4.5 vs. 5.2, p=0.042 MPQ evaluative: 21.9 vs. 26.2, p=0.022 <u>12 months</u> FIQ Total: 62.8 vs. 65.0, p=0.329 FIQ pain: 8.8 vs. 8.7, p=0.519 VAS pain: 8.74 vs. 8.92, p=0.306 MPQ sensory: 18.2 vs. 21.2, p=0.038 MPQ affective: 4.8 vs.	A vs. B <u>6 months</u> Clinical Global Impression Severity Scale (Likert, 1-7): 5.3 vs. 6.0, p=0.048 Clinical Global Impression Improvement Scale (Likert, 1-7): 5.6 vs. 6.3, p=0.046 <u>12 months</u> Clinical Global Impression Severity Scale: 5.5 vs. 6.2 p=0.147 Clinical Global Impression Improvement Scale: 5.8 vs. 6.5, p=0.049 p-values are from authors' ANOVA[b]

			5.1, p=0.232 MPQ evaluative: 23.2 vs. 26.7, p=0.036 p-values are from authors' ANOVA[b]	
Castro-Sanchez, 2011b[148] 1 and 6 months Duration of pain, NR *Poor*	A. Massage-Myofascial Release (n=32): Massage-Myofascial release therapy (across 18 pain regions) administered by a physiotherapist; weekly 90-minute session for 20 weeks. B. Sham magnotherapy (n=32): weekly 30-minute session of disconnected magnotherapy (applied on cervical and lumbar area for 15 minutes each) for 20 weeks.	A vs. B Age: 49 vs. 46 years Female: 94% vs. 96% Race: NR Mean duration of pain: NR Pain Intensity (VAS, 0-10)[c]: 9.1 vs. 9.6	A vs. B <u>1 month</u> VAS pain[c]: 8.4 vs. 9.4, p<0.043 <u>6 months</u> VAS pain[c]: 8.8 vs. 9.7, p=ns p-values are from authors' ANOVA[b]	A vs. B <u>1 month</u> STAI state anxiety (20-80)[c]: 21.5 vs. 22, p=ns STAI trait anxiety (20-80)[c]: 25.1 vs. 26.3, p=ns BDI (0-63)[c]: 2.1 vs. 2.5, p=ns SF-36 physical function (0-100): 46.8 vs. 49.6, p=0.049 SF-36 physical role (0-100): 24.6 vs. 29.0, p=0.047 SF-36 bodily pain (0-100): 75.1 vs. 89.9, p=0.046 SF-36 general health (0-100): 66.8 vs. 68.4, p=0.093 SF-36 vitality (0-100): 61.6 vs. 59.2, p=0.055 SF-36 social function (0-100): 60.6 vs. 63.6, p=0.081 SF-36 emotional role (0-100): 50.5 vs. 47.0, p=0.057 SF-36 mental health (0-100): 75.0 vs. 78.3, p=0.082 PSQI, sleep duration, p=0.041[d]: patients with severe problems, 60% vs. 83%; moderate problems, 37% vs. 10%; and no problems, 3% vs. 7% <u>6 months</u> BDI[c]: 2.3 vs. 2.5, p=ns STAI state anxiety[c]: 22.0 vs. 23.0, p=ns STAI trait anxiety[c]: 25.8 vs. 26.2, p=ns SF-36 physical function: 48.2 vs. 51.2, p=0.281 SF-36 physical role: 25.5 vs. 27.5, p=0.213 SF-36 body pain: 75.6 vs. 77.8, p=0.293 SF-36 general health: 67.5 vs. 68.1, p=0.401 SF-36 vitality: 62.2 vs. 58.9, p=0.312 SF-36 social function: 61.3 vs. 63.9, p=0.088

				SF-36 emotional role: 49.1 vs. 46.9, p=0.219 SF-36 mental health: 76.5 vs. 80.0, p=0.126 PSQI, sleep duration, p=0.047[d]: patients with severe problems, 57% vs. 93%; moderate problems, 37% vs. 0%; and no problems, 7% vs. 7% p-values are from authors' ANOVA[b]

ANOVA, repeated-measures analysis of variance; BDI: Beck Depression Inventory; FIQ, Fibromyalgia Impact Questionnaire; MPQ: McGill Pain Questionnaire; NR, not reported; ns, not statistically significant; PSQI, Pittsburgh sleep quality index; SF-36, Short-Form 36 health questionnaire; STAI, State-Trait Anxiety Inventory; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Changes in scores were analyzed by using a 2 (groups: experimental and placebo) X 4 (time points: baseline, immediately postintervention, at 1 and 6 months) repeated-measures analysis of variance

[c] Values estimated from figures in the article.

[d] For all other dimensions of the PSQI (subjective sleep quality, sleep latency, habitual sleep efficiency, sleep disturbance, daily dysfunction), there were no statistically significant difference between groups in the proportion of patients experiencing severe, moderate or no problems in the authors' ANOVA.

Myofascial Release Therapy Compared With Sham

Myofascial release therapy was associated with a slightly greater effect on intermediate-term function compared with sham as measured by the FIQ (58.6 ± 16.3 vs. 64.1 ± 18.1 on a 100 point scale, p=0.048 for group by time repeated measures ANOVA) in one fair-quality trial[147]; this effect did not persist to the long term (62.8 ± 20.1 vs. 65.0 ± 19.8, p=0.329, at 12 months). Function was not reported over the short term.

Regarding pain outcomes, one poor-quality trial reported a small effect for myofascial release compared with sham therapy over the short-term (8.4 vs. 9.4 on a 0-10 VAS at 1 month, p=0.048 for group by time repeated measures ANOVA).[148] Intermediate-term results were inconsistent across the trials as measured on a 0 to 10 VAS pain scale with one fair-quality trial reporting a slightly greater effect for myofascial release versus sham (8.25 ± 1.13 vs. 8.94 ± 1.34, p=0.043)[147] at 6 months and the other (poor quality) reporting no significant difference between groups (8.8 vs. 9.7, p=ns) (Figure 43).[148] Additional pain measures were reported over the intermediate-term by the fair-quality trial, all of which showed a small benefit in favor of myofascial release: FIQ pain (8.5 ± 0.7 vs. 8.0 ± 1.3, p=0.042 for group by time repeated measures ANOVA) and the McGill Pain Questionnaire (MPQ) sensory (17.3 ± 7.8 vs. 20.7 ± 7.1 on a 0-33 scale, p=0.04), affective (4.5 ± 2.9 vs. 5.2 ± 3.8 on a 0-12 scale, p=0.04) and evaluative (21.9 ± 7.2 vs. 26.2 ± 6.8 on a 0-42 scale, p=0.02) dimensions.[147] This effect persisted at long-term followup for the sensory and evaluative dimension of the MPQ only; no differences were seen between groups regarding VAS pain of the affective dimension of the MPQ at long term following in this trial (Table 38).

Depression, anxiety, and sleep outcomes were evaluated in one poor-quality trial, with significant improvement seen in the myofascial release versus the sham group on some subscales of the Short-Form-36 and on the sleep duration subscale of the Pittsburgh Sleep Quality Index (PSQI) over the short-term,[148] but no differences between groups on the STAI or Beck

Depression Index (Table 38); at intermediate followup, only PSQI sleep duration remained significantly improved following myofascial release versus sham.

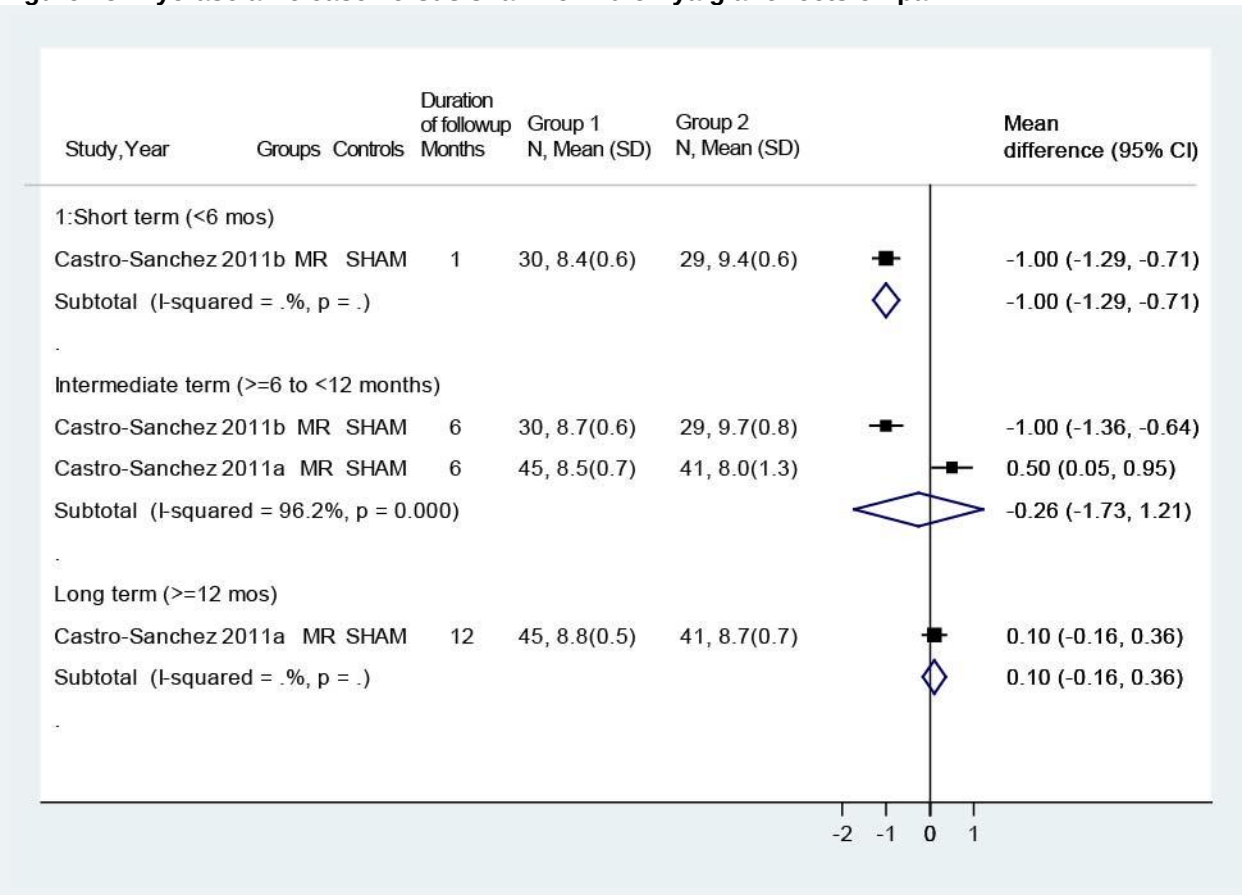
Manual Therapy Compared With Pharmacological Therapy or Exercise

No trial of manual therapy versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

In one trial, no patient experienced an adverse effect (details not reported).[147] There was no information on harms reported by the other trial.

Figure 43. Myofascial release versus sham for fibromyalgia: effects on pain



CI = confidence interval; MR = myofascial release; SD = standard deviation.

Mindfulness-Based Stress Reduction Therapy for Fibromyalgia

Key Points

- No clear short-term effects of mindfulness-based stress reduction (MBSR) were seen on function compared with waitlist or attention control (difference 0 to 0.06 on a 0-10 scale) in two trials (one fair and one poor quality) (SOE: Moderate).
- No clear short-term effects of MBSR on pain (difference 0.1 on a 0-100 VAS pain scale in one poor quality trial; difference -1.38 to -1.59 on the affective and -0.28 to -0.71 on the sensory dimension [scales not reported] of the Pain Perception Scale in one fair-quality trial) compared with waitlist or attention control in two trials (SOE: Moderate). Intermediate-term and long-term outcomes were not reported.
- No trial of MBSR versus pharmacological therapy or versus exercise met inclusion criteria.
- Harms were not reported.

Detailed Synthesis

We identified two trials (3 publications) of mindfulness-based stress reduction (MBSR) for fibromyalgia that met inclusion criteria (Table 39 and Appendix D).[159-161] One study was conducted in the United States[159,161] and the other in Germany.[160] In both trials, MBSR was modeled after the program developed by Kabat-Zinn. The intervention lasted 8 weeks, with weekly 2.5-hour sessions, daily homework assignments, and a single 7-hour session. Sample sizes ranged from 91 to 177, age ranged from 48 to 53 years, and all participants were female. Both studies compared MBSR versus waitlist control; the German study[160] also compared MBSR to an attention control group that consisted of education, relaxation, and stretching. Both studies reported only short-term outcomes.

One study was considered fair quality[160] and the other was considered poor quality[159,161] (Appendix E). Methodological shortcomings in both studies were the lack of long-term followup and the inability to blind patients and providers. The poor-quality study also had a high rate of overall attrition as well as differential attrition between the groups.

Table 39. Summary of results for fibromyalgia: mindfulness-based stress reduction therapy

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Cash 2015,[159] Sephton 2007[b] 2 months Duration of pain NR *Poor*	<u>A. Mindfulness-based Stress Reduction (n=51)</u> 8-week group-based program with one 2.5 hour session/week including instruction in techniques, meditation, and simple yoga positions to encourage relaxation. Participants were asked to complete daily practices with	A vs. B Age: 48 vs. 48 years Female: 100% vs. 100% Caucasian: 94% vs. 93% FIQ Physical Functioning (0-10): 1.3 vs. 1.2 Pain VAS (0-100): 68.1 vs. 69.2 FIQ Severity (0-100)[c]: 67.5 vs. 62.5	A vs. B <u>2 months:</u> FIQ Physical Functioning: 1.2 vs. 1.2; difference 0.0 (95% CI -0.32 to 0.32) Pain VAS: 65.2 vs. 65.1; difference 0.1 (95% CI -9.96 to 10.16) FIQ Severity[c]: 62.0 vs. 66.7; difference -4.7 (95% CI -12.24 to 2.84)	A vs. B <u>2 months</u> BDI Total[b]: 13.3 vs. 14.8; difference -1.5 (95% CI -4.76 to 1.76) BDI Cognitive Subscale[b]: 5.3 vs. 6.4; difference -1.1 (95% CI -2.98 to 0.78) BDI Somatic Subscale[b]: 7.4 vs. 7.7; difference -0.3 (95% CI -1.73 to 1.13) PSS: 20.2 vs. 20.8; difference -0.60 (95%

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	workbook and audiotapes for 45 min a day for 6 days a week. <u>B. Waitlist control group (n=39)</u> Participants were offered the intervention program only after the conclusion of the study and followup.			CI -3.37 to 2.17) SDQ: 8.4 vs. 9.5; difference -1.10 (95% CI -2.58 to 0.38) FSI: 5.5 vs. 6.0; difference -0.50 (95% CI -1.28 to 0.28)
Schmidt, 2011[160] 2 months Duration of fibromyalgia, years: 14 years *Fair*	<u>A. Mindfulness-based Stress Reduction (n=53)</u> 8-week group-based program with one 2.5 hour session/week and one 7 hour all-day session covering training in specific exercises and topics of mindfulness practices. Participants were asked to complete daily practices of 45-60 minutes each <u>B. Active-control Intervention (n=56)</u> Controlled for nonspecific aspects of the MBSR program with similar meeting structure and format to MBSR treatment arm. Equivalent levels of social support and weekly topical education was provided along with Jacobson Progressive Muscle Relaxation training and fibromyalgia-specific gentle stretching exercises. Participants were asked to complete daily homework assignments with the same duration as MBSR group. <u>C. Waitlist (n=59)</u> Received no active	A vs. B vs. C Age: 53 vs. 52 years Female: 100% (all female study) Race: NR A vs. C FIQ Total (0-10): 5.8 vs. 5.7; PPS Affective (scale unclear): 35.5 vs. 34.8 PPS Sensory (scale unclear): 22.4 vs. 22.6	A vs. B <u>2 months</u> Proportion of Patients with >14% improvement in FIQ scores (MCID): 30% vs. 25%; RR 1.21 (95% CI 0.79 to 1.82) FIQ: 5.23 vs. 5.33; MD -0.10 (95% CI -0.84 to 0.64) PPS Affective: 30.79 vs. 32.17; MD -1.38 (95% CI -4.79 to 2.03) PPS Sensory: 21.16 vs. 21.87; MD -0.71 (95% CI -2.77 to 1.34) A vs. C <u>2 months</u> Proportion of Patients with >14% improvement in FIQ scores (MCID): 30% vs. 22%; RR 1.37 (95% CI 0.83 to 1.94) FIQ: 5.23 vs. 5.29; MD -0.06 (95%CI -0.75 to 0.63) PPS Affective: 30.79 vs. 32.38; MD -1.59 (95%CI -5.01 to 1.83) PPS Sensory: 21.16 vs. 21.44; MD -0.28 (95%CI -2.30 to 1.74)	A vs. B <u>2 months</u> Proportion of Patients who saw Clinically Relevant Improvement (score of <23) in CES-D scores: 28% vs. 23%; RR 0.53 (95% CI 0.54 to 1.12) CES-D: 21.70 vs. 22.55; MD -0.85 (95%CI -4.66 to 2.96) STAI Trait Subscale: 47.86 vs. 48.44; MD -0.58 (95%CI -4.42 to 3.26) Proportion of Patients with PSQI score <5 indicates good sleep): 17%vs. 7%; RR 2.38 (95% CI 0.85 to 2.34) PSQI: 10.01 vs. 10.25; MD -0.24 (95%CI -1.71 to 1.23) FMI: 37.66 vs. 35.14; MD 2.52 (95%CI 0.04 to 5.00) GCQ: 42.63 vs. 43.91; MD -1.28 (95% CI -6.51 to 3.95) PLC: 12.83 vs. 12.16; MD 0.67 (95% CI -0.60 to 1.94) A vs. C <u>2 months</u> Proportion of Patients who saw Clinically Relevant Improvement (score of <23) in CES-D scores: 28% vs. 19%; RR 1.52 (95% CI 0.85 to 2.04)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	treatment but were offered either intervention at the conclusion of the followup period.			CES-D: 21.7 vs. 24.0; MD -2.3 (95% CI -5.96 to 1.36) STAI Trait Subscale: 47.9 vs. 49.2; MD -1.32 (95% CI -5.02 to 2.38 Proportion of Patients with PSQI score <5 indicates good sleep): 17% vs. 10%; RR 1.67 (95% CI 0.80 to 2.14) PSQI: 10.0 vs. 10.4; MD -0.36 (95%CI -1.8 to 1.1) FMI: 37.7 vs. 36.1; MD 1.5 (95%CI -0.9 to 3.91) GCQ: 42.6 vs. 45.3; MD -2.7 (95%CI -7.8 to 2.5) PLC: 12.8 vs. 12.3; MD 0.5 (95% CI -0.7 to 1.7)

BDI = Beck Depression Inventory; CES-D = Center for Epidemiological Studies Depression Scale; CI = confidence interval; FSI= Fatigue Symptom Inventory; FIQ = Fibromyalgia Impact Questionnaire; FMI = Freiburg Mindfulness Inventory; GCQ = Giessen Complaint Questionnaire; PLC = Profile for the Chronically Ill; PPS = Pain Perception Scale; PSQI = Pittsburgh Sleep Quality Index; PSS = Perceived Stress Scale; SDQ = Stanford Sleep Disorders Questionnaire; STAI = State-Trait-Anxiety-Inventory; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Sephton is the same population as Cash 2015 but the focus of the study was on depression (Beck Depression Inventory).

[c] FIQ symptom severity is comprised of visual analog ratings of pain, fatigue, morning sleepiness, stiffness, anxiety, and depression.

Mindfulness-Based Stress Reduction Therapy Compared With Waitlist or Attention Control

There were no clear short-term effects of MBSR on any function or pain measure reported compared with waitlist or attention control. Both trials compared MBSR to waitlist and reported function using the FIQ, one reporting the physical function subscale (difference 0 on a 0-10 scale, 95% CI -0.32 to 0.32)[159] and the other reporting the total score (difference -0.06 on a 0-10 scale, 95% CI -0.75 to 0.63).[160] The latter fair-quality trial also reported the proportion of patients who achieved a 14percent or greater improvement in FIQ total scores: 30 percent versus 22 percent, RR 1.37 (95% CI 0.83 to 1.94).[160] Regarding pain, one trial reported a mean difference of 0.1 (95% CI -9.96 to 10.16) on a 0 to 100 VAS pain scale[159] between the MBSR and waitlist groups, while the other reported pain using the affective (difference -1.59, 95% CI -5.01 to 1.83) and sensory (difference -0.28, 95% CI -2.30 to 1.74) domains of the Pain Perception Scale (scale not reported).[160] Estimates for function and pain were similar for the comparison of MBSR versus attention control in the fair-quality trial[160] (Table 39).

Secondary outcomes (measures of depression, anxiety, sleep, fatigue) did not differ significantly between MBSR and waitlist or attention control in either trial[159-161] (Table 39). The fair-quality trial compared medication use (pain killers, anti-depressants, and sleep medication) between baseline and short-term followup; only antidepressant medication was reduced significantly from baseline (46% to 35%, p=0.01) but there was no group effect (data not reported).[160]

Mindfulness-Based Stress Reduction Therapy Compared With Pharmacological Therapy or Exercise

No trial of MBSR versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

Neither trial reported harms.

Mind-Body Therapy for Fibromyalgia

Key Points

- Over the short-term, two trials of mind-body practices reported slight improvement in function for qigong (mean difference -7.5, 95% CI -13.3 to -1.68) and substantial improvement for tai chi (mean difference -23.5, 95% CI -30 to -17) based on 0 to 100 scale total FIQ score; heterogeneity may be explained by duration and intensity of intervention and control condition (SOE: Low).
- Qigong and tai chi were associated with moderately greater improvement in pain (0-10 scale) compared with waitlist and attention control in the short term (2 trials, pooled MD -1.54, 95% CI -2.67, -0.41, I^2=75%) (SOE: Low).
- No evidence in the intermediate or long term.
- Data for harms were insufficient. However, one trial reported two adverse events (in two patients) judged to be possibly related to qigong practice: an increase in shoulder pain and plantar fasciitis; neither participant withdrew from the study. In the trial of tai chi, no adverse events were reported. (SOE: Insufficient)

Detailed Synthesis

Two trials[171,172] that evaluated mind-body therapies for fibromyalgia met inclusion criteria (Table 40 and Appendix D). Across trials, the participants were predominately female (87% to 96%), with mean ages between 51 to 52 years. Prior to study enrollment, participants in both trials were being treated with several drugs from major analgesic and adjuvant drug groups such as analgesics/NSAIDs (53% to 73%), antidepressants (35% to 48%), and anticonvulsants (21% to 27%); in one trial, approximately 30 percent of participants were taking opioids and many participants had tried a variety of other therapies (including acupuncture, chiropractic, naturopathic/homeopathic/osteopathic therapies, massage therapy, and psychological therapies).[171]

One trial compared qigong (3 consecutive half-day training sessions, then weekly practice/review sessions for 8 weeks plus daily at-home practice for 45 to 60 minutes) to a waiting list control condition.[171] The other trial compared tai chi (60-minute sessions twice per week for 12 weeks) to an attention control condition (40 minutes of wellness education and 20 minutes of supervised stretching exercises).[172] In both trials, patients were instructed to continue

the practice at home throughout the followup period. In the tai chi study, the average percent of sessions attended during the 12-week intervention was 77 percent for the tai chi group and 70 percent for the control group.[172] In the qigong trial, the mean self-reported practice time per week for all participants who completed the trial was 4.9 hours at 2 months, 2.9 hours at 4 months, and 2.7 hours at 6 months. [171]

Both trials were rated fair quality (Appendix E). Due to the nature of the intervention and control groups, blinding was not possible in the two studies. Other methodological concerns included differential attrition between groups in the qigong trial (qigong 19% vs. waitlist 4% at 6 months).[171]

Table 40. Summary of results for fibromyalgia: mind-body therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Lynch, 2012[171] (N=100) 4 months Duration of fibromyalgia, mean: 9.6 years *Fair*	A. Qigong (n=53) Chaoyi Fanhuan Qigong; Three consecutive half-day training sessions then weekly practice/review sessions for 8 weeks plus daily at-home practice for 45 to 60 minutes. B. Waitlist (n=47) Continued with usual care; offered qigong after the trial ended	A vs. B Age: 53 vs. 52 years Female: 94% vs. 98% Previous opioid therapy: 42% vs. 30% Current opioid therapy: 36% vs. 23% Current NSAID therapy: 49% vs. 57% FIQ (0-100): 65.5 vs. 61.8 NRS pain (0-10): 6.5 vs. 6.6 SF-36 PCS (0-100): 30.0 vs. 32.6 SF-36 MCS (0-100): 38.1 vs. 40.4 PSQI (0-21): 13.8 vs. 13.1	A vs. B 4 months Mean change from baseline: FIQ: -16.1 vs. -4.8; difference -11.3 (95% CI -19.3 to -3.3) NRS pain: -1.21 vs. -0.27; difference -0.9 (95% CI -1.7 to -0.1)	A vs. B 4 months Mean change from baseline: SF-36 PCS: 4.6 vs. 0.2; difference 4.4 (95% CI 1.5 to 7.3) SF-36 MCS: 4.4 vs. 0.7; difference 3.7 (95% CI -0.3 to 7.7) PSQI: -3.3 vs. -1.1; difference -2.2 (95% CI -3.6 to -0.8)
Wang, 2010[172] (N=66) 3 months Duration of fibromyalgia pain: 11 years *Fair*	A. Tai chi (n=33) Classic Yang style tai chi; at home practice for at least 20 minutes a day; encouraged to maintain tai chi practice using an instructional video. B. Attention control (n=33) 40 minutes of education then 20 minutes of supervised stretching (upper body, trunk, and lower body); plus 20 minutes of daily at-home stretching Both groups had 60-minute sessions twice a week for 12 weeks and	A vs. B Age: 50 vs. 51 years Female: 85% vs. 88% Analgesic use: 88% vs. 73% FIQ (0-100): 62.9 vs. 68.0 VAS pain (0-10): 5.8 vs. 6.3 CES-D (0-60): 22.6 vs. 27.8 SF-36 PCS (0-100): 28.5 vs. 28.0 SF-36 MCS (0-100): 42.6 vs. 37.8 PSQI (0-21): 13.9 vs. 13.5	A vs. B 3 months Proportion with clinically meaningful improvement: FIQ[b]: 81.8% vs. 51.5%; RR 1.6 (95% CI 1.1 to 2.3) VAS pain[c]: 54.5% vs. 27.3%; RR 2.0 (95% CI 1.1 to 3.8) Mean change from baseline: FIQ: -28.6 vs. -10.2; difference -18.3 (95% CI -27.1 to -9.6) VAS pain: -2.4 vs. -0.7; difference -1.7 (95% CI -2.7 to -0.8)	A vs. B 3 months Proportion with clinically meaningful improvement: CES-D[d]: 69.7% vs. 39.4%; RR 1.8 (95% CI 1.1 to 2.9) SF-36 PCS[e]: 51.5% vs. 15.2%; RR 3.4 (95% CI 1.4 to 8.1) SF-36 MCS[f]: 48.5% vs. 24.2%; RR 2.0 (95% CI 1.0 to 4.0) PSQI[g]: 45.5% vs. 18.2%; RR 2.5 (95% CI 1.1 to 5.6) Mean change from baseline: CES-D: -6.5 vs. -2.4;

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	continued regular medications and routine activities.			difference -4.1 (95% CI -8.2 to 0.1) SF-36 PCS: 8.4 vs. 1.5; difference 7.0 (95% CI 2.9 to 11.0) SF-36 MCS: 8.5 vs. 1.2; difference 7.3 (95% CI 1.9 to 12.8) PSQI: -4.2 vs. -1.2; difference -3.0 (95% CI -5.2 to -0.9)

CES-D, Center for Epidemiologic Studies Depression index; CI: confidence interval; FIQ, Fibromyalgia Impact Questionnaire; NRS, numerical rating scale; NSAIDs: Non-steroidal anti-inflammatory drugs; PSQI, Pittsburgh Sleep Quality Index; RR: risk ratio; SF-36 MCS, Short-Form-36 Mental Component Summary; SF-36 PCS, Short-Form-36 Physical Component Summary; VAS, visual analog scale.

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] A reduction of ≥8.1 points from baseline on the FIQ was considered a clinically meaningful improvement.

[c] A reduction of ≥2 points from baseline on the VAS was considered a clinically meaningful improvement.

[d] A reduction of ≥6 points from baseline on the CES-D was considered a clinically meaningful improvement.

[e] An increase of ≥6.5 points from baseline on the SF-36 PCS was considered a clinically meaningful improvement.

[f] An increase of ≥7.9 points from baseline on the SF-36 MCS was considered a clinically meaningful improvement.

[g] A reduction of >5 points from baseline on the PSQI was considered a clinically meaningful improvement.

Mind-Body Therapies Compared With Waitlist or Attention Control

Short-term improvement in function on 0 to 100 scale total FIQ score was reported for qigong (slight improvement, mean difference -7.5, 95% CI -13.3 to -1.68)[171] and for tai chi (substantial improvement, mean difference -23.5, 95% CI -30 to -17)[172] compared with waitlist or attention control. Substantial heterogeneity (I^2 = 92%), precluded meaningful pooling for this outcome (Figure 44). Significantly more participants in the tai chi group also showed clinically meaningful improvement (reduction of ≥8.1 points from baseline) on total FIQ (RR 1.6, 95% CI 1.1 to 2.3). Tai chi and qigong were associated with a moderate improvement in NRS pain (0 to 10 scale) compared with wait list or attention control (2 trials pooled MD -1.54 (95%CI -2.67, -0.41, I^2 = 75%) (Figure 45). Heterogeneity may in part be due to differences in duration and intensity of the intervention.

Mind-body therapy resulted in significant improvement in most secondary outcomes measured. Participants who received tai chi group showed clinically meaningful improvement in depressive symptoms as measured by the Center for Epidemiological Studies-Depression scale (RR 1.8, 95% CI 1.1 to 2.9), in sleep quality as measured by the Pittsburg Sleep Quality Index (PSQI) (RR 2.5, 95% CI 1.1 to 5.6), and in quality of life as measured by the SF-36 PCS (RR 3.4, 95% CI 1.4 to 8.1) and MCS (RR 2.0, 95% CI 1.0 to 4.0) compared with controls; similar results were seen for mean followup scores on these measures (Table 40).[172] In the second trial,[171] compared to a waitlist control, qigong resulted in significantly improved quality of life as measured by the SF-36 PCS (difference in change from baseline 4.4, 95% CI 1.5 to 7.3] and in

sleep quality as measured by the PSQI (difference in change from baseline -2.2, 95% CI -3.6 to -0.8). The change in SF-36 MCS scores did not differ between groups.

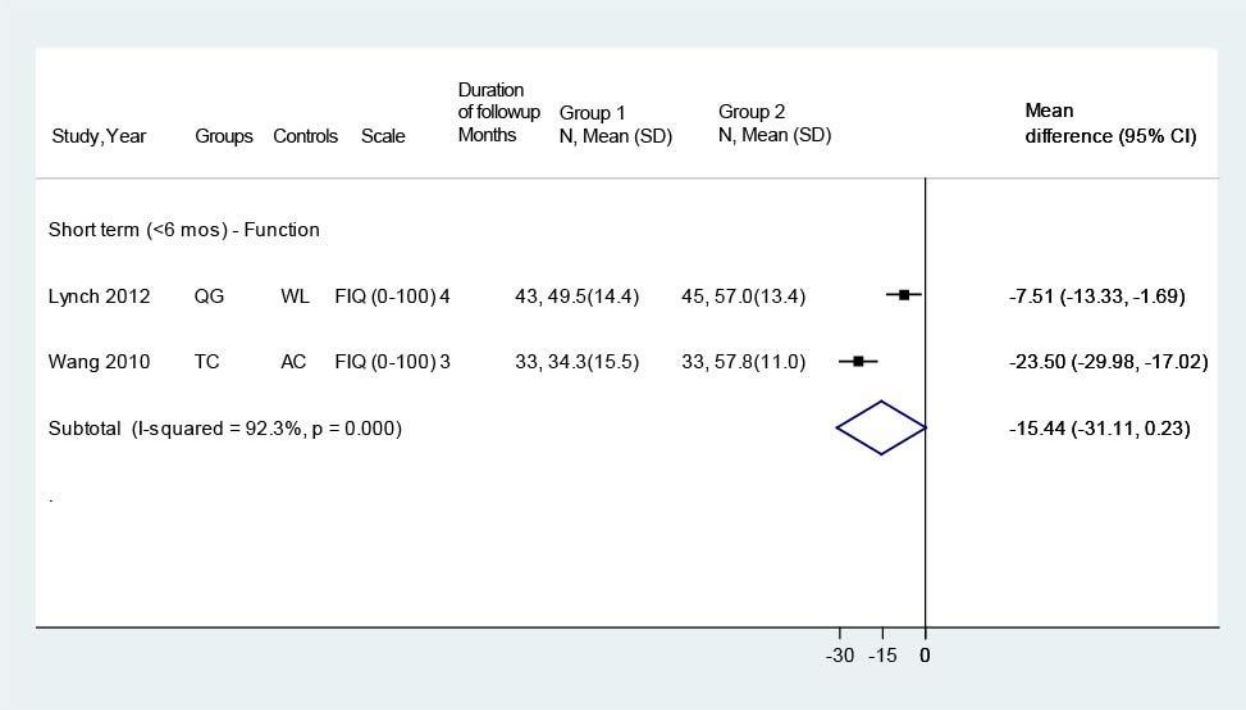
Mind-Body Therapies Compared With Pharmacological Therapy or Exercise

No trials comparing mind-body therapies with pharmacological therapy or with exercise met inclusion criteria.

Harms

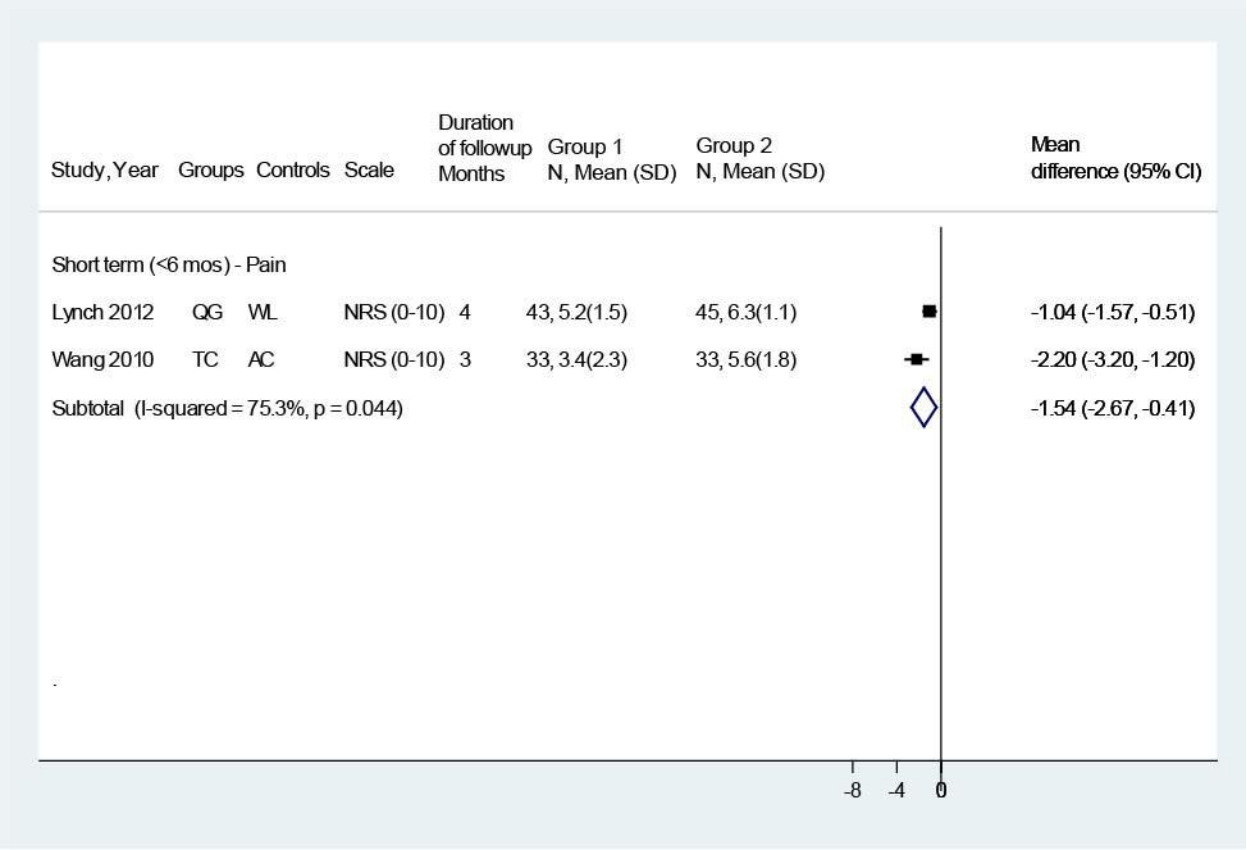
In the trial of qigong,[171] there were two adverse events judged to be possibly related to the practice. One participant reported an increase in shoulder pain and another experienced plantar fasciitis; neither participant withdrew from the study. In the trial of tai chi, no adverse events were reported.[172]

Figure 44. Mind-body therapies for fibromyalgia: effects on function



AC = attention control; CI = confidence interval; FIQ = Fibromyalgia Impact Questionnaire; QG = qigong; SD = standard deviation; TC = tai chi; WL = waitlist

Figure 45. Mind-body therapies for fibromyalgia: effects on pain



AC = attention control; CI = confidence interval; QG = qigong; SD = standard deviation; TC = tai chi; WL = waitlist.

Acupuncture for Fibromyalgia

Key Points

- Acupuncture was associated with slightly greater effects on function based on 0 to 100 FIQ Total Score in patients with fibromyalgia than sham acupuncture in the short-term (2 trials, pooled difference -8.63, 95% CI =12.12 to -5.13, I^2 = 0%) and intermediate-term (2 trials, pooled difference -9.41, 95% CI -13.96 to -4.85, I^2 = 27.4%) (SOE: Moderate).
- There was no clear effect of acupuncture on pain (0 to 10 scale) versus sham acupuncture in the short term (3 trials, pooled difference -0.13, 95% CI -1.06 to 0.79, I^2 =72%) or intermediate term (3 trials, pooled difference – 0.53, 95%CI -1.15 to 0.09 I^2 =45.5%) (SOE: Low).
- No data on long-term effects were reported (SOE: Insufficient).
- Discomfort and bruising were the most common adverse events. Discomfort was substantially more common for acupuncture or sham needling (61% to 70%) compared with simulated acupuncture (29%). Vasovagal symptoms and aggravation of fibromyalgia symptoms were less common (4%, 2.5 of sessions) (SOE: Moderate).

Detailed Synthesis

Three trials of acupuncture for fibromyalgia were identified that met inclusion criteria; one was conducted in Spain[199] and two were conducted in the United States[197,198] (Table 41 and Appendix D). Two trials evaluated traditional Chinese needle acupuncture[197,199] and the third evaluated acupuncture with electrical stimulation.[198] All three studies compared acupuncture to sham. One study[197] employed three different types of sham treatments (needling for an unrelated condition, sham needling, and simulated acupuncture), one used sham needling[198] and one used simulated acupuncture.[199] Sample sizes ranged from 50 to 164 (total sample=314), mean ages from 47 to 53 years, and the proportion of females ranged from 95 percent to 100 percent. The duration of acupuncture treatment ranged from 3 to 12 weeks, with the total number of sessions ranging from 6 to 24. All studies reported short-term and intermediate-term outcomes; no trial had long-term followup.

All three studies were considered good quality (Appendix E). Limitations of all studies were lack of long-term followup and that the person administering acupuncture was not blinded to treatment allocation.

Table 41. Summary of results for fibromyalgia: acupuncture

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Assefi, 2005[197] 3 and 6 months Mean duration of pain: 9 to 12 years *Good*	A. Acupuncture (n=25): in accordance with Traditional Chinese Medicine B. Sham Acupuncture (n=24): Needling for Unrelated Condition C. Sham Acupuncture (n=24): Sham Needling D. Sham Acupuncture (n=23): Simulated Acupuncture Treatment protocol: 24 sessions (2/week for 12 weeks)	A vs. B vs. C vs. D Mean age: 46 vs. 46 vs. 49 vs. 48 years Female: 88% vs. 96% vs. 100% vs. 96% Race (white): 96% vs. 88% vs. 96% vs. 92% Mean duration of pain: 12 vs. 9 vs. 9 vs. 10 years Pain Intensity VAS (0-10): 7.0 vs. 6.9 vs. 6.8 vs. 7.3	A. vs. B vs. C vs. D <u>3 months</u> Pain Intensity VAS[b]: 6.0 vs. 5.4 vs. 5.4 vs. 4.5 <u>6 months</u> Pain Intensity VAS[b]: 5.7 vs. 6.0 vs. 5.2 vs. 5.2 A vs. B+C+D <u>Across all time-points[c]</u> Pain intensity VAS: adjusted MD 0.5, (95% CI -0.3 to 1.2)	A. vs. B vs. C vs. D <u>3 months</u> SF-36 PCS (0-100)[b]: 31 vs. 39 vs. 31.5 vs. 40 SF-36 MSC (0-100)[b]: 46 vs. 46.5 vs. 48.5 vs. 47 Sleep Quality VAS (0-10)[a]: 4.3 vs. 4.1 vs. 5.2 vs. 5.5 Overall Well-Being VAS (0-10)[b]: 4.9 vs. 4.9 vs. 5.0 vs. 6.3 <u>6 months</u> SF-36 PCS[b]: 31 vs. 36 vs. 31. vs. 39 SF-36 MCS[b]: 43 vs. 45 vs. 50 vs. 46.5 Sleep Quality VAS[b]: 4.3 vs. 3.4 vs. 5.4 vs. 5.5 Overall Well-Being VAS[b]: 4.6 vs. 4.6 vs. 5.7 vs. 5.7 A vs. B+C+D <u>Across all time-points[c]</u> SF-36 PCS: adjusted MD -0.4 (95% CI -2.3 to 1.5) SF-36 MCS: adjusted MD -1.5, (95% CI -4.0 to 1.0) Sleep Quality VAS: adjusted MD -0.5, (95% CI -1.3 to 0.2) Overall Well-Being VAS: adjusted MD -0.3, (95% CI

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				-1.0 to 0.3)
Martin, 2006[198] 1 and 7 months Duration of pain: NR *Good*	A. Acupuncture (n=25) B. Sham Acupuncture: Sham Needling (n=25) Treatment protocol: 6 treatments over 2 to 3 weeks	A vs. B Age: 48 vs. 52 years Female: 100% vs. 96% Race: 96% vs. 100% white FIQ total (0-80): 42.4 vs. 44.0 FIQ Physical Function (0-10): 4.1 vs. 3.6 MPI Interference (scale NR): 42.6 vs. 36.9 MPI General Activity Level (scale NR): 55.7 vs. 56.6 MPI Pain Severity (scale NR): 40.4 vs. 43.0 FIQ Pain (0-10): 6.2 vs. 6.5	A vs. B <u>1 month</u> FIQ Total: 34.8 vs. 42.2, MD -4.9 (95% CI -8.7 to -1.2) FIQ Physical Function: 3.7 vs. 3.3, MD –0.4 (95% CI –1.1 to 0.3) MPI Interference: 38.3 vs. 34.9, MD 0.1 (95% CI –3.4 to 3.6) MPI General Activity Level: 55.4 vs. 58.3, MD –1.2, (95% CI –3.8 to 1.4) MPI Pain Severity: 34.2 vs. 41.6, MD –4.6 (95% CI –8.7 to –0.5) FIQ pain: 4.7 vs. 5.9, MD –0.8, (95% CI –1.8 to 0.2) <u>7 months</u> FIQ Total: 38.1 vs. 42.7, MD –4.3 (95% CI –7.7 to –0.9) FIQ Physical Function: 3.5 vs. 3.3, MD –0.3 (95% CI –0.9 to 0.3) MPI Interference: 37.7 vs. 35.5, MD 0.1 (95% CI –3.2 to 3.4) MPI General Activity Level: 58.1 vs. 59.5, MD –0.6 (95% CI –3.1 to 1.8) MPI Pain Severity: 37.3 vs. 41.4, MD –3.8 (95% CI –7.5 to –0.2) FIQ Pain: 5.5 vs. 6.4, MD –0.7 (95% CI –1.5 to 0.3)	A vs. B <u>1 month</u> FIQ Anxiety (0-10): 2.6 vs. 5.1, MD –1.1 (95% CI –2.0 to –0.2) FIQ Depression (0-10): 2.0 vs. 3.7, MD –0.7 (95% CI –1.6 to 0.3) FIQ Sleep (0-10): 5.9 vs. 6.8, MD –0.7 (95% CI –1.8 to 0.5) FIQ Well-Being (0-10): 4.6 vs. 3.1, MD 0.8 (95% CI –0.4 to 2.0) <u>7 months</u> FIQ Anxiety: 3.3 vs. 4.8, MD –1.1 (95% CI –1.9 to –0.2) FIQ Depression: 2.2 vs. 3.6, MD –0.7 (95% CI –1.6 to 0.2) FIQ Sleep: 6.1 vs. 6.3, MD –0.3 (95% CI –1.3 to 0.6) FIQ Well-Being: 3.8 vs. 3.6, MD 0.4 (95% CI –0.6 to 1.4)
Vas,[199] 2016 3.75 and 9.75 months Duration of pain: NR *Good*	A. Acupuncture (n=82) B. Sham Acupuncture: Simulated Acupuncture (n=82) Treatment protocol: One 20 min session per week for 9 weeks. Participants also received pharmacological	A vs. B Age: 52.3 vs. 53.2 years Female: 100% vs. 100% FIQ (0-100): 71.7 vs. 70.1 Pain Intensity VAS (0-100): 79.3 vs. 75.8	A vs. B <u>3.75 months</u> FIQ % mean relative change: −25.0 vs. −11.2, Cohen's d=0.58 Pain Intensity VAS % mean relative change: -23.6 vs. −16.6, Cohen's d = 0.28 <u>9.75 months</u> FIQ % mean relative	A vs. B <u>3.75 months</u> HDRS % mean relative change: NR SF-12 MCS % mean relative change: 30.6 vs. 13.9, Cohen's d = 0.38 SF-12 PCS % mean relative change: 37.0 vs. 15.5, Cohen's d = 0.56 <u>9.75 months</u>

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	treatment as prescribed by GP.		change (%): −22.2 vs. −4.9, Cohen's d = 0.80, Pain intensity VAS % mean relative change: −19.9 vs. −6.2, Cohen's d = 0.62	HDRS % mean relative change: −19.1 vs. −5.9, Cohen's d = 0.22 SF-12 PCS % mean relative change: 37.2 vs. 11.4, Cohen's d = 0.58 SF-12 MCS % mean relative change: 23.0 vs. 9.4, Cohen's d = 0.36

CI = confidence interval; FIQ = Fibromyalgia Impact Questionnaire; GP = general practitioner; HDRS = Hamilton Depression Rating Scale; MCS = Mental Component Score; MD = mean difference; MPI = Multidimensional Pain Inventory; NR = not reported; PCS = Physical Component Score; SF-12 = Short-Form-12; SF-36 = Short-Form 36; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Outcome values were estimated from graphs.

[c] Authors combined the three sham control groups and calculated the adjusted least-square mean difference between the acupuncture group and combined control groups. Treatment-by-time interaction was not included in the models; therefore data reflects results across all time-points.

Acupuncture Compared With Sham

Acupuncture was associated with slightly greater improvement in function compared with sham acupuncture based on the FIQ Total Score (0 to 100) at short-term followup (2 trials, pooled difference -8.63, 95% CI =12.12 to -5.13, I^2=0%,) and intermediate-term followup (2 trials, pooled difference on 0-100 scale, -9.41, 95% CI -13.96 to -4.85, I^2=27.4%) across the same trials[198,199] (Figure 46). There was, however, no clear effect of acupuncture on pain (0 to 10 scale) versus sham acupuncture in the short term (3 trials, pooled difference -0.13, 95% CI -1.06 to 0.79, I^2=72%) or intermediate term (3 trials, pooled difference – 0.53, 95%CI -1.15 to 0.09, I^2=45.5%)[198,199] (Figure 47). All trials were considered good quality.

Results across two trials of acupuncture versus sham were inconsistent, with each reporting effects in the opposite direction. In the trial of acupuncture versus three different types of sham acupuncture,[197] there was no significant benefit of acupuncture versus the combined sham groups on the SF-36 MCS score, a measure of sleep quality, or a measure of overall well-being. In the trial of six acupuncture treatments over 2 to 3 weeks, there was a benefit for true versus sham acupuncture at 1 and 7 months on the FIQ subscale of anxiety, but not depression, sleep, or well-being.[198] In the trial of one 20-minute session per week for 9 weeks plus pharmacological treatment as prescribed by a general practitioner, there was a benefit for true versus sham acupuncture at 1 month for the SF-12 MCS scale (mean relative change 30.6%, 95% CI 19.7% to 41.5% vs. 13.9%, 95% CI 5.4 to 22.5), Cohen's d = 0.38, p=0.01), and at 9.75 months for the Hamilton Rating Scale for Depression (mean relative change -19.1%, 95% CI -34.2% to -3.9% vs. -5.9%, 95% CI -16.6% to -4.8%, Cohen's d = 0.22, p=0.01) and the SF-12 Mental Component scale (mean relative change, 23.0%, 95% CI 13.7% to 32.4% vs. 9.4%, 95% CI 1.9% to 16.9%, Cohen's d = 0.36, p=0.01).[199]

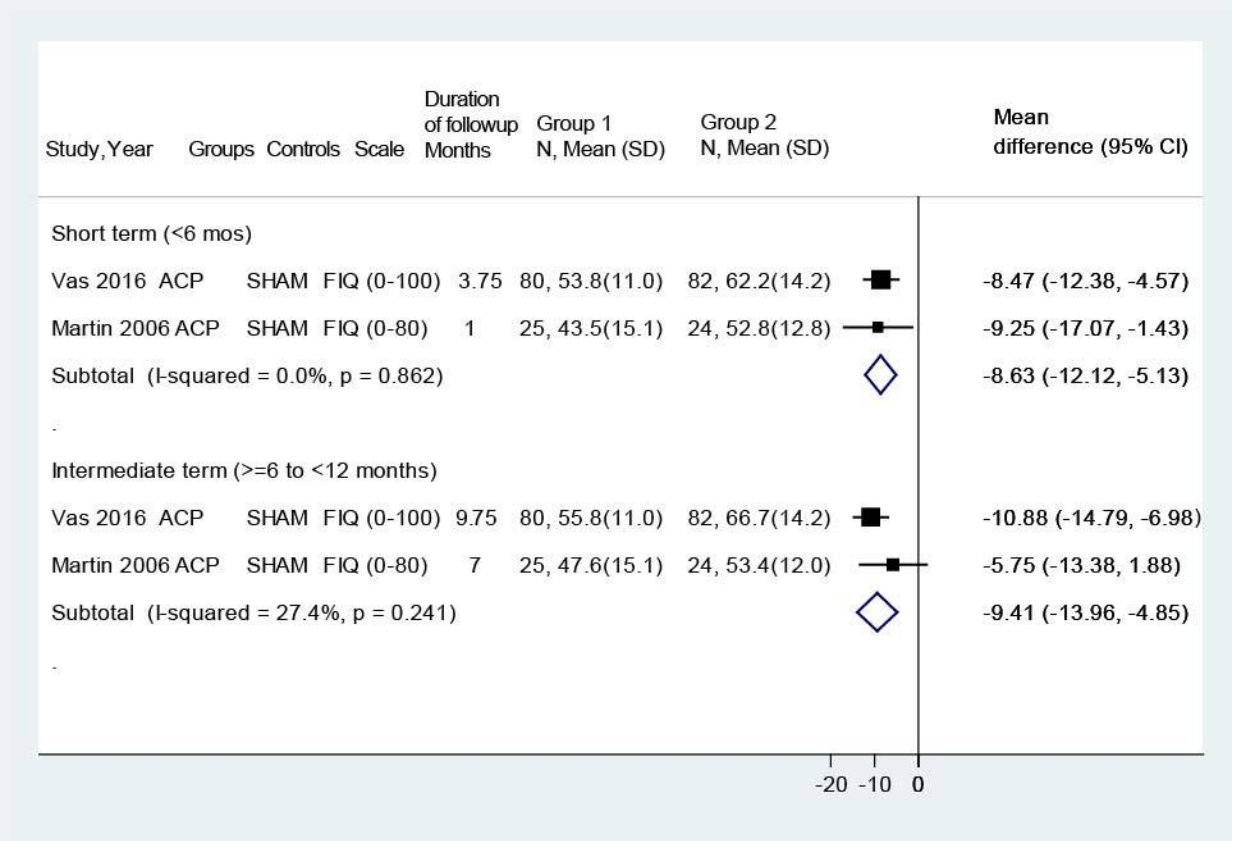
Acupuncture Compared With Pharmacological Therapy or Exercise

No trial of acupuncture versus pharmacological therapy or versus exercise met inclusion criteria.

Harms

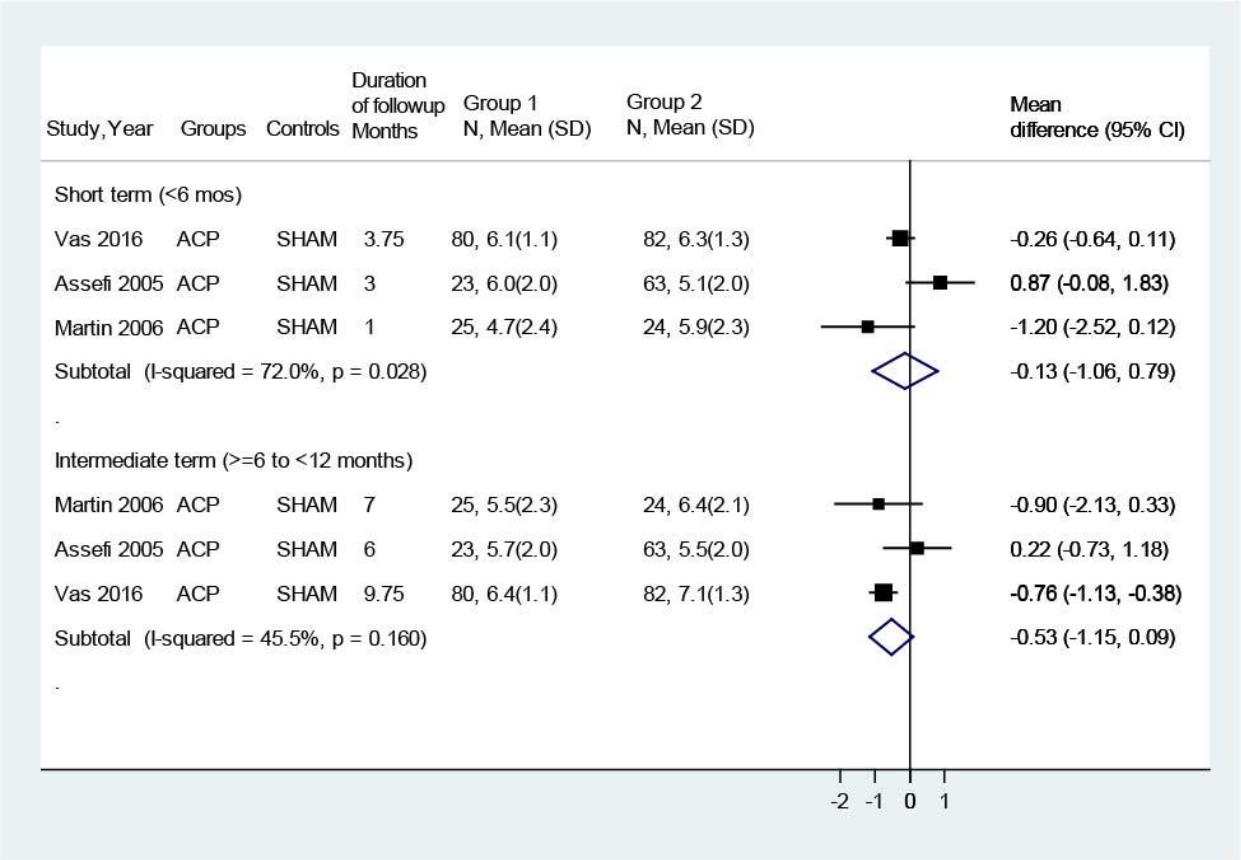
Discomfort and bruising were the most common reported adverse events. In one trial,[197] 89 of 96 treated (true or sham acupuncture) participants reported adverse events; 35 of 96 (37%) reported discomfort at needle insertion sites, 29 of 96 (30%) reported bruising, 3 of 96 (3%) reported nausea, and 1 of 96 (0.3%) felt faint at some point during the study. For patients assigned to simulated acupuncture, 5 of 19 (29%) had significantly less discomfort than those in directed acupuncture (14 of 23, 61%), acupuncture for unrelated condition (15 of 22, 70%) or sham needling (14 of 22, 64%); P = 0.02. In one trial,[198] 2 of 50 (4%) experienced mild vasovagal symptoms and 1 of 50 (2%) experienced a pulmonary embolism believed to be unrelated to treatment. Mild bruising and soreness were reported to be more common in the true acupuncture group, but rates were not reported. In one study,[199] 2.6 percent of sessions led to aggravation of fibromyalgia symptoms and 0.5 percent led to headache. In the true acupuncture group, pain, bruising, and vagal symptoms presented after 4.7 percent of sessions.

Figure 46. Acupuncture versus sham for fibromyalgia: effects on function



ACP = acupuncture; CI = confidence interval; FIQ = Fibromyalgia Impact Questionnaire; SD = standard deviation.

Figure 47. Acupuncture versus sham for fibromyalgia: effects on pain



ACP = acupuncture; CI = confidence interval; SD = standard deviation.

Multidisciplinary Rehabilitation for Fibromyalgia

Key Points

- There were no clear effects of multidisciplinary treatment for fibromyalgia on function versus usual care based on a 0 to 100 FIQ total score in the short-term (2 trials, pooled mean difference -5.06, 95% CI -12.38 to 2.25, I^2 =76.2%); however, it was associated with a slightly greater effect at the intermediate term (3 trials, pooled difference -7.84, 95% CI -11.43 to -4.25, I^2 = 18.2%) and long term (2 trials, pooled difference -8.42, 95%CI -13.76 to -3.08, I^2 = 24.9%). More multidisciplinary treatment participants experienced a clinically meaningful improvement in FIQ total score compared with usual care at short (OR 3.1, 95% CI 1.6 to 6.2), intermediate (1 trial OR 3.1, 95% CI 1.5 to 6.4) and long term (OR 8.8, 95% CI 2.5 to 30.9) (SOE: Low for short, intermediate and long term).
- There were no clear effects of multidisciplinary treatment for fibromyalgia on pain versus usual care or waitlist in the short term (2 trials pooled difference on 0-10 scale -0.24, 95% CI -0.63 to 0.15, I^2 = 0%); however multidisciplinary treatment was associated with a slightly greater effect on pain compared with usual care or waitlist at the intermediate

term (3 trials, pooled difference -0.68, 95%CI -1.07 to -0.30, $I^2 = 0\%$), but there were no clear differences compared with usual care long term (2 trials, pooled difference -0.25, 95% CI -0.68 to 0.17, $I^2 = 0\%$) (SOE: Low for short, intermediate and long-term).

- There was no clear effect of multidisciplinary pain treatment versus aerobic exercise at long term in one fair-quality trial for function the FIQ total score (difference -1.10, 95% CI -8.40 to 6.20, 0-100 FIQ total score) or pain (difference 0.10, 95% CI -0.67 to 0.87, 0-10 FIQ pain scale) (SOE: Low).
- Data were insufficient for harms. However, one poor-quality study reported on adverse events stating that 19% of participants randomized to multidisciplinary treatment withdrew (versus 0% for waiting list) and 2 of these 16 patients gave increased pain as the reason. Reasons for other withdrawals were not given and there was not systematic reporting of adverse events (SOE: Insufficient).

Detailed Synthesis

We identified five trials of multidisciplinary treatments that met inclusion criteria (Table 42 and Appendix D); all were conducted in Europe.[80,211,212,214,215] Across trials, sample sizes ranged from 155 to 203 (total randomized=893) and participants were predominantly (>90%) female with mean ages between 40 to 50 years. The multidisciplinary treatments included physical therapy in all trials, as well as pharmacological therapy and cognitive-behavioral therapy (2 trials),[212,215] sociotherapy", psychotherapy, and creative arts therapy (1 trial),[80] relaxation exercises (1 trial),[214] and education and group discussions (1 trial).[211] All trials compared multidisciplinary treatment with usual care or waitlist; in addition, one trial compared it with exercise.[80] Treatment duration ranged from 2 to 12 weeks and the frequency of sessions from twice a week to daily (total number of sessions ranged from 12 to 24 with durations between 1.5 to 5 hours). Two trials reported outcomes over the short term (3 and 5.5 months),[211,212] three over the intermediate term (6 months)[212,214,215] and two over the long term (12 and 18 months).[80,212] Two trials were judged to be of fair quality;[80,211] three trials were rated poor quality[212,214,215] (Appendix E). The nature of the intervention precluded blinding of participants and of people administering the treatments. Additional methodological shortcomings in the poor quality trials included unclear allocation concealment methods and unacceptable rates of overall attrition (21% to 43%) and differential attrition between groups (12% to 13%).

Table 42. Summary of results for fibromyalgia: multidisciplinary rehabilitation

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Amris, 2014[211] 5.5 months Duration of pain: median 10 to 11 years	A. Multidisciplinary treatment (n=84), 3 to 5 hours of education, sleep hygiene, group discussions, and physical therapy per day over 2 weeks B. Wait list (n=86)	A vs. B Age: 44 vs. 44 years Female: 100% vs. 100% Baseline Fibromyalgia Impact Questionnaire Total (FIQ, 0-100):	A vs. B <u>5.5 months</u> Change in FIQ total from baseline: -1.3 vs. -1.4, difference 0.1 (95% CI -3.6 to 3.8) Change in FIQ pain VAS from baseline: 0.1 vs. -0.1, difference 0.2 (95% CI -0.3 to 0.7)	A vs. B <u>5.5 months</u> Change in Generalized Anxiety Disorder-10 from baseline (scale NR): -0.8 vs. -0.5, difference -0.2 (95% CI -2.0 vs. 1.5) Change in Major

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fair		64.0 vs. 65.7 Baseline FIQ pain VAS (0-10): 7.1 vs. 7.4		Depression Inventory from baseline (0-50): -1.7 vs. -0.5, difference -1.3 (95% CI -3.3 to 0.8) Change in SF-36 physical component score from baseline (0-100): 1.4 vs. 0.8, difference 0.6 (95% CI -1.0 to 2.1) Percent responders in SF-36 physical component score: 27% vs. 23% Change in SF-36 mental component score from baseline (0-100): 2.3 vs. 1.2, difference 1.1 (95% CI -1.5 to 3.8) Percent responders in SF-36 mental component score: 27% vs. 27% Change in SF-36 physical functioning from baseline (0-100): 1.1 vs. 1.6, difference -0.5 (95% CI -3.9 to 3.0)
Castel, 2013[212] 3, 6 and 12 months Duration of pain: Mean 10.8 to 12.5 years *Poor*	A. Multidisciplinary treatment (n=53), conventional pharmacological treatment, 24 sessions of group CBT and physical therapy over 12 weeks. B. Usual care (conventional pharmacological treatment) (n=35), including analgesics, antidepressants, benzodiazepines, and non-benzodiazepine hypnotics	A vs. B Age: 49 vs. 49 years Female: 100% vs. 100% Baseline FIQ (0-100): 64.6 vs. 66.6 Baseline pain NRS (0-10): 6.8 vs. 7.1	A vs. B <u>3 months</u> FIQ: 55.5 vs. 64.6, difference -9.1 (95% CI -14.9 to -3.3) Proportion with clinically significant FIQ improvement (≥14% change): 48% vs. 23%, OR 3.1 (95% CI 1.6 to 6.2) Pain NRS: 6.4 vs. 6.8, difference -0.40 (95% CI -0.98 to 0.18) Proportion with clinically significant NRS pain improvement (≥30% change): 14% vs. 11% <u>6 months</u> FIQ: 55.8 vs. 67.8, difference -12.0 (95% CI -18.2 to -5.8) Proportion with clinically significant FIQ improvement (≥14% change): 42% vs. 19%, OR 3.1 (95% CI 1.5 to	A vs. B <u>3 months</u> HADS (0-42): 15.2 vs. 20.6, difference -5.4 (95% CI -8.2 to -2.6) MOS sleep scale (scale NR): 40.5 vs. 31.2, difference 9.3 (95% CI 6.1 to 12.5) <u>6 months</u> HADS: 16.2 vs. 21.5, difference -5.3 (95% CI -8.1 to -2.5) MOS sleep scale: 38.7 vs. 29.0, difference 9.7 (95% CI 6.6 to 12.8) <u>12 months</u> HADS: 17.1 vs. 22.8, difference -5.7 (95% CI -8.7 to -2.7) MOS sleep scale: 36.3 vs. 28.8, difference 7.5

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			6.4) Pain NRS: 6.4 vs. 7.0, difference -0.60 (95% CI -1.2 to 0) Proportion with clinically significant NRS pain improvement (≥30% change): 16% vs. 5%, OR 3.3 (95% CI 1.0 to 10.8) <u>12 months</u> FIQ: 58.8 vs. 69.6, difference -10.8 (95% CI -16.8 to -4.8) Proportion with clinically significant FIQ improvement (≥14% change): 27% vs. 4%, OR 8.8 (95% CI 2.5 to 30.9) Pain NRS: 6.7 vs. 7.1, difference -0.40 (95% CI -0.94 to 0.14) Proportion with clinically significant NRS pain improvement (≥30% change): 8.6% vs. 0%, OR 0.5 (95% CI 0.4 to 0.6)	(95% CI 4.3 to 10.7)
Cedraschi, 2004[214] 6 months Duration of pain: Mean 8.4 to 9.5 years *Poor*	A. Multidisciplinary treatment (n=84): 12 group pool sessions of physiotherapy, relaxation exercises, and exercise over 6 weeks B. Usual care (n=80) Regular care, including physical therapy, drug treatment and, in some cases, psychotherapy.	A vs. B Age: 49 vs. 50 years Female: 93% vs. 93% FIQ total (0-10): 5.5 vs. 5.6 FIQ physical function (0-10): 4.2 vs. 4.5 FIQ pain (0-10): 6.3 vs. 6.0 FIQ depression (0-10): 5.5 vs. 5.9 FIQ anxiety (0-10): 6.4 vs. 7.1 Regional Pain Score (0-105): 63.9 vs. 67.0	A vs. B <u>6 months</u> FIQ total: 4.9 vs. 5.5, difference -0.6 (95% CI -1.1 to -0.09) FIQ physical function: 4.3 vs. 4.8, difference -0.5 (95% CI -1.3 to 0.3) FIQ pain: 6.1 vs. 6.6, difference -0.5 (95% CI -1.2 to 0.2) Regional Pain Score: 62.6 vs. 68.4, difference -5.8 (95% CI -12.1 to 0.5) FIQ depression: 4.6 vs. 6.1 FIQ anxiety: 5.1 vs. 6.7, difference -1.6 (95% CI -2.6 to -0.6)	A vs. B <u>6 months</u> Psychological General Wellbeing Index total (0-110): 51.1 vs. 43.8, difference 7.3 (95% CI 0.2 to 14.3) Psychological General Wellbeing Index anxiety (0-25): 13.0 vs. 10.3, difference 2.7 (95% CI 0.6 to 4.8) Psychological General Wellbeing Index depression (0-15): 9.0 vs. 7.7, difference 1.3 (95% CI -0.1 to 2.7) SF-36 physical function (0-100): 42.2 vs. 43.9, difference -1.7 (95% CI -8.6 to 5.2)
Martin, 2012[215] 6 months Duration of pain: Mean 14 to 15	A. Multidisciplinary treatment (n=54), conventional pharmacological treatment, 12 sessions of CBT, education, and physiotherapy over 6	A vs. B Age: 49 vs. 52 years Female: 91% vs. 91% FIQ total (0-100): 76.3 vs. 76.2 FIQ physical	A vs. B <u>6 months</u> FIQ total: 70.3 vs. 76.8, difference -6.5 (95% CI -12.3 to -0.7) FIQ physical function: 5.2 vs. 5.9, difference -0.7 (95% CI -1.4 to -0.04)	A vs. B <u>6 months</u> Hospital Anxiety and Depression Scale anxiety (HADS, 0-21): 13.4 vs. 12.8, difference 0.66 (95% CI -1.02 to 2.34)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
years *Poor*	weeks B. Usual care (conventional pharmacological treatment) (n=56), included amitriptyline, paracetamol, and tramadol	functioning (0-10): 5.5 vs. 5.4 FIQ pain (0-10): 7.5 vs. 7.5	FIQ pain: 7.2 vs. 8.2, difference -1.0 (95% CI -1.7 to -0.3)	HADS depression (0-21): 9.8 vs. 10.2, difference -0.43 (95% CI -2.00 to 1.14)
Van Eijk-Hustings, 2013[80] 18 months Duration of pain: Mean of 6.1 to 7.1 years *Fair*	A. Multidisciplinary intervention (n=108), 36 days of sessions of sociotherapy, physiotherapy, psychotherapy, and creative arts therapy over 12 weeks B. Aerobic exercise (n=47): 24 sessions over 12 weeks C. Usual care (n=48), education and lifestyle advice in addition to usual care	A vs. B vs. C Age: 41 vs. 39 vs. 43 years Female: 93% vs. 100% vs. 98% FIQ total (0-100): 64.5 vs. 60.0 vs. 55.4 FIQ pain (0-10): 6.3 vs. 6.2 vs. 5.5 FIQ depression (0-10): 5.2 vs. 4.8 vs. 4.2 FIQ anxiety (0-10): 5.9 vs. 4.9 vs. 4.8	A vs. B[b] <u>18 months</u> FIQ total: 50.9 vs. 52.0, difference -1.10 (95% CI -8.40 to 6.20) FIQ physical function: 3.6 vs. 3.6, difference 0 (95% CI -0.79 to 0.79) FIQ pain: 5.3 vs. 5.2, difference 0.10 (95% CI -0.67 to 0.87) A vs. C <u>18 months</u> FIQ physical function: 3.6 vs. 3.9, ES 0.12 (-0.22 to 0.46) FIQ total: 50.9 vs. 56.2, ES 0.25 (95% CI -0.09 to 0.59) FIQ pain: 5.3 vs. 5.3, ES -0.01 (95% CI -0.35 to 0.34)	A vs. B[b] <u>18 months</u> FIQ Depression: 3.9 vs. 5.0, difference -1.1 (95% CI -2.2 to 0.01) FIQ Anxiety: 4.7 vs. 5.0, difference -0.30 (95% CI -1.41 to 0.81) EQ-5D (-0.59 to 1): 0.6 vs. 0.5, difference 0.01 (95% CI -0.10 to 0.12) GP consultations[c]: 0.9 vs. 1.0, difference -0.10 (95% CI -0.89 to 0.69) Medical specialist consultations[c]: 0.3 vs. 0.4, difference -0.10 (95% CI -0.43 to 0.23) Physiotherapist consultations[c]: 2.6 vs. 0.4, difference 2.20 (95% CI 0.69 to 3.71) Other paramedical professional consultations[c]: 1.0 vs. 2.1, difference -1.10 (95% CI -2.21 to 0.01) A vs. C <u>18 months</u> FIQ depression: 3.9 vs. 4.2, ES 0.10 (95% CI -0.24 to 0.44) FIQ anxiety: 4.7 vs. 4.8, ES 0.03 (95% CI -0.31 to 0.37 EQ-5D: 0.55 vs. 0.51, ES 0.12 (95% CI -0.22 to 0.46) GP consultations[c]: 0.9 vs. 0.7, ES = -0.11 (95% CI -0.45 to 0.23) Medical specialist

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
				consultations[c]: 0.3 vs. 0.2, ES = -0.14 (95% CI -0.48 to 0.20) Physiotherapist consultations[c]: 2.6 vs. 2.8, ES = 0.04 (95% CI -0.30 to 0.38) Other paramedical professional consultations[c]: 1.0 vs. 0.2, ES = -0.28 (95% CI -0.62 to 0.06)

CBT, cognitive- behavioral therapy; ES, effect size; EQ-5D, EuroQol-5D; FIQ, Fibromyalgia Impact Questionnaire; GP, general practitioner; HADS, Hospital Anxiety and Depression Scale; MOS, Medical Outcomes Study; NR, not reported; OR, odds ratio; SF-36, Short-Form 36; VAS, visual analog scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

[b] Authors did not provide effect estimates for the comparison of multidisciplinary rehabilitation versus exercise; mean differences were calculated by the EPC.

[c] Total number of consultations over a period of 2 months prior to measurement.

Multidisciplinary Rehabilitation Compared With Usual Care or Waitlist

There were no clear effects of multidisciplinary treatment for fibromyalgia on function versus usual care or waitlist based on a 0 to 100 FIQ total score in the short term (2 trials, pooled mean difference -5.06, 95% CI -12.38 to 2.25, I^2 =76.2%);[211,212] however, it was associated with a slightly greater effect at the intermediate term across three different poor quality trials (3 trials, pooled difference -7.84, 95% CI -11.43 to -4.25, I^2 = 18.2%)[212,214,215] (Figure 48). The intermediate-term estimate for trials of multidisciplinary treatment versus usual care only was similar (2 trials, pooled difference -9.12, 95%CI -15.92 to -2.48).[212,215] Clinically significant FIQ improvement (≥14% change) was significantly more common for multidisciplinary treatment compared with usual care at both short-term (OR 3.1, 95% CI 1.6 to 6.2) and intermediate-term followup (OR 3.1, 95% CI 1.5 to 6.4) in one poor-quality trial.[212] One of the pooled trials reported a slightly greater effect of multidisciplinary treatment on the FIQ physical function scale versus usual care[215] while another one did not.[214] The slightly greater effect of multidisciplinary rehabilitation versus usual care persisted over the long term (2 trials, pooled difference on 0-100 scale -8.42, 95% CI -13.76 to -3.08, I^2 = 24.9%).[80,212] Clinically significant FIQ improvement (≥14% change) was more common for multidisciplinary treatment compared with usual care at long-term followup (OR 8.8, 95% CI 2.5 to 30.9) in one poor-quality trial.[212] Only one poor-quality trial reported short-term, intermediate-term, and long-term effects on function, showing a significant result for each time frame.[212]

There were no clear effects of multidisciplinary treatment for fibromyalgia on pain versus usual care or waitlist in the short term (2 trials pooled difference on -0.24, 95%CI -0.63 to 0.15, I^2 = 0%)[211,212] (Figure 49). At intermediate term, multidisciplinary treatment was associated with a slightly greater effect on pain compared with usual care or waitlist (3 trials, pooled difference

0-10 scale -0.68, 95%CI -1.07 to -0.30, $I^2 = 0\%$);[212,214,215] the estimate versus usual care only was similar (2 trials, pooled difference -0.76, 95%CI -1.37 to -0.19).[212,215] Clinically important improvement was more common for multidisciplinary treatment compared with usual care in one poor-quality trial (≥30% change, OR 3.4 (95% CI 1.0 to 10.8).[212] Long term, there were no clear effects of multidisciplinary treatment for fibromyalgia on pain versus usual care (2 trials, pooled difference -0.25, 95% CI -0.68 to 0.17, $I^2 = 0\%$).[80,212] Over the long term, clinically important improvement in pain remained more common for multidisciplinary treatment compared with usual care in one poor-quality trial (≥30% change on NRS): 8.6% vs. 0%, $p< 0.5$.[212]

Only one poor-quality trial reported short-, intermediate-, and long-term effects on pain, showing a significant result for each time frame.[212]

Data on secondary outcomes were limited and the results were mixed across five trials. Compared with usual care, one trial found a significant benefit of multidisciplinary treatment for depression and sleep at intermediate-term and long-term followup.[212] Another trial comparing multidisciplinary treatment versus a waiting list at intermediate-term followup found that results differed for different measures.[214] There was a significant benefit of multidisciplinary treatment for depression but not anxiety as measured by FIQ subscales, and for anxiety but not depression when measured by the Psychological General Wellbeing Index subscales. Three trials found no significant difference between multidisciplinary treatment and usual care or waitlist on measures of anxiety and depression (as well as quality of life and mental health in one trial each) at short-term,[211] intermediate-term,[215] and long-term[80] followup. One trial reported no difference in health care utilization between groups during the 2 months prior to the final measurement at 18 months.[80]

Multidisciplinary Rehabilitation Compared With Pharmacological Therapy

No trial of multidisciplinary rehabilitation versus pharmacological therapy met inclusion criteria.

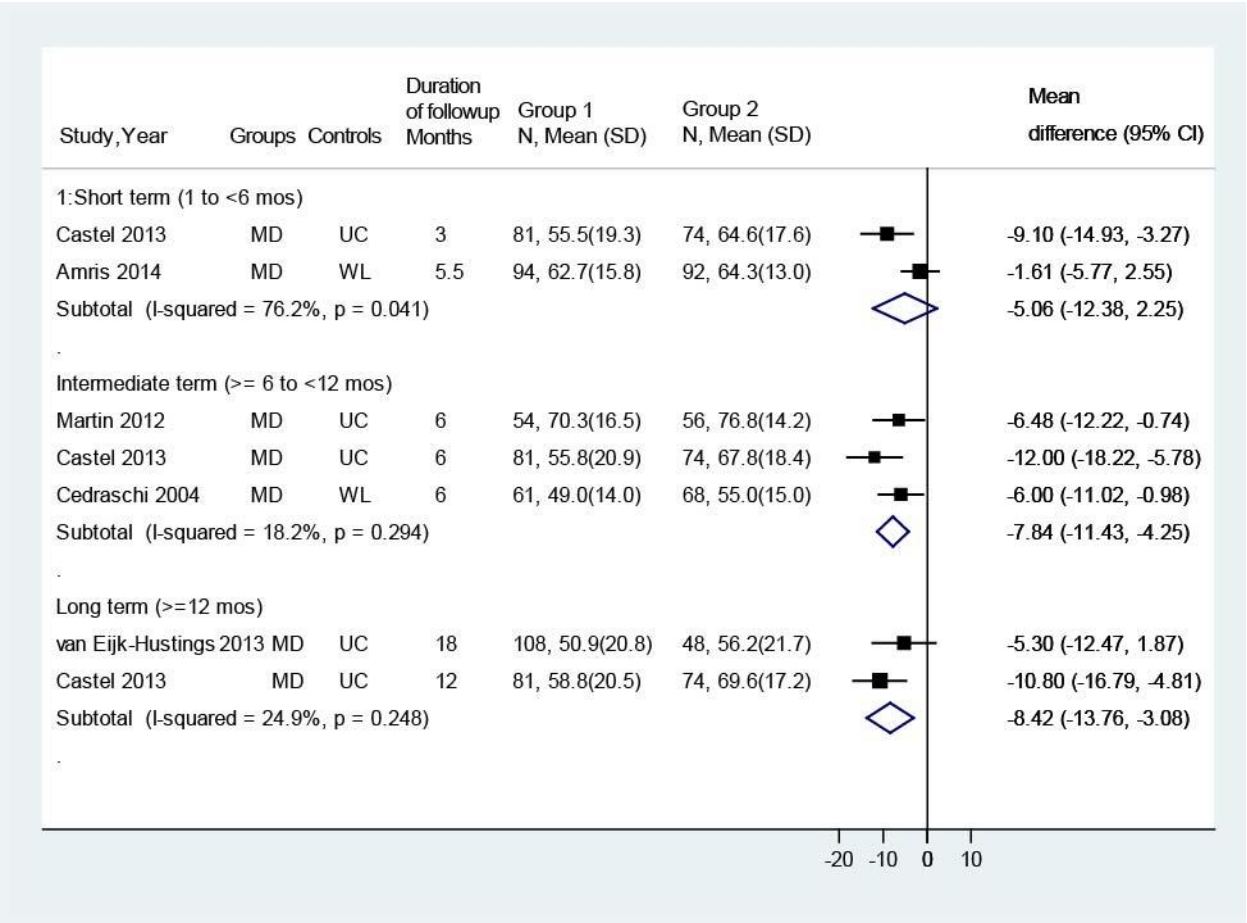
Multidisciplinary Rehabilitation Compared With Exercise

There was no clear effect of multidisciplinary pain treatment versus aerobic exercise at long term in one fair-quality trial[80] for physical function on the FIQ physical function scale (difference 0 on a 0-10 scale, 95% CI -0.79 to 0.79) or the FIQ total score (difference -1.10 on a 0-100 scale, 95% CI -8.40 to 6.20). Similarly, there were no significant differences on the FIQ pain scale (difference 0.10 on a 0-10 scale, 95% CI -0.67 to 0.87), secondary outcomes of quality of life, depression or anxiety, or health care utilization, with the exception of physiotherapist consultations, which was higher for the multidisciplinary group in the 2 months prior to the final measurement at 18 months (Table 42).

Harms

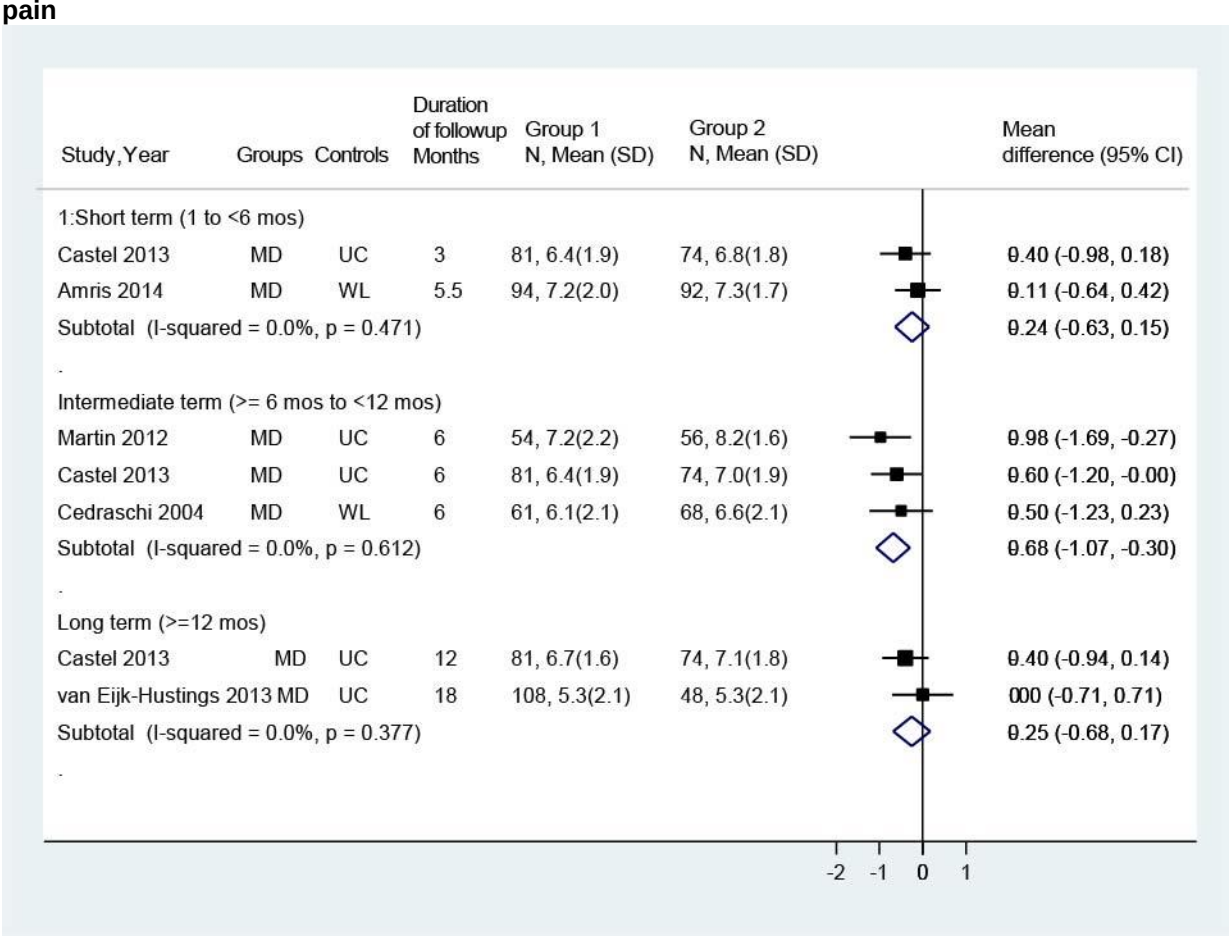
Only one study reported on adverse events and compared multidisciplinary treatment (group pool sessions of physiotherapy, relaxation exercises, and exercise) with usual care (physical therapy, drug treatment and, in some cases, psychotherapy).[214] This trial reported that 16 of 84 (19%) multidisciplinary participants withdrew (versus 0% for waiting list) and two of these gave increased pain as the reason. Reasons for other withdrawals were not given and there was not systematic reporting of adverse events.

Figure 48. Multidisciplinary rehabilitation versus usual care or waitlist for fibromyalgia: effects on function



CI = confidence interval; MD = multidisciplinary rehabilitation; SD = standard deviation; UC = usual care; WL = waitlist

Figure 49. Multidisciplinary rehabilitation versus usual care or waitlist for fibromyalgia: effects on pain



CI = confidence interval; MD = multidisciplinary rehabilitation; SD = standard deviation; UC = usual care; WL = waitlist

Key Question 5: Chronic Tension Headache

Psychological Therapies for Chronic Tension Headache

Key Points

- There is insufficient evidence from three poor quality trials to determine the effects of psychological therapies (CBT, relaxation) on short-term or intermediate-term function or pain compared with waitlist, placebo or attention control (SOE: Insufficient).
- There is insufficient evidence from two poor-quality trials to determine the effects of CBT on short-term or intermediate-term function or pain compared with antidepressant medication (SOE: Insufficient).
- No long-term outcomes were reported and no trials comparing psychological therapies to biofeedback were identified that met inclusion criteria.
- Data were insufficient for harms. Results were mixed across two poor-quality trials comparing CBT with antidepressant medication, with one trial reporting a lower risk of "at least mild" adverse events in the CBT group (0% vs. 59%), four of which led to

withdrawal from the trial, and the second trial reporting a similar low risk of withdrawal due to adverse events (2% to 6% across groups to include placebo) (SOE Insufficient).

Detailed Synthesis

Three trials, all conducted in the United States,[99,100,103] of CBT for chronic tension headache met inclusion criteria (Table 43 and Appendix D). Sample sizes ranged from 41 to 150; the mean age across trials varied from 32 to 42 years and most participants were female (56% to 80%). Duration since the onset of headache pain ranged from 10.7 to 14.5 years. All trials either excluded patients with concomitant migraines or required that they suffer from no more than one migraine per month. Two trials also specifically excluded patients with medication overuse (analgesic-abuse) headaches and required that patients be free from prophylactic headache medication upon study entry.[100,103]

All three trials evaluated some variation of stress management therapy/cognitive coping skills training with a relaxation component; one trial (n=77) also included an additional relaxation only arm.[99] In two trials (n=41, 150), patients received three 60-minute sessions of CBT and training in home-based relaxation,[100,103] and in the third trial (n=77), patients underwent 11 sessions (1-2 per week) of CBT plus progressive muscle relaxation training (session duration varied from 45 to 90 minutes).[99] In all trials, the interventions were administered by a psychologist or counselor over a 2-month period. Two trials compared CBT with placebo (placebo pill),[100] attention control (pseudomediation/body awareness training)[99] and waitlist (monitoring via phone and clinical visits) control groups.[99] Two trials compared CBT with amitriptyline (25-75 mg/day).[100,103] All trials reported short-term results; one trial also provided outcomes at intermediate-term followup.[100]

All three trials were considered poor quality (Appendix E) due to lack of blinding and large differential attrition between groups (in one trial, overall attrition was also substantial[100]). Additionally, randomization, concealment, and intention-to-treat processes were unclear in one trial.[103]

Table 43. Summary of results for chronic tension headache: psychological therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Blanchard, 1990[99] (n=77) 1 month Duration of pain: mean 14.2 years *Poor*	A. Cognitive Stress Coping Training + PMR (n=17): 11, 45-90 minute sessions once or twice per week for 8 weeks B. PMR alone (n=22): 10, 30-70 minute sessions twice weekly for 3 weeks followed by once weekly for 3 weeks with a final session at week 8 C. Pseudomeditation	A vs. B vs. C vs. D Age: 38 vs. 43 vs. 39 vs. 37 years Female: 56% vs. 58% vs. 45% vs. 66% Mean duration of chronicity: 13.0 vs. 13.9 vs. 15.3 vs. 14.3 years Headache Index Scores: mean 5.82 vs. 5.63 vs. 5.23 vs. 5.05 Medication Index Scores: mean 39.8	A vs. C <u>1 month</u> ≥50% improvement (i.e., reduction) in headache frequency: 62.5% vs. 43.7%; RR 1.43 (95%CI 0.81 to 1.97) Headache Index Scores: 3.2 vs. 4.6; difference -1.4 (95% CI -4.3 to 1.5) A vs. D <u>1 month</u> ≥50% improvement (i.e., reduction) in headache frequency: 62.5% vs. 20.0%; RR 3.13 (95% CI 0.91 to	A vs. C <u>1 month</u> Medication Index Scores: 20.7 vs. 8.3; difference 12.4 (95% CI -6.8 to 31.6) A vs. D <u>1 month</u> Medication Index Scores: 20.7 vs. 22.5; difference -1.8 (95% CI -23.8 to 20.2) B vs. C

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	(attention control) (n=19): body awareness and mental control training; 11 sessions over 8 weeks, 40-45 minutes each D. Waitlist (n=19): monitoring via phone, clinical visits and patient diaries.	vs. 16.9 vs. 12.1 vs. 24.0	2.45) Headache Index Scores: 3.2 vs. 4.5; difference -1.3 (95% CI -3.9 to 1.4) B vs. C <u>1 month</u> ≥50% improvement (i.e., reduction) in headache frequency: 31.6% vs. 43.7%; RR 0.72 (95% CI 0.65 to 1.69) Headache Index Scores: 3.8 vs. 4.6; difference -0.8 (95% CI -3.2 to 1.6) B vs. D <u>1 month</u> ≥50% improvement (i.e., reduction) in headache frequency: 31.6% vs. 20%; RR 1.58 (95% CI 0.75 to 2.11) Headache Index Scores: 3.8 vs. 4.5; difference -0.6 (95% CI -2.7 to 1.5)	<u>1 month</u> Medication Index Scores: 9.8 vs. 8.3; difference 1.5 (95% CI -6.8 to 9.8) B vs. D <u>1 month</u> Medication Index Scores: 9.8 vs. 22.5; difference -12.7 (95% CI -25.6 to 0.21)
Holroyd, 1991[103] (n=41) 1 month Duration of pain: mean 10.7 years *Poor*	A. CBT (n=19): three, 1 hour sessions over 8 weeks B. Amitriptyline therapy (n=17): Individualized dosage at 25, 50, or 75 mg/day for 8 weeks	A + B Age: 32.3 years Female: 80% A vs. B % of Headache-free days: 18.0 vs. 18.5 Headache Index scores (0-10): 2.17 vs. 2.04 Headache Pain Peak scores (0-10): 6.41 vs. 6.36	A vs. B <u>1 month</u> Proportion with >66% reduction in headaches (substantial improvement): 37% vs. 18%; RR 2.09 (95% CI 0.79 to 2.23) Proportion with 33-66% reduction in headaches (moderate improvement): 53% vs. 35%; RR 1.49 (95% CI 0.80 to 2.03) % of Headache-free days: 54.7 vs. 42.3; difference 12.4 (95% CI -8.06 to 32.86) Headache Index scores: 0.96 vs. 1.49; difference -0.53 (95% CI -1.14 to 0.08) Headache Peak scores: 4.33 vs. 4.55; difference -0.22 (95% CI -1.70 to 1.26)	A vs. B <u>1 month</u> BDI (0-63): 5.16 vs. 5.56; difference -0.4 (95% CI -3.96 to 3.16) STPI Anxiety (20-80): 18.37 vs. 19.06; difference -0.69 (95% CI -3.99 to 2.62) STPI Anger (20-80): 19.47 vs. 17.44; difference 2.03 (95% CI -1.98 to 6.04) WPSI (scale NR): 16.05 vs. 20.50; difference -4.45 95% CI -9.78 to 0.87) Analgesic Tablets: 0.26 vs. 0.82; difference -0.56 (95% CI -1.16 to 0.04)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Holroyd, 2001[100] (n=150) 1 and 6 months Duration of pain: mean 11.8 years *Poor*	A. Stress Management Therapy + Placebo (n=34): three, 1 hour sessions B. Placebo (n=26) Treatment Protocol: identitcal to group C C. Antidepressant Medications (n=44): Low starting dose (12.5 mg/day increased to 25mg, then 50mg) with the possibility to switch to nortriptyline	A vs. B vs. C Age: 37 vs. 38 vs. 36 years Female: 80% vs. 79% vs. 66% Caucasian: 91% vs. 98% vs. 98% Duration of pain: 12.3 vs. 11.1 vs. 11.9 years Headache frequency, days/month: 26.5 vs. 26.1 vs. 25.1 Headache Index (0-10): 2.8 vs. 2.7 vs. 2.8 Days/month with at least moderately severe headache (≥5 on 0-10 scale): 13.5 vs. 13.5 vs. 14.1	A vs. B <u>1 month</u> Days/month with at least moderately severe headache: mean difference 2.5 (95% CI -0.1 to 5.2) Headache Disability Inventory (0-100): mean difference 7.3 (95% CI 1.6 to 13.0) Headache Index: mean difference 0.46 (95% CI 0.02 to 0.89) <u>6 months</u> Patients who experienced ≥50% reductions in Headache Index Scores: 35% vs. 29%; RR 1.18 (95% CI 0.79 to 1.79) Days/month with at least moderately severe headache: mean difference 5.1 (95% CI 2.3 to 8.0) Headache Disability Inventory: mean difference 9.3 (95% CI 3.5 to 15.1) Headache Index: mean difference 0.79 (95% CI 0.30 to 1.28) A vs. C <u>1 month</u> Days/month with at least moderately severe headache: mean difference -3.5 (95% CI -6.1 to -0.9) Headache Disability Inventory: mean difference 0.1 (95% CI -5.6 to 5.7) Mean Headache Index: mean difference -0.54 (95% CI -0.97 to -0.012) <u>6 months</u> Patients who experienced >50% reductions in Headache Index Scores: 35% vs. 38%; RR 0.92 (95% CI 0.71 to 1.54) Days/month with at least moderately severe headache: mean difference 0.1 (95% CI -2.7 to 2.9) Headache Disability	A vs. B <u>1 month</u> Weighted analgesic use: mean difference -1.7 (95% CI -12.0 to 8.6) <u>6 months</u> Weighted analgesic use: mean difference 11.8 (95% CI 1.5 to 22.1) A vs. C <u>1 month</u> Weighted analgesic use: mean difference -19.4 (95% CI -29.5 to -9.3) <u>6 months</u> Weighted analgesic use: mean difference -6.2 (95% CI -16.2 to 3.8)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
			Inventory: mean difference 2.4 (95% CI -3.3 to 8.0) Headache Index: mean difference -0.13 (95% CI -0.61 to 0.35)	

BDI = Beck Depression Inventory; CBT = cognitive behavioral therapy; CI = confidence interval; MD = mean difference; NR = not reported; PMR = Progressive Muscle Relaxation; RR = risk ratio; SD = standard deviation; STPI = State-Trait Personality Inventory; VAS = Visual Analog Scale; WPSI = Wahler Physical Symptom Inventory

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Psychological Therapy Compared With Waitlist, Placebo or Attention Control

There was insufficient evidence from three poor-quality trials to draw conclusions regarding the effects of psychological therapies compared with waitlist, placebo, or attention control over the short term or intermediate term.

CBT plus placebo was associated with a slightly greater effect on both short-term and intermediate-term function compared with placebo alone as measured by the Headache Disability Inventory (scale 0-100) in one trial (mean difference 7.3, 95% CI 1.6 to 13.0 at 1 month and 9.3, 95% CI 3.5 to 15.1 at 6 months.[100] Long-term function was not reported.

Various pain measures were reported across trials. In general, CBT (plus relaxation), but not relaxation alone, appeared to have a small effect on short-term pain compared with waitlist, placebo, or attention control (Table 43). CBT plus relaxation was associated with a slight improvement in pain on the Headache Index (HI) at 1 month compared with waitlist, attention control, or placebo across two trials (pooled standardized mean difference [SMD] -0.40, 95% CI -0.74 to -0.07, I^2=0%)[99,100] (Figure 51). Relaxation only conferred no benefit for short-term pain compared with waitlist or attention control in one of these trials (difference -0.21 on a 0-20 HI scale, 95% CI -0.78 to 0.36).[99] Almost twice as many patients who received CBT plus relaxation achieved at least a 50 percent improvement in headache frequency compared with usual care or waitlist (RR 1.94, 95% CI 1.03 to 3.66) over the short term in one trial; however, there was no difference between groups when the intervention was relaxation alone (RR 0.98, 95% CI 0.42 to 2.26)[99] (Figure 50). One trial reported similar favorable results regarding pain over the

intermediate-term for CBT plus placebo compared with placebo alone, with the exception of "success" (≥50% improvement from baseline in HI score), which did not differ between groups (Table 43).[100]

Medication use did not differ significantly between the CBT and relaxation therapy groups and waitlist, placebo, or attention control groups over the short-term in two trials.[99,100] Over the intermediate-term, CBT plus placebo resulted in a significant reduction in analgesic use compared with placebo alone (difference 11.8, 95% CI 1.5 to 22.1).[100]

Psychological Therapy Compared With Pharmacological Therapy

There was insufficient evidence from two poor-quality trials to draw conclusions regarding the effect of CBT versus pharmacological therapy through intermediate-term followup.

There was no effect for CBT plus placebo versus antidepressant medication over the short-term or intermediate-term for function as measured by the Headache Disability Inventory (scale 0-100) in one trial (mean difference 0.1, 95% CI -5.6 to 5.7 at 1 month and 2.4, 95% CI -3.3 to 8.0 at 6 months).[100] Long-term function was not reported.

Regarding short-term pain, two trials reported HI index scores with differing results. One trial found that CBT plus placebo resulted in less improvement compared with antidepressant medication at 1 month (SMD 0.50, 95% CI 0.11 to 0.89),[100] whereas the other trial showed an improvement with CBT versus amitriptyline by 1 month, although the difference did not reach statistical significance (SMD -59, 95% CI -1.26 to 0.08)[103] (Figure 52); due to the significant heterogeneity between groups we did not use the pooled estimate. There were no significant differences between CBT and pharmacological treatment for any other pain outcome reported over the short term in both trials[100,103] or over the intermediate-term in one trial[100] (Table 43).

Short-term results were mixed regarding medication use with one trial reporting no difference between CBT and amitriptyline[103] and the other reporting a significant difference between groups favoring antidepressant therapy[100]; however, this difference did not persist to the intermediate term in the latter trial (Table 43).

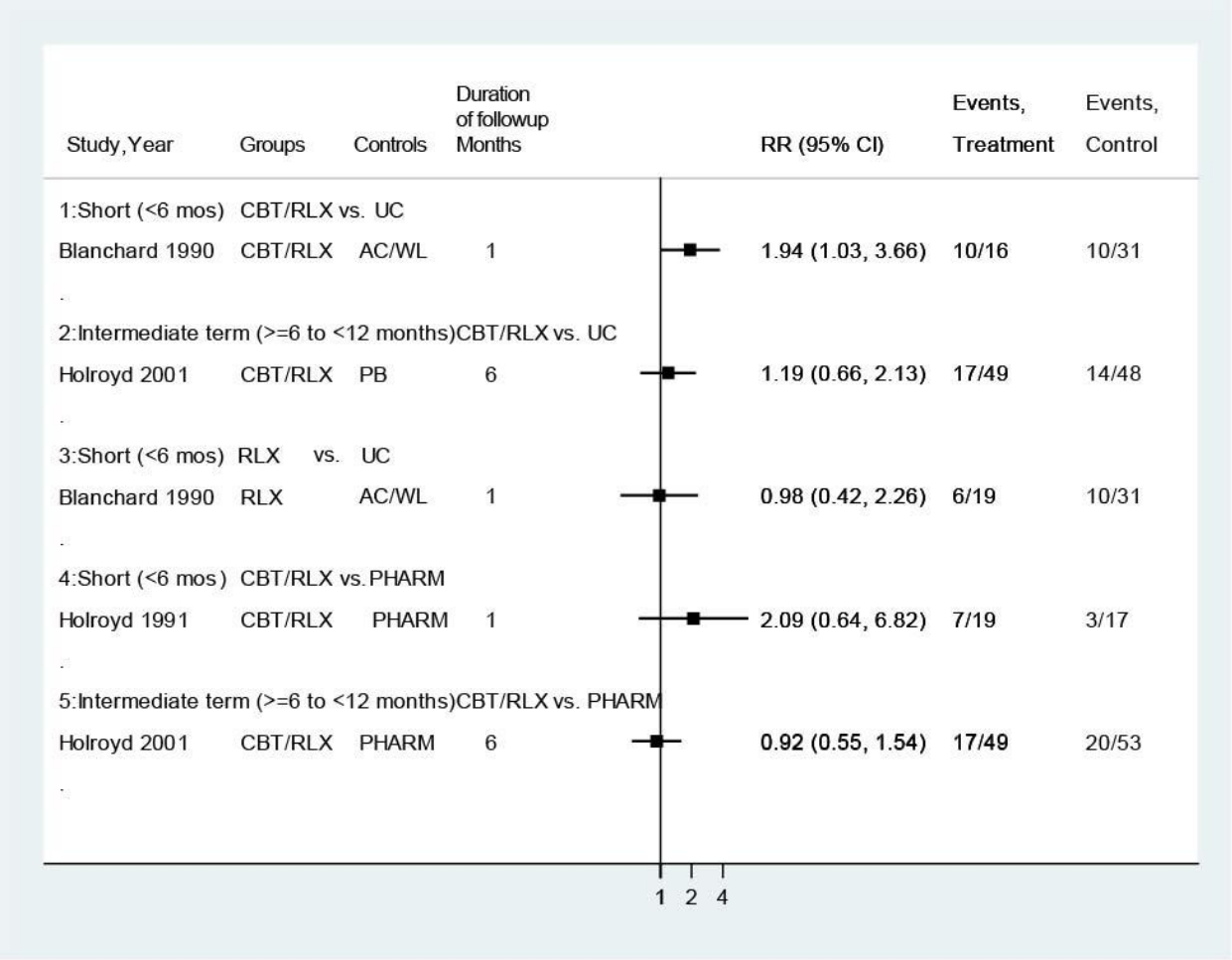
Psychological Therapy Compared With Biofeedback

No trial of psychological therapy versus biofeedback met inclusion criteria.

Harms

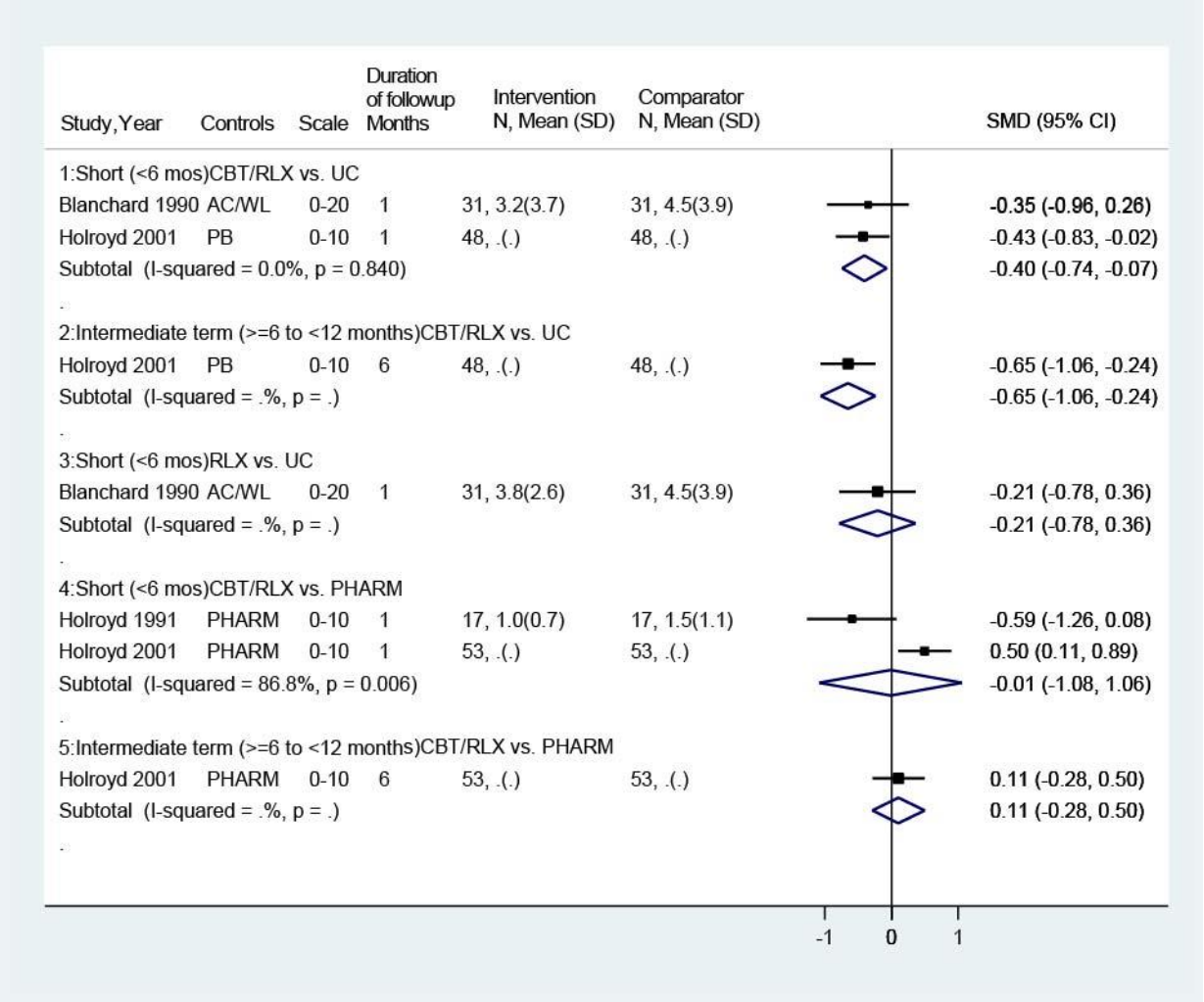
Harms were reported by the two poor-quality trials comparing CBT with antidepressant medication,[103] and with placebo in one.[100] No patient who underwent CBT experienced an adverse effect versus 10 of 17 (59%) of those who took medication in one trial;[103] six events were classified as mild, two as moderate, and two as substantial (no further details provided). Four of these patients withdrew from the trial. The risk of withdrawal due to adverse events was similar across groups in the second trial: CBT (2%) versus antidepressant medication (2%) and placebo (6%); no other information was provided.[100]

Figure 50. Psychological therapies versus waitlist, attention control, placebo intervention, or pharmacological treatment for chronic tension headache: effects on pain (success)



AC/WL = an attention control arm and a waitlist arm; CBT = Cognitive Behavioral Therapy; CBT/RLX = Cognitive Behavioral Therapy with a Relaxation component; CI = confidence interval; PB = placebo (pill); PHARM = standard pharmacological therapy; RLX = Relaxation therapy; RR = risk ratio; UC = usual care

Figure 51. Psychological therapies versus waitlist, attention control, placebo intervention, or pharmacological treatment for chronic tension headache: effects on pain (mean difference)



AC/WL = an attention control arm and a waitlist arm; CBT = Cognitive Behavioral Therapy; CBT/RLX = cognitive behavioral therapy with a relaxation component; CI = confidence interval; PB = placebo (pill); PHARM = standard pharmacological therapy; RLX = Relaxation therapy; SMD = standardized mean difference; UC = usual care

Physical Modalities for Chronic Tension Headache

Key Points

- There is insufficient evidence from one poor-quality trial to determine the effects occipital transcutaneous electrical stimulation (OTES) on short-term term function or pain compared with sham (SOE: Insufficient).
- No longer-term outcomes were reported and no trials comparing physical modalities to pharmacological therapy or to biofeedback were identified that met inclusion criteria.

- Data were insufficient for harms; however, no adverse events occurred in either the real or the sham OTES group in one poor-quality trial (SOE: Insufficient).

Detailed Synthesis

Only one Italian trial[135] was identified that investigated the efficacy of OTES versus sham (Table 44 and Appendix D). Patients were excluded if they had undergone prophylactic treatment in the prior 2 months or had previous treatment with OTES. Acute medications use was permitted during the study period, but other methods of pain control or new preventive treatments were prohibited. At baseline, 46 percent of patients were overusing medications. Identical devices and procedures were used for both the real and the sham OTES, and treatment consisted of 30-minute sessions, three times per day for two consecutive weeks. Limited information on the timing of outcomes was provided, but it was assumed that data was collected at 1 and 2 months post-treatment. This trial was rated poor quality due to unclear randomization sequence, failure to control for dissimilar proportion of females between groups, and no reporting of attrition (Appendix E). The focus of the trial was on allodonia, which was not of interest to this report.

Table 44. Summary of results for chronic tension headache: physical modalities

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Bono, 2015[135] (N=83) 1 month, 2 months Duration of pain: >2 years (mean NR) *Poor*	A. Occipital TES (n=54): Electro-stimulator generated biphasic impulses via electrodes placed on occipital region bilaterally; pulse width: 250 µs; frequency: 40 Hz; intensity 20 mA. B. Sham (n=29): Same device and procedure, but no current was delivered. Treatment protocol: 30 minute sessions 3 times daily for two consecutive weeks (42 sessions total)	A vs. B Age: 42 vs. 40 years Female: 81% vs. 66% Race: NR Headache frequency: mean 29.0 days/month Medication overuse: 43% vs. 52% MIDAS (0-21+): 63 vs. 50 Pain VAS (0-10): 8 vs. 8 BDI-II (0-63): 8 vs. 8 HAM-A: 7 vs. 7	A vs. B <u>1 month</u> Patients who achieved >50% reduction in headache days: 85% vs. 7%; RR 12.4 (95% CI 3.2 to 47.3) <u>2 months</u> MIDAS: 16 vs. 51; difference -35.0 (95% CI -42.6 to -27.4) VAS (0-10): 3 vs. 8; difference -5.0 (95% CI -5.8 to -4.2) Proportion of patients still overusing medications: 7% vs. 48%; RR 0.15 (95% CI 0.06 to 0.42)	A vs. B <u>2 months</u> BDI-II: 7 vs. 8; difference -1.0 (95% CI -2.2 to 0.2) HAM-A: 6 vs. 7; MD -1.0 (95% CI -1.9 to -0.1)

BDI-II = Beck Depression Inventory-II; CI = confidence interval; HAM-A = Hamilton Anxiety Rating Scale; MIDAS = Migraine Disability Assessment Questionnaire; NR = not reported; RR = risk ratio; SD = standard deviation; TES = transcutaneous electrical stimulation; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

Physical Modalities Compared With Sham

There was insufficient data from one poor-quality trial to determine the short-term effects of OTES compared with sham.[135] OTES resulted in greater improvement in function at 2 months as measured by the Migraine Disability Assessment Questionnaire (mean difference -35.0, 95% CI -42.6 to -27.4, scale 0-21+) and in pain intensity as measured by VAS (difference -5.0 on a 0-10 scale, 95% CI -5.8 to -4.2) The proportion of patients who achieved a 50 percent or greater reduction in headache days also favored OTES (RR 12.4; 95% CI 3.2 to 47.3). Measures of depression and anxiety were both somewhat better following OTES compared with sham at 2 months, however, the between-group difference was only statistically significant for anxiety (Table 44). The proportion of patients overusing medications at 2 months was also significantly lower in the OTES group.

Physical Modalities Compared With Pharmacological Therapy or Biofeedback

No trial of physical modalities versus pharmacological therapy and versus biofeedback met inclusion criteria.

Harms

Authors report that neither adverse events nor side effects occurred in either the real or the sham OTES group in one poor-quality trial.[135]

Manual Therapies for Chronic Tension Headache

Key Points

- Spinal manipulation therapy was associated with small to moderate improvements, respectively, in function compared with usual care (difference -5.0, 95% CI -9.02 to -1.16 on the Headache Impact Test, scale 36-78 and difference -10.1, 95% CI -19.5 to -0.64 on the Headache Disability Inventory, scale 0 to 100) and with moderate improvements in pain intensity (difference -1.4 on a 0-10 NRS scale, 95% CI -2.69 to -0.16) over the short term (SOE: Low). Approximately 25 percent of the patients had comorbid migraine.
- There is insufficient evidence from one poor-quality trial to determine the effects of spinal manipulation therapy on short-term pain compared with amitriptyline (SOE: Insufficient).
- No longer-term outcomes were reported and no trials comparing physical modalities to pharmacological therapy or to biofeedback were identified that met inclusion criteria.
- No adverse events occurred in the trial comparing spinal manipulation to usual care, but significantly fewer adverse events were reported following manipulation versus amitriptyline in the other poor-quality trial (4.3% vs. 82.1%; RR 0.05, 95% CI 0.02 to 0.16). The risk of withdrawal due to adverse events was not significantly different (1.4% vs. 8.9%; RR 0.16, 95% CI 0.02 to 1.33). Common complaints were neck stiffness in the manipulation group and dry mouth, dizziness, and weight gain in the medication group (SOE: Low).

Detailed Synthesis

Two trials (n=82 and n=150)[149,150] that evaluated spinal manipulation therapy (SMT) for the treatment of chronic tension headache met inclusion criteria (Table 45 and Appendix D). The majority of patients in both trials were female (61% to 78%) with mean ages ranging from 40 to 42 years and a mean headache duration of 13 years. Both trials included patients with comorbid

migraine as long as their headache problem was determined by a physician to be predominantly tension-type in nature (this included 26% of patients in one trial;[149] proportion not reported in the other trial). In one trial, patients were specifically excluded if they met the criteria for medication overuse or if they had received manual therapy in the 2 months prior to enrollment.[149] At baseline, prophylactic medication use was common. Current or past use of other treatments was not reported.

One Dutch trial compared a maximum of nine, 30-minute sessions of SMT over 8 weeks with usual care (information, reassurance and advice, discussion of lifestyle changes, and analgesics or NSAIDs provided by a general practitioner).[149] The second trial, conducted in the United States, compared 12 SMT sessions of 20 minutes over a 6-week treatment period versus amitriptyline (maximum dose 30 mg/day).[150] Both trials reported only short-term outcomes. One trial was rated fair quality[149] and one poor quality[150] (Appendix E). Due to the nature of the interventions, blinding of patients and researchers was not possible. Additionally, the poor trial had a high rate of differential attrition (7% SMT and 27% amitriptyline).

Table 45. Summary of results for chronic tension headache: manual therapies

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Boline, 1995[150] 1 month Duration of pain: 13.5 years *Poor*	A. Spinal Manipulative Therapy (n=70): short-lever, low-amplitude, high-velocity thrust techniques on cervical, thoracic or lumbar spinal segments. Moist heat and light massage preceded manipulation; 12, 20 minute sessions (2 per week for 6 weeks) B. Amitriptyline (n=56): dose titration of amitriptyline for 6 weeks. Nighttime, daily doses began at 10mg/day for first week, then increased to 20mg/day in the second, followed by 30mg/day in the third week and after; continued use of OTC medications as-needed.	A vs. B Age: 41 vs. 42 years Female: 54% vs. 70% Race: NR Daily headache intensity (0-20)[b]: 5.6 vs. 5.0 Weekly headache frequency (0-28)[c]: 12.4 vs. 10.8	A vs. B <u>1 month</u> Daily headache intensity[b]: adjusted means 3.8 vs. 5.2; difference 1.4 (95% CI 0.3, 2.3) Weekly headache frequency[c]: adjusted means 7.6 vs. 11.8; difference 4.2 (95% CI 1.9, 6.5)	A vs. B <u>1 month</u> SF-36 Function Health Status Global Score (% points): adjusted means 78.8 vs. 73.9; difference 4.9 (95%CI 0.4, 9.4) OTC medication usage: adjusted means 1.3 vs. 2.2; difference 0.9 (95% CI 0.3, 1.5)
Castien, 2011[149] 4.5 months Duration of pain: 13 years	A. Spinal Manipulation (n=38) combination of 3 approaches at the therapist discretion: mobilizations of the cervical and thoracic	A vs. B Age, years: 40 vs. 40 years Female: 78% vs. 78% Race: NR	A vs. B <u>4.5 months</u> Proportion of patients with ≥50% reduction in headache frequency: 81.6% vs. 40.5%; RR 2.01 (95% CI 1.32 to 3.05)	A vs. B <u>4.5 months</u> Resource use, proportion who used: ≥1 sick leave day: 7.9% vs. 32.4%; RR

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Fair	spine, craniocervical muscle exercises and posture correction; maximum of 9, 30-minute sessions over 2 months B. Usual Care (n=37) 2-3 general practitioner visits over 2 months	Mean frequency of headache (days/month): 24 vs. 24 NSAID use: 29% (mean 3 pills/week); Analgesic use: 59% (mean 1.5 pills/week) HIT-6 (36-78): 62.6 vs. 61.2 HDI (0-100): 39.6 vs. 44.2 Pain intensity, NRS (0-10): 6.3 vs. 5.7	HIT-6, mean change from baseline: -10.6 vs. -5.5; difference 5.0 (95% CI -9.02 to -1.16) HDI, mean change from baseline: -20.0 vs. -9.9; difference -10.1 (95% CI -19.5 to -0.64) Headache frequency (days/14 days), mean change from baseline: -9.1 vs. -4.1; difference -4.9 (95% CI -6.95 to -2.98) Pain intensity mean change from baseline: -3.1 vs. -1.7; difference -1.4 (95% CI -2.69 to -0.16) Headache duration (hrs./day), mean change from baseline: -7.0 vs. -3.5; difference -3.5 (95% CI -7.71 to -0.63)	0.23 (95% CI 0.07 to 0.79) Any additional health care: 13.2% vs. 59.4%; RR 0.22 (95% CI 0.09 to 0.52) Additional physical therapy: 2.6% vs. 40.5%; RR 0.06 (95% CI 0.01 to 0.47) Additional medical specialist care: 2.6% vs. 16.2%; RR 0.16 (95% CI 0.02 to 1.28) Additional "other" health care": 7.8% vs. 2.7%; RR 2.9 (95% CI 0.3 to 26.8)

CI = confidence interval; HDI = Headache Disability Index; HIT-6 = Headache Impact Test-6; NR = not reported; NR = not reported; NRS = numerical rating scale; NSAID = nonsteroidal anti-inflammatory drugs OTC = over-the-counter; RR = risk ratio; SF-36 = Short-Form-36 Questionnaire; VAS = Visual Analog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period.

[b] Headache intensity was calculated as the total ratings per period and divided by the number of days per period.

[c] Headache frequency was calculated by summing all headache ratings 2 and above for the month.

Manual Therapies Compared With Usual Care

Only short-term data from one fair-quality trial were reported. SMT resulted in small to moderate improvements in function compared with usual care at 4.5 months post-treatment as measured by the Headache Disability Inventory (HDI, scale 0 to 100) and the Headache Impact Test (HIT-6, scale 36 to 78), respectively (mean difference between groups in change scores from baseline, -10.1, 95% CI -19.5 to -0.64 and -5.0, 95% CI -9.02 to -1.16).[149] Regarding pain outcomes, twice as many patients who received SMT experienced a ≥50% reduction from baseline in the number of headache days (per 2 weeks) compared with usual care: 81.6% versus 40.5%; RR 2.0 (95% CI 1.3, 3.0).[149] Similarly, a statistically greater reduction in the number of headache days (mean difference between groups in change scores from baseline, -4.9; 95% CI -6.95 to -2.98) and in headache pain intensity (mean difference in change scores from baseline, -1.4 on a 0 to 10 NRS scale, 95% CI -2.69 to -0.16) was seen following SMT. Given that 29 percent of SMT patients and 22 percent of usual care patients had comorbid migraine, it is unclear how the coexistence of these headache types may have affected the outcome.

The proportion of patients who used any additional health care services (e.g., physical therapy, medical specialists, other) was statistically lower in the SMT group compared with the usual care group (Table 45).[149] Authors report no statistically significant differences between treatments in analgesic or NSAID use; data were not provided.

Manual Therapies Compared With Pharmacological Therapy

The evidence was insufficient from one poor-quality trial to determine the effects of spinal manipulation compared with amitriptyline over the short term.[150] The spinal manipulation group showed more improvement compared with the amitriptyline group in daily headache intensity (adjusted difference -1.4, 95% CI -2.3 to -0.3), weekly headache frequency (adjusted difference -4.2, 95% CI -6.5 to -1.9), SF-36 Function score (adjusted difference 4.9, 95% CI 0.4 to 9.4), and over-the-counter medication use (difference -0.9, 95% CI -1.5 to -0.3) at 1 month. Attrition in the amitriptyline group was 27 percent, compared with 7 percent in the manipulation group.

Manual Therapies Compared With Biofeedback

No trial of physical modalities versus biofeedback met inclusion criteria.

Harms

No adverse events occurred in the trial comparing spinal manipulation to usual care.[149] The other poor-quality trial reported significantly fewer adverse events following spinal manipulation compared with amitriptyline (4.3% vs. 82.1%; RR 0.05, 95% CI 0.02 to 0.16) but the risk of withdrawal due to adverse events was not significantly different (1.4% vs. 8.9%; RR 0.16, 95% CI 0.02 to 1.33).[150] Patients in the manipulation group complained of neck stiffness which resolved in all cases and common side effects in the amitriptyline group included dry mouth, drowsiness, and weight gain.

Acupuncture for Chronic Tension Headache

Key Points

- There is insufficient evidence from two poor quality trials to determine the effects of Traditional Chinese needle acupuncture on short-term (2 trials), intermediate-term (1 trial) or long-term (1 trial) pain compared with sham acupuncture (SOE: Insufficient).
- Laser acupuncture was associated with slight improvement in pain intensity (median difference -2, IQR 6.3, on a 0-10 VAS scale) and in the number of headache days per month (median difference -8, IQR 21.5) in one fair-quality trial (SOE: Low).
- No trials comparing acupuncture to pharmacological therapy or to biofeedback were identified that met inclusion criteria.
- The fair-quality trial evaluating laser acupuncture reported that no adverse events occurred in either group (SOE: Low).

Detailed Synthesis

Three small trials (n=30 to 50)[200-202] that evaluated acupuncture versus sham treatment for chronic tension headaches met inclusion criteria (Table 46 and Appendix D). Two trials employed traditional Chinese needle acupuncture,[201,202] while one used low-energy laser acupuncture.[200] The number of acupoints ranged from six to ten across studies. The duration of treatment ranged from 5 to 10 weeks, with the total number of sessions ranging from eight to ten (20 to 30 minutes duration, 1 to 3 times per week). Sham treatment consisted of irrelevant acupuncture (superficial needle insertion in areas without acupuncture points) and sham acupuncture (blunt needle that simulates puncturing of the skin, laser power output set to zero).

Across trials, participants were primarily female (49% to 87%), mean ages ranged from 33 to 49 years, and headache frequency from 18 to 27 days per month. Two trials specifically excluded patients with other causes of chronic headache;[200,201] the third trial did not note if any of the

patients had concomitant headaches.[202] One trial required patients to abstain from all other prophylactic therapies (with the exception of rescue analgesics),[202] and one trial excluded patients who had received any treatment for their headache in the 2 weeks prior to enrollment.[200] Concomitant (non-narcotic) medication was permitted in two trials,[201,202] the third stated that no patient took concomitant analgesics.[200] All trials assessed outcomes over the short term; one trial additionally provided intermediate- and long-term data.[202]

One trial was rated fair quality[200] and two poor quality[201,202] (Appendix E). In all three trials, random sequence generation and concealment of allocation were not clearly reported and the care providers were not blinded to treatment. Additional methodological concerns in the poor quality trials included unclear application of intention-to-treat methods, and failure to control for disproportionate baseline characteristics or to account for loss to followup in one trial each.

Table 46. Summary of results for chronic tension headache: acupuncture

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
Ebneshahidi, 2005[200] 3 months Duration of pain: NR *Fair*	A. Low-Energy Laser Acupuncture (n=25): 4 acupoints (two local and two distal), bilaterally (8 total): intensity 1.3J, output 100%, continuous mode, using vertical contact with pressure and a duration of 43 seconds. B. Sham Laser Acupuncture (n=25): Identical procedure to real electroacupuncture except power output set to 0 Treatment Protocol: 3 sessions per week for a total of 10 sessions (session length: NR)	A vs. B Age: 33 vs. 39 years Female: 80% vs. 80% Race: NR Number of headache days per month (0-28), median: 20 vs. 18 Pain intensity on VAS (0-10), median: 10 vs. 10 Duration of attacks, (hours), median: 10 vs. 8, p=0.02	A vs. B <u>3 months</u> Headache Days/Month, median change from baseline: -8 vs. 0, p<0.001 Headache Intensity (VAS), median change from baseline: -2 vs. 0, p<0.001 Duration of attacks (hours), median change from baseline: -4 vs. 0, p<0.001	NR
Karst, 2000[201] 1.5 months Duration of pain: NR *Poor*	A. Acupuncture (n=21) Traditional Chinese acupuncture; maximum of 15 needles, 10 acupoints B. Sham Acupuncture (n=18): blunt placebo needles and elastic foam were used to simulate puncturing and shield needle type.	A vs. B Age: 50 vs. 47 years Female: 38% vs. 61% Race: NR Headache frequency: 27 vs. 27 days/month VAS (0-10): 6.2 vs. 6.3 Analgesic Intake/Month: 8.3 vs. 10.2	A vs. B <u>1.5 months</u> Frequency of headache attacks/month: 22.1 vs. 22.0; difference 0.1 (95% CI -6.6 to 6.8) Headache Severity, VAS: 4.0 vs. 3.9; difference 0.1 (95% CI -11.9 to 12.1)	A vs. B <u>1.5 months</u> Analgesic Intake/Month: 13.7 vs. 21.2; difference -7.5 (95% CI -22.2 to 7.2)

Author, Year, Followup[a], Pain Duration, Study Quality	Intervention	Population	Function and Pain Outcomes	Other Outcomes
	Treatment Protocol: 30-minute sessions twice weekly for 5 weeks (10 sessions total)			
Tavola, 1992[202] (n=30) 1, 6, 12 months Duration of pain: 8 years *Poor*	A. Acupuncture (n=15): Traditional Chinese acupuncture; 6-10 acupoints chosen on an individual basis; insertion depth 10-20 mm; needles were left in place without the use of any manual or electrical stimulation B. Sham Acupuncture (n=15): same number of needles, inserted more superficially (depth 2-4 mm), in the same region used in real acupuncture group but in areas without acupuncture points Treatment Protocol: 20-minute sessions once per week for 8 weeks (8 sessions total)	A vs. B Age: 33 vs. 33 years Female: 87% vs. 87% Mean frequency of headache attacks per month: 18 vs. 17 Mean analgesic use: 12 vs. 12 units/month Mean HI (intensity X duration X frequency/30): 4.3 vs. 4.5 Mean duration of attacks (sum of the hours of headache in a month/number of attacks): 3.3 vs. 4.4	A vs. B <u>1 month</u> Responders, ≥33% improvement in HI: 86.7% vs. 60.0%; RR 1.44 (95% CI 0.91 to 2.28) Responders, ≥50% improvement in HI: 53.3% vs. 46.7%; RR 1.14 (95% CI 0.56 to 2.35) HI, mean[b]: 2.4 vs. 3.0; MD -0.60 (95% CI -6.12 to 4.92) Mean decrease in HI from baseline: 58.3% vs. 27.8% Mean decrease in headache attack frequency from baseline: 44.3% vs. 21.4% <u>6 months</u> HI, mean[b]: 2.2 vs. 3.1; MD -0.90 (95% CI -7.15 to 5.35), <u>12 months</u> Responders, ≥33% improvement in HI: 53.3% vs. 46.7%; RR 1.14 (95% CI 0.56 to 2.35) Responders, ≥50% improvement in HI: 40.0% vs. 26.7%; RR 1.50 (95% CI 0.53 to 4.26) HI, mean[b]: 3.2 (2.1) vs. 3.7 (2.2); MD -0.50 (95% CI -6.73 to 5.73)	A vs. B <u>1 month</u> Mean decrease in analgesic consumption from baseline: 57.7% vs. 21.7%

CI = confidence interval; HI = headache index; MD = mean difference; NR = not reported; RR = risk ratio; VAS = VisualAnalog Scale

[a] Unless otherwise noted, followup time is calculated from the end of the treatment period

[b] Means and standard error of the means (not shown) estimated from graphs.

Acupuncture Compared With Sham

None of the trials reported on function. All three trials reported pain outcomes, although the specific measures varied across the trials. The results were mixed depending on the type of acupuncture used. No significant differences were found between needle acupuncture and sham for any pain outcome evaluated during the short term in two small poor-quality trials,[201,202] or at intermediate and long-term followup in one of these trials[202] (Table 46). In the third small fair-

quality trial,[200] laser acupuncture resulted in a significant reduction in the number of headache days per month (median -8, IQR 21.5), in pain intensity on a 0 to 10 VAS scale (median -2, IQR 6.3), and in the duration of attacks (median -4 hours, IQR 7.5) over the short term compared with the sham group, which reported no improvement from baseline on any outcome at the 3-month followup ($p<0.001$ for all) (Figure 52).

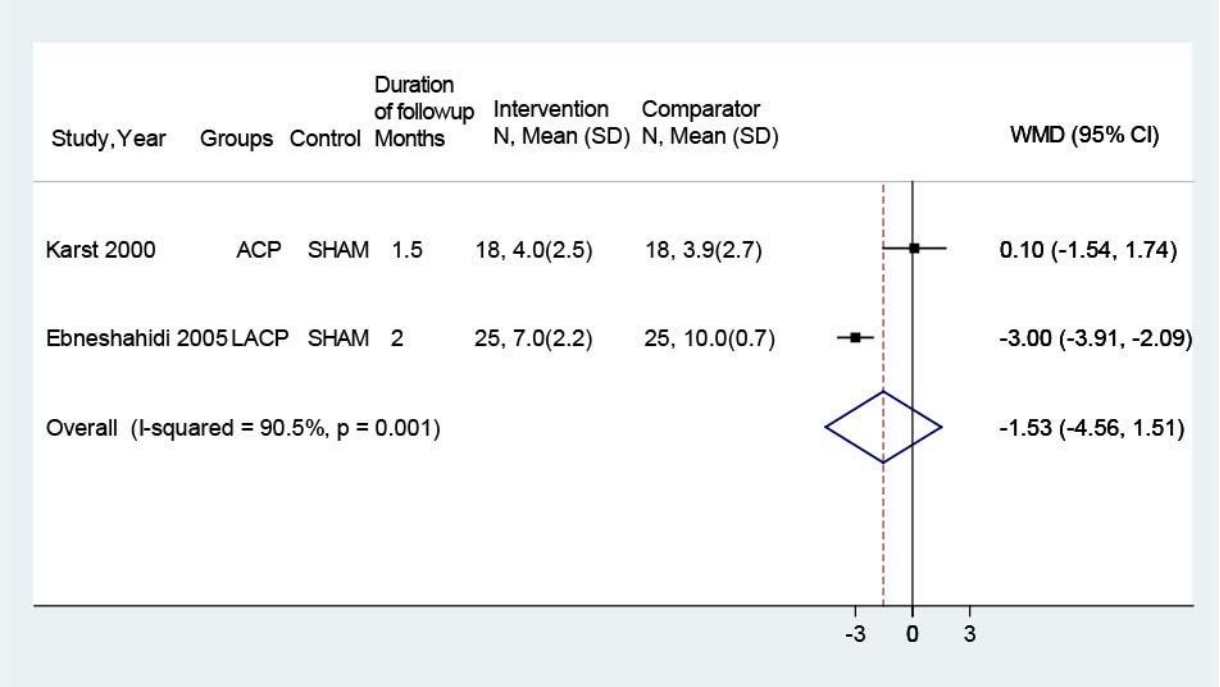
Acupuncture Compared With Pharmacological Therapy or Biofeedback

No trial of acupuncture versus pharmacological therapy and versus biofeedback met inclusion criteria.

Harms

Harms were generally not reported. The trial evaluating laser acupuncture reported that no adverse events occurred in either group.[200]

Figure 52. Acupuncture versus sham for chronic tension headache: effects on pain



ACP = standard needle acupuncture; CI = confidence interval; LACP = laser acupuncture; SD = standard deviation; WMD = weighted mean difference

Key Question 6: Differential Efficacy

RCTs that stratified on patient characteristics of interest, permitting evaluation of factors that might modify the effect of treatment, were considered for inclusion. Factors included age, sex, and presence of comorbidities (e.g., emotional or mood disorders). If a comparison is not listed below there was either no evidence identified that met the inclusion criteria or the included trials did not provide information on differential efficacy or harms. Studies likely had insufficient sample size to evaluate differential efficacy or harms, and evidence was considered insufficient.

Osteoarthritis of the Knee

Key Points

- There is insufficient evidence from one fair-quality trial (across 3 publications) that age, sex, race, BMI, baseline disability, pain, or depression status modify the effects of exercise in patients with OA of the knee. Sample sizes in the subgroup analyses from the FAST trial were likely inadequate to effectively test for modification.

Exercise Compared With Attention Control

One fair-quality trial (n=439) reported across three publications of the Fitness, Arthritis and Seniors Trial (FAST)[40,46,47] included in Key Question 3 compared muscle performance (i.e., resistance training) and aerobic exercise programs to an attention control and formally evaluated factors that may modify treatment. Details regarding these study populations are available in the Results section for Key Question 3 and in Appendix D. Two of the reports performed formal tests for interaction; none of the demographic or clinical variables evaluated were found to modify the effect of either type of exercise.[46,47] One trial explored whether age, sex, race, BMI, baseline disability, or baseline pain modified the effects of exercise on function based on ADL disability measures in a subgroup of patients who were free of ADL disability upon enrollment; however, no data were provided for evaluation.[46] A second publication looked at whether the effects of exercise on pain, disability, and depression were modified by baseline depression status, that is, high versus low depressive symptomology according to the Center for Epidemiologic Studies Depression scale over time (using an adjusted repeated measures analysis of variance). However, the authors do not provide results that directly examined modification by baseline depression without the time component.[47] The third FAST publication stratified on age, sex, race, and BMI and did not perform a formal statistical test for interaction.[40] Upon visual inspection, the point estimates across groups and strata are similar, suggesting that the effect of exercise on physical disability and knee pain was not modified by any patient characteristic evaluated.

Osteoarthritis of the Hip

Key Points

- There is insufficient evidence from one fair-quality trial that age, sex, baseline pain, and the presence of radiographic OA modify the effects of exercise in patients with OA of the hip. Study authors only reported on effects that include evaluation of these factors over time. Sample size was likely inadequate to effectively test for modification

Exercise Compared With Usual Care

One fair-quality trial (n=203) included for Key Question 3 compared combination exercise therapy (strengthening, stretching, and endurance exercises) to usual care and stratified on age, sex, race, and BMI, but it did not formally test for interaction.[59] Details regarding this study population are available in the Results section for Key Question 3 and in Appendix D. Age, sex, education, self-reported knee OA, and baseline pain and Kellgren & Lawrence radiographic OA scores were defined *a priori* as subgroups of interest. Although older patients (age ≥65 years), women, patients with a lower NRS pain score at baseline, and patients with radiographic OA showed somewhat larger effects of exercise therapy on function and pain, data were not

systematically reported and, based on the data provided, overlapping confidence intervals suggest that the effect of exercise was not modified by any of these variables.

Fibromyalgia

Key Points

- There is insufficient evidence from one poor-quality trial that baseline BMI (normal, overweight, obese) modifies the effects of multidisciplinary rehabilitation in patients with fibromyalgia. Study authors only report on effects that include evaluation of these factors over time. Sample size was likely inadequate to effectively test for modification.

Multidisciplinary Rehabilitation Compared With Usual Care

An additional publication (n=130)[213] of a poor-quality trial[212] included for Key Question 4 that compared multidisciplinary rehabilitation to usual care assessed potential modification of treatment based on baseline BMI (normal, overweight, obese). No significant interactions were found for the effect of BMI on exercise over time for any pain or function measure evaluated; however, the authors do not provide results that exclude effects of time. Details regarding this study population are available in the section on efficacy and in Appendix D.

Discussion

Key Findings and Strength of Evidence

The key findings of this review, including strength of evidence ratings, are summarized for each chronic pain condition in Tables 47-61 (interventions and comparators with no evidence for either function or pain outcomes are not shown); domains used to determine the overall strength of evidence are shown in Appendix G. The strength of evidence was low or insufficient for many interventions and was limited by small numbers of trials available for specific comparisons and for our specified time frames, particularly for long term. We focused on evaluating the persistence of effects for therapies beyond the course of treatment, using the following definitions for post-intervention followup: short-term (1 to 6 months), intermediate-term (6 to 12 months) and long-term (≥12 months). Little long-term evidence was available across conditions and interventions. The majority of trials compared interventions with usual care with very few trials employing pharmacological treatments or exercise as comparators. In general, effect sizes for most interventions were small, based on mean differences. There tended to be more evidence for the effects of interventions on pain than for function, and the effects on function were generally smaller or not clearly present. No trials directly compared interventions with opioids. Our previous reports suggest small to moderate effects of opioids on pain during treatment only (effects would not be expected to persist) with evidence primarily from short-term trials.[9,21,230,231] Information on adherence to interventions was not well-reported; poor adherence may have impacted some of our findings. Harms were poorly reported across interventions. No serious intervention-related adverse events requiring medical attention were identified; reported adverse events were generally minor (e.g., muscle soreness with exercise, bruising with acupuncture) and time-limited (e.g., temporary worsening of pain).

Table 47. Nonpharmacological interventions for chronic low back pain compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+		x	+	++	++
	SOE	✓	✓	✓	✓✓	✓	✓
Psychological Therapies	Effect size	+	+	+	+	+	+
	SOE	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Physical Modalities: Short-Wave Diathermy	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Physical Modalities: Ultrasound	Effect size	?	○	⊘	x	⊘	⊘
	SOE	○	○	⊘	✓	⊘	⊘
Physical Modalities: Low Level Laser Therapy	Effect size	+	x	⊘	++	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Manual Therapies: Spinal Manipulation	Effect size	+	+	⊘	x	+	⊘
	SOE	✓	✓	⊘	✓	✓✓	⊘
Manual Therapies: Massage	Effect size	+	x	⊘	+	x	⊘
	SOE	✓✓	✓	⊘	✓✓	✓	⊘
Manual Therapies: Traction	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Mindfulness Practices: MBSR	Effect size	+	+	⊘	+	+	⊘
	SOE	✓	✓	⊘	✓✓	✓	⊘
Mind-Body Practices: Yoga	Effect size	+	+	⊘	++	++	⊘
	SOE	✓✓	✓	⊘	✓	✓✓	⊘
Acupuncture	Effect size	+	x	x	+	x	+
	SOE	✓	✓	✓	✓✓	✓	✓
Multidisciplinary Rehabilitation	Effect size	+	+	x	+	+	x
	SOE	✓	✓	✓	✓✓	✓✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
MBSR = mindfulness-based stress reduction; SOE = strength of evidence

Table 48. Nonpharmacological interventions for chronic <u>low back pain</u> compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies	Effect size	⊘	?	?	⊘	?	?
	SOE	⊘	○	○	⊘	○	○
Physical Modalities: Low Level Laser Therapy	Effect size	⊘	x	⊘	⊘	+	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Manual Therapies: Spinal Manipulation	Effect size	x	x	⊘	x	+	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Manual Therapies: Massage	Effect size	⊘	+	⊘	⊘	+	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Mind-Body Practices: Yoga	Effect size	x	x	⊘	+	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Mind-Body Practices: Qigong	Effect size	x	+	⊘	+	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Multidisciplinary Rehabilitation	Effect size	+	+	x	+	+	x
	SOE	✓✓	✓✓	✓	✓✓	✓✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 49. Nonpharmacological interventions for chronic <u>neck pain</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Psychological Therapies	Effect size	x	x	⊘	x	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Physical Modalities: Electromagnetic Field	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Physical Modalities: Low Level Laser Therapy	Effect size	++	⊘	⊘	++	⊘	⊘
	SOE	✓✓	⊘	⊘	✓✓	⊘	⊘
Manual Therapies: Traction	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Manual Therapies: Massage	Effect size	x	x	⊘	x	x	⊘
	SOE	✓	✓	⊘	✓	✓	⊘
Mind-Body Practices: Alexander Technique	Effect size	+	+	⊘	⊘	⊘	⊘
	SOE	✓	✓	⊘	⊘	⊘	⊘
Acupuncture	Effect size	+	+	x	x	x	x
	SOE	✓	✓	✓	✓	✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 50. Nonpharmacological interventions for chronic <u>neck pain</u> compared with pharmacological treatments: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	?	?	⦸	?	?	⦸
	SOE	○	○	⦸	○	○	⦸
Acupuncture	Effect size	?	⦸	⦸	⦸	⦸	⦸
	SOE	○	⦸	⦸	⦸	⦸	⦸

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⦸ No evidence
SOE = strength of evidence

Table 51. Nonpharmacological interventions for chronic <u>neck pain</u> compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies	Effect size	?	?	⦸	⦸	⦸	⦸
	SOE	○	○	⦸	⦸	⦸	⦸
Manual Therapies: Massage	Effect size	⦸	⦸	⦸	?	⦸	⦸
	SOE	⦸	⦸	⦸	○	⦸	⦸
Mind-Body Practices: Body Awareness Therapy	Effect size	?	⦸	⦸	⦸	⦸	⦸
	SOE	○	⦸	⦸	⦸	⦸	⦸

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⦸ No evidence
SOE = strength of evidence

Table 52. Nonpharmacological interventions for <u>knee osteoarthritis</u> compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	+	+	++	x
	SOE	✓✓	✓	✓	✓✓	✓	✓
Psychological Therapies	Effect size	x	x	x	x	x	x
	SOE	✓	✓	✓	✓	✓	✓
Physical Modalities: Microwave Diathermy	Effect size	?	⦸	⦸	?	⦸	⦸
	SOE	○	⦸	⦸	○	⦸	⦸
Physical	Effect size	?	⦸	?	?	⦸	?

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Modalities: Pulsed Short-Wave Diathermy	SOE	○	⦸	○	○	⦸	○
Physical Modalities: Ultrasound	Effect size	+	x	⦸	+	x	⦸
	SOE	✓	✓	⦸	✓	✓	⦸
Physical Modalities: TENS	Effect size	⦸	x	⦸	⦸	x	⦸
	SOE	⦸	✓	⦸	⦸	✓	⦸
Physical Modalities: Low-Level Laser Therapy	Effect size	?	?	⦸	?	?	⦸
	SOE	○	○	⦸	○	○	⦸
Physical Modalities: Electromagnetic Field	Effect size	x	⦸	⦸	x	⦸	⦸
	SOE	✓	⦸	⦸	✓	⦸	⦸
Physical Modalities: Superficial Heat	Effect size	⦸	⦸	⦸	?	⦸	⦸
	SOE	⦸	⦸	⦸	○	⦸	⦸
Physical Modalities: Braces	Effect size	⦸	?	?	?	?	⦸
	SOE	⦸	○	○	○	○	⦸
Manual Therapies: Joint Manipulation	Effect size	⦸	?	⦸	⦸	⦸	⦸
	SOE	⦸	○	⦸	⦸	⦸	⦸
Manual Therapies: Massage	Effect size	⦸	?	⦸	⦸	?	⦸
	SOE	⦸	○	⦸	⦸	○	⦸
Mind-Body Practices: Tai Chi	Effect size	?	?	⦸	?	?	⦸
	SOE	○	○	⦸	○	○	⦸
Acupuncture	Effect size	x	x	⦸	x	x	⦸
	SOE	✓✓	✓✓	⦸	✓✓	✓✓	⦸

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⦸ No evidence
TENS = transcutaneous electrical nerve stimulation; SOE = strength of evidence

Table 53. Nonpharmacological interventions for <u>knee osteoarthritis</u> compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Manual Therapies: Joint Manipulation	Effect size	⦸	?	⦸	⦸	⦸	⦸
	SOE	⦸	○	⦸	⦸	⦸	⦸
Acupuncture	Effect size	?	⦸	⦸	⦸	⦸	⦸
	SOE	○	⦸	⦸	⦸	⦸	⦸

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⦸ No evidence
SOE = strength of evidence

Table 54. Nonpharmacological interventions for hip osteoarthritis compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	?	+	x	?
	SOE	✓	✓	○	✓	✓	○
Manual Therapies	Effect size	⊘	?	⊘	⊘	⊘	⊘
	SOE	⊘	○	⊘	⊘	⊘	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 55. Nonpharmacological interventions for hip osteoarthritis compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Manual Therapies	Effect size	+	+	⊘	+	?	⊘
	SOE	✓	✓	⊘	✓	○	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 56. Nonpharmacological interventions for hand osteoarthritis compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	?	⊘	⊘	⊘	⊘	⊘
	SOE	○	⊘	⊘	⊘	⊘	⊘
Physical Modalities: Low Level Laser Therapy	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Physical Modalities: Heat Therapy	Effect size	x	⊘	⊘	⊘	⊘	⊘
	SOE	○	⊘	⊘	⊘	⊘	⊘
Multidisciplinary Rehabilitation	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 57. Nonpharmacological interventions for fibromyalgia compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Exercise	Effect size	+	+	x	+	x	x
	SOE	✓	✓✓	✓	✓✓	✓✓	✓✓
Psychological Therapies: CBT	Effect size	+	+	?	+	+	?
	SOE	✓	✓	○	✓	✓	○
Psychological Therapies: Biofeedback, Imagery	Effect size	?	?	?	+	+	?
	SOE	○	○	○	✓	✓	○
Physical Modalities: Magnetic Pads	Effect size	⊘	x	⊘	⊘	x	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘
Manual Therapies: Massage (Myofascial Release)	Effect size	⊘	+	x	?	?	+
	SOE	⊘	✓	✓	○	○	✓
Mindfulness Practices: MBSR	Effect size	x	⊘	⊘	x	⊘	⊘
	SOE	✓✓	⊘	⊘	✓✓	⊘	⊘
Mind-Body Practices: Qigong, Tai Chi	Effect size	+	⊘	⊘	++	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘
Acupuncture	Effect size	+	+	?	x	x	?
	SOE	✓✓	✓✓	○	✓	✓	○
Multidisciplinary Rehabilitation	Effect size	x	+	+	x	+	x
	SOE	✓	✓	✓	✓	✓	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
CBT = cognitive behavioral therapy; SOE = strength of evidence

Table 58. Psychological therapy for fibromyalgia compared with pharmacological treatments: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
CBT vs. plus amitriptyline	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Biofeedback vs. escitalopram	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
CBT vs. pregabalin; duloxetine	Effect size	⊘	+	⊘	⊘	x	⊘
	SOE	⊘	✓	⊘	⊘	✓	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
CBT = cognitive behavioral therapy; SOE = strength of evidence

Table 59. Nonpharmacological interventions for fibromyalgia compared with exercise: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapy	Effect size	?	?	?	?	?	?
	SOE	○	○	○	○	○	○
Multidisciplinary Rehabilitation	Effect size	⊘	⊘	x	⊘	⊘	x
	SOE	⊘	⊘	✓	⊘	⊘	✓

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; **x** No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Table 60. Nonpharmacological interventions for chronic tension headache compared with usual care, placebo, sham, attention control, or waitlist: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies: CBT plus Relaxation	Effect size	?	?	⊘	?	?	⊘
	SOE	○	○	⊘	○	○	⊘
Psychological Therapies: Relaxation Only	Effect size	⊘	⊘	⊘	?	⊘	⊘
	SOE	⊘	⊘	⊘	○	⊘	⊘
Physical Modalities: OTES	Effect size	?	⊘	⊘	?	⊘	⊘
	SOE	○	⊘	⊘	○	⊘	⊘
Manual Therapies: Manipulation	Effect size	+	⊘	⊘	++	⊘	⊘
	SOE	✓	⊘	⊘	✓	⊘	⊘

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Acupuncture	Effect size	⊘	⊘	⊘	+ (laser) ? (needle)	?	?
	SOE	⊘	⊘	⊘		○	○

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
OTES = occipital transcutaneous electrical stimulation; SOE = strength of evidence

Table 61. Nonpharmacological interventions for chronic tension headache compared with pharmacological treatments: summary of effects on pain and function

Intervention	Effect Size and SOE	Function *Short-term*	Function *Intermediate-term*	Function *Long-term*	Pain *Short-term*	Pain *Intermediate-term*	Pain *Long-term*
Psychological Therapies: CBT	Effect size	?	?	⊘	?	?	⊘
	SOE	○	○	⊘	○	○	⊘
Manual Therapies: Manipulation	Effect size	⊘	⊘	⊘	?	⊘	⊘
	SOE	⊘	⊘	⊘	○	⊘	⊘

+ Small magnitude of effect; ++ Moderate magnitude of effect; +++ Large magnitude of effect; x No effect/no statistically significant effect; ? Unclear; ✓Low strength of evidence; ✓✓ Moderate strength of evidence; ✓✓✓ High strength of evidence; ○ Insufficient evidence; ⊘ No evidence
SOE = strength of evidence

Low Back Pain. For chronic low back pain, there was moderate evidence of slight improvement in function, at least in the short term, for massage, yoga, psychological therapies (SOE: Moderate) and low evidence for exercise, acupuncture, low-level laser therapy, mindfulness-based stress reduction (MSBR), spinal manipulation, multidisciplinary rehabilitation (SOE: Low), compared with usual care, attention control, sham, or placebo. With the exception of spinal manipulation, these interventions also showed small effects (exercise, acupuncture, massage, MBSR, psychological therapies, multidisciplinary rehabilitation, SOE: Low) or moderate improvements (yoga, low-level laser therapy, SOE: Low) in pain short term. The small improvements in function compared with controls were sustained into the intermediate term for yoga, MBSR, spinal manipulation, psychological therapies, and multidisciplinary rehabilitation, with low strength of evidence for all but the psychological therapies, which was moderate. No clear improvement in function was seen at intermediate term for exercise, acupuncture, massage or low-level laser therapy (SOE: Low for all). Improvements in pain persisted into the intermediate term for exercise and yoga (moderate effect, SOE: Low), MBSR (small effect, SOE: Low) as well as spinal manipulation, psychological therapies and multidisciplinary rehabilitation (small effects, SOE: Moderate). Long-term evidence was available for four intervention categories: psychological therapies, multidisciplinary rehabilitation, exercise, and acupuncture. The strongest evidence was for psychological therapies, which were associated with slightly greater effects than usual care or attention control on both function and pain at short, intermediate, and long term (SOE: Moderate for all time frames). Neither exercise nor

acupuncture was associated with improved function long term, even though both demonstrated continued pain improvement (SOE: Low for all). For multidisciplinary rehabilitation, effects on function from earlier time frames were not sustained in the long term versus usual care (SOE: Low). High intensity multidisciplinary rehabilitation (≥20 hours/week or >80 hours total) was not clearly better than non-high intensity programs. Short-term effects on function and pain were somewhat larger with high intensity multidisciplinary rehabilitation than with non-high intensity interventions but the tests for interaction were not statistically significant. At intermediate term, estimates were similar for high intensity and non-high intensity programs.

In persons with chronic low back pain, there were no clear differences in short-term function for comparisons of qigong, yoga, or spinal manipulation with exercise even though small improvements in pain were seen for qigong and yoga (SOE: Low for all). Multidisciplinary rehabilitation was associated with small effects on function short term as well as pain (SOE: Moderate). Qigong and massage were associated with small effects on function at intermediate term compared with exercise, but improvement in pain was only seen for massage (SOE: Low for all). Again, multidisciplinary rehabilitation was associated with small effects on function and pain at intermediate term (SOE: Moderate), but this was not sustained in the long term (SOE: Low). Long-term data were only available for multidisciplinary rehabilitation.

Neck Pain. For chronic neck pain, in the short term, moderate effects on function and pain were seen for low-level laser therapy (SOE: Moderate). In the short term and intermediate term, acupuncture and Alexander Technique were associated with slightly greater effect on function compared with usual care (both interventions), sham acupuncture or sham laser (SOE: Low). The effect of acupuncture was not sustained long term (SOE: Low) compared with sham acupuncture, sham laser, or usual care, and no improvement in pain was seen at any time frame (SOE: Low). There were no clear differences in function or pain across types of exercise (short term) or for psychological therapies or massage compared with usual care, sham procedures, or attention controls (SOW: Low for all).

Knee Osteoarthritis. For knee osteoarthritis, only exercise and ultrasound were associated with functional improvement in the short term compared with usual care, attention control, or sham procedure; the effect size was small (SOE: Moderate for exercise, Low for ultrasound). While the small effects of exercise on function persisted into the intermediate and long term (SOE: Low for both), there were no clear benefits to ultrasound at intermediate term (SOE: Low). Similarly, small short-term effects of ultrasound on pain did not persist to intermediate term (SOE: Low) in contrast to moderate improvement in pain for exercise (SOE: Low). Long term, the small improvement in function seen with exercise was sustained, but there was no clear effect on pain (SOE: Low). There were no clear differences in function or pain associated with electromagnetic fields (short-term SOE: Low), with psychological therapies for any time frame (SOE: Low), or with acupuncture at short (SOE: Moderate) or intermediate term (SOE: Low) versus usual care, attention control, or sham procedure.

Hip and Hand Osteoarthritis. Evidence was sparse on interventions for hip and hand osteoarthritis. Exercise was associated with slightly greater function than usual care at short and intermediate-term (SOE: Low), but data were in sufficient to determine long-term effects. For pain, a small effect was seen only at short term; no differences were seen at the other time points (SOE: Low for short term and intermediate term, Insufficient for long term). Compared with

exercise, a small effect on function was seen with manual therapy in the short and intermediate term, and small improvement in pain short term (SOE: Low for all). For hand osteoarthritis, no clear differences were seen for low-level laser therapy versus sham or for multidisciplinary rehabilitation versus waitlist control at short term for either function or pain (SOE: Low).

Fibromyalgia. Short-term, in patients with fibromyalgia, there was low-quality evidence that small effects on function were associated with exercise, CBT, and mind-body practices of tai chi and qigong (SOE: Low) compared with wait list and attention control, and moderate-quality evidence for slight functional improvement acupuncture compared with sham acupuncture (SOE: Moderate). Improvements in short-term pain were seen with exercise (SOE: Moderate) and mind body practices (SOE: Low), but not with acupuncture. No clear differences in function or pain outcomes were seen for mindfulness-based stress reduction (SOE: Moderate) or multidisciplinary rehabilitation (SOE: Low). Small effects on function continued into the intermediate erm for acupuncture and cognitive behavioral therapy (SOE: Low), and small functional improvement was seen at intermediate term for myofascial release massage and multidisciplinary rehabilitation; there was no clear effect of magnetic mattress pads versus sham pad (SOE: Low for all). Small effects on pain intermediate-term were seen for psychological therapies and multidisciplinary rehabilitation (SOE: Low), but not for exercise (SOE: Moderate), acupuncture, or magnetic mattress pads (SOE: Low). Long term, small improvements in function continued for multidisciplinary rehabilitation but not for exercise or massage (SOE: Low for all), and there was no clear impact on pain for exercise (SOE: Moderate) or multidisciplinary rehabilitation (SOE: Low). No clear differences were seen between multidisciplinary rehabilitation and exercise for the long term on function or pain (SOE: Low). Cognitive behavioral therapy was associated with a small benefit for function but not for pain compared with pregabalin over the intermediate term (SOE: Low).

Chronic Tension Headache. Only nine trials of nonpharmacological treatments for chronic tension headache met the inclusion criteria and all but one was considered poor quality, resulting in a rating of insufficient evidence for comparisons of psychological therapies with waitlist or attention control, electrical stimulation versus sham, and acupuncture versus sham. One fair-quality trial of laser acupuncture versus sham suggests moderate improvement in pain short term (SOE: Low), and another fair-quality trial of spinal manipulation versus usual care suggests a small effect on short-term function based on the Headache Impact Test (SOE: Low). Approximately 25 percent of the patients in the trial had comorbid migraine headache.

Usual Care/Wait-List and Non-Active Comparators. For comparisons involving usual care/wait-list or non-active comparators (placebo, sham, attention control), there were some differences depending on the specific comparator evaluated. For some interventions results different by control type. For example, acupuncture was associated with greater effects on pain in patients with chronic low back pain when compared with usual care than when compared with sham acupuncture, suggesting that much of the benefit may be due to placebo or other non-specific effect.

Harms. Harms were poorly reported across interventions. No serious intervention-related adverse events requiring medical attention were identified; reported adverse events were generally minor

(e.g., muscle soreness with exercise, bruising with acupuncture) and time-limited (e.g., temporary worsening of pain).

Medication Use. Few trials compared opioid use pre- and post-intervention, and medication use in general was not well reported across trials.

Subgroups. One fair-quality trial in persons with knee osteoarthritis formally examined factors that might modify the effect of exercise on disability; the effect of exercise on activities of daily living disability did not appear to be modified by age, sex, baseline disability, knee pain score, BMI, or race.[46] The few trials that reported subgroup analyses either did not provide sufficient data to assess modification by demographic or other factors or did not formally test for modification; trials were generally too small to effectively evaluate outcomes in subgroups.

Findings in Relationship to What Is Already Known

Many reviews have addressed the effects of interventions for chronic pain management during or immediately following treatments. We focused on evaluating the sustainability of effects for at least 1 month post-intervention.

This review updates our previous review on low back pain[21] by incorporating new evidence on nonpharmacological treatments for chronic low back pain. The current review is based on primary literature and gives more attention to describing effects over the short, intermediate, and long terms. Consistent with the prior review, we identified small to moderate effects of exercise, yoga, various psychological therapies, acupuncture, spinal manipulation, and low-level laser therapy. This review suggests that most effects are at short or intermediate term; long-term data are sparse.

The recent Institute for Clinical and Economic Review (ICER) review[232] on chronic low back pain and neck pain used relevant portions of our previous review for chronic low back pain and updated it with new publications. The ICER review also adopted our previous review's approach to evaluation of cognitive and mind-body therapies for chronic neck pain. New publications listed in the ICER report were included in our current review if they met the inclusion criteria. Our findings are generally consistent with the ICER report relative to chronic low back pain. Differences between the reports on chronic neck pain findings are likely due to differences in inclusion criteria (particularly related to followup post intervention) and thus differences in the literature assessed. The ICER report suggests that cognitive and mind-body therapies for treatment of chronic low back pain and chronic neck pain would be cost-effective, would meet value-based price benchmarks, and may result in only a small increase ($0.75) per member per month for a hypothetical payer plan covering 1 million members, compared with approximately $4.46 per member per month for pain medication.

Our findings indicate that a number of nonpharmacological treatments improve pain and/or function. This is consistent with other reviews, including recent reviews on exercise[233] and complementary health approaches[234] for chronic pain management, an AHRQ report on knee osteoarthritis treatment,[235] and a review of chronic pain treatment guidelines on the use of manual and physical therapies.[236]

The protocol for a systematic review and network meta-analysis of interventions for fibromyalgia was identified;[237] no publication timeline for this review is currently available.

Applicability

The applicability of our findings may be impacted by a number of factors. Included trials provided limited information on diagnostic criteria, symptom duration, clinical characteristics, comorbid conditions, and concomitant treatments. Thus it is not clear to what extent these trials reflect the populations seen in clinical practice or how the variation in these factors may impact our results. Information on overlapping chronic pain conditions or psychosocial factors was generally not provided in included trials, and the extent to which these characteristics were present in trial populations and their impact on our results is not clear. Across conditions, the majority of trial participants were female, with trials of fibromyalgia and many chronic neck pain studies enrolling female participants exclusively. There were fewer female participants in trials of low back pain, osteoarthritis, and headache (57%, 68% 75% respectively). Trials also included a broad range of ages, with trials of headache and fibromyalgia generally enrolling younger participants (30 to 50 years old) than those with osteoarthritis (52 to 76 years old). Within each condition, symptom duration varied (e.g., 3 months to 15 years for low back pain, 9 months to 15 years for neck pain). Pain severity for most conditions appeared to be moderate. While we excluded trials where <90% of study sample had the defined condition, there was still likely heterogeneity. For example, some low back pain trials included a small number of patients with radiculopathy, and in some trials of chronic tension headache, a large proportion of participants had concomitant migraine headache and likely medication overuse headache. Most trials excluded persons with serious medical and psychological comorbidities. Our findings are generally most applicable to persons without such comorbidities who have moderate or severe intensity pain that has persisted for more than 1 year. The heterogeneity in populations across included trials likely is consistent with the heterogeneity seen in clinical practice, so our findings are likely applicable to most primary care clinical settings.

Variability in interventions and comparators may impact our findings. For interventions, there was variability in the numbers of sessions, length of sessions, and duration of treatment, as well as in methods of delivering the intervention. In addition, there was heterogeneity within intervention categories with regard to techniques or methods used. For example, across trials of exercise in individuals with knee osteoarthritis, duration of programs ranged from 2 to 24 weeks; the number of exercise sessions ranged from 4 to 36, each lasting 20 to over 60 minutes and including diverse types of exercise alone or in combination (e.g., aerobics, strength training, stretching, core strengthening). In general, there were no clear differences in effects based on differences in techniques, numbers of sessions, etc. However, conclusions are limited by the relatively small numbers of trials available for stratified analysis. Heterogeneity across and within comparators is also a consideration: details of usual care were rarely provided, details of co-interventions were rarely reported and likely varied across trials, and we can assume that all patients received some sort of "usual care." While the heterogeneity in interventions and comparators precludes drawing conclusions regarding specific optimal techniques and their delivery, the review findings likely represent the conditions under which the various interventions are currently delivered and speak to the need for research to identify what may be optimal.

To facilitate interpretation of results across trials and interventions, we categorized the magnitude of effects for function and pain outcomes using the system described in our previous review.[21] Using this system, beneficial effects identified were generally in the small or moderate range. We recognize that effects that we classified as small (e.g., 5 to 10 points on a 0 to 100 scale for pain or function) may be below some proposed thresholds for minimum clinically

important differences (e.g., 2 points on a 0 to 10 VAS pain scale). However, our classification provides some consistent and objective benchmarks to assess magnitude of smaller effects across trials and interventions. Interpretation of clinically important differences in mean change for continuous variables is challenging. Thus, if data were provided we also evaluated the proportion of patients who experienced a clinically important improvement in pain or function. This provides valuable insight regarding clinically important improvement. For example, one trial[84] of MBSR versus usual care in low back pain reported a small improvement in function on the RMD (1.87, 9%% CI -3.14 to -0.60 on 0-23 scale); however, the 20 percent absolute difference between treatments on the percentage of patients meeting a clinically meaningful (≥30%) improvement (68.8%- 48.6%, RR 1.56, 95% CI 1.14 to 2.14) suggests that the benefits may be more substantial.

Implications for Clinical and Policy Decisionmaking

Our review provides evidence that an array of nonpharmacological treatments provide small to moderate improvements in function and pain that are durable for more than 1 month for the five conditions addressed in this review. These encompass the vast majority of chronic pain conditions for which people seek treatment in the United States. The evidence synthesized in this review may help inform guidelines and health care policy (including reimbursement policy) related to use of noninvasive, nonpharmacological treatments as alternatives to opioids for these conditions, and inform policy decisions regarding funding priorities for future research. Recent guidelines[11] from the Centers for Disease Control and Prevention (CDC) in the United States and the Canadian Guideline for Opioid Use in Chronic Non-Cancer Pain[236] recommend non-opioid treatment as the preferred treatment for chronic pain. Further, guidelines from the American College of Physicians recommend nonpharmacological therapies over medications for chronic back pain.[22] Our findings confirm the feasibility of implementing these guideline recommendations by showing that there are some nonpharmacological treatments for chronic pain that have evidence of sustained effectiveness after the completion of therapy. Importantly, some interventions, such as exercise, multidisciplinary rehabilitation, mind-body interventions, and some complementary and integrative medicine therapies, such as acupuncture and spinal manipulation, also were associated with some sustained effects on function. There was no evidence suggesting serious harms from these interventions, although harms data were limited.

Our report reviewed evidence that may also help inform decisions regarding prioritization of nonpharmacological therapies by clinicians selecting therapy. Consistent with a biopsychosocial understanding of chronic pain,[3,5] evidence was somewhat more robust for "active" interventions that engage patients in movement and address psychological contributors to pain, particularly at longer-term followup, versus more "passive" treatments focused on symptom relief. Active interventions include exercise, multidisciplinary rehabilitation, cognitive behavioral therapy, and mind-body interventions. This provides some support for clinical strategies that focus on "active" interventions as primary therapies, with "passive" interventions used in a more adjunctive or supplementary role.

Our review also has policy implications related to access to treatment and reimbursement. Given heterogeneity in chronic pain, variability in patient preferences for treatments,[3,5] and differential responses to specific therapies in patients with a given chronic pain condition, policies that broaden access to a wider array of effective nonpharmacological treatments may have greater impact than those that focus on one or a few therapies. Several considerations could inform policy decisions regarding access to and coverage of nonpharmacological therapies.

Policymakers could prioritize access to interventions with evidence of persistent effectiveness across different pain conditions, such as exercise, multidisciplinary rehabilitation, mind-body interventions, and acupuncture. Because the level of supporting evidence varies from condition to condition, policymakers may need to consider the degree to which evidence may be reasonably extrapolated across conditions (e.g., effectiveness of psychological therapies for back pain to neck pain). Although the Affordable Care Act has improved access to complementary and integrative therapies, variability in reimbursement and authorization procedures remain a potential barrier. Although evidence supports the use of multidisciplinary rehabilitation over exercise therapy or usual care, primarily for low back pain, cost and availability remain important barriers. Our report suggests that less-intensive multidisciplinary rehabilitation may be similarly effective to high-intensity multidisciplinary rehabilitation, which could inform decisions about how such interventions are designed and delivered. In addition, not all patients may require multidisciplinary rehabilitation.[238] Policy efforts that focus on use of multidisciplinary rehabilitation in individuals more likely to benefit (e.g., severe functional deficits, failure to improve on standard nonmultidisciplinary therapies, significant psychosocial contributors to pain) could also inform efforts to deliver this modality efficiently.

Limitations of the Evidence Base

A number of limitations to the evidence base should be noted. First, evidence was relatively sparse for most interventions, particularly with regard to long-term outcomes. Data on outcomes other than pain and function were very limited. In almost all trials reporting of effects was confined to mean differences, with few reporting on the likelihood of participants successfully achieving a clinically meaningful response. Only 5 percent of included trials across conditions were considered to be of good quality; the majority were considered fair quality (58%). No trial of treatment for chronic tension headache was considered to be of good quality. A primary limitation of most trials of nonpharmacological treatments is the inability to effectively blind participants and in many cases providers; thus, observed effects may be due in part to placebo, attentional, or other non-specific effects and results may have been susceptible to performance and other biases. Many included trials were small (fewer than 70 participants) and only few or single trials were available for some interventions (e.g., some physical modalities, acupuncture). The combination of these factors for many interventions and comparators led to a determination that evidence was insufficient to determine the effects on function, pain, or harms. There was no or little includable evidence for a number of interventions, including electromuscular simulation, traction, superficial heat or cold, bracing, use of magnets, interferential therapy, transcutaneous electrical nerve stimulation, manual therapies (other than for low back pain), and functional restoration training. For most conditions, evidence was also sparse for mindfulness and mind-body practices. Evidence on interventions for hip and hand osteoarthritis and chronic tension headache was very limited.

Heterogeneity in clinical diagnosis and presentation was present for most of the conditions. In addition, it is likely that while studies reported a focus on a specific chronic pain condition (e.g., fibromyalgia or chronic tension headache), included patients may have additional conditions and/or psychological comorbidities that are not described in the trials. Details provided by trials were insufficient to conduct meaningful subanalyses.

Some of the limitations described for the review process reflect limitations of the evidence base, including those related to heterogeneity within and across interventions and heterogeneity within a given condition. Details of concurrent interventions and components of usual care were

generally not reported or poorly reported. Additionally, it is assumed that most patients with chronic pain likely continued medications and other therapies or practices during the trials. These factors likely resulted in substantial mixing of effects, so that the effects of the intervention studied could not be separated out based on information presented in the trials. These factors possibly attenuated observed effects.

As previously noted, data on long-term sustainability of the intervention effects is sparse as are data on potential harms, although serious harms are not generally expected with the interventions included in this review. Serious treatment-related adverse events were not reported in any of the trials.

Limitations of the Systematic Review Process

Our analysis was restricted to trials that reported outcomes after at least 1 month following the end of therapy. We did not attempt to address outcomes experienced during or immediately after the end of treatment or comparisons involving other active treatments. Based on input from the Technical Expert Panel (TEP), we focused on comparisons involving usual care/nonactive therapies and to active treatments of pharmacological interventions and exercise or biofeedback to provide a common point of comparison for individual interventions. Exercise was chosen as the common comparator for low back pain, neck pain, fibromyalgia, and osteoarthritis, as it is widely available, it is the most frequent comparison in the literature, and it is recommended in most guidelines for these conditions. For chronic tension headache, biofeedback was considered a more appropriate common comparator based on TEP input.

To maintain a reasonable scope for this review, inclusion was limited to trials for five common chronic pain conditions: chronic low back pain, chronic neck pain, osteoarthritis of the knee, hip or hand, fibromyalgia, and chronic tension headache, which comprise the vast majority of chronic pain conditions. There may be a need to examine the evidence on nonpharmacological treatments for other chronic pain conditions. Similarly, we focused on common noninvasive, nonpharmacological interventions determined through a protocol development process with the input of stakeholders. Examination of a number of other noninvasive interventions for chronic pain, such as topical treatments, chronic pain self-management education programs, and cannabis use was beyond the scope of this review, as were comparisons with surgical or minimally invasive procedures, such as various injection therapies.

A broad range of possible intervention types and methods of delivery exist within each category of intervention. For example, psychological therapies cover a range of different types of cognitive behavioral therapy, as well as stress reduction training, guided imagery, and biofeedback. There are numerous types of exercise and a number of different styles and components (including breathing and meditation) for the practices of yoga, tai chi and qigong. We attempted to examine the impact of different types of exercise, various styles of yoga, intensity of multidisciplinary treatment, and different control conditions. However, sparse literature for many of the interventions precluded extensive examination of the heterogeneity within a given intervention category.

Our meta-analyses sometimes included only two or three trials. DerSimonian-Laird methods may give an overly narrow confidence interval when combining small numbers of trials, particularly when homogeneity is present. We therefore also examined more conservative results from proximal likelihood methods to confirm findings; there were no substantial differences in effect sizes or conclusions identified. Nonetheless, meta-analyses based on small numbers of trials must be interpreted with caution.

Another limitation of the review process is that we excluded non-English-language articles; however, we did not identify large numbers of non-English-language articles in our review of bibliographies. We searched ClinicalTrials.gov and identified some potentially relevant studies, but none had results available. We did not formally search conference proceedings or other sources. We were unable to assess for publication bias using graphical or statistical methods to evaluate any potential impact of small samples, methodological limitations in trials, or heterogeneity in interventions, populations or outcomes. Based on hand searches of reference lists, searches of ClinicalTrials.gov, and suggestions from technical experts, we did not find evidence indicating the presence of unpublished literature sufficient to impact conclusions.

Research Recommendations

A number of evidence gaps preclude full understanding of the effectiveness, comparative effectiveness, and harms of noninvasive, nonpharmacological treatments for chronic low back pain, chronic neck pain, osteoarthritis, fibromyalgia, and chronic tension headache. Lack of followup beyond the intermediate intervention period was the most common reason for trial exclusion across interventions. Comparative trials examining the effects beyond the immediate post-intervention period, including considerations of attrition and motivation to continue, are needed to better understand whether benefits are sustained over time. Incorporation of pragmatic trial designs that incorporate strategies to improve participant recruitment, adherence, and continuation could improve retention in trials and facilitate understanding of sustainability of effects. Many excluded trials tested nonpharmacological treatments as adjuncts to other treatments to assess the incremental value of a given treatment. Among included trials for a given intervention type, there was substantial heterogeneity related to techniques used and to regimens (i.e., timing, duration, intensity). Research to identify optimal techniques and their delivery would help define more standardized interventions to evaluate in future trials. Standardization of protocols and outcomes measures would facilitate comparison of results across trials. Routine collection of common or known harms associated with interventions is needed in future trials.

Outcome measures such as the Visual Analog Scale or the Numerical Rating Scale may not fully capture the impact of pain or allow for accurate classification or evaluation of changes in chronic pain. Inclusion of recommendations[239] for pain assessment that incorporate understanding of pathophysiological mechanisms and address multiple domains of pain (including temporal dimensions, sensory and affective qualities of pain, and the location and bodily distribution of pain) in trial planning and execution may facilitate more accurate classification and longitudinal tracking of response to interventions. Reporting the proportions of patients achieving a clinically meaningful improvement in pain, function, or quality of life as measures of "success" may provide additional clinical information to complement data on average changes in continuous measures of pain, function, and quality of life, for which there is difficulty describing clinically important effects.

There is heterogeneity with regard to research design, execution, and outcomes reporting in trials of interventions included in this review, compared with well-funded trials of devices or pharmacological agents. Lack of funding to design methodologically sound studies of nonpharmacological interventions with reasonable sample sizes may have contributed to the general low quality of evidence described in this review. Education of researchers examining nonpharmacological approaches to pain management on clinical trial design, execution, and

analysis may also assist with improving the quality of the evidence base for many of the interventions.

Conclusions

Our review provides evidence that an array of nonpharmacological treatments provide small to moderate improvements in function and pain that are durable for more than 1 month for the five common chronic pain conditions addressed. Our findings provide some support for clinical strategies that prioritize use of nonpharmacological therapies for chronic pain, including "active therapies" such as exercise, psychological therapies, multidisciplinary rehabilitation, and mind-body interventions. Comparative research on the sustainability of effects beyond the immediate post-treatment period has been conducted for low back pain interventions; such research is needed for other chronic pain conditions.

References

1. Ballantyne JC, Shin NS. Efficacy of opioids for chronic pain: a review of the evidence. Clin J Pain. 2008 Jul-Aug;24(6):469-78. doi: 10.1097/AJP.0b013e31816b2f26. PMID: 18574357.
2. Eriksen J, Sjogren P, Bruera E, et al. Critical issues on opioids in chronic non-cancer pain: an epidemiological study. Pain. 2006 Nov;125(1-2):172-9. doi: 10.1016/j.pain.2006.06.009. PMID: 16842922.
3. Institute of Medicine. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington, DC: The National Academies Press; 2011.
4. [No authors listed]. Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms. Prepared by the International Association for the Study of Pain, Subcommittee on Taxonomy. Pain Suppl. 1986;3:S1-226. PMID: 3461421.
5. National Pain Strategy Task Force. National Pain Strategy: A Comprehensive Population Health-Level Strategy for Pain. Interagency Pain Research Coordinating Committee (IPRCC), National Institutes of Health (NIH); 1-83. 2015. https://iprcc.nih.gov/National_Pain_Strategy/NPS_Main.htm.
6. Boudreau D, Von Korff M, Rutter CM, et al. Trends in long-term opioid therapy for chronic non-cancer pain. Pharmacoepidemiol Drug Saf. 2009 Dec;18(12):1166-75. doi: 10.1002/pds.1833. PMID: 19718704.
7. Olsen Y, Daumit GL, Ford DE. Opioid prescriptions by U.S. primary care physicians from 1992 to 2001. J Pain. 2006 Apr;7(4):225-35. doi: 10.1016/j.jpain.2005.11.006. PMID: 16618466.
8. Sullivan MD, Edlund MJ, Fan MY, et al. Trends in use of opioids for non-cancer pain conditions 2000-2005 in commercial and Medicaid insurance plans: the TROUP study. Pain. 2008 Aug 31;138(2):440-9. doi: 10.1016/j.pain.2008.04.027. PMID: 18547726.
9. Chou R, Turner JA, Devine EB, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a National Institutes of Health Pathways to Prevention Workshop. Ann Intern Med. 2015 Feb 17;162(4):276-86. doi: 10.7326/M14-2559. PMID: 25581257.
10. Centers for Disease Control and Prevention. Vital signs: overdoses of prescription opioid pain relievers---United States, 1999--2008. MMWR Morb Mortal Wkly Rep. 2011 Nov 4;60(43):1487-92. PMID: 22048730.
11. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain--United States, 2016. JAMA. 2016 Apr 19;315(15):1624-45. doi: 10.1001/jama.2016.1464. PMID: 26977696.
12. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
13. Headache Classification Committee of the International Headache Society. The international classification of headache disorders, 3rd edition (beta version). Cephalalgia. 2013;33(9):629-808.
14. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from http://handbook.cochrane.org.; 2011.
15. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated method guideline for systematic reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015 Nov;40(21):1660-73. doi: 10.1097/BRS.0000000000001061. PMID: 26208232.
16. McCracken LM, Vowles KE. Acceptance and commitment therapy and mindfulness for chronic pain: model, process, and progress. Am Psychol. 2014 Feb-

Mar;69(2):178-87. doi: 10.1037/a0035623. PMID: 24547803.

17. A collection of R functions supporting the text book Modern Epidemiology, Second Edition, by Kenneth J. Rothman and Sander Greenland. GitHub, Inc.; 2017. https://github.com/epijim/episheet. Accessed October 13, 2017.

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

19. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. Stat Med. 1996 Mar 30;15(6):619-29. doi: 10.1002/(SICI)1097-0258(19960330)15:6<619::AID-SIM188>3.0.CO;2-A. PMID: 8731004.

20. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 06;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

21. Chou R, Deyo R, Friedly J, et al. Noninvasive treatment for low back pain. Comparative effectiveness review No. 169. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. HHSA 290-2012-00014-I.) AHRQ Publication No. 16-EHC004-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2016. PMID: 26985522.

22. Chou R, Deyo R, Friedly J, et al. Nonpharmacologic therapies for low back pain: A systematic review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2017 Feb 14;166:[Epub ahead of print]. doi: 10.7326/M16-2459. PMID: 28192793.

23. Berkman ND, Lohr KN, Ansari M, et al. Grading the strength of a body of evidence when assessing health care interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. PMID: 24404627.

24. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2015 Nov;68(11):1312-24. doi: 10.1016/j.jclinepi.2014.11.023. PMID: 25721570.

25. Costa LO, Maher CG, Latimer J, et al. Motor control exercise for chronic low back pain: a randomized placebo-controlled trial. Phys Ther. 2009 Dec;89(12):1275-86. doi: 10.2522/ptj.20090218. PMID: 19892856.

26. Goldby LJ, Moore AP, Doust J, et al. A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. Spine (Phila Pa 1976). 2006 May 01;31(10):1083-93. doi: 10.1097/01.brs.0000216464.37504.64. PMID: 16648741.

27. Kankaanpaa M, Taimela S, Airaksinen O, et al. The efficacy of active rehabilitation in chronic low back pain. Effect on pain intensity, self-experienced disability, and lumbar fatigability. Spine (Phila Pa 1976). 1999 May 15;24(10):1034-42. PMID: 10332798.

28. Nassif H, Brosset N, Guillaume M, et al. Evaluation of a randomized controlled trial in the management of chronic lower back pain in a French automotive industry: an observational study. Arch Phys Med Rehabil. 2011 Dec;92(12):1927-36.e4. doi: 10.1016/j.apmr.2011.06.029. PMID: 22133239.

29. Miyamoto GC, Costa LO, Galvanin T, et al. Efficacy of the addition of modified Pilates exercises to a minimal intervention in patients with chronic low back pain: a randomized controlled trial. Phys Ther. 2013 Mar;93(3):310-20. doi: 10.2522/ptj.20120190. PMID: 23064732.

30. Natour J, Cazotti Lde A, Ribeiro LH, et al. Pilates improves pain, function and quality of life in patients with chronic low back pain: a randomized controlled trial. Clin Rehabil. 2015 Jan;29(1):59-68. doi: 10.1177/0269215514538981. PMID: 24965957.

31. Andersen LL, Jorgensen MB, Blangsted AK, et al. A randomized controlled intervention trial to relieve and prevent neck/shoulder pain. Med Sci Sports Exerc. 2008 Jun;40(6):983-90. doi:

10.1249/MSS.0b013e3181676640. PMID: 18461010.

32. Lauche R, Stumpe C, Fehr J, et al. The Effects of Tai Chi and neck exercises in the treatment of chronic nonspecific beck pain: A randomized controlled trial. J Pain. 2016 Sep;17(9):1013-27. doi: 10.1016/j.jpain.2016.06.004. PMID: 27345663.

33. Stewart MJ, Maher CG, Refshauge KM, et al. Randomized controlled trial of exercise for chronic whiplash-associated disorders. Pain. 2007 Mar;128(1-2):59-68. PMID: 17029788.

34. Viljanen M, Malmivaara A, Uitti J, et al. Effectiveness of dynamic muscle training, relaxation training, or ordinary activity for chronic neck pain: randomised controlled trial. BMJ. 2003 Aug 30;327(7413):475. PMID: 12946968.

35. Waling K, Jarvholm B, Sundelin G. Effects of training on female trapezius Myalgia: An intervention study with a 3-year follow-up period. Spine. 2002 Apr 15;27(8):789-96. PMID: 11935098.

36. Abbott JH, Robertson MC, Chapple C, et al. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee: a randomized controlled trial. 1: clinical effectiveness. Osteoarthritis Cartilage. 2013 Apr;21(4):525-34. doi: 10.1016/j.joca.2012.12.014. PMID: 23313532.

37. Bennell KL, Hinman RS, Metcalf BR, et al. Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. Ann Rheum Dis. 2005 Jun;64(6):906-12. PMID: 15897310.

38. Chen TW, Lin CW, Lee CL, et al. The efficacy of shock wave therapy in patients with knee osteoarthritis and popliteal cyamella. Kaohsiung J Med Sci. 2014 Jul;30(7):362-70. doi: 10.1016/j.kjms.2014.03.006. PMID: 24924842.

39. Dias RC, Dias JM, Ramos LR. Impact of an exercise and walking protocol on quality of life for elderly people with OA of the knee. Physiother Res Int. 2003;8(3):121-30. PMID: 14533368.

40. 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). JAMA. 1997 Jan 1;277(1):25-31. PMID: 8980206.

41. 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.

42. Huang MH, Lin YS, Yang RC, et al. A comparison of various therapeutic exercises on the functional status of patients with knee osteoarthritis. Semin Arthritis Rheum. 2003 Jun;32(6):398-406. PMID: 12833248.

43. 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.

44. 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. doi: 10.2340/16501977-0134. PMID: 18509579.

45. 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. doi: 10.1002/art.20256. PMID: 15146420.

46. 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. Arch Intern Med. 2001 Oct 22;161(19):2309-16. PMID: 11606146.

47. 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.

48. Quilty B, Tucker M, Campbell R, et al. Physiotherapy, including quadriceps exercises and patellar taping, for knee

osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. J Rheumatol. 2003 Jun;30(6):1311-7. PMID: 12784408.

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

50. Rosedale R, Rastogi R, May S, et al. Efficacy of exercise intervention as determined by the McKenzie System of Mechanical Diagnosis and Therapy for knee osteoarthritis: a randomized controlled trial. J Orthop Sports Phys Ther. 2014 Mar;44(3):173-81, A1-6. doi: 10.2519/jospt.2014.4791. PMID: 24450370.

51. Segal NA, Glass NA, Teran-Yengle P, et al. Intensive gait training for older adults with symptomatic knee osteoarthritis. Am J Phys Med Rehabil. 2015 Oct;94(10 Suppl 1):848-58. doi: 10.1097/PHM.0000000000000264. PMID: 25768068.

52. 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.

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

54. 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 Musculoskelet Disord. 2005;6:27. PMID: 15924620.

55. 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 Jun;25(6):306-15. doi: 10.1016/S1607-551X(09)70521-2. PMID: 19560995.

56. Williamson L, Wyatt MR, Yein K, et al. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. Rheumatology. 2007 Sep;46(9):1445-9. PMID: 17604311.

57. Juhakoski R, Tenhonen S, Malmivaara A, et al. A pragmatic randomized controlled study of the effectiveness and cost consequences of exercise therapy in hip osteoarthritis. Clinical Rehabilitation. 2011 Apr;25(4):370-83. doi: 10.1177/0269215510388313. PMID: 21078702.

58. Tak E, Staats P, Van Hespen A, et al. The effects of an exercise program for older adults with osteoarthritis of the hip. J Rheumatol. 2005 Jun;32(6):1106-13. PMID: 15940775.

59. Teirlinck CH, Luijsterburg PA, Dekker J, et al. Effectiveness of exercise therapy added to general practitioner care in patients with hip osteoarthritis: a pragmatic randomized controlled trial. Osteoarthritis Cartilage. 2016 Jan;24(1):82-90. doi: 10.1016/j.joca.2015.07.023. PMID: 26254237.

60. Osteras N, Hagen KB, Grotle M, et al. Limited effects of exercises in people with hand osteoarthritis: results from a randomized controlled trial. Osteoarthritis Cartilage. 2014 Sep;22(9):1224-33. doi: 10.1016/j.joca.2014.06.036. PMID: 25008206.

61. Altan L, Korkmaz N, Bingol U, et al. Effect of pilates training on people with fibromyalgia syndrome: a pilot study. Arch Phys Med Rehabil. 2009 Dec;90(12):1983-8. doi: 10.1016/j.apmr.2009.06.021. PMID: 19969158.

62. Baptista AS, Villela AL, Jones A, et al. Effectiveness of dance in patients with fibromyalgia: a randomized, single-blind, controlled study. Clin Exp Rheumatol. 2012 Nov-Dec;30(6 Suppl 74):18-23. Epub 2012 Dec 14. PMID: 23020850.

63. Buckelew SP, Conway R, Parker J, et al. Biofeedback/relaxation training and exercise interventions for fibromyalgia: a prospective trial. Arthritis Care Res. 1998 Jun;11(3):196-209. PMID: 9782811.

64. Clarke-Jenssen AC, Mengshoel AM, Strumse YS, et al. Effect of a fibromyalgia rehabilitation programme in warm versus cold climate: a randomized controlled study.

J Rehabil Med. 2014 Jul;46(7):676-83. doi: 10.2340/16501977-1819. PMID: 24788929.

65. Da Costa D, Abrahamowicz M, Lowensteyn I, et al. A randomized clinical trial of an individualized home-based exercise programme for women with fibromyalgia. Rheumatology (Oxford). 2005 Nov;44(11):1422-7. Epub 2005 Jul 19. PMID: 16030079.

66. Fontaine KR, Conn L, Clauw DJ. Effects of lifestyle physical activity on perceived symptoms and physical function in adults with fibromyalgia: results of a randomized trial. Arthritis Res Ther. 2010;12(2):R55. doi: 10.1186/ar2967. PMID: 20353551.

67. Fontaine KR, Conn L, Clauw DJ. Effects of lifestyle physical activity in adults with fibromyalgia: results at follow-up. J Clin Rheumatol. 2011 Mar;17(2):64-8. doi: 10.1097/RHU.0b013e31820e7ea7. PMID: 21325963.

68. Giannotti E, Koutsikos K, Pigatto M, et al. Medium-/long-term effects of a specific exercise protocol combined with patient education on spine mobility, chronic fatigue, pain, aerobic fitness and level of disability in fibromyalgia. BioMed Res Int. 2014;2014:474029. doi: 10.1155/2014/474029. PMID: 24616894.

69. Gowans SE, deHueck A, Voss S, et al. Effect of a randomized, controlled trial of exercise on mood and physical function in individuals with fibromyalgia. Arthritis Rheum. 2001 Dec;45(6):519-29. PMID: 11762688.

70. Gusi N, Tomas-Carus P, Häkkinen A, et al. Exercise in waist-high warm water decreases pain and improves health-related quality of life and strength in the lower extremities in women with fibromyalgia. Arthritis Care Res. 2006;55(1):66-73. PMID: 16463415.

71. Kayo AH, Peccin MS, Sanches CM, et al. Effectiveness of physical activity in reducing pain in patients with fibromyalgia: a blinded randomized clinical trial. Rheumatol Int. 2012 Aug;32(8):2285-92. doi: 10.1007/s00296-011-1958-z. PMID: 21594719.

72. King SJ, Wessel J, Bhambhani Y, et al. The effects of exercise and education, individually or combined, in women with fibromyalgia. J Rheumatol. 2002 Dec;29(12):2620-7. PMID: 12465163.

73. Mannerkorpi K, Nordeman L, Ericsson A, et al. Pool exercise for patients with fibromyalgia or chronic widespread pain: a randomized controlled trial and subgroup analyses. J Rehabil Med. 2009 Sep;41(9):751-60. doi: 10.2340/16501977-0409. PMID: 19774310.

74. Paolucci T, Vetrano M, Zangrando F, et al. MMPI-2 profiles and illness perception in fibromyalgia syndrome: The role of therapeutic exercise as adapted physical activity. J Back Musculoskelet Rehabil. 2015;28(1):101-9. PMID: 25061029.

75. Sanudo B, Carrasco L, de Hoyo M, et al. Vagal modulation and symptomatology following a 6-month aerobic exercise program for women with fibromyalgia. Clin Exp Rheumatol. 2015 Jan-Feb;33(1 Suppl 88):S41-5. Epub 2015 Mar 17. PMID: 25786042.

76. Sanudo B, Carrasco L, de Hoyo M, et al. Effects of exercise training and detraining in patients with fibromyalgia syndrome: a 3-yr longitudinal study. Am J Phys Med Rehabil. 2012 Jul;91(7):561-9; quiz 70-3. doi: 10.1097/PHM.0b013e31824faa03. PMID: 22469880.

77. Sanudo B, Galiano D, Carrasco L, et al. Aerobic exercise versus combined exercise therapy in women with fibromyalgia syndrome: a randomized controlled trial. Arch Phys Med Rehabil. 2010 Dec;91(12):1838-43. doi: 10.016/j.apmr.2010.09.006. PMID: 21112423.

78. Tomas-Carus P, Gusi N, Hakkinen A, et al. Eight months of physical training in warm water improves physical and mental health in women with fibromyalgia: a randomized controlled trial. J Rehabil Med. 2008 Apr;40(4):248-52. doi: 10.2340/16501977-0168. PMID: 18382819.

79. Tomas-Carus P, Gusi N, Hakkinen A, et al. Improvements of muscle strength predicted benefits in HRQOL and postural balance in women with fibromyalgia: an 8-month randomized controlled trial. Rheumatology (Oxford). 2009 Sep;48(9):1147-51. doi: 10.093/rheumatology/kep208. Epub 2009 Jul 14. PMID: 19605373.

80. van Eijk-Hustings Y, Kroese M, Tan F, et al. Challenges in demonstrating the effectiveness of multidisciplinary treatment on quality of life, participation and health care utilisation in patients with fibromyalgia: a randomised controlled trial. Clin Rheumatol. 2013 Feb;32(2):199-209. doi: 10.1007/s10067-012-2100-7. Epub 2012 Oct 10. PMID: 29053692.

81. van Santen M, Bolwijn P, Verstappen F, et al. A randomized clinical trial comparing fitness and biofeedback training versus basic treatment in patients with fibromyalgia. J Rheumatol. 2002 Mar;29(3):575-81. PMID: 11908576.

82. Wigers SH, Stiles TC, Vogel PA. Effects of aerobic exercise versus stress management treatment in fibromyalgia. A 4.5 year prospective study. Scand J Rheumatol. 1996;25(2):77-86. PMID: 8614771.

83. Aslan Telci E, Karaduman A. Effects of three different conservative treatments on pain, disability, quality of life, and mood in patients with cervical spondylosis. Rheumatology International. 2012 Apr;32(4):1033-40. doi: 10.1007/s00296-010-1751-4. PMID: 21246365.

84. Cherkin DC, Sherman KJ, Balderson BH, et al. Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: a randomized clinical trial. JAMA. 2016 Mar 22-29;315(12):1240-9. doi: 10.1001/jama.2016.2323. PMID: 27002445.

85. Johnson RE, Jones GT, Wiles NJ, et al. Active exercise, education, and cognitive behavioral therapy for persistent disabling low back pain: a randomized controlled trial. Spine (Phila Pa 1976). 2007 Jul 01;32(15):1578-85. doi: 10.1097/BRS.0b013e318074f890. PMID: 17621203.

86. Lamb SE, Hansen Z, Lall R, et al. Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. Lancet. 2010 Mar 13;375(9718):916-23. doi: 10.1016/s0140-6736(09)62164-4. PMID: 20189241.

87. Lamb SE, Mistry D, Lall R, et al. Group cognitive behavioural interventions for low back pain in primary care: extended follow-up of the Back Skills Training Trial (ISRCTN54717854). Pain. 2012 Feb;153(2):494-501. doi: 10.1016/j.pain.2011.11.016. PMID: 22226729.

88. Poole H, Glenn S, Murphy P. A randomised controlled study of reflexology for the management of chronic low back pain. Eur J Pain. 2007 Nov;11(8):878-87. doi: 10.1016/j.ejpain.2007.01.006. PMID: 17459741.

89. Helminen EE, Sinikallio SH, Valjakka AL, et al. Effectiveness of a cognitive-behavioural group intervention for knee osteoarthritis pain: a randomized controlled trial. Clinical Rehabilitation. 2015 Sep;29(9):868-81. doi: 10.1177/0269215514558567. PMID: 25413168.

90. Somers TJ, Blumenthal JA, Guilak F, et al. Pain coping skills training and lifestyle behavioral weight management in patients with knee osteoarthritis: a randomized controlled study. Pain. 2012 Jun;153(6):1199-209. doi: 10.1016/j.pain.2012.02.023. PMID: 22503223.

91. Alda M, Luciano JV, Andres E, et al. Effectiveness of cognitive behaviour therapy for the treatment of catastrophisation in patients with fibromyalgia: a randomised controlled trial. Arthritis Res Ther. 2011;13(5):R173. doi: 10.1186/ar3496. PMID: 22018333.

92. Ang DC, Chakr R, Mazzuca S, et al. Cognitive-behavioral therapy attenuates nociceptive responding in patients with fibromyalgia: a pilot study. Arthritis Care Res. 2010 May;62(5):618-23. doi: 10.1002/acr.20119. PMID: 20191481.

93. Castel A, Cascon R, Padrol A, et al. Multicomponent cognitive-behavioral group therapy with hypnosis for the treatment of fibromyalgia: long-term outcome. J Pain. 2012 Mar;13(3):255-65. doi: 10.1016/j.jpain.2011.11.005. Epub 2 Jan 29. PMID: 22285609.

94. Jensen KB, Kosek E, Wicksell R, et al. Cognitive Behavioral Therapy increases pain-evoked activation of the prefrontal cortex in patients with

fibromyalgia.[Erratum appears in Pain. 2012 Sep;153(9):1982]. Pain. 2012 Jul;153(7):1495-503. doi: 10.1016/j.pain.2012.04.010. PMID: 22617632.

95. Thieme K, Flor H, Turk DC. Psychological pain treatment in fibromyalgia syndrome: efficacy of operant behavioural and cognitive behavioural treatments. Arthritis Res Ther. 2006;8(4):R121. doi: 10.1186/ar2010. PMID: 16859516.

96. Verkaik R, Busch M, Koeneman T, et al. Guided imagery in people with fibromyalgia: a randomized controlled trial of effects on pain, functional status and self-efficacy. J Health Psychol. 2014 May;19(5):678-88. PMID: 23520350.

97. Wicksell RK, Kemani M, Jensen K, et al. Acceptance and commitment therapy for fibromyalgia: a randomized controlled trial. Eur J Pain. 2013 Apr;17(4):599-611. doi: 10.1002/j.1532-2149.2012.00224.x. PMID: 23090719.

98. Williams DA, Cary MA, Groner KH, et al. Improving physical functional status in patients with fibromyalgia: a brief cognitive behavioral intervention. J Rheumatol. 2002 Jun;29(6):1280-6. PMID: 12064847.

99. Blanchard EB, Appelbaum KA, Radnitz CL, et al. Placebo-controlled evaluation of abbreviated progressive muscle relaxation and of relaxation combined with cognitive therapy in the treatment of tension headache. J Consult Clin Psychol. 1990 Apr;58(2):210-5. PMID: 2186066.

100. Holroyd KA, O'Donnell FJ, Stensland M, et al. Management of chronic tension-type headache with tricyclic antidepressant medication, stress management therapy, and their combination: a randomized controlled trial. JAMA. 2001 May 2;285(17):2208-15. PMID: 11325322.

101. Falcão DM, Sales L, Leite JR, et al. Cognitive behavioral therapy for the treatment of fibromyalgia syndrome: a randomized controlled trial. J Musculoskelet Pain. 2008;16(3):133-40.

102. Kayiran S, Dursun E, Dursun N, et al. Neurofeedback intervention in fibromyalgia syndrome; a randomized, controlled, rater blind clinical trial. Appl Psychophysiol Biofeedback. 2010 Dec;35(4):293-302. doi: 10.1007/s10484-010-9135-9. PMID: 20614235.

103. Holroyd KA, Nash JM, Pingel JD, et al. A comparison of pharmacological (amitriptyline HCL) and nonpharmacological (cognitive-behavioral) therapies for chronic tension headaches. J Consult Clin Psychol. 1991 Jun;59(3):387-93. PMID: 2071723.

104. Turner JA, Clancy S, McQuade KJ, et al. Effectiveness of behavioral therapy for chronic low back pain: a component analysis. J Consult Clin Psychol. 1990 Oct;58(5):573-9. PMID: 2147702.

105. Larsson A, Palstam A, Lofgren M, et al. Resistance exercise improves muscle strength, health status and pain intensity in fibromyalgia--a randomized controlled trial. Arthritis Res Ther. 2015 Jun 18;17:161.(doi)doi: 10.1186/s13075-015-0679-1. PMID: 26084281.

106. Redondo JR, Justo CM, Moraleda FV, et al. Long-term efficacy of therapy in patients with fibromyalgia: a physical exercise-based program and a cognitive-behavioral approach. Arthritis Rheum. 2004 Apr 15;51(2):184-92. PMID: 15077258.

107. Beurskens AJ, de Vet HC, Koke AJ, et al. Efficacy of traction for nonspecific low back pain. 12-week and 6-month results of a randomized clinical trial. Spine (Phila Pa 1976). 1997 Dec 01;22(23):2756-62. PMID: 9431610.

108. Schimmel JJ, de Kleuver M, Horsting PP, et al. No effect of traction in patients with low back pain: a single centre, single blind, randomized controlled trial of Intervertebral Differential Dynamics Therapy. Eur Spine J. 2009 Dec;18(12):1843-50. doi: 10.1007/s00586-009-1044-3. PMID: 19484433.

109. Ebadi S, Ansari NN, Naghdi S, et al. The effect of continuous ultrasound on chronic non-specific low back pain: a single blind placebo-controlled randomized trial. BMC Musculoskelet Disord. 2012 Oct 02;13:192. doi: 10.1186/1471-2474-13-192. PMID: 23031570.

110. Licciardone JC, Minotti DE, Gatchel RJ, et al. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. Ann

Fam Med. 2013 Mar-Apr;11(2):122-9. doi: 10.1370/afm.1468. PMID: 23508598.

111. Soriano F, Ríos R. Gallium arsenide laser treatment of chronic low back pain: a prospective, randomized and double blind study. Laser Ther. 1998;10(4):175-80.

112. Basford JR, Sheffield CG, Harmsen WS. Laser therapy: a randomized, controlled trial of the effects of low-intensity Nd:YAG laser irradiation on musculoskeletal back pain. Arch Phys Med Rehabil. 1999 Jun;80(6):647-52. PMID: 10378490.

113. Gibson T, Grahame R, Harkness J, et al. Controlled comparison of short-wave diathermy treatment with osteopathic treatment in non-specific low back pain. Lancet. 1985 Jun 01;1(8440):1258-61. PMID: 2860453.

114. Altan L, Bingol U, Aykac M, et al. Investigation of the effect of GaAs laser therapy on cervical myofascial pain syndrome. Rheumatol Int. 2005 Jan;25(1):23-7. doi: 10.1007/s00296-003-0396-y. PMID: 14673617.

115. Chiu TT, Ng JK, Walther-Zhang B, et al. A randomized controlled trial on the efficacy of intermittent cervical traction for patients with chronic neck pain. Clinical Rehabilitation. 2011 Sep;25(9):814-22. doi: 10.1177/0269215511399590. PMID: 21427150.

116. Chow RT, Heller GZ, Barnsley L. The effect of 300 mW, 830 nm laser on chronic neck pain: a double-blind, randomized, placebo-controlled study. Pain. 2006 Sep;124(1-2):201-10. doi: 10.1016/j.pain.2006.05.018. PMID: 16806710.

117. Gur A, Sarac AJ, Cevik R, et al. Efficacy of 904 nm gallium arsenide low level laser therapy in the management of chronic myofascial pain in the neck: a double-blind and randomize-controlled trial. Lasers Surg Med. 2004;35(3):229-35. doi: 10.1002/lsm.20082. PMID: 15389743.

118. 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.

119. Al Rashoud AS, Abboud RJ, Wang W, et al. Efficacy of low-level laser therapy applied at acupuncture points in knee osteoarthritis: a randomised double-blind comparative trial. Physiotherapy. 2014 Sep;100(3):242-8. doi: 10.1016/j.physio.2013.09.007. PMID: 24418801.

120. Battisti E, Piazza E, Rigato M, et al. Efficacy and safety of a musically modulated electromagnetic field (TAMMEF) in patients affected by knee osteoarthritis. Clin Exp Rheumatol. 2004 Sep-Oct;22(5):568-72. PMID: 15485009.

121. 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 Aug;14(8):777-83. PMID: 16563810.

122. Cakir S, Hepguler S, Ozturk C, et al. Efficacy of therapeutic ultrasound for the management of knee osteoarthritis: a randomized, controlled, and double-blind study. Am J Phys Med Rehabil. 2014 May;93(5):405-12. PMID: 24322433.

123. Fary RE, Carroll GJ, Briffa TG, et al. The effectiveness of pulsed electrical stimulation in the management of osteoarthritis of the knee: results of a double-blind, randomized, placebo-controlled, repeated-measures trial. Arthritis Rheum. 2011 May;63(5):1333-42. doi: 10.1002/art.30258. PMID: 21312188.

124. 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. doi: 10.2522/ptj.20100306. PMID: 21642511.

125. Giombini A, Di Cesare A, Di Cesare M, et al. Localized hyperthermia induced by microwave diathermy in osteoarthritis of the knee: a randomized placebo-controlled double-blind clinical trial. Knee Surg Sports Traumatol Arthrosc. 2011 Jun;19(6):980-7. PMID: 21161171.

126. Hegedus B, Viharos L, Gervain M, et al. The effect of low-level laser in knee osteoarthritis: a double-blind, randomized, placebo-controlled trial. Photomed Laser Surg. 2009 Aug;27(4):577-84. doi: 10.1089/pho.2008.2297. PMID: 19530911.

127. 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. Clinical rehabilitation. 2005 May;19(3):255-63. PMID: 15859526.

128. Mazzuca SA, Page MC, Meldrum RD, et al. Pilot study of the effects of a heat-retaining knee sleeve on joint pain, stiffness, and function in patients with knee osteoarthritis. Arthritis Rheum. 2004 Oct 15;51(5):716-21. PMID: 15478166.

129. Tascioglu F, Armagan O, Tabak Y, et al. Low power laser treatment in patients with knee osteoarthritis. Swiss Med Wkly. 2004 May 01;134(17-18):254-8. PMID: 15243853.

130. 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 Jul;13(7):575-81. PMID: 15979009.

131. Yildiz SK, Ozkan FU, Aktas I, et al. The effectiveness of ultrasound treatment for the management of knee osteoarthritis: a randomized, placebo-controlled, double-blind study. Turk J Med Sci. 2015;45(6):1187-91. PMID: 26775369.

132. Brosseau L, Wells G, Marchand S, et al. Randomized controlled trial on low level laser therapy (LLLT) in the treatment of osteoarthritis (OA) of the hand. Lasers Surg Med. 2005 Mar;36(3):210-9. doi: 10.1002/lsm.20137. PMID: 15704096.

133. Dilek B, Gozum M, Sahin E, et al. Efficacy of paraffin bath therapy in hand osteoarthritis: a single-blinded randomized controlled trial. Arch Phys Med Rehabil. 2013 Apr;94(4):642-9. doi: 10.1016/j.apmr.2012.11.024. PMID: 23187044.

134. Alfano AP, Taylor AG, Foresman PA, et al. Static magnetic fields for treatment of fibromyalgia: a randomized controlled trial. J Altern Complement Med. 2001 Feb;7(1):53-64. PMID: 11246937.

135. Bono F, Salvino D, Mazza MR, et al. The influence of ictal cutaneous allodynia on the response to occipital transcutaneous electrical stimulation in chronic migraine and chronic tension-type headache: a randomized, sham-controlled study. Cephalalgia. 2015 Apr;35(5):389-98. doi: 10.1177/0333102414544909. PMID: 25078717.

136. Djavid GE, Mehrdad R, Ghasemi M, et al. In chronic low back pain, low level laser therapy combined with exercise is more beneficial than exercise alone in the long term: a randomised trial. Aust J Physiother. 2007;53(3):155-60. PMID: 17725472.

137. Haas M, Vavrek D, Peterson D, et al. Dose-response and efficacy of spinal manipulation for care of chronic low back pain: a randomized controlled trial. Spine J. 2014 Jul 01;14(7):1106-16. doi: 10.1016/j.spinee.2013.07.468. PMID: 24139233.

138. Hondras MA, Long CR, Cao Y, et al. A randomized controlled trial comparing 2 types of spinal manipulation and minimal conservative medical care for adults 55 years and older with subacute or chronic low back pain. J Manipulative Physiol Ther. 2009 Jun;32(5):330-43. doi: 10.1016/j.jmpt.2009.04.012. PMID: 19539115.

139. Senna MK, Machaly SA. Does maintained spinal manipulation therapy for chronic nonspecific low back pain result in better long-term outcome? Spine (Phila Pa 1976). 2011 Aug 15;36(18):1427-37. doi: 10.1097/BRS.0b013e3181f5dfe0. PMID: 21245790.

140. UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ. 2004 Dec 11;329(7479):1377. doi: 10.1136/bmj.38282.669225.AE. PMID: 15556955.

141. Ajimsha MS, Daniel B, Chithra S. Effectiveness of myofascial release in the management of chronic low back pain in nursing professionals. J Bodyw Mov Ther. 2014 Apr;18(2):273-81. doi: 10.1016/j.jbmt.2013.05.007. PMID: 24725797.

142. Cherkin DC, Eisenberg D, Sherman KJ, et al. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic

low back pain. Arch Intern Med. 2001 Apr 23;161(8):1081-8. PMID: 11322842.

143. Quinn F, Hughes CM, Baxter GD. Reflexology in the management of low back pain: a pilot randomised controlled trial. Complement Ther Med. 2008 Feb;16(1):3-8. doi: 10.1016/j.ctim.2007.05.001. PMID: 18346622.

144. Rudolfsson T, Djupsjobacka M, Hager C, et al. Effects of neck coordination exercise on sensorimotor function in chronic neck pain: a randomized controlled trial. J Rehabil Med. 2014 Oct;46(9):908-14. doi: 10.2340/16501977-1869. PMID: 25182501.

145. Sherman KJ, Cherkin DC, Hawkes RJ, et al. Randomized trial of therapeutic massage for chronic neck pain. Clin J Pain. 2009 Mar-Apr;25(3):233-8. doi: 10.1097/AJP.0b013e31818b7912. PMID: 19333174.

146. Perlman AI, Ali A, Njike VY, et al. Massage therapy for osteoarthritis of the knee: a randomized dose-finding trial. PLoS ONE [Electronic Resource]. 2012;7(2):e30248. doi: 10.1371/journal.pone.0030248. PMID: 22347369.

147. Castro-Sanchez AM, Mataran-Penarrocha GA, Arroyo-Morales M, et al. Effects of myofascial release techniques on pain, physical function, and postural stability in patients with fibromyalgia: a randomized controlled trial. Clin Rehabil. 2011 Sep;25(9):800-13. doi: 10.1177/0269215511399476. PMID: 21673013.

148. Castro-Sanchez AM, Mataran-Penarrocha GA, Granero-Molina J, et al. Benefits of massage-myofascial release therapy on pain, anxiety, quality of sleep, depression, and quality of life in patients with fibromyalgia. Evid Based Complement Alternat Med. 2011;2011:561753. doi: 10.1155/2011/561753. PMID: 21234327.

149. Castien RF, van der Windt DA, Grooten A, et al. Effectiveness of manual therapy for chronic tension-type headache: a pragmatic, randomised, clinical trial. Cephalalgia. 2011 Jan;31(2):133-43. doi: 10.1177/0333102410377362. PMID: 20647241.

150. Boline PD, Kassak K, Bronfort G, et al. Spinal manipulation vs. amitriptyline for the treatment of chronic tension-type headaches: a randomized clinical trial. J Manipulative Physiol Ther. 1995 Mar-Apr;18(3):148-54. PMID: 7790794.

151. Little P, Lewith G, Webley F, et al. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain. BMJ. 2008 Aug 19;337:a884. doi: 10.1136/bmj.a884. PMID: 18713809.

152. Bronfort G, Maiers MJ, Evans RL, et al. Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. Spine J. 2011 Jul;11(7):585-98. doi: 10.1016/j.spinee.2011.01.036. PMID: 21622028.

153. Ferreira ML, Ferreira PH, Latimer J, et al. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: A randomized trial. Pain. 2007 Sep;131(1-2):31-7. doi: 10.1016/j.pain.2006.12.008. PMID: 17250965.

154. Gudavalli MR, Cambron JA, McGregor M, et al. A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. Eur Spine J. 2006 Jul;15(7):1070-82. doi: 10.1007/s00586-005-0021-8. PMID: 16341712.

155. Hoeksma HL, Dekker J, Ronday HK, et al. Comparison of manual therapy and exercise therapy in osteoarthritis of the hip: a randomized clinical trial. Arthritis Rheum. 2004 Oct 15;51(5):722-9. PMID: 15478147.

156. Banth S, Ardebil MD. Effectiveness of mindfulness meditation on pain and quality of life of patients with chronic low back pain. Int J Yoga. 2015 Jul-Dec;8(2):128-33. doi: 10.4103/0973-6131.158476. PMID: 26170592.

157. Morone NE, Greco CM, Moore CG, et al. A mind-body program for older adults with chronic low back pain: a randomized clinical trial. JAMA Intern Med. 2016 Mar;176(3):329-37. doi: 10.1001/jamainternmed.2015.8033. PMID: 26903081.

158. Morone NE, Rollman BL, Moore CG, et al. A mind-body program for older adults with

chronic low back pain: results of a pilot study. Pain Med. 2009 Nov;10(8):1395-407. doi: 10.1111/j.1526-4637.2009.00746.x. PMID: 20021599.

159. Cash E, Salmon P, Weissbecker I, et al. Mindfulness meditation alleviates fibromyalgia symptoms in women: results of a randomized clinical trial. Ann Behav Med. 2015 Jun;49(3):319-30. doi: 10.1007/s12160-014-9665-0. PMID: 25425224.

160. Schmidt S, Grossman P, Schwarzer B, et al. Treating fibromyalgia with mindfulness-based stress reduction: results from a 3-armed randomized controlled trial. Pain. 2011 Feb;152(2):361-9. doi: 10.1016/j.pain.2010.10.043. PMID: 21146930.

161. Sephton SE, Salmon P, Weissbecker I, et al. Mindfulness meditation alleviates depressive symptoms in women with fibromyalgia: results of a randomized clinical trial. Arthritis Rheum. 2007 Feb 15;57(1):77-85. doi: 10.1002/art.22478. PMID: 17266067.

162. Saper RB, Lemaster C, Delitto A, et al. Yoga, physical therapy, or education for chronic low back pain: A randomized noninferiority trial. Ann Intern Med. 2017 Jul 18;167(2):85-94. doi: 10.7326/m16-2579. PMID: 28631003.

163. Sherman KJ, Cherkin DC, Erro J, et al. Comparing yoga, exercise, and a self-care book for chronic low back pain: a randomized, controlled trial. Ann Intern Med. 2005 Dec 20;143(12):849-56. PMID: 16365466.

164. Sherman KJ, Cherkin DC, Wellman RD, et al. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. Arch Intern Med. 2011 Dec 12;171(22):2019-26. doi: 10.1001/archinternmed.2011.524. PMID: 22025101.

165. Tilbrook HE, Cox H, Hewitt CE, et al. Yoga for chronic low back pain: a randomized trial. Ann Intern Med. 2011 Nov 01;155(9):569-78. doi: 10.7326/0003-4819-155-9-201111010-00003. PMID: 22041945.

166. Williams K, Abildso C, Steinberg L, et al. Evaluation of the effectiveness and efficacy of Iyengar yoga therapy on chronic low back pain. Spine (Phila Pa 1976). 2009 Sep 01;34(19):2066-76. doi: 10.1097/BRS.0b013e3181b315cc. PMID: 19701112.

167. Williams KA, Petronis J, Smith D, et al. Effect of Iyengar yoga therapy for chronic low back pain. Pain. 2005 May;115(1-2):107-17. doi: 10.1016/j.pain.2005.02.016. PMID: 15836974.

168. MacPherson H, Tilbrook H, Richmond S, et al. Alexander Technique Lessons or Acupuncture Sessions for Persons With Chronic Neck Pain: A Randomized Trial.[Summary for patients in Ann Intern Med. 2015 Nov 3;163(9):I30; PMID: 26524582]. Ann Intern Med. 2015 Nov 3;163(9):653-62. doi: 10.7326/M15-0667. PMID: 26524571.

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

170. Wang C, Schmid CH, Hibberd PL, et al. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. Arthritis Rheum. 2009 Nov 15;61(11):1545-53. doi: 10.1002/art.24832. PMID: 19877092.

171. Lynch M, Sawynok J, Hiew C, et al. A randomized controlled trial of qigong for fibromyalgia. Arthritis Res Ther. 2012;14(4):R178. doi: 10.1186/ar3931. PMID: 22863206.

172. Wang C, Schmid CH, Rones R, et al. A randomized trial of tai chi for fibromyalgia. N Engl J Med. 2010 Aug 19;363(8):743-54. doi: 10.1056/NEJMoa0912611. PMID: 20818876.

173. Blodt S, Pach D, Kaster T, et al. Qigong versus exercise therapy for chronic low back pain in adults--a randomized controlled non-inferiority trial. Eur J Pain. 2015 Jan;19(1):123-31. doi: 10.1002/ejp.529. PMID: 24902673.

174. Nambi GS, Inbasekaran D, Khuman R, et al. Changes in pain intensity and health related quality of life with Iyengar yoga in nonspecific chronic low back pain: A randomized controlled study. Int J Yoga.

2014 Jan;7(1):48-53. doi: 10.4103/0973-6131.123481. PMID: 25035607.

175. Lansinger B, Larsson E, Persson LC, et al. Qigong and exercise therapy in patients with long-term neck pain: a prospective randomized trial. Spine. 2007 Oct 15;32(22):2415-22. PMID: 18090079.

176. Seferiadis A, Ohlin P, Billhult A, et al. Basic body awareness therapy or exercise therapy for the treatment of chronic whiplash associated disorders: a randomized comparative clinical trial. Disabil Rehabil. 2016;38(5):442-51. doi: 10.3109/09638288.2015.1044036. PMID: 25955823.

177. Brinkhaus B, Witt CM, Jena S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. Arch Intern Med. 2006 Feb 27;166(4):450-7. doi: 10.1001/archinte.166.4.450. PMID: 16505266.

178. Carlsson CP, Sjolund BH. Acupuncture for chronic low back pain: a randomized placebo-controlled study with long-term follow-up. Clin J Pain. 2001 Dec;17(4):296-305. PMID: 11783809.

179. Cherkin DC, Sherman KJ, Avins AL, et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. Arch Intern Med. 2009 May 11;169(9):858-66. doi: 10.1001/archinternmed.2009.65. PMID: 19433697.

180. Cho YJ, Song YK, Cha YY, et al. Acupuncture for chronic low back pain: a multicenter, randomized, patient-assessor blind, sham-controlled clinical trial. Spine (Phila Pa 1976). 2013 Apr 01;38(7):549-57. doi: 10.1097/BRS.0b013e318275e601. PMID: 23026870.

181. Haake M, Muller HH, Schade-Brittinger C, et al. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. Arch Intern Med. 2007 Sep 24;167(17):1892-8. doi: 10.1001/archinte.167.17.1892. PMID: 17893311.

182. Kerr DP, Walsh DM, Baxter D. Acupuncture in the management of chronic low back pain: a blinded randomized controlled trial. Clin J Pain. 2003 Nov-Dec;19(6):364-70. PMID: 14600536.

183. Thomas KJ, MacPherson H, Thorpe L, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. BMJ. 2006 Sep 23;333(7569):623. doi: 10.1136/bmj.38878.907361.7C. PMID: 16980316.

184. Birch S, Jamison RN. Controlled trial of Japanese acupuncture for chronic myofascial neck pain: assessment of specific and nonspecific effects of treatment. Clin J Pain. 1998 Sep;14(3):248-55. PMID: 9758075.

185. Liang Z, Zhu X, Yang X, et al. Assessment of a traditional acupuncture therapy for chronic neck pain: a pilot randomised controlled study. Complement Ther Med. 2011 Jan;19 Suppl 1:S26-32. doi: 10.1016/j.ctim.2010.11.005. PMID: 21195292.

186. Sahin N, Ozcan E, Sezen K, et al. Efficacy of acupunture in patients with chronic neck pain--a randomised, sham controlled trial. Acupunct Electrother Res. 2010;35(1-2):17-27. PMID: 20578644.

187. Vas J, Perea-Milla E, Mendez C, et al. Efficacy and safety of acupuncture for chronic uncomplicated neck pain: a randomised controlled study. Pain. 2006 Dec 15;126(1-3):245-55. PMID: 16934402.

188. White P, Lewith G, Prescott P, et al. Acupuncture versus placebo for the treatment of chronic mechanical neck pain: a randomized, controlled trial.[Summary for patients in Ann Intern Med. 2004 Dec 21;141(12):I26; PMID: 15611483]. Ann Intern Med. 2004 Dec 21;141(12):911-9. PMID: 15611488.

189. Zhang SP, Chiu TT, Chiu SN. Long-term efficacy of electroacupuncture for chronic neck pain: a randomised controlled trial. Hong Kong Med J. 2013 Dec;19 Suppl 9:36-9. PMID: 24473589.

190. Berman BM, Singh BB, Lao L, et al. A randomized trial of acupuncture as an adjunctive therapy in osteoarthritis of the knee. Rheumatology. 1999 Apr;38(4):346-54. PMID: 10378713.

191. Hinman RS, McCrory P, Pirotta M, et al. Acupuncture for chronic knee pain: a randomized clinical trial.[Summary for patients in JAMA. 2014 Oct 1;312(13):1365; PMID: 25268455]. JAMA. 2014 Oct 1;312(13):1313-22. doi: 10.1001/jama.2014.12660. PMID: 25268438.

192. Jubb RW, Tukmachi ES, Jones PW, et al. A blinded randomised trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupunct Med. 2008 Jun;26(2):69-78. PMID: 18591906.

193. Lansdown H, Howard K, Brealey S, et al. Acupuncture for pain and osteoarthritis of the knee: a pilot study for an open parallel-arm randomised controlled trial. BMC Musculoskelet Disord. 2009;10:130. doi: 10.1186/1471-2474-10-130. PMID: 19852841.

194. Suarez-Almazor ME, Looney C, Liu Y, et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. Arthritis Care Res. 2010 Sep;62(9):1229-36. doi: 10.1002/acr.20225. PMID: 20506122.

195. Witt C, Brinkhaus B, Jena S, et al. Acupuncture in patients with osteoarthritis of the knee: a randomised trial. Lancet. 2005 Jul 9-15;366(9480):136-43. PMID: 16005336.

196. Yurtkuran M, Alp A, Konur S, et al. Laser acupuncture in knee osteoarthritis: a double-blind, randomized controlled study. Photomed Laser Surg. 2007 Feb;25(1):14-20. PMID: 17352632.

197. Assefi NP, Sherman KJ, Jacobsen C, et al. A randomized clinical trial of acupuncture compared with sham acupuncture in fibromyalgia.[Summary for patients in Ann Intern Med. 2005 Jul 5;143(1):I24; PMID: 15998747]. Ann Intern Med. 2005 Jul 5;143(1):10-9. PMID: 15998750.

198. Martin DP, Sletten CD, Williams BA, et al. Improvement in fibromyalgia symptoms with acupuncture: results of a randomized controlled trial. Mayo Clin Proc. 2006 Jun;81(6):749-57. doi: 10.4065/81.6.749. PMID: 16770975.

199. Vas J, Santos-Rey K, Navarro-Pablo R, et al. Acupuncture for fibromyalgia in primary care: a randomised controlled trial. Acupunct Med. 2016 Aug;34(4):257-66. doi: 10.1136/acupmed-2015-010950. Epub 2016 Feb 15. PMID: 26879181.

200. Ebneshahidi NS, Heshmatipour M, Moghaddami A, et al. The effects of laser acupuncture on chronic tension headache--a randomised controlled trial. Acupunct Med. 2005 Mar;23(1):13-8. PMID: 15844435.

201. Karst M, Rollnik JD, Fink M, et al. Pressure pain threshold and needle acupuncture in chronic tension-type headache--a double-blind placebo-controlled study. Pain. 2000 Nov;88(2):199-203. PMID: 11050375.

202. Tavola T, Gala C, Conte G, et al. Traditional Chinese acupuncture in tension-type headache: a controlled study. Pain. 1992 Mar;48(3):325-9. PMID: 1594255.

203. Cho JH, Nam DH, Kim KT, et al. Acupuncture with non-steroidal anti-inflammatory drugs (NSAIDs) versus acupuncture or NSAIDs alone for the treatment of chronic neck pain: an assessor-blinded randomised controlled pilot study. Acupunct Med. 2014 Feb;32(1):17-23. doi: 10.1136/acupmed-2013-010410. PMID: 24171895.

204. Bendix AF, Bendix T, Vaegter K, et al. Multidisciplinary intensive treatment for chronic low back pain: a randomized, prospective study. Cleve Clin J Med. 1996 Jan-Feb;63(1):62-9. PMID: 8590519.

205. Harkapaa K, Jarvikoski A, Mellin G, et al. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part I. Pain, disability, compliance, and reported treatment benefits three months after treatment. Scand J Rehabil Med. 1989;21(2):81-9. PMID: 2526364.

206. Von Korff M, Balderson BH, Saunders K, et al. A trial of an activating intervention for chronic back pain in primary care and physical therapy settings. Pain. 2005 Feb;113(3):323-30. doi: 10.1016/j.pain.2004.11.007. PMID: 15661440.

207. Lambeek LC, van Mechelen W, Knol DL, et al. Randomised controlled trial of integrated care to reduce disability from chronic low back pain in working and private life. BMJ.

2010 Mar 16;340:c1035. doi: 10.1136/bmj.c1035. PMID: 20234040.

208. Abbasi M, Dehghani M, Keefe FJ, et al. Spouse-assisted training in pain coping skills and the outcome of multidisciplinary pain management for chronic low back pain treatment: a 1-year randomized controlled trial. Eur J Pain. 2012 Aug;16(7):1033-43. doi: 10.1002/j.1532-2149.2011.00097.x. PMID: 22337646.

209. Strand LI, Ljunggren AE, Haldorsen EM, et al. The impact of physical function and pain on work status at 1-year follow-up in patients with back pain. Spine (Phila Pa 1976). 2001 Apr 01;26(7):800-8. PMID: 11295903.

210. Stukstette MJ, Dekker J, den Broeder AA, et al. No evidence for the effectiveness of a multidisciplinary group based treatment program in patients with osteoarthritis of hands on the short term; results of a randomized controlled trial. Osteoarthritis Cartilage. 2013 Jul;21(7):901-10. doi: 10.1016/j.joca.2013.03.016. PMID: 23583457.

211. Amris K, Waehrens EE, Christensen R, et al. Interdisciplinary rehabilitation of patients with chronic widespread pain: primary endpoint of the randomized, nonblinded, parallel-group IMPROvE trial. Pain. 2014 Jul;155(7):1356-64. doi: 10.1016/j.pain.2014.04.012. PMID: 24727345.

212. Castel A, Fontova R, Montull S, et al. Efficacy of a multidisciplinary fibromyalgia treatment adapted for women with low educational levels: a randomized controlled trial. Arthritis Care Res. 2013 Mar;65(3):421-31. doi: 10.1002/acr.21818. PMID: 22899402.

213. Castel A, Castro S, Fontova R, et al. Body mass index and response to a multidisciplinary treatment of fibromyalgia. Rheumatol Int. 2015 Feb;35(2):303-14. doi: 10.1007/s00296-014-3096-x. Epub 2014 Aug 1. PMID: 25080875.

214. Cedraschi C, Desmeules J, Rapiti E, et al. Fibromyalgia: a randomised, controlled trial of a treatment programme based on self management. Ann Rheum Dis. 2004 Mar;63(3):290-6. PMID: 14962965.

215. Martin J, Torre F, Padierna A, et al. Six-and 12-month follow-up of an interdisciplinary fibromyalgia treatment programme: results of a randomised trial. Clin Exp Rheumatol. 2012 Nov-Dec;30(6 Suppl 74):103-11. Epub 2012 Dec 14. PMID: 23261008.

216. Tavafian SS, Jamshidi AR, Montazeri A. A randomized study of back school in women with chronic low back pain: quality of life at three, six, and twelve months follow-up. Spine (Phila Pa 1976). 2008 Jul 01;33(15):1617-21. doi: 10.1097/BRS.0b013e31817bd31c. PMID: 18580739.

217. Bendix AF, Bendix T, Ostenfeld S, et al. Active treatment programs for patients with chronic low back pain: a prospective, randomized, observer-blinded study. Eur Spine J. 1995;4(3):148-52. PMID: 7552649.

218. Bendix T, Bendix A, Labriola M, et al. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. Spine (Phila Pa 1976). 2000 Oct 01;25(19):2494-500. PMID: 11013502.

219. Jousset N, Fanello S, Bontoux L, et al. Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. Spine (Phila Pa 1976). 2004 Mar 01;29(5):487-93; discussion 94. PMID: 15129059.

220. Nicholas MK, Wilson PH, Goyen J. Operant-behavioural and cognitive-behavioural treatment for chronic low back pain. Behav Res Ther. 1991;29(3):225-38. PMID: 1831972.

221. Nicholas MK, Wilson PH, Goyen J. Comparison of cognitive-behavioral group treatment and an alternative non-psychological treatment for chronic low back pain. Pain. 1992 Mar;48(3):339-47. PMID: 1534400.

222. van der Roer N, van Tulder M, Barendse J, et al. Intensive group training protocol versus guideline physiotherapy for patients with chronic low back pain: a randomised controlled trial. Eur Spine J. 2008 Sep;17(9):1193-200. doi: 10.1007/s00586-008-0718-6. PMID: 18663487.

223. Monticone M, Ferrante S, Rocca B, et al. Effect of a long-lasting multidisciplinary program on disability and fear-avoidance

behaviors in patients with chronic low back pain: results of a randomized controlled trial. Clin J Pain. 2013 Nov;29(11):929-38. doi: 10.1097/AJP.0b013e31827fef7e. PMID: 23328343.

224. Monticone M, Ambrosini E, Rocca B, et al. A multidisciplinary rehabilitation programme improves disability, kinesiophobia and walking ability in subjects with chronic low back pain: results of a randomised controlled pilot study. Eur Spine J. 2014 Oct;23(10):2105-13. doi: 10.1007/s00586-014-3478-5. PMID: 25064093.

225. Roche G, Ponthieux A, Parot-Shinkel E, et al. Comparison of a functional restoration program with active individual physical therapy for patients with chronic low back pain: a randomized controlled trial. Arch Phys Med Rehabil. 2007 Oct;88(10):1229-35. doi: 10.1016/j.apmr.2007.07.014. PMID: 17908562.

226. Roche-Leboucher G, Petit-Lemanac'h A, Bontoux L, et al. Multidisciplinary intensive functional restoration versus outpatient active physiotherapy in chronic low back pain: a randomized controlled trial. Spine (Phila Pa 1976). 2011 Dec 15;36(26):2235-42. doi: 10.1097/BRS.0b013e3182191e13. PMID: 21415807.

227. Bendix AF, Bendix T, Lund C, et al. Comparison of three intensive programs for chronic low back pain patients: a prospective, randomized, observer-blinded study with one-year follow-up. Scand J Rehabil Med. 1997 Jun;29(2):81-9. PMID: 9198257.

228. Bendix AE, Bendix T, Haestrup C, et al. A prospective, randomized 5-year follow-up study of functional restoration in chronic low back pain patients. Eur Spine J. 1998;7(2):111-9. PMID: 9629934.

229. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clin Psychol Sci Pract. 2003;10(2):144-56. doi: 10.1093/clipsy.bpg016.

230. Chou R, Deyo R, Friedly J, et al. Systemic pharmacologic therapies for low back pain: A systematic review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2017 Feb 14;166:[Epub ahead of print]. doi: 10.7326/M16-2458. PMID: 28192790.

231. Chou R, Deyo R, Devine B, et al. The Effectiveness and Risks of Long-Term Opioid Treatment of Chronic Pain. Evidence Report/Technology Assessment No. 218. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 14-E005-EF. Agency for Healthcare Research and Quality; September 2014. doi: 10.23970/AHRQEPCERTA218.

232. Tice J, Kumar V, Otunoye I, et al. Cognitive and Mind-Body Therapies for Chronic Low Back and Neck Pain: Effectiveness and Value. Evidence Report. Prepared for The California Technology Assesment Forum. Boston, MA: The Institute for Clinical and Economic Review; 2017. https://icer-review.org/wp-content/uploads/2017/03/CTAF_Chronic_Pain__Evidence_Report_100417.pdf Accessed October 13, 2017.

233. Geneen LJ, Moore RA, Clarke C, et al. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. The Cochrane Library. 2017.

234. Nahin RL, Boineau R, Khalsa PS, et al. Evidence-based evaluation of complementary health approaches for pain management in the United States. Mayo Clin Proc. 2016 Sep;91(9):1292-306. doi: 10.1016/j.mayocp.2016.06.007. PMID: 27594189.

235. Newberry SJ, FitzGerald J, SooHoo NF, et al. Treatment of Osteoarthritis of the Knee: An Update Review. Comparative Effectiveness Review No. 190. (Prepared by the RAND Southern California Evidence-based Practice Center under Contract No. 290-2015-00010-I.) AHRQ Publication No.17-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2017. doi: 10.23970/AHRQEPCCER190. PMID: 28825779.

236. Busse JW, Craigie S, Juurlink DN, et al. Guideline for opioid therapy and chronic noncancer pain. CMAJ. 2017 May 8;198(18):E659-E66. doi: 10.1503/cmaj.170363. PMID: 28483845.

237. Busse JW, Ebrahim S, Connell G, et al. Systematic review and network meta-analysis of interventions for fibromyalgia: a protocol. Syst Rev. 2013 Mar 13;2:18. doi: 10.1186/2046-4053-2-18. PMID: 23497523.

238. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet. 2011 Oct 29;378(9802):1560-71. doi: 10.1016/S0140-6736(11)60937-9. PMID: 21963002.

239. Fillingim RB, Loeser JD, Baron R, et al. Assessment of chronic pain: Domains, methods, and mechanisms. J Pain. 2016 Sep;17(9 Suppl):T10-20. doi: 10.1016/j.jpain.2015.08.010. PMID: 27586827.

Acronyms and Abbreviations

Acronym/Abbreviation	Term
AC	Attention Control
ADL	Activities of daily living
AIMS	Arthritis Impact Measurement Scale
AQoL 6D	Assessment of Quality of Life version 6D
AUSCAN	Australia Canadian Osteoarthritis Hand Index
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
BMI	Body mass index
BPI	Brief Pain Inventory
BPI-SF	Brief Pain Inventory-Short Form
CBT	Cognitive behavioral therapy
CDC HRQOL-4	Centers for Disease Control and Prevention's Health-Related Quality of Life Questionnaire
CES-D	Center for Epidemiologic Studies Depression Scale
CGI-I	Clinical Global Impressions of Improvement Scale
CGI-S	Clinical Global Impressions of Severity Scale
CI	Confidence interval
CSQ	Coping Strategies Questionnaire
DASS	Depression Anxiety Stress Scales
DPQ	Dallas Pain Questionnaire
DFI	Dreiser Functional Index
EEG	Electroencephalography
EMG	Electromyography
EQ-5D	EuroQoL-5D
FABQ	Fear Avoidance Beliefs Questionnaire
FIHOA	Functional Index for Hand Osteoarthritis
FIQ	Fibromyalgia Impact Questionnaire
FMI	Freiburg Mindfulness Inventory
FRI	Functional Rating Index
FSI	Fatigue Symptom Inventory
GAR	Groningen Activity Restriction Scale
GCQ, GBB-24	Giessen Complaint Questionnaire
GDS	Geriatric Depression Scale
GPE	Global Perceived Effect Scale
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
GSI	Global Severity Index (Symptom Checklist-90-Revised)
HADS	The Hospital Anxiety and Depression Scale
HAM-A	Hamilton Anxiety Rating Scale (HAM-A)
HAM-D	Hamilton Depression Rating Scale (HAM-D)
HAQ	Health Assessment Questionnaire
HDI	Headache Disability Inventory
HHS	Harris Hip Score
HIT-6	Headache Impact Test-6
HRQoL	Health-related quality of life
HSCL-25	Hopkin’s Symptom Checklist
HSS	Hospital for Special Surgery Knee Function
IPAQ	International Physical Activity Questionnaire
IPQ(-R)	Illness Perception Questionnaire(-Revised)
IQR	Interquartile range
ITT	Intention-to-treat
KPS	Knee Pain Scale
JLEQ	Japan Low Back Pain Evaluation Questionnaire

Acronym/Abbreviation	Term
JOA	Japanese Orthopedic Association
LBP	Low back pain
LBPOI	Low Back Pain Outcome Instrument
LBPRS	Low back pain rating scale
LLFDI	Late Life Function and Disability Instrument
MACTAR	McMaster Toronto Arthritis patient preference questionnaire
MASS	Mindful Attention Awareness Scale
MBSR	Mindfulness-based Stress Reduction
MCE	Motor control exercise
MCS-12	Mental component score of the SF-12
MD	Mean difference
MIDAS	Migraine Disability Assessment questionnaire
MOS	Medical Outcome Study
MPI	Multidimensional Pain Inventory
MPQ(-SF)	McGill Pain Questionnaire(-Short Form)
NDI	Neck Disability Index
NHP	Nottingham Health Profile
NIAMS	National Institute of Arthritis and Musculoskeletal and Skin Diseases
NIH	National Institute of Health
NPAD	Neck Pain and Disability Index
NR	Not reported
NRS	Numeric rating scale
NS	Not statistically significant
NSAID	Nonsteroidal anti-inflammatory drug
NT	No treatment
OA	Osteoarthritis
OARSI-OMERACT	Osteoarthritis Research Society International – Outcome Measures in Rheumatology
ODI	Oswestry Disability Index
OKS	Oxford Knee Score
PDI	Pain Disability Index
PPS	Pain Perception Scale
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSEQ	Chronic Pain Self Efficacy Scale
PSFS	Patient-Specific Functional Scale
PSQI	Pittsburgh Sleep Quality Index
PSS	Perceived Stress Scale
PT	Physical therapy
QBPDS	Quebec Back Pain Disability Scale
QHS	Each night at bedtime
QOL	Quality of life
RAND-36 QoL	Quality of Life RAND-36
QoL VAS	Quality of Life Visual Analog Scale
RCT	Randomized controlled trail
RDMQ	Roland-Morris Disability Questionnaire
RR	Relative risk
SD	Standard Deviation
SA	Sham acupuncture
SCL-90	Symptom Checklist 90
SF-12, SF-12 MCS/PCS	Short Form-12, Physical Component Score/Mental Component Score
SF-36, SF-36 MCS/PCS	Short Form-36, Physical Component Score/Mental Component Score
SF-MPQ	McGill Pain Questionnaire Pain Rating Index-Short-Form
SHCI	Subjective Health Complaint Inventory
SIP	Sickness Impact Profile
SKFS	Saudi version of the Knee Function Scale
SMD	Standardized mean difference
SOE	Summary of evidence

Acronym/Abbreviation	Term
SSDQ	Stanford Sleep Disorders Questionnaire
SSS	Swiss Spinal Stenosis Questionnaire
STAI	State-Trait Anxiety Inventory
TENS	Transcutaneous electrical nerve stimulation
UC	Usual Care
VAS	Visual analog scale
VKPS	Von Korff pain scale
WHOQOL-BREF	World Health Organization Quality of Life-BREF instrument
WL	Waitlist
WMD	Weighted mean difference
WPAI	Work activity impairment subscale
WPSI	Wahler Physical Symptoms Inventory
ZPS	Zung Self-Rating Depression Scale