Guidelines for the Management of Acute Low Back Pain: A Systematic Review

Prepared by: Pacific Northwest Evidence-based Practice Center

Investigators:
Andrea C. Skelly, PhD, MPH[1,2]
Roger Chou, MD[2]
Erika D. Brodt, BS[1,2]
Ian Blazina, MPH[2]
Esther K. Choo, MD, MPH[3]
Tracy Dana, MLS[2]
Andrea Diulio-Nakamura, PhD[4]
Rongwei F. Fu, PhD[2,5]
Yun Yu, MS[2,6]
Nancy A. Carney, PhD[2]
Annette Totten, PhD[2]
Kim Mauer, MD[4]
Dakota Riopelle, MPH[1,2]
Shay Stabler-Morris, MSc[1,2]

Affiliations:
[1] Aggregate Analytics, Inc. Fircrest, Washington
[2] Pacific Northwest Evidence-based Practice Center, Department of Medicine, Oregon Health & Science University, Portland Oregon
[3] Department of Emergency Medicine, Oregon Health & Science University, Portland Oregon
[4] Department of Anesthesiology and Perioperative Medicine, School of Medicine, Oregon Health & Science University, Portland Oregon
[5] School of Public Health, Oregon Health & Science University, Portland, Oregon
[6] Biostatistics Shared Resources, Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon

Corresponding author: Andrea Skelly, Ph.D., M.P.H., Aggregate Analytics, Inc., Fircrest WA; email: andrea@aggregate-analytics.com

Funding: This project was financially supported by the U.S. Food and Drug Administration of the U.S. Department of Health and Human Services as part of a financial assistance award totaling $1,999,980.00, with 100 percent funded by U.S. Food and Drug Administration/U.S. Department of Health and Human Services. The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement by, U.S. Food and Drug Administration, U.S. Department of Health and Human Services, or the U.S. Government. The funders had no decision-making role in designing and conducting systematic review, data collection, analysis, and interpretation of the data. As requested, the U.S. Food and Drug Administration provided nonbinding feedback and technical support to the guideline panel and methodological team.

Conflicts of Interest: The EPC team members (i.e., all authors) had no conflicts of interest.

Structured Abstract

Objectives. To synthesize evidence on the evaluation and management of acute low back pain (ALBP).

Data sources. Electronic databases (Ovid® MEDLINE®, PsycINFO®, Embase®, the Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews) to March 2025, and reference lists.

Review methods. We selected studies on diagnostic accuracy or predictive value of patient assessment and randomized controlled trials (RCTs) of outpatient management of ALBP (<6 weeks duration) using predefined criteria and dual review. Dual risk of bias assessment was performed. Systematic reviews were utilized for diagnostic accuracy of red flags for serious spine pathology and prediction of chronicity. We reported positive and negative likelihood ratios for red flags and prognostic factors. When sufficient RCT evidence was available, we conducted meta-analyses to improve the precision of effect estimates. We classified magnitude of effects as small, moderate, or large using predefined criteria. Strength of evidence was assessed for the primary outcomes of function, pain, and adverse events. There were too few RCTs to evaluate small study effects.

Results. We included 27 studies evaluating assessment in of ALBP patients and 111 treatment studies, including 103 RCTs. Few patient history or clinical examination factors (red flags) for serious spine pathology were highly informative for identifying or ruling out such pathologies. Stronger predictors included cancer history, saddle anesthesia, urinary retention, procedure-related infection, and age. High levels of maladaptive pain coping behaviors, nonorganic signs, substantial baseline functional impairment, poor general health status, and psychiatric comorbidities were associated with an increased likelihood of developing persistent, disabling low back pain (LBP). Risk prediction instruments incorporating these factors have not been extensively validated; validation evidence was most robust for the Acute Low Back Pain Screening Questionnaire and related modifications. RCTs comparing early imaging with usual care without imaging found no difference in function or pain in patients with ALBP who do not have signs of serious spine pathology.

Oral nonsteroidal anti-inflammatory drugs (NSAIDs) and muscle relaxants likely improve function and pain, although NSAID use was associated with gastrointestinal side effects. Acetaminophen did not improve pain or function compared with placebo. In patients with radiculopathy, systemic steroids may improve function, and epidural steroid injections may improve LBP. Evidence on opioids was sparse. Although one trial reported a small improvement in pain for a limited time, other trials did not; no improvement in function was reported. Opioids were associated with an increased risk of adverse events and withdrawal due to adverse events compared with placebo or nonopioids. Observational studies suggested that opioid therapy for ALBP increased the likelihood of longer-term opioid use.

Several nonpharmacologic treatments may improve function and/or pain, including manipulation and mobilization, acupuncture, exercise, and advice to stay active. Early intervention strategies likely improve pain and function compared with usual care. Most studies had methodological limitations, and effect sizes were generally small for function and pain.

Serious treatment-related harms were uncommon for all interventions; however, reporting of harms was limited, particularly for nonpharmacologic therapies. Studies were not designed to

assess long-term harms, risk of overdose, or opioid use disorder. We did not identify trials that compared opioids with nonpharmacologic treatments, and evidence comparing nonopioids with nonpharmacologic treatments was mostly insufficient. Evidence was insufficient to determine whether treatment benefits and harms vary across patient, clinical, or intervention characteristics.

Conclusions: Psychosocial factors were most predictive of developing persistent LBP. Early imaging did not improve patient outcomes in patients with ALBP who did not have signs of serious spine pathology. Oral NSAIDs, alone or in combination with muscle relaxants, were associated with improvements in pain and function. Evidence on opioids was limited, and opioids were not consistently associated with pain improvement; trials did not find functional improvement with opioids. Several nonpharmacologic therapies, including manual therapy and acupuncture, improved pian and/or function; however, evidence was limited for many interventions. Further research in patients with ALBP is needed to evaluate the comparative effectiveness of pharmacologic and nonpharmacologic treatments, the impact of therapies on long-term outcomes including risk of opioid use disorder, how to best assess patient risk for chronic LBP, and how benefits and harms may vary across clinical and patient factors, including social risk factors. Additional research is also needed to identify factors related to shared decision-making and to determine how to best tailor treatments to patient needs.

Contents

Tables

Figures

Appendices

Introduction

Background

Among musculoskeletal conditions, low back pain (LBP) accounts for the highest prevalence. Among all health conditions, LBP is the leading cause of disability globally and within the United States.[1-3] Treatment for LBP and spine disorders accounts for the highest costs among medical problems in the United States. Although LBP can be encountered in a variety of clinical settings, acute low back pain (ALBP) is most often managed in outpatient primary and ambulatory care settings.[4,5]

Acute pain is usually characterized as lasting 7 days but often extends up to 30 days.[6] For the purposes of this review, ALBP will be defined as lasting less than 6 weeks. Although prognosis for ALBP is favorable for most patients,[7,8] there is some evidence to challenge the assumption that ALBP typically completely resolves within 3 months.[9,10] A systematic review reported that a median of 26% of patients (range 2% to 48%) with ALBP transition to chronic LBP.[11] Chronic LBP tends to be persistent, can be challenging to manage, and accounts for a majority of costs associated with treatment of LBP.[12,13] Therefore, reducing the likelihood of transitioning from acute to chronic LBP is an important goal.

Pain is complex. It substantially impacts quality of life as well as physical and mental function. Pain is greatly influenced by a variety of factors including psychosocial factors, general health status, and environmental factors, which may predict who will develop chronic disabling pain as well as treatment response. Thus, a biopsychosocial framework to understanding LBP that considers biological as well as psychological and social factors is important.[11] Guideline concordant management of ALBP involves assessment of biological as well as psychological and social factors to rule out serious or specific causes of ALBP (e.g., cancer, infection, vertebral fracture) and to consider next steps for assessment (e.g., imaging) and treatment.[12] The majority of patients who present to primary care have LBP that cannot be attributed to a specific disease or spinal abnormality. Imaging of patients without factors suggesting a serious or specific cause of LBP has been discouraged in clinical guidelines and has not been associated with improved outcomes,[14] but may still be common.

The overarching goal of pain management is to relieve pain and improve function. A number of approaches and treatments are available for ALBP management. A guideline-supported initial approach for ALBP management is provision of information to reassure patients of the likelihood of resolution and encouragement to maintain physical activity.[15-17] Systematic review evidence supports a variety of pharmacological and noninvasive, nonpharmacological approaches for further treatment options for LBP.[12,18-20] Pharmacological treatments include simple analgesics such as acetaminophen or nonsteroidal anti-inflammatory drugs (NSAIDs), as well as skeletal muscle relaxants, topical medications, systemic corticosteroids, opioids, herbal medications, and others. Nonpharmacologic treatments often include heat/cold therapies, exercise therapy, psychological therapies, massage, acupuncture, spinal manipulation, and mind-body treatments (e.g., mindfulness-based relaxation, others). Interventional procedures are rarely used in patients with acute nonspecific LBP but may be considered on a case-by-case basis. Recent guidelines have prioritized use of nonpharmacological therapies for ALBP due to a more favorable balance of benefits to harms compared to pharmacological therapies.[21,22] The 2017 American College of Physicians (ACP) LBP guideline did not recommend opioids for ALBP and the 2022 Centers for Disease Control and Prevention (CDC) guideline recommends judicious,

short-term use of opioids for acute pain based on an individualized assessment of potential benefits and harms.[21,22]

In general, current evidence-based guidelines do not describe additional important clinical considerations regarding opioid use, including optimal dosing, duration of treatment, and use of risk mitigation strategies while optimizing and individualizing patient care based on the range of available options. Additional evidence has been published subsequent to many of these guidelines. A 2020 National Academies of Science, Engineering and Medicine (NASEM) Consensus Report[23] highlights some of these concerns and identifies gaps in current evidence and clinical guidelines related to ALBP management, identifying acute LPB as a priority for developing updated guidelines.

Purpose of the Review

The key decisional management dilemma for ALBP is selection of appropriate interventions to provide adequate pain relief, improve quality of life, improve function, and facilitate recovery, while minimizing adverse effects and judiciously using opioid and nonopioid pharmacologic treatments. To address this dilemma, the purpose of this systematic review is to provide an updated evidence base on the assessment of patients presenting with ALBP and on usual treatment options for ALBP. This systematic review will form the basis of new clinical practice guidelines on the assessment and management of ALBP. The review process will be conducted according to rigorous and accepted standards developed by Agency for Healthcare Research and Quality (AHRQ).[24] The Key Questions and scope of the review (populations, interventions, comparators, outcomes, timing, and settings – PICOTS) serve as the basis for the review and are described below.

Methods

Review Approach

The methods for this systematic review followed Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (hereafter referred to as "AHRQ Methods Guide".[24] This systematic review is in accordance with the Preferred Items for Reporting in Systematic Reviews and Meta-Analyses (PRISMA).[25]

Key Questions

An initial set of Key Questions were updated with input from the project sponsor (U.S. Food and Drug Administration), Advisory Board members, Guideline Development Group members, Technical Expert Panel members, and American Academy of Pain Medicine partners, as well as application of Evidence-based Practice Center (EPC) team expertise. The final protocol was registered on PROSPERO (CRD42024537276). The following Key Questions and inclusion criteria reflect these discussions.

History/Physical, Imaging (Key Questions 1 and 2)

Key Question 1. In adults presenting with acute (<6 weeks duration) low back pain (ALBP) in any setting

a. How accurate is a focused patient history and physical exam (including risk factor and symptom assessment, presence and severity neurologic deficit and psychological risk factors) for identifying serious underlying conditions, specifically cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture), or low back pain associated with severe or progressive neurological deficits including cauda equina syndrome?

b. What is the association between a focused patient history and physical exam (or components of these) and the likelihood of chronic LBP and long-term disability?

c. Does the use of screening tools (e.g., Keele STarT Back Screening Tool) improve short or long-term patient outcomes (versus not using such a tool) based on prognosis?

d. Does accuracy or predictive utility differ based on (1) patient demographics or characteristics including those related to social risk factors for health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., duration, severity, recurrent, radicular, non-radicular); (4) clinical setting, provider type?

Key Question 2. In adults with ALBP (<6 weeks duration), with or without radiculopathy who *do not* present with historical or clinical features suggestive of serious low back problems (e.g., absence of "red flag" symptoms"),

a. Does immediate/early lumbar spine imaging improve patient health outcomes versus usual care without imaging in the short term or long term?

b. What harms are associated with immediate lumbar spine imaging versus usual care without imaging?

c. Does outcome differ based on (1) patient demographics or characteristics including those related to social risk factors for health, access to imaging; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) imaging modality used (e.g., radiography, magnetic resonance imaging [MRI], computed tomography [CT]); (5) clinical setting, provider type?

d. What is the cost-effectiveness of immediate/early lumbar spine imaging versus usual care without imaging?

Treatment (Key Questions 3-8)

Opioid Therapy

Key Question 3. In adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy

a. What is the comparative effectiveness of systemic opioid therapy versus (1) placebo; (2) nonopioid pharmacologic therapy; or (3) noninvasive nonpharmacologic therapy on health outcomes at short term and long term?

b. What is the comparative effectiveness of systemic opioid therapy combined with a *nonopioid pharmacologic* intervention (e.g., hydrocodone with acetaminophen, codeine with acetaminophen) versus (1) placebo; (2) the same opioid alone; (3) the same nonopioid intervention alone; (4) a different nonopioid pharmacologic therapy; or (5) a noninvasive nonpharmacologic therapy on health outcomes at short term and long term?

c. What is the comparative effectiveness of systemic opioid therapy combined with a *nonpharmacologic* intervention versus (1) placebo or sham; (2) the same opioid alone; (3) the same nonpharmacologic intervention alone; or (4) a different nonpharmacologic therapy on health outcomes at short term and long term?

d. What are the effects of prescribing opioid therapy versus not prescribing opioid therapy for ALBP on continued need for prescription pain relief, such as need for opioid refills, short-term and long-term opioid use?

e. What are the comparative harms of opioid therapy (alone or in combination with a nonopioid or in combination with a nonpharmacologic treatment) versus placebo, no opioid, nonopioid pharmacologic therapy, or nonpharmacologic therapy including (1) overdose; (2) misuse, withdrawal, opioid use disorder or other substance use disorder, and related outcomes; (3) diversion; (4) serious adverse events; (5) other harms (e.g., gastrointestinal-related harm, sedation/fatigue, pruritus, dizziness, falls, fractures, motor vehicle accidents, endocrinological harms, cardiovascular events, cognitive harms); (6)

withdrawals due to adverse events; (7) psychological harms (e.g., depression, anxiety, suicidal ideation or suicidal behavior)?

f. Does effectiveness of a systemic opioid (alone or in combination with a nonopioid) vary based on prescribing strategy (dose, frequency, dosing interval, duration of therapy, quantity dispensed, type of opioid dispensed [e.g., long-acting, sustained release]), and what is the impact of different prescribing strategies on refill requests and quantity of pills?

g. Do the harms of a systemic opioid (alone or in combination with a nonopioid) vary based on prescribing strategy (dose, frequency, dosing interval, duration of therapy, quantity dispensed, type of opioid dispensed [e.g., long-acting, sustained release]), and what is the impact of different prescribing strategies on (1) long-term opioid use; (2) overdose; (3) misuse, withdrawal, opioid or other substance use disorder, and related outcomes; (4) diversion; (5) serious adverse events; (6) other harms (e.g., gastrointestinal-related harm, sedation/fatigue, pruritus, dizziness, nausea, falls, fractures, motor vehicle accidents, endocrinological harms, cardiovascular events, cognitive harms); (7) withdrawals due to adverse events; (8) psychological harms (e.g., depression, anxiety, suicidal ideation or behavior)?

h. Do effectiveness or harms (questions a-g) differ based on (1) patient demographics or characteristics including those related to social risk factors related to health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular, history of prior LBP), coexistent lower extremity pain without progressive/severe neurologic deficit; (4) dose of opioids, frequency, interval between doses; (5) timing and duration of therapy; (6) type of opioid; (7) current treatment for opioid use disorder; (8) opioid use history; (9) substance use history; (10) use of concomitant therapies; (11) clinical setting, provider type?

Key Question 4. In adults with ALBP (<6 weeks duration), including back pain with radiculopathy, being considered for opioid therapy or who have been prescribed opioids for ALBP

a. What is the effect of the following factors on the decision to prescribe opioids:
   Policy related factors including (1) existing opioid management plans; (2) patient education; (3) urine drug screening; (4) use of prescription drug monitoring program data; (5) availability of close follow-up; (6) prescribing/provision of naloxone (or other opioid antagonists);
   Patient related factors including (1) patient presentation (e.g., pain severity, etiology); (2) prior opioid prescription and experience; (3) patient expectations for pain control; (4) contraindications to other treatment options (e.g., to nonsteroidal anti-inflammatory drugs [NSAID]s), 5) history of substance use disorder?

b. What is the effect of the following factors on patient health outcomes (1) existing opioid management plans; (2) patient education; (3) urine drug screening; (4) use of prescription

drug monitoring program data; (5) availability of close follow-up; (6) prescribing/provision of naloxone (or other opioid antagonists)?

c. What is the impact of shared decision making on opioid prescription strategy, health outcomes, continued opioid use or harms?

d. What is the accuracy of instruments for predicting risk of opioid misuse, withdrawal, opioid use disorder, or overdose in patients with ALBP?

   [Notes: Interventions: Instruments, genetic/metabolic tests for predicting risk of misuse, opioid use disorder, and overdose; Comparison to reference standard for misuse, opioid use disorder, or overdose; or other benchmarks]

e. What is the effectiveness of instruments for predicting risk of long-term use of opioids, opioid misuse, withdrawal, opioid use disorder, or overdose in patients with ALBP versus usual care (i.e., not using a formal instrument)?

Nonopioid Pharmacologic Therapy

Key Question 5. In adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy

a. What is the comparative effectiveness of a single nonopioid pharmacologic therapy (e.g., acetaminophen, NSAIDs, antidepressants, anticonvulsants, systemic corticosteroids, benzodiazepines, muscle relaxants, lidocaine, ketamine, cannabis, common herbal remedies [e.g., willow-bark]) versus: (1) placebo; (2) other nonopioid pharmacologic treatments (e.g., those in a different medication class); or (3) nonpharmacologic therapy on health outcomes at short term and long term?

b. What is the comparative effectiveness of a combination of a maximum of two nonopioid pharmacologic therapies, one of which must be an NSAID or acetaminophen, versus: (1) placebo; (2) an NSAID or acetaminophen alone; or (3) nonpharmacologic therapy on health outcomes at short term and long term?

c. What is the comparative effectiveness of nonopioid pharmacologic therapy combined with a *nonpharmacologic* intervention versus (1) placebo or sham; (2) the same nonopioid pharmacologic therapy alone; (3) the same *nonpharmacologic* intervention alone; (4) a different *nonpharmacologic* therapy on health outcomes at short term and long term?

d. What are the comparative harms of nonopioid therapy (alone or in combination with a nonpharmacologic treatment or other nonopioid) versus placebo, a different class of nonopioid medication or nonpharmacologic therapy including (1) overdose; (2) misuse, withdrawal, opioid use disorder or other substance use disorder, and related outcomes; (3) serious adverse events; (4) other harms (e.g., gastrointestinal-related harm, sedation/fatigue, pruritus, dizziness, nausea, falls, fractures, motor vehicle accidents, endocrinological harms, cardiovascular events, cognitive harms); (5) withdrawal due to

adverse events; (6) psychological harms (e.g., depression, anxiety, suicidal ideation or behavior)?

e. Do effectiveness or harms (questions a-d) differ based on (1) patient demographics or characteristics including those related to social risk factors related to health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) timing of treatment; (5) dose, frequency/dosing interval, and route of administration of nonopioid; (6) duration of therapy; (7) type of nonopioid or type of drug combination; (8) current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type?

Noninvasive, Nonpharmacologic Therapy

Key Question 6. In adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy

a. What is the comparative effectiveness of noninvasive nonpharmacologic therapies (e.g., exercise, cognitive behavioral therapy, acupuncture) versus: (1) sham; (2) waitlist, usual care, attention control, no treatment; or (3) other included noninvasive, nonpharmacologic treatments on health outcomes at short term and long term?

b. What are the comparative harms of noninvasive, nonpharmacologic therapy versus sham, waitlist, usual care, attention control, no treatment or other included noninvasive, nonpharmacologic treatments including (1) harms specific to the treatment used; (2) withdrawal due to adverse events; (3) psychological harms (e.g., depression, suicidal ideation, suicidal behavior); (4) misuse, withdrawal, opioid use disorder, substance use disorder and related outcomes; (6) serious adverse events?

c. Do effectiveness or harms (questions a, b) differ based on (1) patient demographics or characteristics including those related to social risk factors related to health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) timing of treatment; (5) dose, frequency of treatment; (6) duration of treatment; (7) type of nonpharmacologic, noninvasive treatment; (8) current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type?

d. What is the effectiveness of policy approaches for reducing barriers and improving access to noninvasive, nonpharmacologic treatments?

Selected Interventional Procedures

Key Question 7. In adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy

a. What is the comparative effectiveness of selected interventional procedures (e.g., botulinum toxin, trigger point injection, epidural steroids/peri-radicular injections, SI joint/extra-articular injections, dry needling) compared with: (1) placebo or sham; (2) with each other; (3) systemic opioid therapy; (4) nonopioid pharmacologic treatments; or (5) nonpharmacologic therapy on health outcomes at short term and long term?

b. What are the comparative harms of selected interventional procedures (e.g., botulinum toxin, trigger point injection, epidural steroids, SI joint/extra-articular injections, peri-radicular injections, dry needling) compared with placebo or sham, another included procedure, systemic opioid therapy, nonopioid pharmacologic treatments or noninvasive, nonpharmacologic therapy including (1) harms specific to the treatment used; (2) withdrawal due to adverse events; (3) psychological harms (e.g., depression, suicidal ideation, suicidal behavior); (4) misuse, withdrawal, opioid use disorder, substance use disorder and related outcomes; (5) serious adverse events?

c. Do effectiveness or harms (questions a, b) differ based on 1) patient demographics or characteristics including those related to social risk factors related to health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) timing of treatment; (5) dose, frequency of treatment; (6) duration of treatment; (7) type of procedure; (8) current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type?

Management Strategies or Pathways

Key Question 8. In adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy

a. What is the comparative effectiveness of specific management strategies (e.g., stratification by risk for poor prognosis, early referral to therapy [e.g., manipulation therapies, exercise, physical therapy modalities)], matching therapies through treatment-based classification of patients, stepped care models, interdisciplinary models) versus (1) waitlist, usual care, or no treatment; or (2) other strategies on health outcomes short term and long term?

b. What are the comparative harms of management strategies versus waitlist, usual care, no treatment or other strategies including (1) harms specific to the strategy/pathway; (2) withdrawal due to adverse events; (3) psychological harms (e.g., depression, suicidal ideation, suicidal behavior); (4) overdose; (5) misuse, withdrawal, opioid use disorder/substance use disorder and related outcomes; (6) serious adverse events?

c. Do effectiveness or harms (questions a, b) differ based on (1) patient demographics or characteristics including those related to risk factors related to of health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) timing of strategy; (5) dose, frequency of strategy; (6) duration of strategy; (7) type or components of the strategy; (8)

current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type?

Cost-Effectiveness of Treatment (Key Question 9)

Key Question 9. What is the cost-effectiveness of pharmacologic, nonpharmacologic, interventional procedures and management strategies for treatment of ALBP (<6 weeks duration), including back pain with radiculopathy? (Include only full economic studies.)

Population, Intervention, Comparator, Outcome, Timing, Setting, Study Design (PICOTS)

The criteria for inclusion and exclusion of studies for the systematic review were based on the Key Questions and on the specific criteria for population, interventions, comparators, outcomes, timing, and settings (PICOTS), listed in Tables 1-7.

Table 1. Key Questions 1 and 2: history/physical, imaging

PICOTS Element	Include	Exclude
Population	**Assessment KQ 1** Adults presenting with ALBP (<6 weeks duration), including back pain with radiculopathy Note: Including pregnant/breastfeeding adults **Assessment KQ 2** Adults with ALBP (< 6 weeks duration) with or without radiculopathy who do not present with historical or clinical features suggestive of serious low back problems (e.g., absence of "red flag" symptoms") Note: including pregnant/breastfeeding adults **Special populations/factors** (1) patient demographics or characteristics including those related to social risk factors for health, access to imaging; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, non-radicular); (4) provider type, clinical setting; (5) type of imaging (KQ 2 only)	**Assessment KQ 1** • Adults with subacute (6 to 12 weeks) or chronic (>12 weeks) LBP • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) **Assessment KQ 2** • Adults with subacute (6 to 12 weeks) or chronic LBP (>12 weeks) • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) • Patients with ALBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome
Intervention	**Assessment KQ 1** • Focused patient history and physical exam to identify serious underlying conditions • Use of screening tools for patient stratification **Assessment KQ 2** • Early/immediate lumbar imaging in patients	

PICOTS Element	Include	Exclude
Comparators	**Assessment KQ 1** • Appropriate reference standard for condition (e.g., imaging findings) • Usual care (not using screening tool) **Assessment KQ 2** • Usual care (no use of early/immediate imaging)	
Outcomes	**Assessment KQ 1** • Diagnostic accuracy metrics • Prognostic factors, factors predicting chronic LBP, long term disability • Patient health outcomes (e.g., pain, function, quality of life) **Assessment KQ 2** • Patient health outcomes (e.g., pain, function, quality of life) • KQ 2d: ICER or similar outcome comparing treatment options and describing change in costs per change in benefits or harms, cost per specific beneficial outcome or harm, etc. **Harms** • Assessment related • Imaging related • KQ 2: Harms related to subsequent additional testing or treatment	
Timing	At presentation, follow-up appropriate to determine outcomes	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport for KQ1)	• Inpatient setting
Study design	**Assessment KQ 1** • Studies of diagnostic accuracy • Prognostic, predictive studies that control for confounding • RCTs of stratification tools **Assessment KQ 2** • RCTs • Full economic studies (KQ2d) that include modeling of downstream utilization	• Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed or are not part of a government technology assessment, duplicate publications of the same study that do not report on different outcomes or preliminary reports when results are published in later versions • Costing studies

ALBP = acute low back pain; ICER = incremental cost-effectiveness ratio; KQ = Key Question; LBP = low back pain; RCT = randomized controlled trial.

Special Populations and Factors to Evaluate for Modification

The following variables are relevant to all Key Questions related to treatment (Key Questions 3-8; Tables 2-6): (1) patient demographics or characteristics including those related to social risk factors related to health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset/acuity, duration, severity, recurrent LBP, prior history of LBP, radicular, non-radicular, coexistent lower extremity pain without progressive/severe neurologic deficit); (4) timing of treatment (e.g., relative peak pain); (5) dose, frequency, interval between doses of medication (or frequency/intensity of nonpharmacologic treatment); (6) duration of therapy; (7) type of treatment (e.g., type of opioid, nonopioid, nonpharmacologic,

etc.); (8) current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type.

Table 2. Key Questions 3 and 4: opioid therapy

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks), including back pain with radiculopathy Note: including pregnant/breast-feeding adults	• Adults with subacute (6 to 12 weeks) or chronic (>12 weeks) LBP • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) • Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome

PICOTS Element	Include	Exclude
Intervention	**KQ 3 a-c, e, g** • Systemic opioid therapy including agonists, partial agonists, and mixed mechanism opioids (tapentadol or tramadol); transdermal patches, topical opioids • Systemic opioid combined with a nonopioid pharmacologic therapy • Systemic opioid combined with nonpharmacologic therapy **KQ 3d** • Opioid prescription **KQ 3f** • Opioid prescribing strategy **KQ 4a** • Factors: (1) existing opioid management plans; (2) patient education; (e) urine drug screening; (4) use of prescription drug monitoring program data; (5) availability of close follow-up; (6) prescribing, provision of naloxone (or other opioid antagonists); 7) patient presentation (e.g., pain severity, etiology); 8) prior opioid prescription and experience; 9) patient expectations for pain control; 10) contraindications to other treatment options (e.g., to NSAIDs) **KQ 4b** • Factors: (1) existing opioid management plans; (2) patient education; (e) urine drug screening; (4) use of prescription drug monitoring program data; (5) availability of close follow-up; (6) prescribing, provision of naloxone (or other opioid antagonists) **KQ 4c** • Shared decision-making strategy, tool **KQ 4 d** • Instruments, genetic/metabolic tests for predicting risk of misuse, opioid use disorder, and overdose **KQ 4 e** • Instruments for predicting risk of long-term use of opioids, opioid misuse, opioid use disorder, or overdose	• Interventions to treat opioid use disorder, misuse, or overdose • Opioids not available in the United States • Non-FDA-approved opioids

PICOTS Element	Include	Exclude
Comparators	**KQ 3 a-c, e, g** • Placebo • Nonopioid pharmacologic therapy • Noninvasive, nonpharmacologic therapy • For opioids combined with a nonopioid compare with the nonopioid alone or another nonopioid and with opioid alone • For opioids combined with a nonpharmacologic compare with the nonpharmacologic therapy alone or another nonpharmacologic treatment and with opioid alone **KQ 3d** • No opioid prescription **KQ 3f** • Different opioid prescribing strategy vs. intervention strategy; stronger vs. weaker opioids; different administration routes **KQ 4a, b** • Not utilizing decision factors in 4a, b above (in interventions) **KQ 4c** • No shared decision-making strategy • Different tool **KQ 4d** • Reference standard for misuse, opioid use disorder, or overdose; or other benchmarks **KQ 4e** • Usual care (no tool) • Different tool	• Included therapies vs. excluded therapies • Opioids not available in the United States • Non-FDA-approved opioids

PICOTS Element	Include	Exclude
Outcomes	**KQ 3, KQ 4b, 4c** **Primary health outcomes (validated measures)** • Function • Pain • Pain relief satisfaction, completeness of pain relief • Quality of life • Sleep quality, sleep disturbance • Psychological distress (including depression, anxiety, etc.) • Return to work • Recurrence of LBP • Patient perception of improvement • Use of rescue medication • Continued opioid use (KQ 3a, b, c) **KQ 4a** • Opioid prescribing rates **KQ 4d** • Measures of diagnostic accuracy **KQ 4e** • Persistent opioid use **KQ 3, KQ 4b, 4c** **Harms** • Shorter term harms (e.g., sedation, fatigue, pruritus, dizziness, nausea, etc.) • Gastrointestinal-related harms (e.g., including opioid induced constipation) • Other harms (e.g., falls, fractures, motor vehicle accidents) • Endocrinological harms • Cardiovascular events • Cognitive harms, etc. • Serious AEs • Withdrawal due to AEs • Psychological harms (e.g., depression, suicidal ideation, suicidal behavior, etc.) • Overdose • Misuse, withdrawal, opioid use disorder, substance use disorder and related outcomes • Diversion	• Nonclinical outcomes (e.g., non-harm lab measures) • Non-validated measures or non-validated instruments for health outcomes (e.g., pain, function, quality of life) or psychological measures (e.g., depression, anxiety)
Timing	<1 day; 1 day to <1 week (also divided into 2-3 days and 4 days to 1 week); 1 week to <2 weeks; 2 weeks to <4 weeks; ≥4 weeks	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting

PICOTS Element	Include	Exclude
Study design	• RCTs to be the focus for effectiveness, benefits (including cross-over, cluster and pragmatic trials) • Comparative NRSI that control for confounding; Preference for what is best evidence (e.g., RCTs, prospective studies with least potential for bias) • NRSI for harms/AEs • Pre-post studies were be considered for KQ 4a	• Case series, pre-post, single arm studies for effectiveness outcomes; case reports • Case-control studies for effectiveness outcomes • Cross-sectional studies • Uncontrolled NRSI • Studies with historic controls • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed, duplicate publications of the same study that do not report on different outcomes or follow-up times, preliminary reports when results are published in later versions • Studies with fewer than 20 patients per treatment arm or 40 patients total

AE = adverse event; ALBP = acute low back pain; FDA = US Food and Drug Administration; KQ = Key Question; LBP = low back pain; NRSI = nonrandomized studies of interventions; RCT = randomized controlled trial.

Table 3. Key Question 5: nonopioid pharmacologic therapy

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks), including back pain with radiculopathy Note: including pregnant/breast-feeding adults	• Adults with subacute (6 to 12 weeks) or chronic LBP (>12 weeks) • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years old) • Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome
Intervention	Oral, parenteral (IV, IM), transmucosal, or topical nonopioid pharmacological therapy used for acute pain (e.g., acetaminophen, nonsteroidal anti-inflammatory drugs, systemic corticosteroids, skeletal muscle relaxants, benzodiazepines, antidepressants, anticonvulsants); lidocaine, sub-dissociative does of ketamine, cannabis, common herbal remedies (e.g., willow bark) Combination of two non-opioid drugs, one of which must be an NSAID or acetaminophen	• Combination of more than 2 non-opioid drugs • Combinations that do not include either an NSAID or acetaminophen

PICOTS Element	Include	Exclude
Comparators	• Placebo or sham • Other nonopioid pharmacological therapy or noninvasive nonpharmacological therapy • For nonopioids combined with another therapy (e.g., nonpharmacologic therapy) compare the combination with the same individual therapy alone • Compare combination of 2 non-opioid drugs vs. NSAID or acetaminophen ***NOTE:*** *Include oral vs. topical NSAID studies as well as aspirin vs. NSAID studies; include comparisons between different classes of nonopioids*	• Included therapies vs. excluded therapies. • For comparison of combination of nonopioid drugs, comparators other than and NSAID or acetaminophen
Outcomes	**Primary health outcomes (validated measures)** • Function • Pain • Pain relief satisfaction, completeness of pain relief • Quality of life • Sleep quality, sleep disturbance • Psychological distress (including depression, anxiety, etc.) • Return to work • Recurrence of LBP • Patient perception of improvement • Use of rescue medication • Opioid use, continued opioid use **Harms** • Short term harms (e.g., sedation, fatigue, pruritus, dizziness, nausea, etc.) • Gastrointestinal-related harms • Other harms (e.g., falls, fractures, motor vehicle accidents, endocrinological harms, cardiovascular events, cognitive harms, etc.) • Serious AEs • Withdrawal due to AEs • Psychological harms (e.g., depression, suicidal ideation, suicidal behavior, etc.) • Overdose • Misuse, withdrawal, opioid use disorder, substance use disorder and related outcomes • Diversion	• Nonclinical outcomes (e.g., non-harm lab measures) • Non-validated measures or instruments for health outcomes (e.g., pain, function, quality of life, depression)
Timing	<1 day; 1 day to <1 week (also divided into 2-3 days and 4 days to 1 week); 1 week to <2 weeks; 2 weeks to <4 weeks; ≥4 weeks	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting

PICOTS Element	Include	Exclude
Study design	• RCTs to be the focus for effectiveness, benefits (including cross-over, cluster and pragmatic trials) • Comparative NRSI that control for confounding; Preference for what is best evidence (e.g., RCTs, prospective studies with least potential for bias) • NRSI for harms/AEs	• Case series, pre-post, single arm studies for effectiveness outcomes; case reports • Case-control studies for effectiveness outcomes • Cross-sectional studies • Uncontrolled NRSI • Studies with historic controls • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed, duplicate publications of the same study that do not report on different outcomes or follow-up times, preliminary reports when results are published in later versions • Studies with fewer than 20 patients per treatment arm or 40 patients total

AE = adverse event; ALBP = acute low back pain; IM = intramuscular; IV = intravenous; LBP = low back pain; NRSI = nonrandomized studies of interventions; NSAID = nonsteroidal anti-inflammatory drug; RCT = randomized controlled trial.

Table 4. Key Question 6: noninvasive, nonpharmacologic treatments

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks), including back pain with radiculopathy Note: including pregnant/breast-feeding adults	• Adults with subacute (6 to 12 weeks) or chronic (>12 weeks) LBP • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) • Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome

PICOTS Element	Include	Exclude
Intervention	• Exercise (and related therapies, including self-guided exercise e.g., from a physical therapist, stretching) • Psychological therapies (e.g., CBT, operant therapy, others) • Hypnosis • Eye movement desensitization and reprocessing therapy • Mindfulness practices (e.g., Mindfulness Base Stress Reduction) • Manual therapies (e.g., musculoskeletal manipulation/mobilization) • Physical modalities [transcutaneous electrical nerve stimulation, ultrasound, braces, traction, heat, cold] • Transcranial magnetic stimulation • Meditation • Relaxation • Music therapy • Virtual reality • Acupuncture • Massage • Cupping • Mind-body practices (e.g., Yoga, Tai Chi) • Energy therapy (e.g., Reiki) • Advice to stay active, patient education • Standardized/structured self-management (e.g., Back School) • Support groups • Self-management methods (e.g., use of back support, postural changes, bedding support, etc.)	• Studies evaluating incremental value of adding a noninvasive nonpharmacological intervention to another noninvasive nonpharmacological intervention • Invasive nonsurgical treatments (e.g., injections, nerve block, parenterally administered medications) • Others not listed for inclusion
Comparators	• Sham • Waitlist, usual care, attention control, no treatment • Other included noninvasive, nonpharmacologic treatments listed	• Comparisons to interventions not listed • Comparisons within nonpharmacological intervention types (e.g., comparisons of different types of exercise with each other, different types of massage with each other)

PICOTS Element	Include	Exclude
Outcomes	**Primary health outcomes (validated measures)** • Function • Pain • Pain relief satisfaction, completeness of pain relief • Quality of life • Sleep quality, sleep disturbance • Psychological distress (including depression, anxiety, etc.) • Return to work • Recurrence of LBP • Patient perception of improvement • Use of rescue medication • Opioid use, continued opioid use **Harms** • Harms specific to the treatment used • Serious AEs • Withdrawal due to AEs • Psychological harms (e.g., depression, suicidal ideation, suicidal behavior, etc.) • Misuse, withdrawal, opioid use disorder/substance use disorder and related outcomes	• Nonclinical outcomes (e.g., non-harm lab measures) • Non-validated measures or instruments for health outcomes (e.g., pain, function, quality of life, depression)
Timing	<1 day; 1 day to <1 week (also divided into 2-3 days and 4 days to 1 week); 1 week to <2 weeks; 2 weeks to <4 weeks; ≥4 weeks	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting
Study design	• RCTs to be the focus for effectiveness, benefits (including cross-over, cluster and pragmatic trials) • Comparative NRSI that control for confounding; Preference for what is best evidence (e.g., RCTs, prospective studies with least potential for bias) • NRSI for harms/AEs	• Case series, pre-post, single arm studies for effectiveness outcomes; case reports • Case-control studies for effectiveness outcomes • Cross-sectional studies • Uncontrolled NRSI • Studies with historic controls • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed, duplicate publications of the same study that do not report on different outcomes or follow-up times, preliminary reports when results are published in later versions • Studies with fewer than 20 patients per treatment arm or 40 patients total

AE = adverse event; ALBP = acute low back pain; CBT = cognitive behavioral therapy; LBP = low back pain; NRSI = nonrandomized studies of interventions; RCT = randomized controlled trial.

Table 5. Key Question 7: selected interventions and procedures

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks) LBP, including back pain with radiculopathy Note: including pregnant/breast-feeding adults	• Adults with subacute (6 to 12 weeks) or chronic LBP (>12 weeks) • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome
Interventions	• Trigger point injection, fascial plane block • Botulinum toxin • Epidural steroid injection, peri-radicular injection • SI joint injection • Dry needling	• Surgical procedures • Others not listed for inclusion
Comparators	• Sham • Waitlist, attention control, no treatment • Usual care • Included interventions	• Non-included procedures
Outcomes	**Primary health outcomes (validated measures)** • Function • Pain • Pain relief satisfaction, completeness of pain relief • Quality of life • Sleep quality, sleep disturbance • Psychological distress (including depression, anxiety, etc.) • Return to work • Recurrence of LBP • Patient perception of improvement • Use of rescue medication • Opioid use, continued opioid use **Harms** • Harms specific to the treatments used • Serious AEs • Withdrawal due to AEs • Psychological harms (e.g., depression, suicidal ideation, suicidal behavior, etc.) • Misuse, withdrawal, opioid use disorder/substance use disorder and related outcomes	• Nonclinical outcomes (e.g., non-harm lab measures) • Non-validated measures or instruments for health outcomes (e.g., pain, function, quality of life, depression)
Timing	<1 day; 1 day to <1 week (also divided into 2-3 days and 4 days to 1 week); 1 week to <2 weeks; 2 weeks to <4 weeks; ≥4 weeks	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting

PICOTS Element	Include	Exclude
Study design	• RCTs to be the focus for effectiveness, benefits (including cross-over, cluster and pragmatic trials) • Comparative NRSI that control for confounding; Preference for what is best evidence (e.g., RCTs, prospective studies with least potential for bias) • NRSI for harms/AEs	• Case series, pre-post, single arm studies for effectiveness outcomes; case reports • Case-control studies for effectiveness outcomes • Cross-sectional studies • Uncontrolled NRSI • Studies with historic controls • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed, duplicate publications of the same study that do not report on different outcomes or follow-up times, preliminary reports when results are published in later versions • Studies with fewer than 20 patients per treatment arm or 40 patients total

AE = adverse event; ALBP = acute low back pain; LBP = low back pain; NRSI = nonrandomized studies of interventions; RCT = randomized controlled trial; SI = sacroiliac.

Table 6. Key Question 8: management strategies or pathways

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks), including back pain with radiculopathy Note: including pregnant/breastfeeding adults	• Adults with subacute (6 to 12 weeks) or chronic LBP (>12 weeks) • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) • Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome
Intervention	• Stratification by risk for poor prognosis • Early referral to therapy (e.g., manipulation therapies, exercise, physical therapy modalities) • Matching therapies through treatment-based classification of patients, • Stepped care strategies, interdisciplinary models	
Comparators	• Waitlist or no treatment • Usual care • Another included strategy	

PICOTS Element	Include	Exclude
Outcomes	**Primary health outcomes (validated measures)** • Function • Pain • Pain relief satisfaction, completeness of pain relief • Quality of life • Sleep quality, sleep disturbance • Psychological distress (including depression, anxiety, etc.) • Return to work • Recurrence of LBP • Patient perception of improvement • Use of rescue medication • Opioid use continued, opioid use **Harms** • Harms specific to the strategy/pathway • Serious AEs • Withdrawal due to AEs • Psychological harms (e.g., depression, suicidal ideation, suicidal behavior, etc.) • Overdose • Misuse, withdrawal, opioid use disorder/substance use disorder and related outcomes	• Nonclinical outcomes (e.g., non-harm lab measures) • Non-validated measures or instruments for health outcomes (e.g., pain, function, quality of life, depression)
Timing	<1 day; 1 day to <1 week (also divided into 2-3 days and 4 days to 1 week); 1 week to <2 weeks; 2 weeks to <4 weeks; ≥4 weeks	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting
Study design	• RCTs to be the focus for effectiveness, benefits (including cross-over, cluster and pragmatic trials) • Comparative NRSI that control for confounding; Preference for what is best evidence (e.g., RCTs, prospective studies with least potential for bias) • NRSI for harms/AEs	• Case series, pre-post, single arm studies for effectiveness outcomes; case reports • Case-control studies for effectiveness outcomes • Cross-sectional studies • Uncontrolled NRSI • Studies with historic controls • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed, duplicate publications of the same study that do not report on different outcomes or follow-up times, preliminary reports when results are published in later versions • Studies with fewer • Studies with fewer than 20 patients per treatment arm or 40 patients total

AE = adverse event; ALBP = acute low back pain; LBP = low back pain; NRSI = nonrandomized studies of interventions; RCT = randomized controlled trial.

Table 7. Key Question 9: cost-effectiveness of treatment

PICOTS Element	Include	Exclude
Population	Adults with ALBP (<6 weeks), including back pain with radiculopathy Note: including pregnant/breastfeeding adults	• Adults with subacute (6 to 12 weeks) or chronic LBP (>12 weeks) • Mixed chronic, subacute, and/or ALBP populations if study does not report results separately for ALBP or if <80% of the population does not have ALBP • Children and adolescents (age <18 years) • Patients with LBP due to tumor, cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture); or LBP associated with severe or progressive neurological deficits including cauda equina syndrome
Intervention	• Systemic opioid therapy • Nonopioid therapy • Nonpharmacologic, noninvasive therapy • Selected interventional procedures • Management strategies, pathways	• Treatments on included for the review
Comparators	• Placebo/sham • Usual care • Included treatment(s) vs. other included treatment(s)	• Included therapies vs. excluded therapies
Outcomes	• ICER or similar outcome comparing treatment options and describing change in costs per change in benefits or harms, cost per specific beneficial outcome or harm, etc.	• Cost of treatment only
Timing	Any	
Settings	Any outpatient, urgent care, or emergency care setting, pre-hospital settings (e.g., during ambulance transport)	• Inpatient setting
Study design	• Full economic studies (e.g., cost-utility, cost-benefit) • Preference given to United States studies, populations	• Costing studies • Editorials, letters, white papers, conference proceedings, citations that have not been peer-reviewed or are not part of a government technology assessment, duplicate publications of the same study that do not report on different outcomes or preliminary reports when results are published in later versions

ALBP = acute low back pain; ICER = incremental cost-effectiveness ratio; LBP = low back pain; RCT = randomized controlled trial.

Non-English Language Studies: To assess likelihood of language bias, we restricted to English-language articles but reviewed English-language abstracts of non-English-language articles to identify studies that would otherwise meet inclusion criteria.

Study Selection

We searched PubMed® MEDLINE, Embase®, APA PsycINFO® (Key Question 6 only), Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews from inception through January 2024. All searches were conducted by a qualified medical librarian and were peer reviewed (See Appendix A). Reference lists of included articles and relevant systematic reviews were also reviewed for includable literature.

We used the pre-established criteria above (Tables 1-7) to screen citations (titles and abstracts) identified through our searches to determine eligibility for full-text review. We used DistillerSR® to improve efficiency in screening articles. We followed a "best evidence" approach[26] and sought randomized controlled trials (RCTs) initially, as well-conducted RCTs have the least risk of bias. Nonrandomized studies of interventions (NRSIs) in pain can be misleading, due to the subjective nature of pain, which may exacerbate effects of confounding, selection bias, and attentional and other nonspecific effects. When included, to the extent possible, we focused on comparative NRSIs with concurrent controls, and which control for confounding as appropriate to the Key Question. We excluded uncontrolled observational studies, case series, and case reports. For Key Question 1, we included studies of diagnostic accuracy that used an appropriate reference standard (questions a and d) and prognostic or predictive studies that controlled for confounding (questions b and d). For Key Question 4a, we included pre-post studies. For Key Question 4c we included studies that evaluated the performance of a risk prediction instrument against a reference standard for opioid misuse, opioid use disorder, or overdose. We included systematic reviews as primary sources of evidence if they were a strong match to our Key Questions and PICOTS criteria, and were assessed as being at least moderate risk of bias using the AMSTAR-2 quality tool on factors such as the methods used to conduct searches, select studies, abstract data, assess risk of bias, and synthesize data.[24,27] We updated the findings of included systematic reviews with new primary studies identified in our searches.

All abstracts were dual reviewed. All citations deemed appropriate for inclusion by at least one reviewer were retrieved. Each full-text article was independently reviewed for eligibility by a minimum of two team members, including any articles suggested by the Guideline Development Group, Advisory Board, Technical Expert Panel members, or that arose from the public posting of the protocol. Any disagreements regarding inclusion were resolved by consensus. Appendix B provides a list of all the included studies. A record of studies excluded at the full-text level with reasons for exclusion is included in Appendix C.

Data Extraction and Data Management

After studies were selected for inclusion, data was abstracted using standardized templates into categories that include but are not limited to: study design, year, setting, country, sample size, eligibility criteria, attrition, participant enrollment methods, prior and concurrent treatments, population and clinical characteristics including age, gender, presence of radiculopathy, history of opioid use or substance use disorder, medical and psychiatric comorbidities, inclusion or exclusion of pregnant or breastfeeding women, duration of LBP episode, severity of back pain, history of prior LBP, intervention characteristics (e.g., dose, frequency, duration), results relevant to each Key Question as outlined in the previous PICOTS section, and information related to special populations and factors which may impact treatment effectiveness or harms. Information on confounders (in addition to those already identified for

abstraction) related to patient and intervention characteristics and methods of adjustment for them were also abstracted. Information relevant for assessing applicability was abstracted and included the number of patients randomized relative to the number of patients enrolled, use of run-in or wash-out periods, and characteristics of the population, intervention, and care settings. We categorized types of exercise into the same categories (e.g., mobility and flexibility) used for our prior report on nonpharmacologic treatments.[19,20] All study data were verified for accuracy and completeness by a second team member.

As stated in our protocol, if information regarding methods or results appeared to be omitted from the published results of a study, or if we were aware of unpublished data, we contacted authors to obtain this information. The authors of the Opioid Analgesia for Acute Low Back Pain and Neck Pain (OPAL) trial[28] (included in Key Question 3) were contacted and kindly provided patient characteristics and results data for the subgroup of patients with acute low back only. The detailed data tables include outcomes for the entire population as well as the subgroup with back pain only. While this trial evaluated an opioid-naloxone combination which is no longer available in the United States, the opioid used was oxycodone, which is available in the United States, and naloxone is not absorbed systemically and has no analgesic properties.

Assessment of Methodological Risk of Bias of Individual Studies

We used predefined criteria to determine the risk of bias for all included studies. Methods from the AHRQ Methods Guide[24] were used in concordance with the approach recommended in the chapter, Assessing the Risk of Bias of Individual Studies When Comparing Medical Interventions.[24,29] RCTs were evaluated using criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Chapter 8.5 Risk of Bias Tool)[30]and methods developed by the Cochrane Back and Neck Review Group.[31,32] NRSIs were evaluated based on instruments tailored to observational studies.[33,34] and criteria developed by the U.S. Preventive Services Task Force,[35] which includes methods of patient selection (e.g., consecutive patients, use of an inception cohort) and appropriate control for confounding. Studies of diagnostic accuracy was assessed using QUADAS-2/QUADAS-C[36,37] and the Quality in Prognosis Studies (QUIPS) tool for studies evaluating risk factors was used for prognostic studies.[38] Systematic reviews were assessed using the AMSTAR-2 quality rating instrument on factors such as the methods used to conduct searches, select studies, abstract data, assess risk of bias, and synthesize data.[27]

Studies were rated as being "low," "moderate," or "high" risk of bias. Studies rated at low risk of bias and their results are generally considered valid. Good-quality intervention studies include clear descriptions of the population, setting, interventions, and comparison groups; a valid method for allocating patients to treatment; low dropout rates and clear reporting of dropouts; appropriate means for preventing bias; and appropriate measurement of outcomes. Good-quality diagnostic accuracy studies use unbiased methods to select patients; report interpretation of the index test without knowledge of the reference standard; report a predefined threshold for a positive index test; report use of an appropriate reference standard; apply the reference standard to all patients; report interpretation of the reference standard blinded to the results of the index test; and report low attrition.

Studies rated at moderate risk of bias are susceptible to some bias, though not enough to invalidate the results. These studies may not meet all the criteria for a rating of good quality, but no flaw or combination of flaws is likely to cause major bias. The study may be missing information, making it difficult to assess limitations and potential problems. The fair-quality

category is broad, and studies with this rating vary in their strengths and weaknesses. The results of some studies at moderate risk of bias are likely to be valid, while others may be only possibly valid.

Studies rated at high risk of bias have significant flaws that imply biases of various types that may invalidate the results. They have a serious or "fatal" flaw (or combination of flaws) in design, analysis, or reporting; large amounts of missing information; discrepancies in reporting; or serious problems in the delivery of the intervention. The results of these studies are at least as likely to reflect flaws in the study design as to show true difference between the compared interventions. We did not exclude studies rated at high risk of bias a priori, but high risk of bias studies were considered less reliable than higher-quality studies when synthesizing the evidence, particularly if discrepancies between studies are present.

Two team members independently assessed quality. Any disagreements were resolved by consensus.

Data Synthesis

We constructed evidence tables (Appendix E) identifying the study and patient characteristics (as discussed above), results of interest, and quality ratings for all included studies, as well as summary tables and/or figures to highlight the main findings for some sections. We reviewed and highlighted studies by using the best evidence focus for our synthesis for each Key Question. We analyzed RCTs and NRSIs separately and reported them separately unless findings were very consistent across study designs and the studies were clinically homogeneous. Studies with the least risk of bias were summarized separately and compared with summarized results from studies at high risk of bias.

Findings were synthesized qualitatively (e.g., ranges and descriptive analysis, with interpretation of results) and quantitatively (meta-analysis) when appropriate. To address anticipated heterogeneity in reported outcomes, variation in their definitions and criteria for what constitutes response, we focused on validated outcomes such as pain, function, and quality of life. We sought input from the Guideline Development Group, Advisory Board, and Technical Expert Panel regarding outcomes and their prioritization. We classified the magnitude of effects for continuous measures of pain and function using a similar system as in prior AHRQ reviews on pain[18-20,39,40] (Appendix D) and evaluated the proportion of patients meeting thresholds for clinically important differences (e.g., ≥30% pain relief) when reported. Effects below the threshold for small were categorized as no effect. For analysis of continuous measures across the same outcome measures (e.g., visual analog scale [VAS] for pain) we reported mean differences and used standardized mean differences for outcome measures with similar constructs together with 95% confidence intervals.

Pairwise meta-analyses were conducted to obtain more precise effect estimates if there were two to five clinically and methodologically comparable studies.[41-43] For continuous outcomes, mean difference was the effect measure for pain and function outcomes when reported using the same scale. Pain scales were converted to a common 0 to 10 scale. For function, standard mean difference was the effect measure when different scales were used (most commonly, Roland-Morris Disability Questionnaire [RDQ] and Oswestry Disability Index [ODI]). However, we also pooled data separately for RDQ and ODI, using the original scales. To facilitate the interpretation of standardized mean difference (SMD) for function, we converted SMD back to its original scales based on the average SMDs of the same comparison across all timepoints. For both mean difference and SMD, adjusted mean difference from the analysis of covariance or

other appropriate regression models was used if available, followed by difference in follow-up scores and change scores. When standard deviation (SD) was not reported, or could not be calculated from the reported data, it was imputed using the average coefficient of variation from the other included studies reporting the same outcome. Pooled relative risks (RRs) were estimated for binary outcomes including perception of pain improvement or relief, adverse events, withdrawal due to adverse events, global efficacy, gastrointestinal problems, and medication used in the past 24 hours.

A random effects meta-analysis using the profile likelihood method[44] was performed to combine the included randomized trials. Because of the potential differential impact of the type of control group on treatment effects, when applicable, separate analysis was conducted according to whether the control group received a sham intervention or usual care. All analyses were stratified by the prespecified follow-up duration categories (e.g., <1 day, 1 day to <1 week, 1 to <2 weeks, 2 to <4 weeks, ≥4 weeks for most Key Questions and 4 to 6 weeks, 3 months, 6 months, and 12 months [1 year] for Key Question 8). For studies that reported outcomes at more than one timepoint within a duration category, we used data for the longest duration within the category for most cases, or used data from the most comparable timepoints with other studies under same comparison when the longest duration was exceptionally longer than other studies.

For analysis, we considered opioid agonists (e.g., oxycodone, codeine) and mixed agonists (e.g., tramadol, tapentadol) under the general heading of opioids. Trials of common formulations of an opioid with acetaminophen or an NSAID (e.g., Vicodin, Tylenol #3) or in combination with naloxone, which does not have analgesic properties, were also included and considered as trials primarily of the opioid, not as combination treatment. Trials were considered placebo controlled if a placebo was received or if all patients in the placebo group also received the same NSAID that had been combined with the opioid (e.g., tramadol plus diclofenac versus placebo plus diclofenac).

Statistical heterogeneity was assessed using the I^2 statistic and the Cochran's $\chi 2$ test.[45] Sensitivity and subgroup analyses were performed to explore statistical heterogeneity and differences by study quality, study design, intervention differences, patient characteristics, and outcome measurement as data permitted. We focused on RCTs to evaluate differential efficacy and safety as they have the least potential for bias and confounding thus allowing for causal inference. Few of the included RCTs performed formal tests for interaction between subgroups.

Evidence gaps and applicability to U.S. practice settings were assessed based on the AHRQ Methods Guide[24] using the PICOTS framework.[46]

For analyses with at least 10 trials that were sufficiently homogeneous with regard to populations, interventions, and outcomes, small study effects were planned to be evaluated using funnel plots and the Egger test.[47] No meta-analysis had at least 10 studies.

All meta-analyses were conducted using Stata/SE 16.1 (StataCorp, College Station, TX), and all results were provided with 95% confidence intervals.

Grading the Strength of Evidence for Major Comparisons and Outcomes

Outcomes assessed for strength of evidence (SOE) were prioritized based on input from the Guideline Development Group, Advisory Board and Technical Expert Panel. Based on this prioritized list, the strength of evidence for comparison-outcome pairs within each Key Question were initially assessed by one researcher for each clinical outcome (see PICOTS) by using the approach described in the AHRQ Methods Guide.[24] To ensure consistency and validity of the

evaluation, the initial assessment was independently reviewed by at least one other experienced investigator using the following criteria:

- Study limitations (low, medium, or high level of study limitations)
    - This is the degree to which studies for a given outcome are likely to have reduced bias based on study design, analysis, and conduct. The aggregate risk of bias across individual studies reporting an outcome is considered.
- Consistency (consistent, inconsistent, or unknown/not applicable)
    - This is the degree to which studies report similar magnitudes of effect (i.e., range sizes are similar) or same direction of effect (i.e., effect sizes have the same sign).
- Directness (direct or indirect)
    - This is the degree to which the outcome is directly or indirectly related to health outcomes of interest. Patient-centered outcomes are considered direct.
- Precision (precise or imprecise)
    - Describes the level of certainty of the effect estimate for a particular outcome with a precise estimate being one that allows a clinically useful conclusion. This may be based on sample size sufficiency and number of events. If these are adequate, the interpretation of the confidence interval is also considered. When quantitative synthesis is not possible, sample size and assessment of variance within individual studies will be considered.
- Reporting bias (suspected or undetected)
    - Publication bias, selective outcome reporting, and selective analysis reporting are types of reporting bias. Reporting bias is difficult to assess as systematic identification of unpublished evidence is challenging. If sufficient numbers of RCTs (>10) are available, quantitative funnel plot analysis may be done.

The strength of evidence was assigned an overall grade of high, moderate, low, or very low according to a four-level scale (Table 8) by evaluating and weighing the combined results of the above domains.

Table 8. 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.
Very Low	We are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. The body of evidence has unacceptable deficiencies which preclude reaching a firm conclusion. If no evidence is available, it will be noted as “no evidence”

Bodies of evidence consisting of RCTs were initially considered as high strength while bodies of comparative observational studies began as low-strength evidence for benefit and

moderate for harms. The strength of the evidence may be downgraded based on the limitations described above. There are also situations where the observational evidence may be upgraded (e.g., large magnitude of effect, presence of dose-response relationship or existence of plausible unmeasured confounders), if there are no downgrades on the primary domains, as described in the AHRQ Methods Guide.[24,29]

Where both RCTs and observational studies are included for a given intervention-outcome pair, we follow the additional guidance on weighting RCTs over observational studies, assessing consistency across the two bodies of evidence, and determining a final rating.[24] Summary tables include ratings for individual strength of evidence domains (risk of bias, consistency, precision, directness) based on the totality of underlying evidence identified.

Assessing Applicability

Applicability was assessed in accordance with AHRQ Methods Guide,[24] using the PICOTS framework. Applicability refers to the degree to which outcomes associated with the intervention are likely to be similar across patients and settings relevant to the care of patients undergoing treatment for ALBP on the populations, interventions, comparisons, and outcomes synthesized across included studies. Multiple factors identified a priori that are likely to impact applicability may include (but are not limited to) type or characteristic of back pain (e.g., onset/acuity, duration, severity, recurrent LBP, prior history of LBP, radicular, non-radicular, coexistent lower extremity pain without progressive/severe neurologic deficit), clinical setting (e.g., ED vs. primary care), provider type, treatment characteristics (type, dose, frequency, duration, etc.), or characteristics of enrolled patient populations (e.g., sex, age, social risk factors related to health, health and functional status, comorbidities). Review of abstracted information on these factors was used to assess situations for which the evidence is most relevant and to evaluate applicability to real-world clinical practice in typical U.S. settings. We provided a qualitative summary of our assessment.

Specific Methods for Key Question 1b

For Key Question 1b, we updated a prior systematic review[11] led by one of the authors of the current review team. The prior review was based on searches conducted through 2010. We utilized the same search strategy and applied the same inclusion criteria as the original review. Specifically, we included English-language studies of adult patients with <8 weeks of LBP that prospectively evaluated the prognostic accuracy of individual risk factors or risk prediction instruments for persistent, disabling LBP. We defined "persistent" as pain lasting beyond 3 months after the start of symptoms and "disabling" as at least moderately affecting ability to work or function. We accepted studies that enrolled patients with up to 12 weeks of LBP if >80% of patients had <8 weeks pain or average pain duration was <6 weeks. Studies that enrolled patients with up to 12 weeks of LBP but did not report the proportion with <8 weeks pain or average pain duration were included if we judged that the population was predominantly acute (e.g., patients undergoing evaluation in emergency department (ED) settings or initial primary care evaluation). We excluded studies that focused on patients with back pain related to major trauma, infection, cauda equina syndrome, cancer, vertebral compression fracture, inflammatory arthritis, fibromyalgia, significant scoliosis, prior back surgery, and neurogenic claudication. We also excluded studies that did not report sensitivity, specificity, or likelihood ratios; did not provide data with which to calculate them; or only evaluated average changes in pain, function, or work status (i.e., did not define "disabling" back pain using a dichotomous outcome).

We assessed study risk of bias using eight criteria for assessment of diagnostic[48] or prognostic[49] studies (Appendix H, Table H-1). For studies of risk assessment instruments, we also evaluated whether the study evaluated diagnostic test performance in a population other than the one used to derive the instrument (external validation) and utilized predefined thresholds to define a positive result.[50]

We categorized potential yellow flags into one of 18 risk factor domains (Appendix H, Table H-2). These domains were based on potential yellow flags commonly evaluated in prognosis studies, or found to be predictive in other systematic reviews. Two domains (race/ethnicity and lack of early improvement) were added for the update; the other 16 domains were addressed in the original review. We excluded potential yellow flags that could not be categorized into one of these domains.

For all studies, we created 2 x 2 tables for diagnostic accuracy using reported data, or calculated from information provided in the studies. Using the 2 x 2 tables, we calculated sensitivities, specificities, and likelihood ratios[51] with 95% confidence intervals, based on outcomes at 3 to 6 months and at 1 year. In the original review, we used the *diagti* procedure in Stata (Stata version 10, StataCorp, College Station, TX) to calculate these measures and for studies added in the update, we used an online calculator (https://www.medcalc.org/calc/diagnostic_test.php). When there were unexplained discrepancies between results calculated from 2 x 2 tables and results reported in the studies, we utilized the calculated results.

The positive likelihood ratio (LR+) is the odds of developing chronic disabling LBP among subjects with the risk factor present.[52] The negative likelihood ratio (LR-) is the odds of developing chronic disabling LBP among subjects without the risk factor. For consistency in interpretation, when studies used not developing chronic disabling LBP as the outcome, we recalculated results for developing chronic disabling LBP. For baseline pain, impairment in function, and maladaptive pain coping behaviors, which were measured using various numerical or ordinal scales, we categorized results as high (e.g., 8 or higher on a 0 to 10 scale), moderate (e.g., 4 to 6), or low (e.g., <4). We dichotomized other risk factors that were measured using numerical or ordinal scales. Results of multicomponent risk assessment instruments were either dichotomized (e.g., score above or below a certain value; result positive or negative) or evaluated as ≥3 discrete categories (e.g., high, moderate, or low risk, based on score). For all dichotomized risk factors and instruments, we reported the LR+ and LR-. For risk factors and instruments that were evaluated as ≥3 discrete categories. we reported the LR (LR+) for each category.

Because of variability in the thresholds, definitions of risk factors, populations, and outcomes assessed, we summarized results with the median likelihood ratios and the range. The total range, rather than the interquartile range, was chosen because several prognostic factors were reported in few studies and because the summary range highlights the greater uncertainty we have in the estimates. To evaluate the usefulness of individual risk factors and risk prediction instruments, we considered the median LR value, the variability of estimates across studies, and whether the LR range crossed 1.

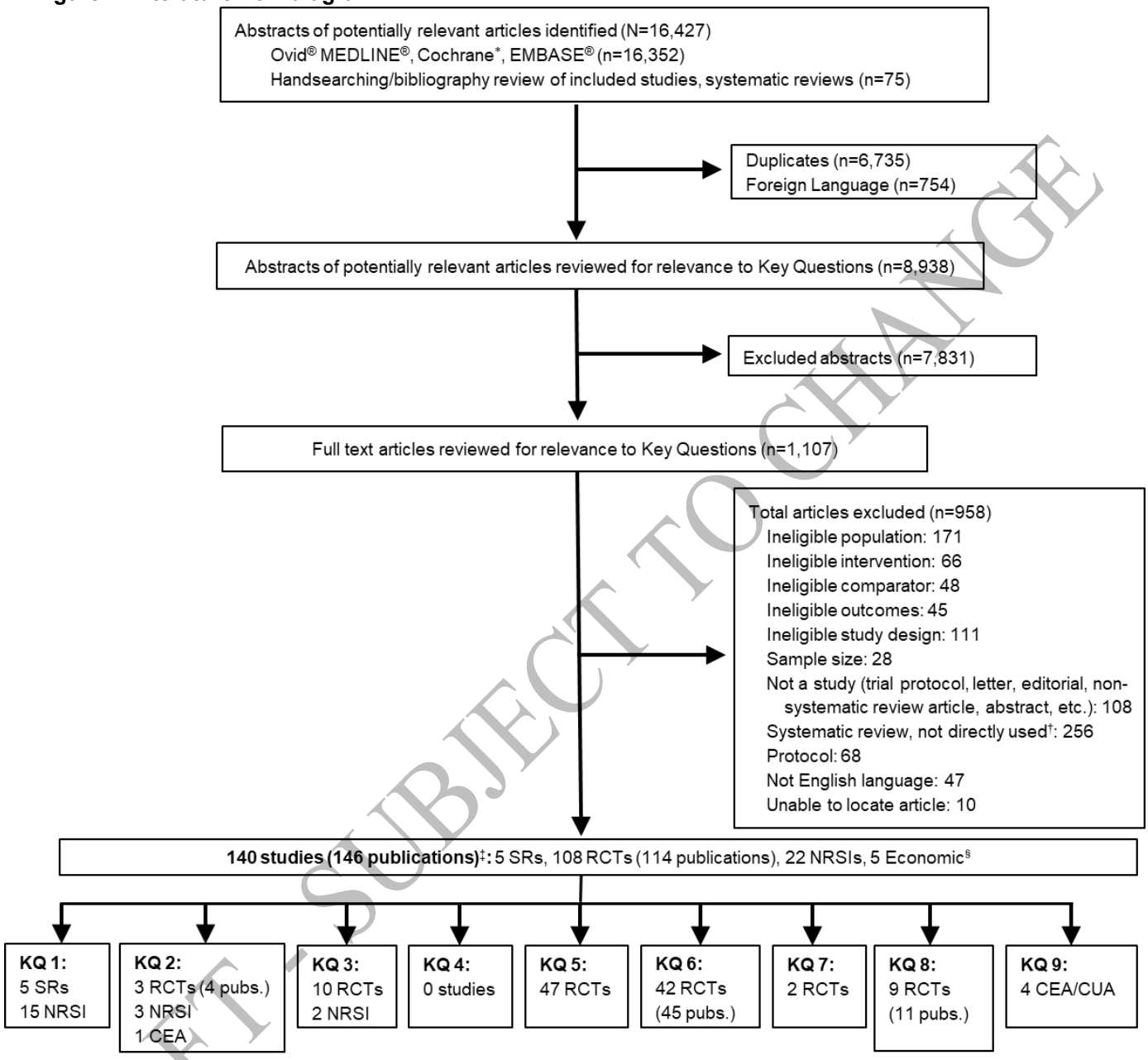
We performed stratified analyses according to whether studies evaluated workers' compensation or other populations to determine if and how results differed. We also performed stratified analyses of studies that focused on outcomes related to return to work versus nonwork outcomes, whether they did or did not meet various risk of bias criteria, longer (>1 year) versus shorter duration of follow-up, and <4 weeks versus 4-8 weeks duration of LBP symptoms.

Results

A total of 16,422 citations were identified, 16,352 from electronic database searches and an additional 70 from handsearching and bibliography review of included studies and systematic reviews. After deduplication, 8,933 citations underwent dual review of titles and abstracts of which 1,102 articles were selected for full-text review. In total, 140 studies (in 146 publications) were included in this review: five systematic reviews,[11,53-56] 108 RCTs (in 114 publications),[28,57-169] 22 nonrandomized studies,[170-191] and five full economic studies.[192-196] The most evidence was identified for Key Question 5 (nonopioid therapy) followed by Key Question 6 (noninvasive nonpharmacologic treatments). Search results and selection of studies are summarized in the literature flow diagram in Figure 1 below, and an overview of the number and design of studies included by Key Question and intervention and comparison can be found in Table 9 and Table 10 below. All systematic reviews (SRs) and all nonrandomized studies/nonrandomized studies of interventions (NRSIs) were assessed as moderate risk of bias. Across the RCTs, 23% were assessed as low, 66% as moderate, and 11% as high risk of bias. The Quality of Health Economic Studies scores across the economic studies ranged from 62 to 93 out of 100.

Appendix B lists included studies, and Appendix C lists excluded studies along with reasons for exclusion. Detailed evidence tables for included studies and risk of bias assessments are available in Appendixes E and F. Appendix G contains details on the strength of evidence (SOE). The definitions of magnitude of effects for continuous measures of pain and function are presented in Appendix D.

Figure 1. Literature flow diagram



KQ = Key Question; NRSI = nonrandomized study of interventions; RCT = randomized controlled trial; SR = systematic review.
* Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews.
† Bibliographies/reference lists were reviewed for relevant studies not captured by the systematic search.
‡ Some studies contributed to more than one Key Question.
§ Only full economic studies were included (e.g., cost-effectiveness, cost utility trials).

Overview of Included Studies

Tables Table 9 and Table 10 below provide an overview of the number of trials included by Key Question, intervention and comparison.

Table 9. Number of studies for Key Question 1 (history/physical) and Key Question 2 (imaging)

Key Question	Subquestion	Intervention	Comparator	Number of Studies and Design (number of publications)
1: History and Physical	1a: Accuracy	NA	NA	4 SRs[53-56] 1 nonrandomized study[182]
	1b: Predictive utility	NA	NA	1 SR[11] 14 nonrandomized studies[170,173,174,177-180,183,185-187,189-191]
	1c: Screening tools	NA	NA	None
	1d: Differential accuracy or predictive utility	NA	NA	None
	Any KQ 1			**5 SRs**[11,53-56] **15 nonrandomized studies**[170,173,174,177-180,182,183,185-187,189-191]
2: Early Imaging	2a, b: Effectiveness and safety	Early imaging (MRI)	Usual Care	1 RCT (2)[62,92] 3 nonrandomized studies[175,181,188]
		Early imaging (lumbar radiograph)	Usual Care	2 RCTs[158,159]
	2c: Differential effectiveness and harms	NA	NA	None
	2d: Cost-effectiveness	Early imaging (lumbar radiograph)	Usual Care	1 CEA[196]
	Any KQ 2	Any	Any	**3 RCTs (4)**[62,92,158,159] **3 nonrandomized studies**[175,181,188] **1 CEA**[196]

CEA = cost effectiveness analysis; KQ = Key Question; MRI = magnetic resonance imaging; NA = not applicable; RCT = randomized controlled trial; SR = systematic review.

Table 10. Number of studies for Key Questions 3-9 (studies of treatment, cost-effectiveness of treatment)

Key Question	Subquestion	Intervention	Comparator	Number of Studies and Design (number of publications)
3: Opioids - Treatment	3a, e: Single Opioid – Effectiveness and Safety	Opioid	Placebo	5 RCTs[28,113,133,168,169]
		Opioid	Nonopioid	3 RCTs[77,138,169]
	3b, e: Opioid + Nonopioid – Effectiveness and Safety	Opioid + Nonopioid	Nonopioid	4 RCTs[89,105,113,138]
		Opioid + Nonopioid	Opioid	2 RCTs[162,168]
	3c, e: Opioid + Nonpharmacologic – Effectiveness and Safety	NA	NA	None
	3d: Prescribing vs. not Prescribing	Early use	Late Use	2 NRSI[172,176]

Key Question	Subquestion	Intervention	Comparator	Number of Studies and Design (number of publications)
	3f, g: Prescribing strategies – Effectiveness and Harms	NA	NA	None
	3h: Differential effectiveness and harms	NA	NA	None
	Any KQ 3	Any	Any	**10 RCTs**[28,77,89,105,113,133,138,162,168,169] **2 NRSI**[172,176]
4: Opioids – Decision Making	**Any KQ 4**	NA	NA	**None**
5: Nonopioid Treatment	5a, d: Single Nonopioid – Effectiveness and Safety	NSAID	Placebo	10 RCTs[68,75,83,91,93,95,100,144,154,155]
			Acetaminophen	3 RCTs[67,77,108]
			Muscle Relaxant	1 RCT[75]
			Lidocaine	1 RCT[99]
			Manual Therapy	3 RCTs[76,83,143]
			Acupuncture	2 RCTs[101,153]
			Physiotherapy	1 RCT[143]
		Muscle Relaxant	Placebo/Sham	8 RCTs[70,74,75,97,104,132,161,163]
			Benzodiazepine	1 RCT[151]
			Manual Therapy	1 RCT[74]
		Steroid	Placebo	6 RCTs[79,85,137,140,156,160]
		Acetaminophen	Placebo	1 RCT[125]
			Electroacupuncture	1 RCT[73]
		Cannabidiol	Placebo	1 RCT[78]
	5b, d: Combination 2 Nonopioids – Effectiveness and Safety	Muscle Relaxant + NSAID	Placebo	1 RCT[75]
			NSAID	8 RCTs[57,58,69,75,87,113,126,128]
		Acetaminophen + NSAID	NSAID	2 RCTs[72,141]
		Benzodiazepine + NSAID	NSAID	1 RCT[71]
		Pregabalin + NSAID	NSAID	1 RCT[142]
	5c, d: Combination Nonopioid + Nonpharmacologic – Effectiveness and Safety	Heat + NSAID	NSAID	1 RCT[86]
		Cold + NSAID	NSAID	1 RCT[86]
		Acupuncture + NSAID	NSAID	1 RCT[153]
			Acupuncture	1 RCT[153]
	5e: Nonopioid – differential effectiveness and harms	Muscle Relaxant	Placebo	1 RCT[161]
		Steroid	Placebo	1 RCT[140]
		Muscle Relaxant + NSAID	NSAID	1 RCT[57]
	Any KQ 5	Any	Any	**47 RCTs**[57,58,67-79,83,85-87,91,93,95,97,99-101,104,108,113,125,126,128,132,137,140-144,151,153-156,160,161,163]
6: Noninvasive, Nonpharmacologic Treatment	6a, b: Effectiveness and Harms	Exercise	Placebo	1 RCT(2)[121,122]
			Usual Care	10 RCTs (12)[59,88,94,96,107,111,119,121,122,127,131,149]
			Bed Rest	3 RCTs[59,127,139]
			Manual Therapy	2 RCTs[84,119]
			Education (Back School)	1 RCT(2) [66,82]
			Advice to Stay Active	1 RCT[139]
			Acupuncture	1 RCT[111]
		Manual Therapy	Placebo/Sham	5 RCTs[74,83,102,106,152]
			Usual Care	5 RCTs[115,116,119,148,166]
			Physiotherapy	2 RCTs[143,146]

Key Question	Subquestion	Intervention	Comparator	Number of Studies and Design (number of publications)
			Education (Back School)	1 RCT[106]
			Massage	1 RCT[109]
		Manual Therapy (Manipulation)	Manual Therapy (Mobilization)	1 RCT[147]
		Acupuncture	Placebo/Sham	2 RCTs[124,130]
			Usual Care	3 RCTs[111,112,130]
			Massage Therapy	1 RCT[136]
		Heat	Placebo	2 RCTs[63,114]
		TENS	Sham	1 RCT[65]
		Massage	Sham	1 RCT[61]
		Advice to Stay Active	Bed Rest	3 RCTs[129,135,139]
		Education	Attention Control	1 RCT[150]
			Usual Care	2 RCTs[64,120]
		Self-management (Back School)	Placebo	1 RCT[106]
			Usual Care	1 RCT[165]
		Minimal Psychosocial Intervention	Usual Care	1 RCT (2)[81,117]
	6c: Differential effectiveness	Exercise	Usual Care	1 RCT (2)[121,122]
		Manual Therapy	Usual Care	2 RCTs[116,166]
		Minimal Psychosocial Intervention	Usual Care	1 RCT[81]
	6d: Policy approaches to improve access	NA	NA	None identified
	Any KQ 6	Any	Any	**42 RCTs (45)**[59,61,63-66,81,82,84,88,94,96,102,106,107,109,111,112,114-117,119-122,124,127,129-131,135,136,139,146-150,152,165,166 74,83,143]
7: Select Interventional Procedures	7a, b: Effectiveness and harms	Trigger Point Injection	IV NSAID	1 RCT[103]
		Epidural Steroid Injection	Usual Care	1 RCT[134]
	7c: Differential effectiveness and harms	NA	NA	None
	Any KQ 7	Any	Any	**2 RCTs**[103,134]
8: Management Strategies or Pathways	8a, b: Effectiveness and harms	Early intervention	Usual Care	6 RCTs (8)[60,80,110,118,123,145,157,167] 1 NRSI[184]
		Treatment-based classification	Usual Care	1 RCT[98] 1 NRSI[171]
		Educational, Goal-directed treatment plan	Usual Care	1 RCT[164]
		Interdisciplinary Treatment	No Treatment	1 RCT[90]
	8c: Differential effectiveness and harms	Early intervention	Usual Care	2 RCTs (4)[60,110,145,157])
	Any KQ 8	Any	Any	**9 RCTs (11)**[60,80,90,98,110,118,123,145,157,164,167] **2 NRSI**[171,184]
9: Cost-effectiveness of treatment	NA	Epidural Steroid Injections	Usual Care	1 CEA[192]
		Exercise	Usual Care	1 CUA[193]
		Acupuncture	Usual Care	1 CUA[195]
		Minimal Psychosocial Intervention	Usual Care	1 CEA[194]
	Any KQ 9	Any	Any	**4 CEA/CUA**[192-195]

Key Question	Subquestion	Intervention	Comparator	Number of Studies and Design (number of publications)
Total Number Studies Included	**Any KQ 3-9**	Any	Any	**105 RCTs (110)**[28,57-61,63-91,93-157,160-169] **4 NRSI**[171,172,176,184] **4 CEA/CUA**[192-195]

CEA = cost effectiveness analysis; CUA = cost-utility analysis; IV = intravenous; KQ = Key Question; MRI = magnetic resonance imaging; NRSI = nonrandomized studies of interventions; NA = not applicable; NSAID = nonsteroidal anti-inflammatory drug; RCT = randomized controlled trial; SR = systematic review; TENS = transcutaneous electrical nerve stimulation.

Key Question 1. In adults presenting with ALBP (<6 weeks duration)in any setting

Key Question 1a. How accurate is a focused patient history and physical exam (including risk factor and symptom assessment, presence and severity neurologic deficit and psychological risk factors) for identifying serious underlying conditions, specifically cancer, infection, inflammatory arthropathy, blunt force trauma, fracture (including vertebral compression fracture), or low back pain associated with severe or progressive neurological deficits including cauda equina syndrome?

Vertebral Fracture

One large SR ((14 NRSIs, N=19,759; including studies identified through July 2022) evaluated the accuracy of factors for identifying presence of vertebral fracture in various (primary, secondary, or tertiary care) settings [53] (Appendix E). A second SR (N=1,377; including studies identified through January 2019) focused on ED settings and included four NRSIs on identification of vertebral fracture.[54] All but one of the studies in were included in the first SR.[197] Apart from not reporting searches for unpublished studies or assessing potential publication bias, both SRs met criteria for well-conducted reviews (Appendix F). We did not identify any primary studies published after these SRs.

The larger SR stratified results by setting (primary care, secondary care, or tertiary care).[53] Across included studies, the prevalence of vertebral fracture ranged from 0.7% to 54%. A meta-analysis was not conducted due to the small number of studies with heterogeneity. The SR also noted methodological limitations in the included studies, most frequently related to patient selection, the index test, the reference standard, and the flow and timing domains.

Several studies included in the SR evaluated different age thresholds for identifying vertebral fractures. Higher age cutoffs were somewhat more informative than lower cutoffs for detecting fractures in primary care settings, but not secondary care (Table 11). In primary care settings (3 NRSI, N=3,899), age >50 or >54 years was associated with LR+ ranging from 1.72 (95% confidence interval [CI] 1.54 to 1.91) to 2.16 (95% CI 1.58 to 2.95) and LR- ranging from 0.33 (95% CI 0.20 to 0.53) to 0.57 (95% CI 0.23 to 1.39). In two NRSIs (N=3,278), age >64 years was associated with LR+ of 2.46 (95% CI 2.16 to 2.8) and 7.13 (95% CI 4.04 to 12.59) and LR- of 0.32 (95% CI 0.21 to 0.48) and 0.41 (95% CI 0.17 to 1.01). Based on three NRSIs (N=3,947), age >74 or ≥75 years was associated with LR+ ranging from 3.08 (95% CI 2.03 to 4.67) to 9.39 (95% CI 2.69 to 32.75) and LR- ranging from 0.49 (95% CI 0.38 to 0.63) and 0.77 (95% CI 0.52 to 1.15).

In secondary care settings (2 NRSIs, N=11,388), age >50 or >52 years was associated with LR+ of 1.52 (95% CI 1.42 to 1.68) and 1.10 (95% CI 1.05 to 1.16) and LR- of 0.14 (95% CI 0.04 to 0.53) and 0.79 (95% CI 0.69 to 0.91). In two NRSIs (N=10,046), age >70 or ≥75 years was associated with LR+ of 1.55 (95% CI 1.36 to 1.76) and 2.51 (95% CI 1.48 to 4.27) and LR- of 0.86 (95% CI 0.82 to 0.91) and 0.60 (95% CI 0.46 to 0.79).

The SR found that results for female sex were highly inconsistent in both primary care (3 studies, N=3,795) and secondary care (2 studies, N=1,549) settings (Table 11). However, combining female sex and age >74 years appeared somewhat more informative than age alone for identifying patients with vertebral fractures in primary care settings. LR+ were 2.75 (95% CI 2.26 to 3.35) and 14.59 (95% CI 8.00 to 26.61) and LR- were 0.52 (95% CI 0.40 to 0.68) and 0.39 (95% CI 0.16 to 0.96) across two NRISs (N=3,278).

The SR found that the utility of trauma for identifying vertebral fracture varied and appeared more useful in primary care settings. In primary care settings(4 NRSIs, N=3,278), LR+ ranged from 1.93 (95% CI 0.67 to 5.53) to 12.85 (95% CI 8.58 to 19.24) and LR- ranged from 0.36 (95% CI 0.22 to 0.62) to 0.88 (95% CI 0.67 to 1.17). One NRSI (N=9,940) conducted in a secondary care setting reported LR+ for trauma of 2.18 (95% CI 1.86 to 2.54) and LR- of 0.85 (95% CI 0.81 to 0.89). In tertiary care settings (3 NRSIs, N=1,259), LR+ for trauma ranged from 0.18 (95% CI 0.07 to 0.48) to 1.93 (95% CI 1.48 to 2.52) and LR- ranged from 0.12 (95% CI 0.01 to 1.79) to 1.54 (95% CI 1.38 to 1.71).

Fewer studies in the SR evaluated the utility of corticosteroid use and history of osteoporosis for fracture in primary care settings. Long-term corticosteroid use was modestly to highly informative for identifying patients with fractures, but absence of corticosteroid use was not very useful for ruling out fracture. Based on three NRSIs (N=2,468), corticosteroid use was associated with LR+ that ranged from 2.46 (95% CI 1.13 to 5.34) to 48.50 (95% CI 11.46 to 204.98) and LR- that ranged from 0.75 (95% CI 0.51 to 1.13) to 0.98 (95% CI 0.89 to 1.07). In one NRSI (n=669), the presence of osteoporosis was associated with LR+ of 3.13 (95% CI 1.90 to 5.15) and a LR- of 0.72 (95% CI 0.56 to 0.93).

The SR included one NRSI conducted in a tertiary care setting (n=552) that found presence of contusion/abrasion was highly informative for identifying fracture (LR 31.09, 95% CI 18.25 to 52.96) and absence of contusion/abrasion was highly informative for ruling out fracture (LR- 0.15, 95% CI 0.07 to 0.32).

The SR also evaluated clusters or combinations of fractures for identifying vertebral fractures. In general, clusters of factors appeared more informative than individual factors. One NRSI conducted in a primary care setting (n=1178) that found that having at least two of four signs (age >70 years, female, prolonged corticosteroid use, or trauma ["Henschke rule"]) was associated with a LR+ of 15.48 (95% CI 8.45 to 28.36) and LR- of 0.39 (95% CI 0.16 to 0.96) for fracture. Another NRSI conducted in a secondary care setting (n=1,448) found that having at least four of five signs (age >52 years, no leg pain, body mass index ≤22 kg/m2, no regular exercise, or female sex ["Roman cluster"]) was associated with a LR+ of 9.62 (95% CI 5.88 to 15.73) and LR- of 0.66 (95% CI 0.52 to 0.84). Increasing the threshold to at least three of five signs was less informative (LR+ 2.45, 95% CI 2.02 to 2.97 and LR- 0.34, 95% CI 0.19 to 0.61).

The SR that was restricted to studies conducted in ED settings found highly inconsistent results for trauma and presence of vertebral fracture (prevalence ranged from 0.3% to 7.2%).[54] Based on three NRSIs (N=1,259), LR+ ranged from 1.12 (95% CI 0.52 to 2.41) to 2.91 (95% CI 2.42 to 3.49) and LR- ranged from 0.00 to 0.93 (95% CI 0.56 to 1.56). All other factors addressed in the SR were assessed by one NRSI each. The presence of trauma plus neurological

signs (n=225) was most informative for identifying fracture, with a LR+ of 31.1 (95% CI 5.10 to 190.25); the LR- was 0.72 (95% CI 0.45 to 1.15). Other factors evaluated in ED settings were modestly informative. The presence of BMI ≥30 kg /m2 (n=118) was associated with a LR+ of 3.07 (95% CI 1.28 to 7.34) and LR- of 0.43 (95% CI 0.09 to 2.12); presence of contusion or abrasion (n=552) was associated with LR+ of 5.49 (95% CI 2.23 to 13.50) and LR- of 0.87 (95% CI 0.77 to 1.10); presence of tenderness on physical exam (n=552) was associated with LR+ of 1.76 (95% CI 1.42 to 2.19) and LR- of -0.47 (95% CI 0.28 to 0.78) and presence of abnormal neurological exam (n=225) was associated with LR+ of 4.15 (95% CI 1.17 to 14.77) and LR- of 0.77 (95% CI 0.48 to 1.23). The presence of two or more physical findings (contusion/abrasion, tenderness, spasm, sensory deficit, motor deficit, deep tendon reflex abnormality, positive straight leg raise test) was not more predictive than individual findings (n=552; LR+ 2.05, 95% CI 1.38 to 3.06 and LR- 0.73, 95% CI 0.55 to 0.95).

Table 11. Summary table: vertebral fracture

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Han, 2023	Vertebral fracture (primary care settings)	3 (N=3,899)	Age >50 or >54 years	Range 0.63 (0.24−0.91) to 0.83 (0.73−0.90)	Range 0.52 (0.49−0.54) to 0.66 (0.63−0.69)	Range 1.72 (1.54−1.91) to 2.16 (1.58−2.95)	Range 0.33 (0.20−0.53) to 0.57 (0.23−1.39)
Han, 2023	Vertebral fracture (primary care settings)	2 (N=3,278)	Age >64 years	0.63 (0.24−0.91) and 0.78 (0.68−0.87)	0.91 (0.89−0.93) and 0.68 (0.66−0.70	2.46 (2.16−2.8) and 7.13 (4.04−12.59)	0.32 (0.21−0.48) and 0.41 (0.17−1.01)
Han, 2023	Vertebral fracture (primary care settings)	3 (N=3,947)	Age >74 or ≥75 years	Ranged from 0.25 (0.03−0.65) to 0.59 (0.48−0.70)	Ranged from 0.84 (0.82−0.86) to 0.97 (0.96−0.98)	Ranged from 3.08 (2.03−4.67) to 9.39 (2.69−32.75);	Ranged from 0.49 (0.38−0.63) to 0.77 (0.52−1.15)
Han, 2023	Vertebral fracture (primary care settings)	3 (N=3,795)	Female	Ranged from 0.43 (0.36−0.49) to 0.67 (0.48−0.82)	Ranged from 0.41 (0.37−0.44) to 0.43 (0.41−0.45)	Ranged from 0.72 (0.61−0.84) to 1.26 (1.10−1.45)	Ranged from 0.65 (0.46−0.92) to 1.41 (1.23−1.63)
Han, 2023	Vertebral fracture (primary care settings)	2 (N=3,278)	Female and >74 years	0.63 (0.24−0.91) and 0.59 (0.48−0.70)	0.96 (0.94−0.97) and 0.79 (0.77−0.80)	2.75 (2.26−3.35) and 14.59 (8.00−26.61)	0.52 (0.40−0.68) and 0.39 (0.16−0.96)
Han, 2023	Vertebral fracture (primary care settings)	4 (N=3,339)	Trauma	Ranged from 0.21 (0.05−0.51) to 0.65 (0.44−0.83)	Ranged from 0.89 (0.85−0.92) to 0.98 (0.96−0.98)	Ranged from 1.93 (0.67−5.53) to 12.85 (8.58−19.24)	Ranged from 0.36 (0.22−0.62) to 0.88 (0.67−1.17)
Han, 2023	Vertebral fracture (primary care settings)	3 (N=2,468)	Corticosteroid use	Ranged from 0.00 (0.00−0.23) to 0.25 (0.03−0.65)	Ranged from 0.93 (0.90−0.95) to 0.99 (0.99−1.00)	Ranged from 2.46 (1.13−5.34) to 48.50 (11.46−204.98)	Ranged from 0.75 (0.51−1.13) to 0.98 (0.89−1.07)
Han, 2023	Vertebral fracture (primary care settings)	1 (N=669)	Osteoporosis	0.36 (0.20−0.55)	0.88 (0.86−0.91)	3.13 (1.90−5.15)	0.72 (0.56−0.93)
Han, 2023	Vertebral fracture (primary care settings)	1 (N=1,178)	Henschke rule ≥2 of 4 positive signs (age >70 years, female, prolonged corticosteroid use, trauma)	0.63 (0.24−0.91)	0.96 (0.95−0.97)	15.48 (8.45−28.36)	0.39 (0.16−0.96)

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Han, 2023	Vertebral fracture (secondary care settings)	2 (N=11,388)	Age >50 or >52 years	0.74 (0.70−0.78) and 0.95 (0.82−0.99)	0.33 (0.32−0.34) and 0.39 (0.36−0.41)	1.52 (1.42−1.68) and 1.10 (1.05−1.16)	0.14 (0.04−0.53) and 0.79 (0.69−0.91)
Han, 2023	Vertebral fracture (secondary care settings)	2 (N=10,076)	Age >70 or ≥75 years	0.31 (0.27−0.35) and 0.53 (0.41−0.64)	0.80 (0.79−0.81) and 0.79 (0.67−0.88)	1.55 (1.36−1.76) and 2.51 (1.48−4.27)	0.86 (0.82−0.91) and 0.60 (0.46−0.79)
Han, 2023	Vertebral fracture (secondary care settings)	2 (N=1,549)	Female	0.25 (0.03−0.65) and 0.59 (0.48−0.70)	0.97 (0.96−0.98) and 0.84 (0.82−0.86)	1.51 (1.35−1.71) and 0.53 (0.27−1.05)	0.26 (0.10−0.65) and 1.26 (0.98−1.61)
Han, 2023	Vertebral fracture (secondary care settings)	1 (N=9,940)	Trauma	0.25 (0.21−0.29)	0.89 (0.88−0.89)	2.18 (1.86−2.54)	0.85 (0.81−0.89)
Han, 2023	Vertebral fracture (secondary care settings)	1 (N=1,448)	Roman cluster ≥3 of 5 positive signs (age >52 years, no leg pain, body mass index ≤22, no regular exercise, female sex)	0.76 (0.60−0.89)	0.69 (0.66−0.71)	2.45 (2.02−2.97)	0.34 (0.19−0.61)
Han, 2023	Vertebral fracture (secondary care settings)	1 (N=1,448)	Roman cluster ≥4 of 5 positive signs (age >52 years, no leg pain, body mass index ≤22, no regular exercise, female sex)	0.37 (0.22−0.54)	0.96 (0.95−0.97)	9.62 (5.88−15.73)	0.66 (95% 0.52−0.84)
Han, 2023	Vertebral fracture (tertiary care settings)	3 (N=1,259)	Trauma	Ranged from 0.07 (0.02−0.18) to 1.00 (0.59−1.00)	Ranged from 0.51 (0.41−0.62) to 0.60 (0.56−0.65)	Ranged from 0.18 (0.07−0.48) to 1.93 (1.48−2.52)	Ranged from 0.12 (0.01−1.79) to 1.54 (1.38−1.71)
Han, 2023	Vertebral fracture (tertiary care settings)	1 (N=552)	Contusion/abrasion	0.85 (0.70−0.94)	0.97 (0.95−0.98)	31.09 (18.25−52.96)	0.15 (0.07−0.32)
Galliker, 2020	Vertebral fracture (ED setting)	3 (N=1,259)	History of trauma	Ranged from 0.40 (0.12−0.74) to 1.00 (0.59−1.00)	Ranged from 0.55 (0.50−0.59) to 0.66 (0.59−0.72)	Ranged from 1.12 (0.52−2.41) to 2.91 (2.42−3.49)	Ranged from 0.00 to 0.93 (0.56−1.56)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=225)	Trauma + neurological signs	0.29 (0.04−0.71)	0.99 (0.97−0.999)	31.1 (5.10−190.25)	0.72 (0.45−1.15)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=118)	Obesity (BMI ≥30 kg/m^2)	0.67 (0.09−0.99)	0.78 (0.70−0.85)	3.07 (1.28−7.34)	0.43 (0.09−2.12)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Contusion/abrasion	0.15 (0.06−0.30)	0.97 (0.95−0.98)	5.49 (2.23−13.50)	0.87 (0.77−1.00)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Tenderness on physical exam	0.72 (0.56−0.85)	0.59 (0.54−0.63)	1.76 (1.42−2.19)	0.47 (0.28−0.78)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=225)	Abnormal neurological exam	0.29 (0.04−0.71)	0.93 (0.89−0.96)	4.15 (1.17−14.77)	0.77 (0.48−1.23)

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Two or more physical findings (contusion/abrasion, tenderness, spasm, sensory deficit, motor deficit, deep tendon reflex abnormality, positive straight leg raising)	0.42 (0.27−0.59)	0.79 (0.76−0.83	2.05 (1.38−3.06)	0.73 (0.55−0.95)
Galliker, 2020	Vertebral fracture (ED setting)	3 (N=1,259)	History of trauma	Ranged from 0.40 (0.12−0.74) to 1.00 (0.59−1.00)	Ranged from 0.55 (0.50−0.59) to 0.66 (0.59−0.72)	Ranged from 1.12 (0.52−2.41) to 2.91 (2.42−3.49)	Ranged from 0.00 to 0.93 (0.56−1.56)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=225)	Trauma + neurological signs	0.29 (0.04−0.71)	0.99 (0.97−0.999)	31.1 (5.10−190.25)	0.72 (0.45−1.15)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=118)	Obesity (BMI ≥30 kg/m^2)	0.67 (0.09−0.99)	0.78 (0.70−0.85)	3.07 (1.28−7.34)	0.43 (0.09−2.12)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Contusion/ abrasion	0.15 (0.06−0.30)	0.97 (0.95−0.98)	5.49 (2.23−13.50)	0.87 (0.77−1.00)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Tenderness on physical exam	0.72 (0.56−0.85)	0.59 (0.54−0.63)	1.76 (1.42−2.19)	0.47 (0.28−0.78)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=225)	Abnormal neurological exam	0.29 (0.04−0.71)	0.93 (0.89−0.96)	4.15 (1.17−14.77)	0.77 (0.48−1.23)
Galliker, 2020	Vertebral fracture (ED setting)	1 (N=552)	Two or more physical findings (contusion/ abrasion, tenderness, spasm, sensory deficit, motor deficit, deep tendon reflex abnormality, positive straight leg raising)	0.42 (0.27−0.59)	0.79 (0.76−0.83	2.05 (1.38−3.06)	0.73 (0.55−0.95)

BMI = body mass index; CI = confidence interval; ED = emergency department; LR = likelihood ratio.

Spinal Malignancy

One SR[55] (8 NRSIs, N=7,361; including studies identified through April 2012) evaluated the accuracy of factors for identifying presence of spinal malignancy in various settings (Appendix E). A second, previously described SR[54] (2 NRSI, N=1,219; including studies identified through January 2019) focused on identifying spinal malignancies in ED settings; one of these studies[198] was not included in the prior review. Apart from not reporting searches for unpublished studies or assessing potential publication bias, both SRs met criteria for well-conducted reviews (Appendix F). We identified no primary studies published subsequent to the SRs.

Of the studies include in the broader SR, six were conducted in primary care settings, one was conducted in an accident or ED setting, and one was conducted in a secondary care setting.[55] The prevalence of spinal malignancy ranged from 0% to 0.66%. Likelihood ratios could not be estimated from one primary care study[199] (n=1,172) because no cases of spinal malignancy occurred and it was excluded from the ranges described below. Meta-analysis was not conducted due to the small number of studies with heterogeneity. In addition, the review noted methodological limitations in the included studies. Only one study reported interpretation of the

reference standard to the factors under investigation; only one study reported uninterpretable results; only two studies explained study withdrawals; and only two studies described an acceptable delay between assessment of the factors under investigation and the reference standard. Additionally, most studies did not use magnetic resonance imaging (MRI), the most accurate imaging test, as the reference standard.

The SR reported highly inconsistent results for age >50 years and identification of spinal malignancy (Table 12). Based on five NRSIs (N=4,473), LR+ ranged from 0.72 (95% CI 0.41 to 1.25) to 2.50 (95% CI 1.40 to 4.46) and LR- ranged from 0.0 to 1.11 (95% CI 0.96 to 1.29). The only NRSI that evaluated age >70 years did not identify any cases of spinal malignancy.

Three NRSIs (N=3,583) found a history of cancer highly informative for presence of spinal malignancy. In the two studies in which likelihood ratios could be calculated, LR+ were 15.27 (95% CI 6.38 to 36.55) and 35.00 (95% CI 20.48 to 59.81) and LR- were 0.71 (95% CI 0.49 to 1.02) and 0.0.

Two NRSIs (N=2,596) evaluated failure to improve after 1 month. This factor was modestly informative for identifying spinal malignancy (LR+ 2.61, 95% CI 0.47 to 14.52 and 3.08, 95% CI 1.35 to 7.04), but absence of this factor was only modestly informative for ruling out spinal malignancy (LR- 0.83, 95% CI 0.47 to 1.46 and 0.77, 95% CI 0.54 to 1.11).

One NRSI (n=1,902) found duration of symptoms >1 month was associated with LR+ of 2.63 (95% CI 1.48 to 4.67) and LR- of 0.62 (95% CI 0.35 to 1.09).

Two NRSIs evaluated absence of relief with bedrest and unexplained weight loss, but one of these studies did not identify any cases of spinal malignancy. In the other study absence of relief with bedrest had a LR+ of 1.85 (95% CI 1.75 to 1.96); because sensitivity was 100%, the LR- was 0.0. For unexplained weight loss, the LR+ was 2.57 (95% CI 0.71 to 9.31) and the LR- was 0.90 (95% CI 0.71 to 1.14).

Two NRSIs (N=2,816) found that presence of neurological symptoms was not informative for identifying spinal malignancy. Because no patients with spinal malignancy had neurological symptoms (sensitivity 0.0), the LR+ were 0. LR- were 1.10 (95% CI 1.08 to 1.12) and 1.03 (95% CI 1.02 to 1.05). One NRSI (n=1959) found that presence of fever was not informative for identification of spinal malignancy, with a LR+ of 0.00 (sensitivity 0.00) and LR- of 1.02 (95% CI 1.01 to 1.03).

Likelihood ratios could not be estimated for the presence of constant progressive pain or being systemically unwell, as both were only evaluated in one study that identified no cases of spinal malignancy (n=1172 and 1169, respectively).

The SR that focused on studies conducted in ED settings included one NRSI (n=482) that found a history or clinical suspicion of cancer highly informative for spinal malignancy (LR+ 27.9, 95% CI 17.52 to 44.56 and LR- 0.0 [sensitivity 1.0]).[54] A history of cancer (1 NRSI, N=737; LR+ 5.86, 95% CI 1.94 to 17.70) and having an abnormal neurological exam (1 NRSI, N=482; LR+ 5.99, 95% CI 2.40 to 14.94) were moderately informative for identifying spinal malignancy, but absence of these factors was only modestly informative for ruling out spinal malignancy (LR- 0.64, 95% CI 0.31 to 1.32 and 0.62, 95% CI 0.32 to 1.17, respectively).

Table 12. Summary table: spinal malignancy

Study, Year	Condition	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Henschke, 2013	Spinal malignancy	5 (N=4,473)	Age >50 years	Ranged from 0.50 (0.01−0.99)* to 1.00 (0.81−1.00);	Ranged from 0.41 (0.35−0.48) to 0.74 (0.70−0.78)	Ranged from 0.72 (0.41−1.25) to 2.50 (1.40−4.46)*	ranged from 0.0 to 1.11 (0.96−1.29)*
Henschke, 2013	Spinal malignancy	1 (N=1,172)	Age >70 years	Not calculable	0.95 (0.94−0.96)	Not calculable	Not calculable
Henschke, 2013	Spinal malignancy	2 (N=2,816)	Neurological symptoms	0.00 (0.00−0.26) and 0.00 (0.00−0.97)	0.91 (0.90−0.92) and 0.97 (0.95−0.98)	0.00 (both studies)	1.10 (1.08−1.12) and 1.03 (1.02−1.05)
Henschke, 2013	Spinal malignancy	1 (N=1,172)	Constant progressive pain	Not calculable	0.97 (0.96−0.98)	Not calculable	Not calculable
Henschke, 2013	Spinal malignancy	1 (N=1,902)	Duration >1 month	0.50 (0.21−0.79)	0.81 (95% 0.79−0.83)	2.63 (1.48−4.67)	0.62 (0.35−1.09)
Henschke, 2013	Spinal malignancy	2 (N=2,596)	Not improved after 1 month	0.25 (0.01−0.81) and 0.31 (0.09−0.61)	0.90 (0.88−0.93) and 0.90 (0.89−0.91)	2.61 (0.47−14.52) and 3.08 (1.35−7.04)	0.83 (0.47−1.46) and 0.77 (0.54−1.11)
Henschke, 2013	Spinal malignancy	3 (N=3,583)	History of cancer	0.31 (0.09−0.61) and 1.00 (0.59−1.00)*	Ranged from 0.96 (0.95−0.97) to 0.98 (0.97−0.99)	15.27 (6.38−36.55) and 35.00 (20.48−59.81)*	0.71 (0.49−1.02) and 0.0*
Henschke, 2013	Spinal malignancy	1 (N=1,169)	Systemically unwell	Not calculable	0.98 (0.97−0.98)	Not calculable	Not calculable
Henschke, 2013	Spinal malignancy	2 (N=2,115)	No relief with bedrest	1.00 (0.40−1.00) and not calculable	0.46 (0.43−0.49) and 0.84 (0.81−0.86)	1.85 (1.75−1.96) and not calculable	0.00 and not calculable
Henschke, 2013	Spinal malignancy	2 (N=3,123)	Unexplained weight loss	0.15 (0.02−0.45) and not calculable	0.94 (0.93−0.95) and 1.00 (0.99−1.00)	2.57 (0.71−9.31) and not calculable	0.90 (0.71−1.14) and not calculable
Henschke, 2013	Spinal malignancy	1 (N=1,959)	Fever	0.0 (0% 0.0−0.25)	0.98 (0.97−0.99)	0.00	1.02 (1.01−1.03)
Galliker, 2020	Cancer (ED setting)	1 (N=482)	History/clinical suspicion of cancer	1.00 (0.59−1.00)	0.96 (0.94−0.98)	27.9 (17.52−44.56)	0.00
Galliker, 2020	Cancer (ED setting)	1 (N=737)	History of cancer	0.40 (0.05−0.85)	0.93 (0.91−0.95)	5.86 (1.94−17.70)	0.64 (0.31−1.32)
Galliker, 2020	Cancer (ED setting)	1 (N=482)	Abnormal neurological exam	0.43 (0.10−0.82)	0.93 (0.90−0.95)	5.99 (2.40−14.94)	0.62 (0.32−1.17)

CI = confidence interval; ED = emergency department; LR = likelihood ratio.
* Excluding one study with no cases of spinal malignancy

Cauda Equina Syndrome

One SR[56] (7 NRSIs, n=869; including studies identified through January 2018) evaluated the accuracy of factors for identifying presence of cauda equina syndrome in any secondary or tertiary care (ED or hospital) setting (Appendix E). A second, previously described SR[54] focused on identifying cauda equina syndrome in ED settings (2 NRSIs, N=324; including studies

identified through January 2019). One of these NRSI[197] was not included in the prior review. Apart from not reporting searches for unpublished studies or assessing for potential publication bias, the SRs both met criteria for well-conducted reviews (Appendix F). We identified no primary studies published after the SRs.

In the larger SR, the prevalence of cauda equina syndrome ranged from 15.3% to 56.5%. Meta-analyses were conducted, although statistical heterogeneity was present in most analyses. MRI was the reference standard in all studies. The SR assessed all included studies as high risk of bias. Methodological limitations were identified with regard to patient selection, index test, reference standard, and flow and timing. The evidence was assessed as low certainty, due to inconsistency, imprecision, and methodological limitations.

The SR identified no factor that was highly informative for identification of cauda equina syndrome (Table 13). Four factors appeared modestly informative. Based on four NRSIs (N=371), saddle anesthesia was associated with a pooled LR+ of 1.73 (95% CI 0.98 to 3.08) and pooled LR- of 0.80 (95% CI 0.61 to 1.05). Based on four NRSIs (N=217), bowel incontinence was associated with a pooled LR+ of 1.60 (95% CI 0.65 to 3.94) and pooled LR- of 0.96 (95% CI 0.78 to 1.18). Based on three NRSIs (N=190), reduced anal tone was associated with a pooled LR+ of 1.72 (95 5CI 0.91 to 3.24) and pooled LR- of 0.86 (95% CI 0.69 to 1.07). Based on five NRSIs (N=240), presence of leg pain was associated with pooled LR+ of 1.45 (95% CI 0.81 to 2.58) and pooled LR- of 0.90 (95% CI 0.60 to 1.35).

Presence of urinary retention or urinary incontinence was not informative for identifying cauda equina syndrome. Based on five NRSIs (N=252), presence of urinary retention was associated with a pooled LR+ of 0.99 (95% CI 0.61 to 1.61) and pooled LR- of 1.06 (95% CI 0.86 to 1.30). Based on five NRSIs (N=251), presence of urinary incontinence was associated with a pooled LR+ of 0.80 (95% CI 0.56 to 1.14) and pooled LR- of 1.13 (95% CI 0.90 to 1.43).

Presence of back pain slightly increased the likelihood of cauda equina syndrome (3 studies, N=253; pooled LR+ 1.23, 95% CI 0.77 to 2.20). Although absence of back pain was modestly informative for ruling out cauda equina syndrome, the pooled estimate was imprecise (pooled LR- 0.40, 95% CI 0.05 to 2.41).

The SR[54] that focused on studies conducted in ED settings evaluated the utility of BMI ≥30 kg/m2, presence of saddle anesthesia, or bowel/bladder dysfunction for identification of spinal cord compression or cauda equina syndrome. Each factor was evaluated in one NRSI. Presence of bowel or bladder dysfunction was modestly informative for identifying spinal cord compression or cauda equina syndrome (N=206; LR+ 2.08, 95% CI 1.41 to 3.08 and LR- 0.60, 95% CI 0.40 to 0.90). Compared to presence of bowel or bladder dysfunction, presence of saddle anesthesia was associated with slightly higher likelihood for presence of spinal cord compression or cauda equina syndrome, but absence of saddle anesthesia was slightly less useful for ruling out these conditions (N=206; LR+ 3.15, 95% CI 1.66 to 5.97 and LR- 0.74, 95% CI 0.57 to 0.95). Notably, the confidence intervals were wide and overlapped for these two factors. Estimates for BMI ≥30 kg/m2 were imprecise (LR+ 2.23, 95% CI 0.54 to 9.29 and LR- 0.16 to 2.59).

Herniated Disc and Spinal Stenosis

One previously described SR[54] included one NRSI[200] (N=100) on identification of herniated disc (prevalence 81%) and one NRSI[197] (n=118) on identification of severe spinal canal stenosis (prevalence 13%) in ED settings (Appendix E, Appendix F). The SR noted methodological

limitations in both studies, including risk of bias related to patient selection, index test, reference standard, and flow and timing. We identified no studies published after the systematic review.

For herniated disc, the SR found that correlation of symptoms with a positive straight leg raise test was highly informative (Table 13). Presence of this factor was associated with a sensitivity of 0.90 (95% CI 0.82 to 0.95) and specificity of 1.00 (95% CI 0.72 to 1.00); LR+ was not calculable due to the specificity of 1.00, and the LR- was 0.10 (95% CI 0.05 to 0.19). A positive crossed straight leg raise test (reproduction of symptoms in the affected leg when then opposite leg is raised) was less informative (LR+ 3.21, 95% CI 0.48 to 21.41 and LR- 0.78, 95% CI 0.62 to 0.98) and a positive straight leg raise was not very useful (LR+ 1.22, 95% CI 0.77 to 1.93 and LR- 0.62, 95% CI 0.26 to 1.48).

For severe spinal stenosis, the SR reported that presence of BMI ≥30 kg/m2 was associated with LR+ of 4.75 (95% CI 2.80 to 8.05) and LR- of 0.28 (95% CI 0.10 to 0.75) (Table 13). Presence of a focal neurologic deficit was less informative (PL+ 2.31, 95% CI 0.53 to 9.96 and LR- -0.91, 95% CI 0.72 to 1.15).

Table 13. Summary table: compression-related conditions

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Dionne, 2019	Cauda equina syndrome	4 (N=371)	Saddle anesthesia	0.43 (0.31−0.55, I^2=92%)	0.79 (0.74−0.83, I2=80%)	1.73 (0.98−3.08)	0.80 (0.61−1.05)
Dionne, 2019	Cauda equina syndrome	4 (N=217)	Bowel incontinence	0.18 (0.09−0.32, I^2=64%)	0.86 (0.80−0.91, I2=0%)	1.60 (0.65−3.94)	0.96 (0.78−1.18)
Dionne, 2019	Cauda equina syndrome	3 (N=190)	Reduced anal tone	0.29 (0.15−0.47, I^2=0%)	0.83 (0.76−0.88, I2=77%)	1.72 (0.91−3.24)	0.86 (0.69−1.07)
Dionne, 2019	Cauda equina syndrome	5 (N=240)	Leg pain	0.46 (0.32−0.59, I^2=68%)	0.65 (0.58−0.72, I2=95%)	1.45 (0.81−2.58)	0.90 (0.60−1.35)
Dionne, 2019	Cauda equina syndrome	5 (N=252)	Urinary retention	0.36 (0.25−0.49, I^2=0%)	0.61 (0.54−0.68, I2=85%)	0.99 (0.61−1.61)	1.06 (0.86−1.30)
Dionne, 2019	Cauda equina syndrome	5 (N=251)	Urinary incontinence	0.36 (0.25−0.48, I^2=0%)	0.56 (0.49−0.64, I2=25%)	0.80 (0.56−1.14)	1.13 (0.90−1.43)
Dionne, 2019	Cauda equina syndrome	3 (N=253)	Back pain	0.92 (0.82−0.98, I^2=38%)	0.30 (0.23−0.37, I2=93%)	1.23 (0.77−2.20)	0.40 (0.05−2.41)
Galliker, 2020	Spinal cord compression or cauda equina syndrome (ED setting)	1 (N=118)	Obesity (BMI ≥30 kg/m^2)	0.50 (0.01−0.99)	0.78 (0.69−0.85)	2.23 (0.54−9.29)	0.64 (0.16−2.59)
Galliker, 2020	Spinal cord compression or cauda equina syndrome (ED setting)	1 (N=206)	Disturbance of saddle sensation on physical examination	0.34 (0.19−0.53)	0.89 (0.83−0.93)	3.15 (1.66−5.97)	0.74 (0.57−0.95)
Galliker, 2020	Spinal cord compression or cauda equina syndrome (ED setting)	1 (N=206)	Bowel/bladder sphincter dysfunction	0.56 (0.38−0.74)	0.73 (0.66−0.79)	2.08 (1.41−3.08)	0.60 (0.40−0.90)

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Galliker, 2020	Severe spinal stenosis (ED setting)	1 (N=118)	Obesity (BMI ≥30 kg/m^2)	0.77 (0.46−0.95)	0.84 (0.75−0.90)	4.75 (2.80−8.05)	0.28 (0.10−0.75)
Galliker, 2020	Severe spinal stenosis (ED setting)	1 (N=118)	Focal neurological deficit	0.15 (0.02−0.45)	0.93 (0.87−0.97)	2.31 (0.53−9.96)	0.91 (0.72−1.15)
Galliker, 2020	Herniated disc (ED setting)	1 (N=100)	Positive straight leg raise test	0.78 (0.67−0.86)	0.36 (0.11−0.69)	1.22 (0.77−1.93)	0.62 (0.26−1.48)
Galliker, 2020	Herniated disc (ED setting)	1 (N=100)	Positive cross straight leg raise test	0.29 (0.20−0.40)	0.91 (0.59−0.998)	3.21 (0.48−21.41)	0.78 (0.62−0.98)
Galliker, 2020	Herniated disc (ED setting)	1 (N=100)	Correlation of symptoms with positive straight leg raise test	0.90 (0.82−0.95)	1.00 (0.72−1.00)	not calculable	0.10 (0.05−0.19)

BMI = body mass index; CI = confidence interval; ED = emergency department; LR = likelihood ratio.

Epidural Abscess

One previously described SR[54] included two NRSI (N=19,222) that investigated factors for identification of epidural abscess in ED settings (Appendix E, Appendix F). Each factor was evaluated in only one study (Table 14). The most predictive factors for presence of epidural abscess were: history of intravenous drug use (N=19,033; LR+ 13.74, 95% CI 11.44 to 16.50 and LR- 0.41, 95% CI 0.32 to 0.54); having an indwelling catheter (N=19,033; LR+ 15.46, 95% CI 7.79 to 30.67 and LR- 0.91, 95% CI 0.85 to 0.98); having another site of infection (N=19,033, LR+ 13.46, 95% CI 9.26 to 19.58 and LR- 0.76, 95% CI 0.67 to 0.86); having a recent spine fracture (N=19,032, LR+ 9.53, 95% CI 5.06 to 17.95 and LR- 0.91, 95% CI 0.84 to 0.97) and systolic blood pressure <90 mm Hg (N=19,033; LR+ 8.28, 95% 3.48 to 19.72 and LR- 0.95, 95% CI 0.90 to 1.00). Factors associated with less informative LR+ (<5) were: being immunocompromised (N=19,033, LR+ 4.98, 95% CI 3.13 to 7.94 and LR- 0.86, 95% CI 0.78 to 0.94); history of liver disease (N=19,033; LR+ 3.10, 95% CI 1.83 to 5.26 and LR- 0.90, 95% CI 0.83 to 0.98); recent spine procedure or surgical implant (N=19,032, LR+ 3.64, 95% CI 2.29 to 5.78 and LR- 0.87, 95% CI 0.79 to 0.96), and temperature ≥100.4º F (N=19,033, LR+ 3.49, 95% CI 1.60 to 7.59 and LR- 0.95, 95% CI 0.90 to 1.01).

The SR also evaluated combinations of factors for identifying epidural abscess. Two NRSIs included in the SR reported that presence of at least one risk factor moderately increased the likelihood of epidural abscess, with absence of any risk factor highly useful for ruling out epidural abscess. In one large NRSI (N=19,033), presence of diabetes, intravenous drug use, a recent spine procedure or indwelling spinal surgical implant, recent spine fracture, indwelling vascular catheter, being immunocompromised, having another site of infection, or having chronic liver disease was associated with a LR+ of 4.28 (95% CI 4.17 to 4.39) and LR- of 0.00 (sensitivity was 1.0). Requiring an erythrocyte sedimentation rate of >20 mm/hour or c-reactive protein of >1.0 plus one of these risk factors did not improve diagnostic accuracy (LR+ 3.03, 95% CI 2.97 to 3.09 and 1.74, 95% CI 1.61 to 1.89, respectively, and LR- 0.00 and 0.26, 95% CI 0.15 to 0.44, respectively). In the second, smaller NRSI (N=189), presence of diabetes, intravenous drug use, recent spine procedure, recent spine fracture, indwelling catheter, being immunocompromised, having a distant site of infection, alcohol abuse, chronic renal failure, or

cancer was associated with LR+ of 4.59 (95% CI 3.28 to 6.43) and LR- of 0.02 (95% CI 0.00 to 0.14). One NRSI (N=19,033) found that presence of the "classical triad" (temperature ≥100.4º F, spine pain, and neurologic deficit) was associated with a similar LR+ compared to the prior rules (5.80, 95% CI 1.45 to 23.23), but absence of the triad was not useful for ruling out epidural abscess (LR- 0.98, 95% CI 0.95 to 1.01).

Table 14. Summary table: epidural abscess

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Any risk factor*	1.00 (0.96−1.00)	0.77 (0.76−0.77)	4.28 (4.17−4.39)	0.00
Galliker, 2020	Epidural abscess (ED setting)	1 (N=189)	Any risk factor†	0.98 (0.91−0.999 6)	0.79 (0.70−0.85)	4.59 (3.28−6.43)	0.02 (0.00−0.14)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Immuno-compromised	0.17 (0.10−0.27)	0.96 (0.96−0.97);	4.98 (3.13−7.94)	0.86 (0.78−0.94)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Diabetes mellitus	0.16 (0.09−0.26)	0.92 (0.92−0.93)	2.17 (1.34−3.51)	0.91 (0.82−0.99)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Intravenous drug use	0.60 (0.49−0.71)	0.96 (0.95−0.96)	13.74 (11.44−16.50)	0.41 (0.32−0.54)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Liver disease	0.14 (0.07−0.23)	0.96 (0.95−0.96)	3.10 (1.83−5.26)	0.90 (0.83−0.98)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Renal disease	0.02 (0.003−0.08)	0.99 (0.988−0.991)	2.33 (0.59−9.24)	0.99 (0.95−1.02)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Recent spine procedure/surgical implants	0.17 (0.10−0.27)	0.95 (0.95−0.96)	3.64 (2.29−5.78)	0.87 (0.79−0.96)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Recent spine fracture	0.10 (0.05−0.19)	0.99 (0.99−0.99)	9.53 (5.06−17.95)	0.91 (0.84−0.97)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Classical triad (temperature ≥100.4 F, spine pain, and neurological deficit)	0.02 (0.003−0.08)	0.996 (0.995−0.997)	5.80 (1.45−23.23)	0.98 (0.95−1.01)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Temperature ≥100.4 F	0.07 (0.03−0.15)	0.98 (0.98−0.98)	3.49 (1.60−7.59)	0.95 (0.90−1.01)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	SBP <90 mm Hg	0.06 (0.02−0.13)	0.993 (0.992−0.944)	8.28 (3.48−19.72)	0.95 (0.90−1.00)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Heart rate >100 beats/minute	0.37 (0.27−0.48)	0.84 (0.83−0.84)	2.27 (1.72−2.99)	0.75 (0.64−0.88)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Abnormal neurological exam	0.24 (0.16−0.35)	0.91 (0.90−0.91)	2.65 (1.83−3.86)	0.83 (0.74−0.94)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Indwelling vascular catheter	0.09 (0.04−0.18)	0.994 (0.993−0.995)	15.46 (7.79−30.67)	0.91 (0.85−0.98)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	Other site of infection	0.26 (0.17−0.36)	0.98 (0.98−0.98)	13.46 (9.26−19.58)	0.76 (0.67−0.86)
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	ESR >20 mm/hour + risk factor‡	1.00 (95% 0.96−1.00)	0.67 (0.66−0.68)	3.03 (2.97−3.09)	0.00
Galliker, 2020	Epidural abscess (ED setting)	1 (N=19,033)	CRP >1.0 + risk factor‡	0.87 (0.78−0.93)	0.50 (0.49−0.51)	1.74 (1.61−1.89)	0.26 (0.15−0.44)

CI = confidence interval; CRP = C-reactive protein; ED = emergency department; ESR = erythrocyte sedimentation rate; LR = likelihood ratio; SBP = systolic blood pressure.

* Diabetes, intravenous drug use, recent spine procedure or indwelling spinal surgical implant, recent spine fracture, indwelling vascular catheter, immunocompromised, other site of infection, chronic liver disease.
† Diabetes, intravenous drug use, recent spine procedure, recent spine fracture, indwelling catheter, immunocompromised, distant site of infection, alcohol abuse, chronic renal failure, cancer.
‡ Intravenous drug use, recent spine procedure or indwelling spinal surgical implant, recent spine fracture, indwelling vascular catheter, immunocompromised, other stie of infection, chronic liver disease.

Any Serious Spinal Pathology

One previously described SR evaluated factors for identification of any serious spinal pathology (defined as including cancer, epidural abscess, spinal cord compression, vertebral fracture, or severe spinal stenosis) in ED settings.[54] Findings were based on a single NRSI[201] (N=329; prevalence of any serious spinal pathology 6.7%) and estimates were imprecise. The most informative factors for identifying presence of any serious spinal condition were new urinary retention (LR+ 6.98, 95% CI 1.87 to 26.04 and LR- 0.88, 95% CI 0.75 to 1.04), new fecal incontinence (LR+ 4.65, 95% CI 0.50 to 42.89 and LR- 0.96, 95% CI 0.88 to 1.06); anticoagulant use (LR+ 8.72, 95% CI 3.11 to 24.44 and LR- 0.79, 95% CI 0.63 to 1.00); saddle anesthesia (LR+ 6.98, 95% CI 1.35 to 36.02 and LR- 0.92, 95% CI 0.81 to 1.05); hemoglobin <100 g/L (LR+ 13.95, 95% CI 2.06 to 94.39 and LR- 0.92, 95% CI 0.80 to 1.04) and international normalized ratio ≥1 (LR+ 6.11, 95% CI 2.81 to 13.26 and LR- 0.72, 95% CI 0.54 to 0.96). However, absence of these factors was either not useful or only modestly useful for ruling out any serious spinal pathology (Likelihood ratios close to 1). Factors that were less informative for identifying any serious spinal pathology were pain duration >365 days (LR+ 3.1, 95% CI 0.71 to 13.49 and LR- 0.94, 95% CI 0.82 to 1.07), pain awakening patients at night (LR+ 2.20, 95% CI 0.71 to 6.88, LR- 0.92, 95% CI 0.78 to 1.09), pain worse at night (LR+ 3.22, 95% CI 0.99 to 10.46 and LR- 0.90, 95% CI 0.76 to 1.07), presence of sensory symptoms (LR+ 2.03, 95% CI 1.11 to 3.71 and LR- 0.78, 95% CI 0.56 to 1.07), having a procedure causing bacteremia (LR+ 3.49, 95% CI 0.79 to 15.45 and LR- 0.93, 95% CI 0.82 to 1.07), a history of cancer (LR+ 2.46, 95% CI 0.78 to 7.77 and LR- 0.91, 95% CI 0.77 to 1.08), age >65 years (LR+ 3.22, 95% CI 2.04 to 5.07 and LR- 0.55, 95% CI 0.35 to 0.87), and decreased sensation (LR+ 3.99, 95% CI 1.80 to 8.85 and LR- 0.78, 95% CI 0.60 to 1.01). All patients who had bladder or suprapubic fullness had serious spinal pathology (specificity 1.00, 95% CI 0.99 to 1.00; LR+ not calculable) but absence of bladder or suprapubic fullness was not informative for ruling out serious spinal pathology (LR- 0.95, 95% CI 0.87 to 1.05).

One additional retrospective cohort study (N=1000) not included in the SR evaluated factors for identifying any serious spinal pathology (fracture, infection, cancer, or inflammatory condition; prevalence 3.3%) or any serious spinal or nonspinal (fracture, infection, cancer, inflammatory, visceral or systemic pathology; prevalence 17.9%) in patients presenting in ED settings (Appendix E). The study was assessed as moderate risk of bias (Appendix F). In addition to the index tests (factors) not being assessed systematically in all patients, other methodological limitations including use of a suboptimal, nonstandardized reference standard (based on primary diagnosis, discharge code, and discharge summary) and failure to blind the reference standard to the factors assessed.

In this study, the most informative factor for identifying presence of any serious spinal pathology was the presence of saddle anesthesia (LR+ 11.0, 95% CI 3.1 to 39.6) (Table 15). Other informative factors for were a history of tuberculosis (LR+ 9.8, 95% CI 1.0 to 91.4); history of intravenous drug use (LR+ 6.9, 95% CI 2.5 to 19.4); acute onset urinary retention (LR+ 6.4, 95% CI 2.6 to 15.7); loss of anal sphincter tone (LR+ 6.3, 95% CI 19.9 to 20.8); prolonged corticosteroids (LR+ 3.9, 95% CI 1.5 to 10.5); and being immune suppressed (LR+

3.0, 95% CI 1.1 to 7.9). Factors that were less informative (likelihood ratios <3) were history of significant trauma (LR+ 2.9, 95% CI 1.6 to 5.3); presence of urinary symptoms (LR+ 2.4, 95% CI 1.4 to 4.3); insidious onset (LR+ 2.2, 95% CI 0.5 to 8.8); anticoagulant use (LR+ 2.0, 95% CI 0.6 to 6.0); fever (LR+ 2.0, 95% CI 0.3 to 14.4); presence of central spinal tenderness (LR+ 2.0, 95% CI 0.9 to 4.1); previous history of cancer (LR+ 1.9, 95% CI 0.8 to 4.5); age >70 years (LR+ 1.9, 95% CI 1.3 to 2.8); widespread of progressive motor weakness (LR+ 1.8, 95% CI 0.3 to 13.4); being systemically unwell (LR+ 1.7, 95% CI 1.1 to 2.6); or having known nephrolithiasis or an abdominal aortic aneurysm (LR+ 1.4, 95% CI 0.3 to 5.4). However, some LR estimates were imprecise. None of these factors were useful for ruling out serious spinal pathology (LR- ranged from 0.8 to 1.0). Other factors assessed in the study were not useful for identifying patients with any serious spinal pathology, including a number of factors that were not present in any person (sensitivity 0.0) with serious spinal pathology (gradual onset before age 40, no relief with bed rest, morning back stiffness, iritis skin rashes, unexplained weight loss, being pregnant, known spinal Paget's disease, writhing in pain, presence of herpes zoster rash, or diffuse sensory level).

For identifying any serious spinal or nonspinal pathology, the most informative factor was the presence of fever (LR+ 68.8, 95% CI 9.2 to 517.5). Other informative factors were: history of tuberculosis (LR+ 13.8, 95% CI 1.4 to 131.5); known nephrolithiasis or abdominal aortic aneurysm (LR+ 10.2, 95% CI 5.5 to 18.7); unexplained weight loss (LR+ 9.2, 95% CI 0.8 to 100.6); writhing in pain (LR+ 6.9, 95% CI 2.5 to 19.9); presence of urinary symptoms (LR+ 5.4, 95% CI 3.9 to 7.5); presence of flank pain (LR+ 5.2, 95 5CI 3.8 to 7.10); presence of iritis, skin rash (psoriasis), colitis, and urethral discharge (LR+ 4.6, 95% CI 0.9 to 22.6); recent bacterial infection (LR+ 4.4, 95% CI 2.5 to 7.7); main pain is chest, thoracic, or abdominal pain (LR+ 4.3, 95% CI 2.6 to 7.0) and being systemically unwell (LR+ 3.9, 95% CI 3.2 to 4.8). Factors that were less informative (LR+ <3.0) were presence of constant progressive nonmechanical pain (LR+ 2.8, 95% CI 0.7 to 11.4); being immune suppressed (LR+ 2.7, 95% CI 1.5 to 4.9); pregnancy (LR+ 2.3, 95% CI 0.2 to 25.2); a history of cancer (LR+ 2.3, 95% CI 1.5 to 3.5); history of intravenous drug use (LR+ 2.3, 95% CI 0.9 to 5.6); history of inflammatory arthritis or osteoporosis or crush fracture (LR+ 2.3, 95% CI 1.5 to 3.7); prolonged corticosteroid use (LR+ 2.2, 95% CI 1.1 to 4.4); acute onset urinary retention or overflow incontinence (LR+ 2.2, 95% CI 1.0 to 4.7); anticoagulant use (LR+ 1.9, 95% CI 1.0 to 3.4); significant trauma (LR+ 1.8, 95% CI 1.2 to 2.7); presence of saddle anesthesia (LR+ 1.7, 95% CI 0.5 to 6.4); and age over 70 years (LR+ 1.4, 95% CI 1.1 to 1.8). Factors that were modestly informative for ruling out any serious spinal or nonspinal pathology were: not being systemically unwell (LR- 0.5, 95% CI 0.4 to 0.6); absence of urinary symptoms (LR- 0.7, 95% CI 0.6 to 0.8); and absence of flank pain (LR- 0.7, 95% CI 0.5 to 0.8). Otherwise, no factor was informative for ruling out any serious spinal or nonspinal pathology (likelihood ratios ranged from 0.8 to 1.0). Other factors, including widespread or progressive motor weakness in legs or gait disturbances; loss of anal sphincter tone or fecal incontinence; central spinal tenderness; recent spinal procedure or surgery; spinal instrumentation or indwelling device; insidious onset; altered sensation below the trunk; known spinal Paget's disease; Herpes zoster rash; gradual onset before age 40; no relief with bedrest; and morning back stiffness ≥30 minutes, were not informative for identifying presence of any serious spinal or nonspinal pathology.

Table 15. Summary table: any serious spinal pathology

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Pain duration >365 days	0.09 (0.01−0.29)	0.97 (0.94−0.99)	3.1 (0.71−13.49)	0.94 (0.82−1.07)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Pain that awakens patients at night	0.14 (0.03−0.35)	0.94 (0.90−0.96)	2.20 (0.71−6.88)	0.92 (0.78−1.09)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Pain that is worse at night	0.14 (0.03−0.35)	0.96 (0.93−0.98)	3.22 (0.99−10.46)	0.90 (0.76−1.07)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Current sensory symptoms	0.36 (0.17−0.59)	0.82 (0.77−0.86)	2.03 (1.11−3.71)	0.78 (0.56−1.07)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	New urinary retention	0.14 (0.03−0.35)	0.98 (0.96−0.99)	6.98 (1.87−26.03)	0.88 (0.75−1.04)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	New fecal incontinence	0.05 (0.01−0.23)	0.99 (0.97−0.998)	4.65 (0.50−42.89)	0.96 (0.88−1.06)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Procedure causing bacteremia	0.09 (0.01−0.29)	0.97 (0.95−0.99)	3.49 (0.79−15.45)	0.93 (0.82−1.07)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	History of cancer	0.14 (0.03−0.35)	0.94 (0.91−0.97)	2.46 (0.78−7.77)	0.91 (0.77−1.08)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	History of osteoporosis	0.05 (0.001−0.23)	0.98 (0.96−0.99)	2.79 (0.34−22.86)	0.97 (0.88−1.06)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	On anticoagulants	0.23 (0.08−0.45)	0.97 (0.95−0.99)	8.72 (3.11−24.44)	LR- 0.79 (0.63−1.00)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Age >65 years	0.55 (0.32−0.76)	0.83 (0.78−0.87)	3.22 (2.04−5.07)	0.55 (0.35−0.87)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Age >75 years	0.41 (0.21−0.64)	0.89 (0.85−0.92)	3.69 (2.04−6.69)	0.66 (0.47−0.94)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Fever (≥38 degrees C)	0.05 (0.001−0.23)	0.97 (0.95−0.99)	1.55 (0.21−11.69)	0.98 (0.90−1.08)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Saddle sensation disturbance	0.09 (0.01−0.29)	0.99 (0.97−0.996)	6.98 (1.35−36.02)	0.92 (0.81−1.05)

Study, Year	Condition Setting	Number of Studies (N)	Factor/ Index Test	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Decreased sensation	0.27 (0.11−0.50)	0.93 (0.90−0.96)	3.99 (1.80−8.85)	0.78 (0.60−1.01)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Bladder/suprapubic fullness	0.05 (0.001−0.23)	1.00 (0.99−1.00)	not calculable	0.95 (0.87−1.05)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Hemoglobin <100 g/L	0.09 (0.01−0.29)	0.99 (0.98−0.999)	13.95 (2.06−94.39)	0.92 (0.80−1.04)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	Hemoglobin <110 g/L	0.18 (0.05−0.40)	0.98 (0.96−0.99)	11.16 (3.23−38.64)	0.83 (0.68−1.01)
Galliker, 2020	Any serious spinal pathology (ED setting)	1 (N=329)	INR ≥1	0.32 (0.14−0.55)	0.95 (0.92−0.97)	6.11 (2.81−13.26)	0.72 (0.54−0.96)

CI = confidence interval; ED = emergency department; LR = likelihood ratio.

Key Question 1b. What is the association between a focused patient history and physical exam (or components of these) and the likelihood of chronic low back pain and long-term disability?

Description of Included Studies

The original review included 23 studies with 10,842 patients (Appendix H, Table H-3). Fourteen studies[202-217] (reported in 16 articles) evaluated individual risk factors and 10 studies evaluated risk prediction instruments.[214,218-226] One study[214,219,225] evaluated both individual risk factors as well as a risk prediction instrument and in three cases, data from the same population were used to evaluate both individual risk factors[203,206,217,225] as well as a risk prediction instrument in separate publications.[222,225,226] Risk prediction instruments evaluated in the original review included the Vermont Disability Prediction Questionnaire[218,223] (2 studies), an earlier version of the Vermont Disability Prediction Questionnaire[220] (1 study), and the Acute Low Back Pain Screening Questionnaire (2 studies); other risk prediction instruments were evaluated in 1 study each. For six studies in the original review, we obtained unpublished data to calculate likelihood ratios.[203,209,213,215,216,225] Duration of follow-up ranged from 3 months to 2 years.

For this update, we added 14 studies[170,173,174,177-180,183,185-187,189-191] with 9,162 patients (Appendix H, Table H-3). Three studies[178,179,185] evaluated individual risk factors (sex, race/ethnicity, and lack of early improvement) and 11 studies[170,173,174,177,180,183,186,187,189-191] evaluated risk prediction instruments (the full or short-form Örebro Musculoskeletal Pain Screening Questionnaire[183,186,187,189-191] [6 studies], the STarT Back instrument or a modified version of it[180,227] [2 studies], a 3-item instrument[173,177] [2 studies], and the Low Back Pain Perceptions scale[189] [1 study]; one additional study evaluated 2 slightly different 8-item instruments for 6 month and 2 year outcomes).[173] In two cases, the same population was used to evaluate different risk prediction instruments in separate publications.[173,174,190,191] Duration of follow-up ranged from 12 weeks to 2 years.

The degree to which the 37 studies met risk of bias criteria varied (Appendix H, Table H-4). No study met all criteria. The most common methodological shortcomings were failure to describe assessment of outcomes blinded to assessment of potential yellow flags (only one study met this criterion[216,217]) and failure to report that interventions were avoided, similar across groups, or allocated randomly (eight studies met this criterion[170,179,180,185,202,211,212,221]). Of the 21 studies evaluating risk prediction instruments, 12[173,177,180,183,186,187,189-191,222,223,225] evaluated an instrument in a population other than the one used to develop the instrument and 9[170,173,214,218-220,224,226,228] were derivation studies only. The cutoffs used to assess the accuracy of risk prediction instruments were predefined in 8 studies.[170,173,177,180,183,186,187,222] Five studies were judged as overall low risk of bias,[178-180,202,212] 20 as overall moderate risk of bias,[170,177,185,186,189,203-211,217,222-224,226,228] and 12[173,174,183,187,190,191,213-215,218-220] as overall high risk of bias.

Seven studies evaluated patients in workers' compensation settings and the others were conducted in clinical settings.[205,215,217,218,223,224,226] The most common clinical setting was primary care[170,174,180,183,185,187,189,202,204,206,209,211-214,219,225,228] [17 studies]; other clinical settings were EDs,[177,179] hospital,[186] physical therapy clinics,[190,191,208] specialty clinics,[210,220] and mixed settings.[203,222] Outcome types evaluated were work disability status[186,187,190,191,203-205,208,211,215,217-220,222-226] (19 studies), pain[170,177,180,189,210,212,213] (7 studies), function[179,206,209] (3 studies), and overall satisfaction[202] (1 study); five studies[173,174,183,185,214,228] evaluated mixed (multicomponent) or multiple outcomes. The proportion of patients experiencing poor outcomes varied, due to variability in patient populations, settings, outcome definitions, duration of follow-up, and other factors. In studies conducted in clinical settings that reported on work absenteeism or compensation status, the median proportion of patients with a poor outcome was 20% at 3 to 6 months[180,203,206,208,219,220,228] (7 studies, range 8% to 28%) and 21% at ≥1 year[186,187,190,191,203,204,222,225] (8 studies, range 9% to 23%). In studies conducted in workers' compensation settings that reported on similar outcomes the median proportion was 24% at 3 to 6 months[205,215,217] (3 studies, range 18% to 42%) and 14% at ≥1 year[217,224,226] (3 studies, range 14% to 26%). In studies conducted in clinical settings that reported on pain, function, or other clinical outcomes, the median proportion of patients with a poor outcome at 3-6 months was 27% (12 studies, range 11% to 48%) and 21% at ≥1 year (11 studies, range 17% to 42%). Two studies conducted in workers' compensation settings reported that the proportion of patients with a poor clinical outcome at 3 to 6 months was 10% and 12%;[218,223] no study in workers' compensation settings reported the proportion of patients with a poor clinical outcome at ≥1 year.

Individual Risk Factors

Demographic and Work-Related Features

Two studies added in this update addressed demographic features[178,185] (sex[178,185] [2 studies] and race/ethnicity[185] [1 study]). The original review found that age, sex, education level, smoking status (described as either "ever smoked" or "current smoker") and overweight (defined by BMI) consistently failed to predict worse outcomes, with likelihood ratios approaching 1 in most or all studies. The two new studies did not change findings regarding sex as a predictor.

No new study evaluated work-related features. The original review found receiving compensation at baseline was associated with slightly increased likelihood of worse outcomes at 1 year (median LR for receiving compensation 1.4 [range 1.2-1.8]). Higher work dissatisfaction and higher physical work demands did not predict worse outcomes at 3 months, but did at 1 year (median LR+ 1.5 [range 1.3-1.8] and median LR+ 1.4 [range 1.2-1.7], respectively).

No study in the original review evaluated race/ethnicity. One study added in this update found Black race increased the likelihood of high impact LBP at 6 months (LR 2.1, 95% CI 1.8 to 2.4) and Hispanic ethnicity slightly increased the likelihood (LR 1.5, 95% CI 1.1 to 2.1); White race was associated with slightly decreased likelihood[185] (LR 0.85, 95% CI 0.82 to 0.89). Results were similar for persistent pain that was not necessarily high impact, though LR estimates were not as pronounced (Table 16, Appendix H, Table H-5).

Table 16. Summary accuracy of demographic variables to predict chronic disabling low back pain

Studies	Number of Studies	Timing of Outcome Assessment	Median LR+ (Range)*	Median LR- (Range)*
Age				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005,[206] Schiottz-Christensen 1999,[211] Truchon 2005,[215] Turner 2006[217]	6	3 to 6 months	≤40, <45, or <46 years old 0.94 (0.74-1.1)	Older 1.1 (0.81-2.0)
Dionne 2007,[203] Grotle 2006,[207] Henschke 2008,[209] Schiottz-Christensen 1999,[211] Thomas 1999,[214] Turner 2008[216]	6	1 year	≤40, <45, or <50 years old 0.93 (0.62-1.0)	Older 1.1 (0.99-1.8)
Sex				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005,[206] Heneweer 2007,[208] Poiraudeau 2006,[210] Roseen 2023,[185] Schiottz-Christensen 1999,[211] Swinkels-Meewisse 2006,[213] Turner 2006,[217] Truchon 2005[215]	10	3 to 6 months	Female 1.1 (0.72-1.4)	Male 0.96 (0.66-1.3)
Cherkin 1996,[202] Dionne 2007,[203] Dixon 1999,[204] Grotle 2006,[207] Henschke 2008,[209] Pozzobon 2019,[178] Schiottz-Christensen 1999,[211] Thomas 1999,[214] Turner 2008[216]	9	1 year	Female 1.3 (1.0-1.7)	Male 0.63 (0.58-1.0)
Race/ethnicity				
Roseen 2023[185]	1	6 months	Black: 2.1 (1.8-2.4) Hispanic: 2.6 (1.1-2.1) White: 0.85 (0.82-0.89)	
Education				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005,[206] Poiraudeau 2006,[210] Swinkels-Meewisse 2006,[213] Truchon 2005,[215] Turner 2006[217]	7	3 to 6 months	No college education or not college graduate 1.0 (0.97-1.3)	More education 0.76 (0.52-1.1)
Dionne 2007,[203] Grotle 2006,[207] Henschke 2008,[209] Turner 2008[216]	4	1 year	No college education or not college graduate 1.1 (1.1-1.2)	More education 0.65 (0.46-0.85)
Smoking status				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005[206]	3	3 to 6 months	Current smoker 1.2 (1.0-1.6)	Not current smoker 0.88 (0.71-0.97)

Studies	Number of Studies	Timing of Outcome Assessment	Median LR+ (Range)*	Median LR- (Range)*
Dionne 2007,[203] Grotle 2006,[207] Henschke 2008,[209] Thomas 1999[214]	4	1 year	Cannot estimate because of overlap in case definitions (current smoker or ever smoked)	
Weight				
Dionne 2007,[203] Fransen 2002,[205] Schiottz-Christensen 1999[211]	3	3 to 6 months	BMI >25 or ≥27 0.91 (0.72-1.2)	Lower BMI 1.0 (0.76-1.2)
Dionne 2007,[203] Schiottz-Christensen 1999[211]	2	1 year	BMI >25 or ≥27 0.84 (0.73-0.97)	Lower BMI 1.1 (1.0-1.2)
Sick leave, off work, or workers' compensation case				
Dionne 2007,[203] Grotle 2005,[206] Schiottz-Christensen 1999,[211] Swinkels-Meewisse 2006[213]	4	3 to 6 months	Compensated work injury or sick leave 1.3 (0.97-2.7)	Not compensated work injury or sick leave 0.88 (0.78-1.0)
Dionne 2007,[203] Dixon 1999,[204] Grotle 2006,[207] Henschke 2008,[209] Schiottz-Christensen 1999[211]	5	1 year	Compensated work injury or seeking compensation 1.4 (1.2-1.8)	Not compensated or seeking compensation 0.86 (0.37-0.93)
Work satisfaction				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005[206]	3	3 to 6 months	Less work satisfaction 1.1 (0.64-1.8)	More work satisfaction 0.98 (0.94-1.2)
Dionne 2007,[203] Grotle 2006,[207] Thomas 1999[214]	3	1 year	Less work satisfaction 1.5 (1.3-1.8)	More work satisfaction 0.88 (0.62-0.94)
Physical work demands				
Dionne 2007,[203] Fransen 2002,[205] Poiraudeau 2006[210]	3	3 to 6 months	Higher physical work demands 1.2 (1.1-1.6)	Lower physical work demands 0.87 (0.85-0.89)
Dionne 2007,[203] Turner 2008[216]	2	1 year	Higher physical work demands 1.4 (1.2-1.7)	Lower physical work demands 0.84 (0.83-0.85)

BMI = body mass index; LR = likelihood ratio.
* See Appendix H, Table H-5 for estimates from individual studies.

Health Status at Onset of Back Pain

No study added in this update assessed health status at onset of back pain as a predictor of worse outcomes. The original review found worse general health status[202,205,214] prior to the onset of pain was associated with worse outcomes at 3 to 6 months (median LR for lower general health status 1.6 [range 1.1-1.7]; median LR for better health status 0.73 [range 0.66-0.88]) and at 1 year (median LR+ 1.8 [range 1.1-2.0]; median LR- 0.85 [range 0.56-0.99]). The presence of higher scores for current psychiatric comorbidity (measured using various scales) had an effect somewhat stronger than poorer overall general health at 3 to 6 months [median LR for higher psychiatric comorbidity 1.9 (range 1.4-2.1); median LR for lower psychiatric comorbidity 0.69 (range 0.55-0.85)] and at 1 year (median LR+ 2.2 [range 1.9-2.3]; median LR- 0.85 [range 0.55-0.93]).[202,203,205,209,215,217] A history of previous or recurrent episodes of LBP was not useful for predicting worse outcomes at 3 to 6 months or 1 year, with likelihood ratios approaching 1 (Table 17; Appendix H, Table H-6).

Table 17. Summary accuracy of general health, psychiatric comorbidities, and prior low back pain episodes for predicting chronic disabling low back pain

Studies	Number of Studies	Timing of Outcome Assessment	LR+ (95% CI)*	LR- (95% CI)*
General health or activity level				
Dionne 2007,[203] Fransen 2002,[205] Poiraudeau 2006[210]	3	3 to 6 months	Lower health status 1.6 (1.1 to 1.7)	Better health status 0.73 (0.66 to 0.88)
Cherkin 1996,[202] Dionne 2007,[203] Henschke 2008[209] Thomas 1999,[214] Turner 2008[216]	5	1 year	Lower health status 1.8 (1.1 to 2.0)	Better health status 0.85 (0.56 to 0.99)
Psychiatric comorbidities				
Dionne 2007,[203] Fransen 2002,[205] Truchon 2005,[215] Turner 2006[217]	4	3 to 6 months	Higher score on psychiatric comorbidity scale 1.9 (1.4 to 2.1)	Lower score 0.69 (0.55 to 0.85)
Cherkin 1996,[202] Dionne 2007,[203] Henschke 2008,[209] Turner 2008[216]	4	1 year	Higher score on psychiatric comorbidity scale 2.2 (1.9 to 2.3)	Lower score 0.85 (0.55 to 0.93)
Prior low back pain episodes				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005,[206] Heneweer 2007,[208] Singer 1987,[212] Swinkels-Meewisse 2006[213]	6	3 to 6 months	More episodes of prior back pain 1.0 (0.90 to 1.2)	Less or no prior back pain 0.88 (0.53 to 1.1)
Dionne 2007,[203] Grotle 2006,[207] Henschke 2008,[209] Singer 1987,[212] Thomas 1999[214]	5	1 year	More episodes of prior back pain 1.1 (0.95 to 1.2)	Less or no prior back pain 0.81 (0.32 to 1.1)

LR = likelihood ratio.
* See Appendix H, Table H-6 for estimates from individual studies.

Symptoms

Pain

No study added in this update evaluated pain intensity as a predictor of worse outcomes. The original review found that high pain intensity predicted worse outcomes at 3 to 6 months (median LR 1.7, range 1.1-3.7), but was a less useful predictor at 1 year (median LR 1.3, range 1.2-2.0). Low pain intensity was associated with a broad range of likelihood ratios at both 3 to 6 months (median LR 0.70, range 0.07-0.86) and at 1 year (median LR 0.33, range 0.08-0.97) (Table 18; Appendix H, Table H-7).

Function

Baseline functional impairment was measured with the Roland Disability Questionnaire,[203,215-217] the Oswestry Disability Index,[205] and various ordinal scales.[209,210] No study added in the update evaluated baseline function as a predictor. In the original review, results using the various measures were categorized into 3 levels of functional impairment ("none/weak", "moderate", and "severe/extreme"). Greater baseline functional impairment progressively increased likelihood of poor outcomes from the lowest functional impairment (median LR 0.52 [range 0.18-0.59]) at 3 to 6 months and median LR 0.40 [range 0.10-0.52] at 1

year) to the highest functional impairment (median LR 1.4 [range 1.3-3.5] at 3 to 6 months and median LR 2.1 [range 1.2-2.7] at 1 year). The Roland-Morris Disability Questionnaire (RDQ) was the most frequently used measure of function[203,213,216,217,220] (Table 18; Appendix H, Table H-7).

Radiculopathy

No study added in this update evaluated radiculopathy as a predictor. The original review found presence of radiculopathy or leg pain slightly increased the likelihood of worse outcomes at 3 to 6 months (median LR 1.4 [range 1.1-1.7]) and at 1 year (median LR 1.4 [range 1.2-2.4]) while its absence slightly reduced the likelihood (median LR 0.63 [range 0.52-0.93] at 3 to 6 months and median LR 0.82 [range 0.54-0.94]) at 1 year) (Table 18; Appendix H, Table H-7).

Maladaptive Pain Coping Behaviors

Maladaptive pain coping behaviors include fear avoidance (avoidance of work, movement or other activities due to fear that they will damage or worsen the back) and catastrophizing (pain coping characterized by excessively negative thoughts and statements about the future[229]). No study added in this update evaluated maladaptive pain coping behaviors as a predictor. The original review found that patients with high maladaptive coping behaviors, as measured by several scales (such as the Fear Avoidance Beliefs Questionnaire [FABQ]),[230] were more likely to have worse outcomes at 3 to 6 months (median LR 2.2 [range 1.5-4.9] and 1 year (median LR 2.5 [range 2.2-2.8]). Patients in the low category using these scales were less likely to have worse outcomes (median LR 0.46 [range 0.30-0.73] at 3 to 6 months and median LR 0.39 [range 0.38-0.40] at 1 year) (Table 18; Appendix H, Table H-7).

Lack of Early Improvement

One study added in this update found presence of LBP at 1 week modestly predicted functional impairment at 3 months (LR+ 1.5, 95% CI 1.3 to 1.7 and LR- 0.38, 95% CI 0.23 to 0.62) (Table 18; Appendix H, Table H-7).[179]

Signs

Nonorganic Signs

Nonorganic signs refer to findings that suggest a strong psychological component to pain, or intentionally false or exaggerated pain symptoms. No study added in the update evaluated nonorganic signs as a predictor. The original review found higher somatization scores predicted failure to return to work at 3 months (LR 2.5 [95% CI 1.8 to 3.4]).[203] Similarly, higher somatization scores or more generalized pain at baseline increased the likelihood of a worse outcome at 1 year (median LR 3.0 [range 1.7 to 4.6]), but the usefulness of lower scores was variable (median LR 0.71 [range 0.31 to 0.76]) (Table 18; Appendix H, Table H-7). No study reported likelihood ratios for Waddell's signs, possibly the best known of the nonorganic signs (Table 19).

Table 18. Summary accuracy of signs and symptoms for predicting chronic disabling low back pain

Studies	Number of Studies	Timing of Outcome Assessment	LR Median (Range)*	
Baseline pain				
Fransen 2002,[205] Poiraudeau 2006,[210] Singer 1987,[212] Swinkels-Meewisse 2006,[213] Truchon 2005,[215] Turner 2006[217]	6	3-6 months	Intensity of pain High 1.7 (1.1-3.7) Medium 0.86 (0.66-1.2) Low 0.70 (0.07-0.86)	
Henschke 2008,[209] Singer 1987,[212] Turner 2008[216]	3	1 year	Intensity of pain High 1.3 (1.2-2.0) Medium 0.78 (0.72-1.0) Low 0.33 (0.08-0.97)	
Baseline function				
Dionne 2007,[203] Fransen 2002,[205] Poiraudeau 2006,[210] Swinkels-Meewisse 2006,[213] Truchon 2005,[215] Turner 2006[217]	6	3 to 6 months	Intensity of impairment High 1.4 (1.3-3.5) Medium 1.3 (0.74-1.5) Low 0.53 (0.18-1.1)	
Dionne 2007,[203] Henschke 2008,[209] Turner 2008[216]	3	1 year	Intensity of impairment High 2.1 (1.2-2.7) Medium 0.86 (0.85-1.7) Low 0.40 (0.10-0.52)	
Radiculopathy				
Dionne 2007,[203] Fransen 2002,[205] Grotle 2005,[206] Schiottz-Christensen 1999,[211] Swinkels-Meewisse 2006[213]	5	3 to 6 months	Leg pain or radiculopathy 1.4 (1.1-1.7)	No leg pain or radiculopathy 0.63 (0.52-0.93)
Cherkin 1996,[202] Dionne 2007,[203] Grotle 2006,[207] Henschke 2008,[209] Schiottz-Christensen 1999,[211] Thomas 1999,[214] Turner 2008[216]	7	1 year	Leg pain or radiculopathy 1.4 (1.2-2.4)	No leg pain or radiculopathy 0.82 (0.54-0.94)
Fear avoidance behaviors or coping strategies				
Dionne 2007,[203] Swinkels-Meewisse 2006,[213] Truchon 2005,[215] Turner 2006[217]	4	3 to 6 months	Intensity of fear avoidance High 2.2 (1.5-4.9) Medium 1.1 (1.0-1.5)Low 0.46 (0.30-0.73)	
Dionne 2007,[203] Turner 2008[216]	2	1 year	Intensity of fear avoidance High 2.5 (2.2-2.8) Medium 1.2 (1.2-1.3) Low 0.39 (0.38-0.40)	
Lack of early improvement				
Friedman 2017[179]	1	3 months	LBP present at 1 week 1.5 (1.3-1.7)	LBP not present at 1 week 0.38 (0.23-0.62)
Nonorganic signs or somatization				
Dionne 2007[203]	1	3 months	More somatization 2.5 (95% CI 1.8-3.4)	Less somatization 0.81 (95% CI 0.74-0.89)

Studies	Number of Studies	Timing of Outcome Assessment	LR Median (Range)*	
Dionne 2007,[203] Thomas 1999,[214] Turner 2008[216]	3	1 year	More widespread pain or somatization 3.0 (1.7-4.6)	Less widespread pain or somatization 0.71 (0.31-0.76)

LBP = low back pain; LR = likelihood ratio.
* See Appendix H, Table H-7 for estimates from individual studies.

Table 19. Waddell's nonorganic signs[231]

Sign	Definition
Superficial or nonanatomic tenderness	Pain with light or superficial palpation of the skin, or widespread deep tenderness that is not localized to one structure, does not follow an anatomic distribution, and often extends to the thoracic spine, sacrum, or pelvis.
Pain on axial loading or simulated rotation	Pain with vertical pressure applied to the standing patient's head, or back pain when the shoulders and pelvis are passively rotated in the same plane as the patient stands relaxed with the feet together.
Nonreproducibility of pain when patient is distracted	A positive physical finding is found on routine formal assessment, but is not present when the patient is distracted and it is checked again, e.g., pain with a standard straight leg raise test (passive flexion at the hip while the patient is lying on the back with knee extended), but not when the examiner passively extends the leg while the patient is sitting and plantar reflex is being checked.
Regional weakness or sensory change	Regional sensory change (stocking sensory loss, or sensory loss in an entire extremity or side of the body) or regional weakness (weakness that is jerky, with intermittent resistance such as cogwheeling or catching). Organic weakness can be overpowered smoothly.
Overreaction	Exaggerated painful response to a stimulus that is not reproduced when the same stimulus is given later, or exaggerated response to a stimulus that should not cause back pain (for example, gently pinching the skin on the back in the area of pain)

Stratified Analyses

As in the original review, conclusions appeared similar in subgroups of studies stratified according to whether they evaluated a workers' or nonworkers' compensation population, whether they evaluated return-to-work versus a nonwork functional outcome, and whether they evaluated patients with acute or subacute LBP. There was no pattern suggesting that results of studies meeting various risk of bias criteria differed from those that did not meet the criteria. In one study that reported LR estimates for 2-year outcomes, results were similar to estimates based on 1-year outcomes.[203]

Risk Assessment Instruments

Details related to the risk assessment instruments described below can be found in Appendix H, Table H-8.

Vermont Disability Prediction Questionnaire

No study added in the update evaluated the Vermont Disability Prediction Questionnaire (VDPQ) or related instruments. The original review included the initial derivation study for an instrument developed at the Vermont Rehabilitation Engineering Center that found it to be highly useful for predicting nonreturn to work at 6 months (LR+ 8.3, 95% CI 5.2 to 13 and LR- 0.17, 95% CI 0.10 to 0.31), based on a cut-off score of 0.5 (0 to 1 scale).[220] A modified version of this instrument, the 11-item VDPQ, was evaluated in one derivation[218] and one validation[223] study. The derivation study found higher scores on the VDPQ increased the likelihood of not returning to work at three months (LR 5.7, 95% CI 3.9 to 8.5, at a cutoff of 0.48 and LR 7.1,

95% CI 3.7 to 13, at a cutoff of 0.65),[218] but in the validation study the VDPQ was not as useful, with the exception of very high (>0.76) scores (LR 7.4, 95% CI 2.8 to 20).[223] Some items in the VDPQ (e.g., previous back problems and doctor visit for back pain) did not predict outcomes in other studies or have not been studied well.

Acute Low Back Pain Screening Questionnaire/Örebro Musculoskeletal Pain Screening Questionnaire

The Acute Low Back Pain Screening Questionnaire (ALBPSQ) is a patient self-administered risk prediction instrument consisting of 21 scored items.[219] The Örebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ) is a modified version of the ALBPSQ, also consisting of 21 scored items, in which one item from the ALBPSQ was modified to address presence of additional nonback pain conditions.[232] A short form (10 item) version of the ÖMPSQ consisting has also been developed, and other versions of the full instrument exist in multiple languages.[187]

Two studies included in the original review evaluated the ALBPSQ.[219,225] In the initial derivation study, higher scores on the ALBPSQ increased the likelihood of >30 days off work through 6 months at various cutoffs from ≥90 to ≥120 (LR+ 2.2-3.9).[219] For decreasing the likelihood of negative work outcomes, having scores below cutoffs of 90 to 105 (LR- 0.06-0.14) was more useful than higher cutoffs (LR- 0.29-0.58). Similar results for predicting worse work outcomes at 6 months were obtained in a subsequent validation study using a Norwegian version of the ALBPSQ (LR+ 4.8, 95% CI 2.0 to 11, for a cutoff score of ≥105; LR+ 7.7, 95% CI 2.8 to 21, for a cutoff score of ≥112), though LR- were less useful (LR 0.62 and 0.60, respectively). In this study, the ALBPSQ was less useful for predicting days off work through 12 months (LR+ 2.1-2.3 and LR- 0.66-0.87 at various cutoffs, with confidence intervals approaching or crossing 1).[225]

Three studies added in the update evaluated the full ÖMPSQ.[186,187,189] Results were generally consistent with results for the ALBPSQ. In one study, higher scores on the ÖMPSQ progressively increased the likelihood of experiencing a suboptimal course of LBP (slightly improved or worse at 2 or more follow-up assessments; LR 1.5 for a cutoff score of ≥90 and LR 2.6 for a cutoff score of >105).[189] However, having scores below these cutoffs only slightly decreased the likelihood of a suboptimal course (LR- 0.73-0.80). A study using a Chinese version of the full ÖMPSQ found a cutoff of >108 moderately useful for not returning to full or part-time work at 1 year (LR+ 1.8, 95% CI 1.4 to 2.2 and LR- 0.43, 95% CI 0.26-0 to 72).[186] The third study evaluated the New Zealand Yellow Flags Screening Instrument, an early version of the ÖMPSQ.[187] A score in the highest category (>105) increased the likelihood of having LBP claims for >90 days (LR 2.8, 95% CI 1.6 to 4.8) and a score in the lowest category (<50) decreased the likelihood (LR 0.39, 95% CI 0.19 to 0.82).

Two studies added in the update evaluated the short form ÖMPSQ.[183,191] One study found a Brazilian version of the short form ÖMPSQ highly useful for predicting functional disability (RDQ ≥15) at 6 months, using a cutoff of >45 (LR+ 8.7 and LR- 0.04, 95% CI not reported).[183] The second study also found a Chinese version of the short form ÖMPSQ useful, although less predictive, for not returning to full or part-time work at 1 year, using a cutoff of >54 (LR+ 1.8, 95% CI 1.4 to 2.2 and LR- 0.43, 95% CI 0.26 to 0.72).[191] The utility of the short form ÖMPSQ was very similar to estimates from the study described above that applied the full ÖMPSQ in the same population.[190]

STarT Back Instrument

The STarT Back instrument is a 9-item patient self-administered risk prediction instrument that uses results to categorize patients as high, medium, or low risk.[233] One study[177] added for the update evaluated the STarT Back Instrument and a second study[173] evaluated a modified version with slight wording differences for 3 items. In both studies, categorization as high, medium, or low risk was not useful for predicting nonrecovery, persistent disability, or persistent pain at 6 months; one of the studies[173] reported similar results for nonrecovery and persistent disability at 2 years. In both studies, being in a higher risk category did not increase likelihood of worse outcomes compared to being in a lower risk category, or LR estimates were close to 1 with confidence intervals that crossed 1.

3-Item Risk Prediction Instrument

Two studies added in this update evaluated a 3-item risk prediction instrument (baseline pain >7, duration of current episode >5 days, and at least one previous episode; score based on number of positive items).[170,180] Both assessed outcomes at 12 weeks. In the derivation study, a higher score increased the likelihood of nonrecovery (LR 0.39, 95% CI 0.05 to 2.8, for score 0, and LR 3.5, 95% CI 1.5 to 8.3, for score 3).[170] A subsequent validation study reported similar results (LR 0.27, 95% CI 0.14-0.50 for score 0 and LR 2.2, 95% CI 1.3-3.5 for score 3).[180]

Other Risk Prediction Instruments

The original review included derivation studies for three other risk prediction instruments. Deyo et al found that the presence of negative responses to all 3 items of an instrument (previous episodes, feel sick all the time, and education level) to be more useful for predicting no improvement in Sickness Impact Profile scores (LR 0.10, 95% CI 0.01 to 0.71) than 1 to 3 positive responses (LR 0.72 to 2.9).[221] Fulton-Kehoe et al evaluated a different three-item instrument (each scored as one point: pain interference with activities ≥5 on a 0 to 10 scale, not working for pay in last week, and presence of leg pain) and found a score of zero (LR 0.01, 95% CI 0.00 to 0.10) to be more useful than scores of 1 to 3 (LR 0.29 to 3.4) for predicting receipt of wage compensation at 1 year.[226] Thomas et al evaluated a 6-item instrument (each scored as one point: female sex, dissatisfaction with employment situation, history of LBP, radiating leg pain, widespread pain, two or more restrictions in spinal movement) and found both high (5 or 6) and low (0-2) scores increased the likelihood of persistent disabling LBP through 12 months (LR 5.3, 95% CI 2.6 to 11 and LR 0.15, 95% CI 0.05 to 0.45, respectively).[214]

A validation study that evaluated failure to return to work at 2 years found that a clinical prediction algorithm for return-to-work was a somewhat weaker predictor of failure to return to work compared to the other instruments (LR+ 2.0, 95% CI 1.6 to 2.3 and LR- 0.41, 95% CI 0.29 to 0.59).[222] A study of an eight-factor model found patients scoring in higher percentiles on a predictive score had progressively greater likelihood of receiving benefits for >3 months, but a pragmatic method of administering and scoring was not provided.[224]

Two studies added in this update evaluated one other risk prediction instrument each.[173,189] In one derivation study, two slightly different 8-item prediction models were developed to predict nonrecovery at 6 months and at 2 years.[173] At 2 years, being in the highest risk category (score ≥8) increased the likelihood of nonrecovery (LR 3.1, 95% CI 2.1 to 4.8) and being in the lowest risk category (score <-4) decreased the likelihood (LR 0.50, 95% CI 0.34 to 0.72). Similar results were reported for 2-year outcomes. The second study validated the previously developed 5-item Low Back Pain Perceptions Scale.[189] A score of ≥4 modestly increased the likelihood of having a

suboptimal LBP course (LR+ 1.6, CI not reported) but having a score <4 was not useful for decreasing the likelihood (LR 0.86, CI not reported).

Key Question 1c. Does the use of screening tools (e.g., Keele STarT Back Screening Tool) improve short or long-term patient outcomes (versus not using such a tool) based on prognosis?

No studies meeting inclusion criteria were identified.

Key Question 1d. Does accuracy or predictive utility differ based on (1) patient demographics or characteristics including those related to social risk factors for health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., duration, severity, recurrent, radicular, nonradicular); (4) clinical setting, provider type?

No studies meeting inclusion criteria were identified.

Key Question 2. In adults with ALBP (<6 weeks duration), with or without radiculopathy who *do not* present with historical or clinical features suggestive of serious low back problems (e.g., absence of "red flag" symptoms"):

a. **Does immediate/early lumbar spine imaging improve patient health outcomes versus usual care without imaging in the short term or long term?**
b. **What harms are associated with immediate lumbar spine imaging versus usual care without imaging?**
c. **Does outcome differ based on (1) patient demographics or characteristics including those related to social risk factors for health, access to imaging; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, nonradicular); (4) imaging modality used (e.g., radiography, magnetic resonance imaging [MRI], computed tomography [CT]); (5) clinical setting, provider type?**
d. **What is the cost-effectiveness of immediate/early lumbar spine imaging versus usual care without imaging? (full economic studies only)**

Three RCTs[62,92,158,159] (N=488) (in four publications) compared early imaging versus usual care in patients with ALBP and provide the primary evidence base for this question. Three NRSI[175,181,188] (N=4800) provide additional information on the consequences of early MRI in patients with ALBP and will be described separately.

Description of Included RCTs

Three RCTs (4 publications, N=488) were included[62,92,158,159] (Table 20, Appendix E). Participants were an average of 38 years with median and 54% were female. Trials either excluded pregnant and breastfeeding women[158,159] or did not report on this.[62,92] One trial reported

81% of participants as Hispanic,[158] and another reported mostly White (70%) and Black (26%) participants.[62,92] One trial did not report on substance use disorder and two trials excluded participants with a history of alcohol use[158] or intravenous[159] drug use. All trials exclude patients with red flag symptoms or signs of serious LBP requiring urgent care (e.g., history of trauma, spinal infection or fracture, malignant neoplasms, etc.). Radiculopathy at baseline was reported in 61% of patients in one RCT.[62,92] Another trial reported 24% of patients had nerve root irritation and 4% neuromotor deficit.[158] Baseline pain duration across trials was <4 weeks. Two trials reported the proportion of patients with prior LBP episodes: 100%[158] and 66.3%.[159] Baseline pain (VAS 0-10 scale) was mean 5.3 in one trial,[62,92] and median 6 in another.[159]

Imaging consisted of MRI[62,92] and lumbar spine radiography.[158,159] Time between injury and imaging was only reported in one study[62,92] (48 hours). Follow-up ranged from 3 weeks to 1 year. Patients in all trials received usual care; two trials[62,92,158] described usual care interventions which included analgesics, physical therapy and bed rest. No imaging was received in the control group in two trials.[158,159] In the third trial, all patients had early imaging; however, the results were withheld from both the patients and their physicians unless the findings were deemed clinically critical to patient care.[62,92] Patients in the "blinded" group and their physicians were made aware of the imaging results at 6-month follow-up. Across trials, it appears that all patients received some form of usual care.

Two trials were considered moderate risk of bias (ROB),[62,92,158] and one was high ROB.[159] A common limitation was the inability to blind patients to the intervention received.[62,92,158,159] In one trial, patients were blinded to results, but not to whether or not they were imaged.[92] One trial had high attrition (≥20%) and did not perform an intention-to-treat analysis.[62,92] In the high-ROB study, allocation concealment was not specified, and overall attrition was unacceptable[159] (Appendix F).

Table 20. Patient and study characteristics of early imaging versus usual care in trials

Study, Year	Modic, 2005; Ash, 2008	Deyo, 1987	Djais, 2005
N Randomized	246	133	92
Mean Age (Years)	43	33	Median 40
% Female	58%	52%	45%
% Radiculopathy	61%	24% (nerve root irritation) 4% (neuromotor deficit)	NR
Pain Duration Inclusion Criteria	<3 weeks	NR	12 weeks
Mean Pain Duration (days)	NR	14 days	Median 28 days
Mean/Median Baseline Pain Severity (0-10)	5.3	NR	Median 6
History substance abuse disorder	NR	0% (Excluded)	0% (Excluded)
Psychiatric Comorbidities	NR	NR	NR
Imaging*	Unblinded MRI†	Lumbar spine radiography	Lumbar spine radiography
Duration between injury and imaging	48 hours	NR	NR
Control	"Blinded" MRI + UC†	UC (Education)	UC
Duration follow-up	1 year	3 months	3 weeks
Country	United States	United States	Indonesia
Funding	None (Modic, 2005) Government (Ash, 2008)	Foundations	NR
ROB	Moderate	Moderate	High

MRI = magnetic resonance imaging; NR = not reported; ROB = risk of bias; UC = usual care.

* Imaging patients also received usual care.
† All patients received imaging those in the 'unblinded' group were notified of their results within 48 hours. The comparison group (patient and treating physician) was "blinded" and only received results if the results were critical to care or at 6 months.

Description of Nonrandomized Studies

Three nonrandomized studies in workers' compensation populations from U.S. administrative databases reported on consequences of early MRI versus no or delayed MRI.[175,181,188] One prospective study[188] (N=1223) from the Washington State workers' compensation program compared early MRI (within 6 weeks of injury) to MRI performed as consistent with local practices (after 6 weeks or not at all). Two retrospective nonrandomized studies,[175,181] sampled a workers' compensation database representing 10% of the U.S. private workers' compensation market between January 1, 2006 and December 31, 2006; there is likely overlap in the sampled population across these two studies from the same author group (Ns ranged from 555 to 3022). In one study, early MRI was done within 30 days of injury and compared to no MRI within 2-year follow-up.[175] The other study compared MRI done within 30 days post-onset of injury to patients with no MRI within 2 years and to delayed imaging between 42 and 180 days post injury.[181] Patients were predominately male (63% to 73%) and 35 to 54 years old. Mean baseline pain was high (7 to 10 on 0-10 scale) in 42% of patients in the prospective study and 20% of patients had ≥1 prior compensable back claim.[188] None of the studies reported mean pain duration, history of substance abuse disorder or psychiatric comorbidities (Appendix E).

Nonrandomized studies evaluated the impact of early imaging in specific subgroups. The prospective study (N=1223) classified back injuries as mild/major sprains (78%) or radiculopathy 22%).[188] Similarly, one retrospective study (N=555) stratified patients on the presence of nonspecific LBP (60%) or radiculopathy (40%).[175] The third study (N=3022) retrospectively classified cases by severity based on receipt of medical services for specific ICD-9 codes into less severe (degenerative changes, nonspecific back pain, or miscellaneous diagnoses, n=2218) and more severe cases (herniated disc, lumbar radiculopathy or neuropathy, spinal stenosis, sciatica, or possible instability, n=804).[181] Over 80% of those in the severe case group had an early diagnosis of radiculopathy.

All nonrandomized studies were at moderate risk of bias. Common limitations included baseline differences between groups on prognostic factors including opioid use, lack of clarity regarding assessor blinding, and in ability to blind patients (Appendix F).

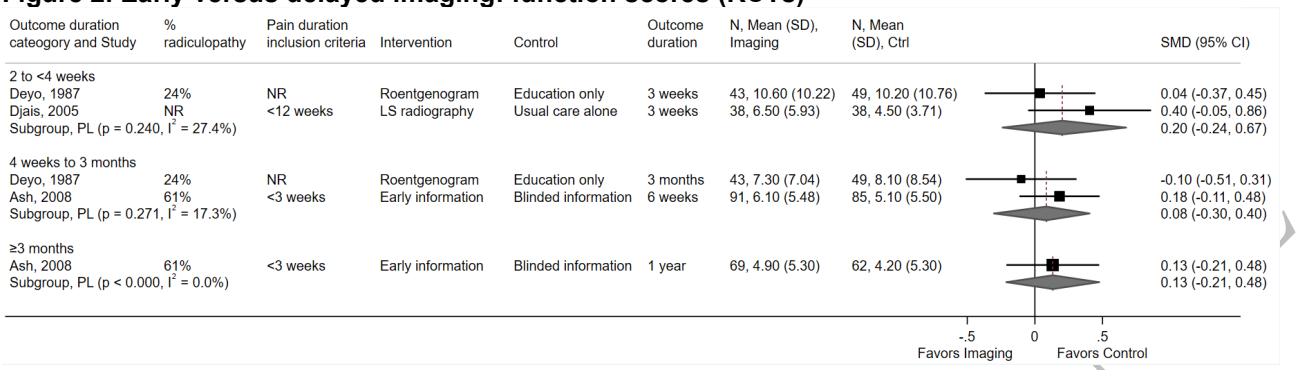
Detailed Synthesis

Randomized Controlled Trial Evidence

Function

Three trials reported measures of function. Two RCTs used the RDQ[62,159] (0-24 scale) and one reported Sickness Impact Profile Physical Dimension Scores (0-100).[158] RCTs found no difference in function scores at 2 to <4 weeks (2 RCTs, N=168)[158,159] between 4 weeks and 6 months (2 RCTs, N=268) or at ≥3 months (1 RCT, N=133),[62] however all estimates are imprecise (Figure 2). One RCT[92] (n=156) found no difference between patients who did and did not receive early imaging in the likelihood of achieving ≥50% improvement in RDQ scores (0-24 scale) at 6 weeks or 2 years (Table 21).

Figure 2. Early versus delayed imaging: function scores (RCTs)



CI = confidence interval; Ctrl = control; LS = lumbar spine; NR = not reported; PL = profile likelihood; SD = standard deviation; SMD = standardized mean difference.

Pain

One trial (N=171) found no difference on average VAS pain scores (0-10) at 6 weeks (N=176) or 1 year between patients who received early MRI and were aware of the results and patients who were blinded to imaging results.[62] In another trial (N=76), there was no difference in median VAS (0-10) scores at 3 weeks between those who had early MRI and those that did not.[159] Knowledge of early MRI results was associated with a small improvement in pain on the 36-item Short Form Questionnaire (SF-36) Bodily Pain Scale (0-100) at 3 weeks in one RCT (N=176, MD -9.40, 95% CI -16.41 to -2.39), but not 1 year (N=131) versus being blinded to the results.[62] Control patients were unblinded to MRI results at 6 months in this trial. The same trial (N=171)[62] found no difference in the likelihood of achieving ≥50% improvement on VAS scores (0-10) at 6 weeks between groups (Table 21).

Table 21. Function and pain outcomes: early imaging versus usual care

Outcome	Author, Year	Follow-up	Early Imaging Mean (SD) or % (n/N)	Usual Care Mean (SD) or % (n/N)	Effect Size (95% CI)
Function Response					
Responders RDQ (≥50% Improvement, 0-24 scale)	Modic 2005*	6 weeks	60.4% (55/91)	67.1% (57/85)	RR 0.90 (0.72 to 1.13)
		2 years	NR	NR	p=0.376†
Pain					
Responders VAS (≥50% Improvement, 0-10 scale)	Ash, 2008)*	6 weeks	48.3% (43/89)‡	53.7% (44/82)‡	RR 0.90 (0.67 to 1.21)‡
Median VAS Pain Score (0-10)	Djais, 2005	3 weeks	Median 4 (IQR 2 to 6) (n=38)	Median 3 (IQR 2 to 5) (n=38)	p=0.07
VAS Pain Score – Average Pain (0-10 scale)	Ash, 2008*	6 weeks	3.50 (2.70) (n=91)	2.96 (2.71) (n=85)	MD 0.54 (-0.26 to 1.34)
		1 year	2.90 (2.40) (n=69)	2.80 (2.80) (n=62)	MD 0.10 (-0.80 to 1.00)
SF-36 Bodily Pain Score (0-100; higher = better)	Ash, 2008*	6 weeks	55.60 (22.70) (n=91)	65.00 (24.40) (n=85)	**MD -9.40 (-16.41 to -2.39)**
		1 year	64.10 (26.5) (n=69)	64.70 (26.80) (n=62)	MD -0.60 (-9.82 to 8.62)

CI = confidence interval; IQR = interquartile range; MD = mean difference; NR = not reported; RDQ = Roland-Morris Disability Questionnaire; RR = risk ratio; SD = standard deviation; SF-36 = 36-item Short Form Questionnaire; VAS = visual analogue scale.

* Modic and Ash: All patients imaged; control group and providers were blinded to results unless clinically important or unblinded at 6 months. Ash reports pain at 1-year follow-up (i.e., after unblinding).
† Adjusted for race, sex, and pain distribution. Data and effect estimates were not reported.
‡ Authors state the proportion and the number of patients experiencing this outcome; denominators were back calculated.

Quality of Life

One RCT reported a small improvements on the SF-36 PCS (0-100) at 6 weeks (N=176) in patients who were unaware of MRI findings (controls blinded to MRI findings) compared with patients who were aware of early MRI findings; there was no difference at 1 year.[62] One trial (N=76) reported no difference between early imaging and usual care at 3 weeks on median EQ-5D scores (0-1) (Table 22).[159]

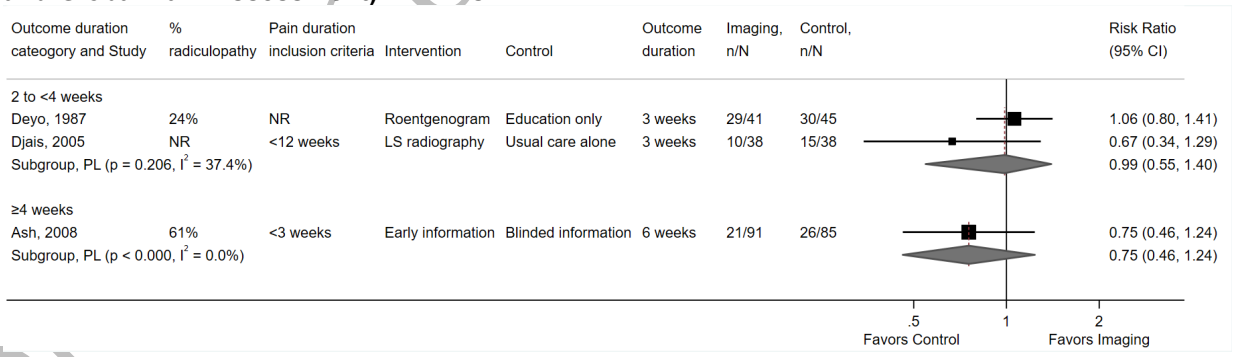
Psychological Distress

No trial explicitly reported on psychological distress. One RCT reported a small improvements on the SF-36 MCS (0-100) at 6 weeks (N=176) and 1 year (N=31) in patients who were unaware of MRI findings (the control group) compared with patients who were aware of early MRI findings.[62]Another RCT (N=93) reported no difference between early imaging and usual care on the Sickness Impact Profile Psychological Dimension Score (0-100) at 3 weeks or 3 months (Table 22).[158]

Overall Satisfaction

Three RCTs reported patient perception of improvement including the proportion of patients that returned to normal activities,[158] were "much improved" (definition NR) on the Global Pain Assessment,[159] or were satisfied with their condition;[92] there was no difference between early imaging and usual care in pooled or individual effect sizes at 3 or 6 weeks[92,158,159] (Figure 3). Additionally, one RCT (N=92) reported overall satisfaction scores (9-27) at 3 weeks, and self-rated improvement in symptoms (1-6) at 3 weeks and 3 months, with no difference between early imaging and usual care for either measure at any timepoint (Table 22).[158]

Figure 3. Perception of improvement (satisfaction with condition, returned to normal activities, and Global Pain Assessment) in RCTs



CI = confidence interval; LS = lumbar spine; NR = not reported; PL = profile likelihood.
Deyo 1987: Return to Normal Activities
Djais 2005: Global Pain Assessment – "Much Improved"
Modic 2005: Satisfaction with Condition

Table 22. Secondary outcomes: early imaging versus usual care

Outcome	Author, Year	Follow-up	Early Imaging Mean (SD) or % (n/N)	Usual Care Mean (SD) or % (n/N)	Effect Size (95% CI)
Quality of Life					
SF-36 PCS (0-100)	Ash, 2008	6 weeks	69.00 (22.30) (n=91)	76.00 (24.70) (n=85)	**MD -7.00 (-13.99 to -0.01)**
		1 year	75.00 (25.40) (n=69)	75.70 (24.20) (n=62)	MD -0.70 (-9.30 to 7.90)
EQ-5D (0-1)	Djais, 2005	3 weeks	Median 0.63 (IQR 0.41 to 0.75) (n=38)	Median 0.74 (IQR 0.53 to 0.81) (n=38)	p=0.15
Mental Health-Related Measures					
SF-36 MCS (0-100)	Ash, 2008	6 weeks	69.40 (21.00) (n=91)	78.70 (20.20) (n=85)	**MD -9.30 (-15.44 to -3.16**)
		1 year	74.30 (20.20) (n=69)	81.00 (16.70) (n=62)	**MD -6.70 (-13.15 to -0.25)**
Sickness Impact Profile – Psychosocial Dimension Score (0-100; lower = better)	Deyo, 1987	3 weeks	20.20 (SD NR) (n=43)	14.70 (SD NR) (n=49)	p>0.05
		3 months	15.70 (SD NR) (n=43)	10.60 (SD NR) (n=49)	p>0.05
Overall Satisfaction					
Self-Rated Improvement in Symptoms (1-6; lower = better)	Deyo, 1987	3 weeks	2.70 (SD NR) (n=43)	2.70 (SD NR) (n=43)	p>0.05
		3 months	2.60 (SD NR) (n=43)	2.60 (SD NR) (n=49)	p>0.05
Overall Satisfaction Score (9-27; higher = better)	Deyo, 1987	3 weeks	23.70 (SD NR) (n=43)	24.00 (SD NR) (n=49)	p>0.05

CI = confidence interval; EQ-5D = EuroQoL 5D; IQR = interquartile range; MCS = mental component scale; MD = mean difference; NR = not reported; PCS = physical component scale; SD = standard deviation; SF-36 = 36-item Short Form Questionnaire.
* Modic 2005 and Ash 2008 imaged all patients; intervention patients were unblinded to the results immediately; controls remained blinded to results unless clinically important or until 6 months.

Opioid Use and Use of Rescue Medication

No RCTs reported on the use of rescue medication explicitly. Two RCTs reported on medication use in general. One trial[62] (N=246) reported no difference at 6 weeks between patients who were unaware of MRI findings (the control group) compared with patients who were aware of early MRI findings on the use of NSAIDS (87% vs. 88%), analgesics (81% vs. 76%), or other medications (not specified, 43% vs. 51 %). Similarly, another trial (N=92) reported no difference in medication use between early imaging and usual care (p>0.02) Appendix E.[158]

Additional Service Use

There was no difference between early imaging and usual care for undergoing surgery (N=246; 0.7% vs. 1.7%)[62] or being hospitalized (N=92; 2.3% vs. 0%)[158] (Appendix E).

Incidental Findings

One RCT (N=246) reported no differences between early imaging and usual care on the frequency of incidental findings, including severe spinal stenosis (9% vs. 11%) or disc herniation (59% vs. 61%)[62] (Appendix E).

Adverse Events

No RCTs reported on adverse events.

Differential Efficacy and Safety

No RCTs evaluated effect modification.

Nonrandomized Studies Evidence

Continued Disability

One prospective study[188] (N=1233) and one retrospective study[175] (N=555) reported on the impact of early MRI on the likelihood of continuing to receive disability (i.e., receiving wage replacement) and rate of going off disability. Early MRI was associated with a substantially higher likelihood of receiving disability at 1 year compared with guideline-consistent MRI done after 6 weeks, (N=955, 30.6% vs. 6.7%, adjusted RR 2.03, 95% CI 1.22 to 3.11) in patients with ALBP due to mild/major sprain, as well as a moderately higher likelihood of patient self-report of not working in the last week (38% vs. 10.9%, adjusted RR 1.98, 95% CI 1.42 to 2.76) and lower rate of ending disability (adjusted HR 0.48, 95% CI 0.38 to 0.60) in the prospective study.[188] Similarly, the retrospective study reported a lower rate of discontinuing disability with early MRI compared with no MRI (N=332, adjusted HR 0.32, 95% CI 0.24 to 0.43) in patients with nonspecific ALBP.[175]Estimates from both studies are imprecise.

In patients with radiculopathy, there was no association between early MRI and continuing to receive disability (N=268, 40.2% vs. 22.6%, adjusted RR 1.31, 95% CI 0.84 to 2.05), or for patient self-report of not working in the previous week (48.5% vs. 29.3%, adjusted RR 1.21, 95% CI 0.85 to 1.74) versus MRI done after 6 weeks in the prospective study.[188] The same study found a lower rate of discontinuing disability with early MRI compared with no MRI (N=268, adjusted HR 0.57, 95% CI 0.40 to 0.81) in patients with radiculopathy as did the retrospective study (N=223, adjusted HR 0.28, 95% CI 0.18 to 0.43) also in patients with radiculopathy.[175] Estimates for both studies are imprecise.

Opioid Use

Two retrospective nonrandomized studies reported descriptive statistics for mean morphine milligram equivalents used (MME)[175,181] and one reported on proportion of patients taking opioids.[181] However, they do not provide adjusted effect estimates or tests of statistical significance comparing early MRI with no MRI or delayed MRI related to opioid use within the first 15 days. Both studies report higher mean MME use in the early MRI groups versus patients not receiving MRI[175] or those who had MRI between 42 and 180 days in patients with and without radiculopathy.[181] This was true for patients with and without radiculopathy. confidence intervals for reported means were wide. One study reported that more patients in the early MRI group (42%) received opioids within 15 days of injury (early use) compared to no MRI (28.2%) or delayed MRI (21.8%) in patients with radiculopathy however adjusted risk ratios were not reported[181] (Appendix F).

Additional Services

One nonrandomized study reported on the use of additional medical services during follow-up of 6 months post-MRI.[181] Early MRI was associated with substantially greater likelihoods for the following versus no MRI in cases where LPB was categorized as less severe (degenerative changes, nonspecific back pain, or miscellaneous diagnoses) and in more severe cases (herniated disc, lumbar radiculopathy or neuropathy, spinal stenosis, sciatica, or possible instability): use of injection therapies, electromyography/nerve conduction velocity, additional advanced imaging, and surgery. Effect estimates (RRs) ranged from 4.4 to 42 and in all cases were imprecise (Appendix F, Utilization of services).

Cost-Effectiveness

One older (1982) economic study compared use of roentgenography at an initial primary care visit versus not doing so in patients 18-60 years old with ALBP (<2 weeks duration)without signs of symptoms consistent with serious disease.[196] Decision analysis was done to determine a cost-effectiveness ratio in terms of dollars per day of suffering averted and a risk-benefit ratio of millirads per day of suffering averted via earlier diagnosis. Various literature sources and expert opinion were used to estimate and model disease probabilities, probability of an abnormal roentgenogram and probability of symptom improvement in general and in patients with various diseases requiring specific treatments. The base case estimate is based on an expert's observational study of 9000 primary care patients in England (literature citation not provided). The study was funded by National Center for Health Services Research and the National Institute of Health (Appendix F).

Authors assumed a $97.50 charge and an exposure of 160 mrad for the x-ray for base case analysis. Modeling with these assumptions, 16.81 days of suffering, a cost of $107.36 and radiation exposure of 165.17 mrad were found for use of x-ray compared with 16.85 days of suffering, a cost of $9.98 and exposure of 15.36 mrad. For this base case, use of x-ray saved an average of 0.04 days (58 minutes) of suffering but a 10-fold higher cost versus not having x-ray.

Sensitivity analyses included varying probabilities of serious disease requiring specific treatment and the probability of an abnormal x-ray at initial visit, 4 weeks, and 8 weeks, and the charge for x-ray. Serious disease probability was the most noticeable driver of cost effectiveness. At the lowest probability of serious disease (0.2%), authors estimate that it would cost $2072 to avert a day of suffering and result in a 3188 mrad exposure in a population with ALBP and $28 to avert a day of suffering at the highest probability (15%) with a radiation exposure 42 mrad. Varying the probability of an abnormal x-ray resulted in outcome values that deviated from the base case by <10%. Use of roentgenography remained 10 times more expensive than not using it regardless of change in absolute charge.

Study limitations included limited modeling of benefits, risk and harms, use of observational data, and reliance on expert opinion. Authors acknowledged the difficulty of estimating risks and benefits such as patient satisfaction, patient apprehension, and provider avoidance of litigation. They indicated that the potential risk from radiation and the dollar cost should not be ignored.

Key Question 3. Systemic Opioid Therapy

A total of 10 RCTs[28,77,89,105,113,133,138,162,168,169] compared an opioid[28,77,113,133,138,168] or an opioid in combination with a nonopioid[89,105,113,138,162,168,169] versus a placebo,[28,113,133,168,169] a nonopioid,[77,89,105,113,138,169] or an opioid.[162,168] Additionally, two cohort studies[172,176] evaluated the effect of prescribing versus not prescribing opioid therapy on long-term opioid use and disability.

Description of Included Studies

Nine RCTs evaluated opioids for ALBP compared to placebo or a nonopioid (Table 23; Appendix E).[28,77,89,105,113,133,138,168,169] Mean age in the trials ranged from 29 to 55 years. One trial was restricted to males; in the other trials, the proportion of females ranged from 21% to 55%. Three trials reported participant race and/or ethnicity. One trial enrolled a predominately White population (92%),[89] whereas White participants comprised approximately half of the sample in the other two trials (50% White, 43% Black, and 5% Asian in one trial[105] and 56%, 21%, and 3%, respectively, in the other[133]). Trials were conducted in the United States,[105,113,138] Canada,[89,133] Australia,[28] Turkey,[77] China,[169] and Europe.[168] Five trials were conducted in EDs, one in primary care outpatient or ED settings, one in outpatient setting, and one in an inpatient and outpatient hospital setting; one trial did not report setting. Patients with radiculopathy or suspected neurologic symptoms, such as leg pain, were excluded from three trials,[77,113,138] and three trials[28,168,169] allowed enrollment of patients with radiculopathy (range: 9.6-61%); three trials[89,105,133] did not report inclusion or exclusion of patients with radiculopathy. Mean pain duration at baseline was 2.3 day,[113] 3.4 days,[105] and 14.8 days[28] in 3 trials; four other trials did not report mean pain duration hours but restricted enrollment to patients with LBP <48 hours,[133,168] <72 hours,[89] ≤1 week,[77] <6 weeks[169] and <12 weeks.[138] One trial did not report mean pain duration or define ALBP.[105] One trial evaluated single-dose intravenous morphine in an ED setting at 30 minutes,[77] and one trial evaluated single-dose tramadol in an ED setting at 8 hours.[168] The other trials evaluated oral opioids at 2.5 days to 6 weeks; outcomes were evaluated at the end of treatment in all trials but one, which evaluated post-treatment outcomes at 1 year.[28] Mean pain intensity at baseline ranged from 4.1 to 8.3 on a 0-10 scale in eight trials that provided this information.

The opioids assessed were intravenous (IV) morphine[77] and oral tramadol,[133,168,169] oxycodone,[28,113,138] and codeine.[89,105,138] One trial of oxycodone administered the drug in combination with naloxone;[28] while this formulation is not available in the U.S. market, we included the study because naloxone is not systemic and has no analgesic properties. This trial also enrolled patients with neck or back pain; we report outcomes for the back pain subgroup provided by the study authors where available, and report outcomes based on the overall trial sample when back pain subgroup results were not available. One trial[77] compared IV morphine 21 mg (morphine equivalent dose [MED]) versus IV acetaminophen 1,000 mg or IV dexketoprofen 50 mg. One trial compared tramadol (30 mg MED) plus acetaminophen versus placebo,[133] one trial compared tramadol alone and plus dexketoprofen versus placebo,[168] one trial compared tramadol plus diclofenac versus placebo (plus diclofenac) and versus tizanidine (plus diclofenac).[169] Three trials evaluated oxycodone. One compared oxycodone (30 mg MED) plus naloxone versus placebo,[28] one compared oxycodone (22.5 mg MED) plus aspirin versus acetaminophen (dosages NR),[138] and one compared oxycodone (60 mg MED) plus acetaminophen versus placebo or cyclobenzaprine.[113] Three trials evaluated codeine; one trial compared codeine (54 mg MED) plus acetaminophen versus diflunisal,[105] one trial compared codeine versus acetaminophen (dosages NR),[138] and one trial compared codeine (up to 54 mg MED) plus acetaminophen versus ketorolac.[89]

Four trials[28,77,89,113] were assessed as low risk of bias, four trials[105,133,168,169] as moderate risk of bias, and one trial[138] as high risk of bias (Appendix F). Methodologic shortcomings in the moderate risk of bias trials included unclear randomization and allocation concealment methods and in one trial[105] lack of blinding. The trial rated high risk of bias was poorly described overall and there was not enough information to judge many of the criteria.

In addition, two moderate risk of bias trials (N=537)[162,168] compared an opioid in combination with a nonopioid versus the same opioid alone. One trial evaluated the addition of acetaminophen 650 mg to tramadol 75 mg (30 mg MED) daily compared to tramadol 50 mg (20 mg MED) daily (with both groups able to increase intake up to 8 doses per day). The second trial compared tramadol (20 mg MED) plus dexketoprofen 25 mg versus tramadol (15 mg MED);[168] outcomes were measured at 8 hours after only a single dose of medication (single-dose phase) and after 8 hours up to 5 days after multiple doses (i.e., every 8 hours; multiple-dose phase). One study was conducted in an outpatient setting in France and excluded patients with radiculopathy[162] while the second trial was conducted in both primary care and hospital (ED) settings in six European countries and included patients with and without radiculopathy.[168] Mean patient age ranged from 43 to 55 years, 52% to 58% were female, and mean baseline pain severity ranged from 6.6 to 7.0 (on a 0-10 scale) across trials. Baseline pain duration was a mean 22 days in one trial[162] and ≤48 hours in the other (mean not reported).[168] Neither trial reported on history of substance abuse disorders or psychiatric comorbidities.

Table 23. Study characteristics of trials that evaluated opioids for acute low back pain

Study, Year	Friedman, 2015	Jones, 2023	Lasko, 2012	Hung, 2024	Varrassi, 2024	Perrot, 2006	Eken, 2014	Wiesel, 1980	Innes, 1998	Brown, 1986
N Randomized	323	277*	277	296	528	119	137	75	123	47
Mean Age (Years)	39	45	42	49	43	55	32	23 (entire trial sample)	35	29
% Female	47%	49%	52%	42%	47%	58%	39%	0%	21%	55%
% Radiculopathy	0 (Excluded)	Back pain: 61% (164/271)*	NR	9.6%	51.3%	0%	0 (Excluded)	0 (Excluded)	NR	NR
Pain Duration Inclusion Criteria	<2 weeks	≤12 weeks	<48 hours	<6 weeks	≤48 hours	≥10 to ≤42 days	≤1 week	<3 months	<72 hours	NR
Mean Pain Duration (days)	2.3	14.8	NR	NR	NR	22	NR	NR	NR	3.4
Mean/Median Baseline Pain Severity (0-10)	NR	5.6	6.9	4.3 (rest) 7.6 (activity)	7.0	NR	8.3	NR	6.7	NR
History Substance Abuse Disorder	NR	NR	Excluded	NR	NR	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Opioid, dose	Oxycodone (5 mg) + acetaminophen (325 mg)	Oxycodone (5-10 mg) + naloxone (2.5 mg)	Tramadol (75 mg) + paracetamol (650 mg)	Tramadol (50 mg) + diclofenac (100 mg)	(1) Tramadol (100 mg) (2) Tramadol (75 mg) + dexketoprofen (25 mg)	Tramadol (37.5 mg) + acetaminophen (325 mg)	IV Morphine (0.1 mg/kg)	(1) Oxycodone + aspirin (doses NR) (2) Codeine (60 mg)	Codeine (60 mg) + acetaminophen (600 mg)	Codeine (30 mg) + acetaminophen (300 mg)
Opioid MED†	22.5 mg	30 mg	30 mg	40 mg	(1) 20 mg (2) 15 mg	Day ≤2: 15 mg Days 3-10: 20 mg	21 mg	Oxycodone: NC Codeine: 36 mg	45 mg	27 mg
Dosage Frequency	Every 8 hours	Twice per day	Every 10 to 12 hours	Every 6 hours	Once (single-dose phase)‡ Every 8 hours (multiple-dose phase)‡	Days ≤2: 4 times per day Days 3-10: Up to 8 times per day	Once	4 times per day	Every 4 to 6 hours	2 tablets initially, then every 4 hours

Study, Year	Friedman, 2015	Jones, 2023	Lasko, 2012	Hung, 2024	Varrassi, 2024	Perrot, 2006	Eken, 2014	Wiesel, 1980	Innes, 1998	Brown, 1986
Treatment Duration	12 weeks	6 weeks	2.5 days	14 days	1 day, 0-8 hrs. (single-dose phase)‡ 1 [>8 hrs.] to 5 days (multiple-dose phase)‡	10 days	Single dose	14 days	1 week	≤15 days
Comparators	(1) Placebo (2) Cyclobenzaprine 5-10 mg	Placebo	Placebo	(1) Placebo (2)Tizanidine 2 mg	Placebo	Tramadol (50 mg)	(1) IV Dexketoprofen 50 mg (2) IV Acetaminophen 1,000 mg	(1) Acetaminophen (dosage NR)	Ketorolac tromethamine 10 mg	Diflunisal 1,000 mg
Duration Follow-up	12 weeks	1 year	2.5 days	28 days	8 hours (single-dose phase)‡ 5 days (multiple-dose phase)‡	10 days	30 minutes	14 days	1 week	15 days
Country	United States	Australia	Canada	Hong Kong	Europe (multi-national)	France	Turkey	United States	Canada	Sweden
Risk of Bias	Low	Low	Moderate	Low	Moderate	Moderate	Low	High	Low	Moderate

IV = intravenous; NR = not reported.

* Subgroup results for patients presenting with acute low back pain; some harms outcomes reported for overall trial sample (including patients presenting with neck pain).

† Calculated by AAI/OHSU.

‡Patients randomized to tramadol/dexketoprofen fixed combination or tramadol alone during the single-dose phase (0-8 hours) continued to receive the same treatment during the multiple-dose phase (>8 hours to 5 days); however, patients randomized to placebo during the single-dose phase received either the tramadol/dexketoprofen fixed combination or tramadol alone during the multiple-dose phase.

Key Question 3a, e, h: Single Systemic Opioid

Detailed Synthesis

Systemic Opioid Versus Placebo

Five RCTs, two rated low[28,113] and three[133,168,169] rated moderate risk of bias, compared a single opioid with a placebo condition (Table 24).[28,113,133,168,169] Opioid treatments and regimens varied across trials. Two trials[28,113] evaluated oxycodone (plus acetaminophen[113] or naloxone[28]) and three trials[133,168,169] evaluated tramadol (alone or plus dexketoprofen[168] or plus acetaminophen[133] or plus diclofenac[169]). The MED for opioids ranged from 15 to 40 mg and treatment duration ranged from 8 hours to 12 weeks. Concomitant and rescue medications were either restricted or not well-reported.

Across trials, mean patient age ranged from 39 to 47 years and 42% to 52% were female. Mean baseline duration of LBP ranged from 48 hours to 14.8 days and baseline pain intensity ranged from 4.3 to 7.6 on a 0 to 10 scale in four RCTs.[28,133,168,169] In two trials[168,169] 10% to 51% patients had radiculopathy, one trial[113] excluded individuals with radiculopathy, and two trials[28,133] did not report on radiculopathy. History of previous LBP episodes, substance use disorder and psychiatric comorbidities were not reported across trials, except for one trial that excluded patients with a history of substance use disorder[133]).

Function

Three trials[28,113,169] (N=966) reported mean differences in measures of function (Table 24). One trial (N=277) found oxycodone plus naloxone (titrated up to 10 mg and 5 mg twice daily) associated with slightly worse back-related physical function based on the RDQ (0-24 scale) compared to placebo at week 6 (adjusted MD 2.09, 95% CI 0.23 to 3.96) and slightly worse generic physical functioning (assessed using Brief Pain Inventory - Interference Subscale [scored 0-10]) at week 6 (end of treatment; adjusted MD 0.70, 95% CI 0.03 to 1.37) and week 12 (adjusted MD 0.83, 95% CI 0.15 to 1.51).[28] Two trials found no difference in RDQ scores for oxycodone 5 mg plus acetaminophen 325 mg every 12 hours at 1 week (N=215, mean 7.8 vs. 8.9; difference -1.1, 95% CI -3.4 to 1.1)[113] and tramadol 200 mg every 6 hours at 1, 3 or 4 weeks (N=192)[169] versus placebo.

No trial reported likelihood of experiencing a positive function outcome (e.g., >30% improvement in RDQ).

Pain Intensity

Four trials reported pain intensity using different measures (Table 24). One trial[133] (N=277) found tramadol 75 mg plus acetaminophen 650 mg every 10-12 hours associated with a small improvement in summed pain intensity difference over 50 hours (SPID50; median -6.0 vs. -4.0, p=0.038) and total pain relief over 50 hours (TOTPAR50; median 13.0 vs. 11.0, p=0.026) versus placebo. One trial[169] of tramadol 200 mg every 6 hours found no differences versus placebo in pain at rest or during activity (NRS 0-10 scale) at 1, 3, or 4 weeks. Another trial[28] (N=277) found no differences between oxycodone plus naloxone (titrated up to 10 mg oxycodone + 5 mg naloxone twice per day) versus placebo in pain severity on the Brief Pain Inventory - Pain Severity (0-10 scale) at 2, 4, 6 or 52 weeks of treatment (Appendix E); this study reported slightly higher pain severity in the opioid group during post-treatment follow-up at 12 weeks

(adjusted MD 0.78, 95% CI 0.18 to 1.38) and 26 weeks (adjusted MD 0.75, 95% CI 0.15 to 1.36).

The fourth trial (n=304)[168] reported no difference in likelihood of experiencing a positive pain outcome (NRS ≤4 or >30% improvement in pain) between tramadol plus dexketoprofen and placebo at 8 hours.

Quality of Life and Pain Relief Satisfaction

Four trials[28,113,133,168] (N=1,073) reported pain relief satisfaction compared to placebo; one of these trials[28] also reported quality of life. One trial (N=277)[133] reported that tramadol 75 mg in combination with acetaminophen 650 mg was rated as more effective than placebo (p=0.005) at 2.5 days, while the other trials found no difference in pain relief satisfaction between oxycodone 5 mg in combination with acetaminophen 325 mg and placebo every 8 hours at 1 week or 3 months (N=323),[113] or oxycodone in combination with naloxone (titrated to 10 mg and 5 mg, respectively, twice daily) and placebo.[28] A trial of single-dose tramadol 100 mg and tramadol 75 mg plus dexketoprofen 25 mg found no differences in patient global evaluation at 8 hours compared with placebo.[168] The trial[28] (N=277) of oxycodone plus naloxone also found opioid treatment associated with slightly worse 36-item Short Form Questionnaire mental component subscale scores (SF-36 MCS) versus placebo at week 6 (scored 0-100; MD -3.65, 95% CI –6.28 to -1.02) and week 12 (MD -4.04, 95% CI -6.09 to -1.38). See Appendix E for more details.

Return to Work

Two trials reported return to work outcomes compared to placebo (Appendix E).[113,169] One trial[113] (N=323) compared oxycodone 5 mg plus acetaminophen 325 mg versus placebo and reported no differences groups in return-to-work outcomes at 1 week. The other trial[169] (N=296) reported no differences in return to work at 7 days or 28 days.

Medication Use

Three trials (N=866)[28,113,168] reported medication use, including rescue medication and long-term opioid use. One trial (N=215) of oxycodone 5 mg in combination with acetaminophen 325 mg every 8 hours found no differences in medication use (type not reported) between patients receiving opioids versus placebo at 1 week or 3 months.[113] Another trial (N=347) found no differences between oxycodone in combination with naloxone (titrated to 10 mg and 5 mg, respectively, twice per day) and placebo at 6 weeks in medication use for back and neck pain (56% vs. 58%) including simple analgesia (non-opioid analgesics, e.g., acetaminophen, NSAIDs) (25% vs. 26%), use of NSAIDs (37% vs. 42%), use of combination opioid (17% vs. 12%), use of strong opioids (9% vs. 11%), or use of weak opioids (2% vs. 1%).[28] A third trial reported no difference in rescue medication use (acetaminophen 500 mg, maximum 2000 mg, daily) between single-dose tramadol or tramadol plus dexketoprofen and placebo at 8 hours.[168]

Harms

Serious Adverse Events

Two trials (N=581) reported serious adverse events for opioids compared to placebo (Table 24).[28,168] One trial (N=277) found no difference between oxycodone plus naloxone versus placebo in likelihood of serious adverse events through 6 weeks (3.7% vs. 2.1%; RR 1.72, 95% CI 0.42 to 7.09).[28] One treatment-related serious adverse event (SAE) (0.7%, 1/136), an acute

mental disorder (not otherwise specified), occurred in the oxycodone group; no treatment-related SAEs were reported in the placebo arm (0%, 0/141). The other trial reported no serious adverse events with single-dose tramadol and one SAE (0.5%, 1/211; urinary calculus) with tramadol plus dexketoprofen versus placebo at 8 hours.[168]

Study Withdrawal Due To Adverse Events

Two trials (N=581) reported withdrawal due to adverse events for opioids versus placebo (Table 24).[133,168] One trial (N=277) found tramadol plus acetaminophen associated with a moderately increased likelihood of withdrawal due to adverse events versus placebo through 2.5 days (12.1% vs. 2.2%; RR 5.46, 95% CI 1.64 to 18.23).[133] The most commonly reported reason for withdrawal in the tramadol group was vomiting. Another trial of single-dose tramadol (n=3326) and tramadol plus dexketoprofen (n=330) reported no difference versus placebo in rates of withdrawal due to adverse events with tramadol at 8 hours.[168] Nervous system disorders and dizziness were the most common reason for withdrawal in both groups.

Any Adverse Event

Four trials (N=1,142) reported rates of overall adverse events (Table 24).[28,113,133,168] One trial (N=347) found oxycodone in combination with naloxone associated with a similar likelihood of adverse events at 6 weeks compared with placebo for the overall population that included back and neck pain patients (35% vs. 30%; RR 1.18, 95% CI 0.87 to 1.61).[28] Another trial found no differences in overall adverse events with single-dose tramadol or tramadol plus dexketoprofen versus placebo at 8 hours.[168] Two other trials[113,133] (N=492) found an opioid in combination with a nonopioid associated with an increased likelihood of adverse events compared with placebo: in the trial of oxycodone plus acetaminophen, the effect was moderate at 3 months (40% vs. 21%; difference 19%, 95% CI 7% to 31%; RR 1.94, 95% CI 1.25 to 3.00)[113] and in the trial of tramadol plus acetaminophen the effect was large at 2.5 days (42% vs. 13%; difference 29%, 95% CI 19% to 39%; RR 3.35, 95% CI 2.06 to 5.44)

Specific Adverse Events

Four trials (N=1,038) reported specific adverse events for the comparison of oral opioids versus placebo. One study (n=215) comparing oxycodone in combination with acetaminophen versus placebo reported significantly greater likelihood of drowsiness (15% vs. 4%; difference 11%, 95% CI 4% to 19%), dizziness (15% vs. 3%; difference 12%, 95% CI 5% to 19%), and nausea or vomiting (18% vs. 6%; difference 12%, 95% CI 4 to 20%).[113] Two trials (N=546) reported that tramadol was associated with higher likelihood of nausea (24.1% vs. 2.2%; RR 10.93, 95% CI 3.44 to 34.76 and 37.5% vs. 14.0%; RR 2.68, 95% CI 1.51 to 4.75) and dizziness (14.9% vs. 1.5%; RR 10.13, 95% CI 2.42 to 42.37 and 51.1% vs. 14.0%; RR 3.66, 95% CI 2.12 to 6.30) compared with placebo.[133,169] One of these trials also reported increased risk of sleepiness (52.3% vs. 30.1%; RR 1.74, 95% CI 1.20 to 2.51) and vomiting (15.9% vs. 3.2%; RR 4.93, 95% CI 1.47 to 16.58) with tramadol compared with placebo.[169] A trial (N=277) of oxycodone in combination with naloxone found no significant difference in nausea and vomiting (7.4% vs. 3.6%; RR 2.07, 95% CI 0.73 to 5.91) or somnolence (2.9% vs. 0.7%; RR 4.15, 95% CI 0.47 to 36.64).[28]

Differential Effectiveness and Safety

No studies were identified that evaluated differential efficacy or safety.

Table 24. Systemic single opioid versus placebo: function and pain and adverse events

Intervention	Outcome	Author Year (Patients) Opioid treatment	Results Opioid vs. Placebo Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
Oxycodone	**Function**	Friedman, 2015 (N=215) *Oxycodone + acetaminophen*	**RDQ scores (0-24)*** 1 week: 7.8 (8.5) vs. 8.9 (8.44), MD –1.1, 95% CI –3.4 to 1.1
		Jones, 2023 (n=277, back pain subgroup) *Oxycodone + naloxone*	**RDQ scores (0-24)†** 6 weeks: 8.43 (7.81) vs. 6.33 (7.96); adjusted MD 2.09, 95% CI 0.23 to 3.96 **BPI-IS scores (0-10)†** 6 weeks: 2.67 (2.80) vs. 1.97 (2.85); adjusted MD 0.70, 95% CI 0.03 to 1.37 12 weeks: 2.48 (2.92) vs. 1.65 (2.85); adjusted MD 0.83, 95% CI 0.15 to 1.51
	Pain	Jones, 2023 (n=277, back pain subgroup) *Oxycodone + naloxone*	**BPI-PS scores (0-10)†** 2 weeks: 3.69 (2.45) vs. 3.44 (2.49); p=NS 4 weeks: 2.93 (2.57) vs. 2.65 (2.49); p=NS 6 weeks: 2.64 (2.57) vs. 2.10 (2.49); adjusted MD 0.54, 95% CI –0.05 to 1.12 12 weeks: 2.63 (2.57) vs. 1.85 (2.49); adjusted MD 0.78, 95% CI 0.18 to 1.38 26 weeks: 2.48 (2.57) vs. 1.73 (2.49); adjusted MD 0.75, 95% CI 0.15 to 1.36 52 weeks: 2.15 (2.57) vs. 1.72 (2.49); adjusted MD 0.44, 95% CI –0.16 to 1.03
	Any AE	Jones, 2023 (N=346) *Oxycodone + naloxone Whole population*	6 weeks: 35% (61/174) vs. 30% (51/172); RR 1.18, 95% CI 0.87 to 1.61
		Friedman, 2015 (N=215) *Oxycodone + acetaminophen*	12 weeks: 40% (43/108) vs. 21% (22/107); RR 1.94, 95% CI 1.25 to 3.00
	Serious AE	Jones, 2023 (n=277, back pain subgroup) *Oxycodone + naloxone*	**Any serious AE** 6 weeks: 3.7% (5/136) vs. 2.1% (3/141), RR 1.72, 95% CI 0.42 to 7.09 **Treatment-related serious AE:** 6 weeks: 0.7% (1/136) vs. 0% (0/141); 1 case of acute mental disorder NOS
Tramadol	**Function**	Hung, 2024 (N=192) *Tramadol + diclofenac‡*	**RDQ scores (0-24)** 1 week: 7.2 (5.8) vs. 8.5 (5.5), adjusted MD in change scores –0.85, 95% CI –2.80 to 1.10 3 weeks: 4.6 (6.5) vs. 4.5 (5.1), adjusted MD in change scores –0.58, 95% CI –4.06 to 2.90 4 weeks: 4.6 (5.7) vs. 4.5 (5.0), adjusted MD in change scores –0.29, 95% CI –2.25 to 1.67

Intervention	Outcome	Author Year (Patients) Opioid treatment	Results Opioid vs. Placebo Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
	Pain	Hung, 2024 (N=192) *Tramadol + diclofenac‡*	**NRS pain at rest (0-10)** 1 week: 2.1 (2.0) vs.1.9 (2.2), adjusted MD in change scores 0.50, 95% CI –0.26 to 1.27 3 weeks: 1.1 (1.4) vs. 1.7 (2.2), adjusted MD in change scores –0.32, 95% CI –1.41 to 0.77 4 weeks: 1.2 (1.8) vs. 1.0 (1.8), adjusted MD in change scores 0.32, 95% CI –0.34 to 0.98 **NRS pain during activity (0-10)** 1 weeks: 3.8 (2.6) vs. 3.9 (2.7), adjusted MD in change scores 0.43, 95% CI –0.53 to 1.41 3 weeks: 2.2 (2.5) vs. 2.6 (2.5), adjusted MD in change scores 0.30, 95% CI –1.93 to 1.33 4 weeks: 1.9 (2.5) vs. 2.2 (2.6), adjusted MD in change scores –0.29, 95% CI –1.28 to 0.69
		Varrassi, 2024 (N=305) *Tramadol*	**TOTPAR pain scores (scale unknown)** 8 hours: 8.5 vs. 8.4, MD 0.1, p=NS (estimated from graph, SD NR)
		Varrassi, 2024 (N=267) *Tramadol + dexketoprofen*	**NRS scores (0-10)** 8 hours: 5.2 vs. 5.65, MD -0.45, p=NS (estimated from graph, SD NR) **Time to achieve pain relief (NRS score <4 or improvement ≥30%)** 8 hours: 46.1% (94/204) vs. 42.6% (43/102); HR 1.11, 95% CI 0.78 to 1.6
		Lasko, 2012 (N=277) *Tramadol + acetaminophen*	**VAS, median SPID50** 50 hours (2 days): –6.0 (min. –22, max. 3) vs. –4.0 (min. –23, max. 10); p=0.038 **VAS, median TOTPAR50** 50 hours (2 days): 13.0 (min. 0, max. 32) vs. 11.0 (min. 0, max. 40); p=0.026
	Any AE	Varrassi, 2024 (N=326) *Tramadol*	**Any AE** 8 hours: 15.0% vs. 9.2%; RR 1.62, 95% CI 0.85 to 3.10 **Any treatment-related AE** 8 hours: 13.0% vs. 9.2%; RR 1.41, 95% CI 0.73 to 2.74
		Varrassi, 2024 (N=330) *Tramadol + dexketoprofen*	**Any AE** 8 hours: 13.3% vs. 9.2%; RR 1.43, 95% CI 0.74 to 2.78 **Any treatment-related AE** 8 hours: 11.8% vs. 9.2%; RR 1.28, 95% CI 0.65 to 2.51
		Lasko, 2012 (N=277) *Tramadol + acetaminophen*	2.5 days: 41.8% (59/141) vs. 12.5% (17/126); RR 3.35, 95% CI 2.06 to 5.44
		Hung, 2024 (N=192) *Tramadol + diclofenac‡*	Timing NR **Dizziness:** 51.1 % vs. 14%; RR 3.66, 95% CI 2.12 to 6.30 **Nausea:** 37.5% vs. 14%; RR 2.68, 95% CI 1.51 to 4.75 **Vomiting:** 15.9% vs. 3.2%; RR 4.93, 95% CI 1.47 to 16.58 **Sleepiness:** 52.3% vs. 30.1%; RR 1.74, 95% CI 1.20 to 2.51 **Stomachache:**13.6% vs. 18.3%; RR 0.75, 95% CI 0.38 to 1.47 **Indigestion** 13.6% vs. 8.6%; RR 1.59, 95% CI 0.68 to 3.69

Intervention	Outcome	Author Year (Patients) Opioid treatment	Results Opioid vs. Placebo Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
	Serious AE	Varrassi, 2024 (N=326) *Tramadol*	8 hours: 0% (0/207) vs. 0% (0/119)
		Varrassi, 2024 (N=330) *Tramadol + dexketoprofen*	8 hours: 0.5% (1/207) vs. 0% (0/119)
	Withdrawal due to AE	Varrassi, 2024 (N=326) *Tramadol*	8 hours: 3.4% (7/207) vs. 3.4% (4/119); RR 1.01, 95% CI 0.30 to 3.37
		Varrassi, 2024 (N=330) *Tramadol + dexketoprofen*	8 hours: 2.8% (6/211) vs. 3.4% (4/119); RR 0.85, 95% CI 0.24 to 2.94
		Lasko, 2012 (N=277) *Tramadol + acetaminophen*	2.5 days: 12.1% (17/141) vs. 2.2% (3/136); RR 5.46, 95% CI 1.64 to 18.23

AE = adverse event; BPI-IS = Brief Pain Inventory - Interference Subscale; BPI-PS = Brief Pain Inventory - Pain Severity; CI = confidence interval; max. = maximum; HR = hazard ratio; MD = mean difference; min. = minimum; NOS = not otherwise specified; NR = not reported; NRS = numerical rating scale; NS = not statistically significant; RDQ = Roland-Morris Disability Questionnaire; RR = risk ratio; SD = standard deviation; SPID50 = sum of pain intensity differences over the 50-hour observation; TOTPAR = total pain relief score; TOTPAR50 = total pain relief score over the 50-hour observation; VAS = visual analog scale.
* 95% confidence intervals were converted to standard deviations by AAI.
† Standard errors were converted to standard deviations by AAI.
‡ The comparison group for Hung, 2024 is placebo + diclofenac.

Systemic Opioid Versus Nonopioid

Three RCTs[77,138,169] compared opioids (IV morphine,[77] oral tramadol,[169] or oral codeine[138]) versus a nonopioid medication (IV dexketoprofen, oral acetaminophen, oral tizandine). Two trials were rated low[77,169] and one was rated high[138] risk of bias. Opioid treatments varied across trials, as did nonopioids. One trial each evaluated tramadol[169] (plus diclofenac 100 mg), IV morphine,[77] and oxycodone[138] (plus aspirin) (doses not reported). In the trial of tramadol plus diclofenac, the comparator group (tizanidine) also received diclofenac therefore we considered this a comparison of tramadol versus tizanidine.[169] The MED ranged from 21 to 49 mg in two trials,[77,169] and was not calculable in one trial (no information on opioid dosage).[138] Treatment duration ranged from a single dose in one trial[77] to 14 days in the other two trials[138,169]

Across trials, mean patient age ranged from 23 to 49 years and populations were predominately male (proportion female: 0%,[138] 39%[77] and 42%[169]). Mean baseline duration of LBP was not reported in any trial. Baseline pain intensity ranged from 7.6 to 8.3 (0 to 10 scale) across two trials (one trial[138] did not report this). Most patients did not have radiculopathy; two trials excluded such patients,[138,169] and only 9.6% of patients in the third trial had radiculopathy. History of substance abuse disorder and psychiatric comorbidities was not reported.

Function

One trial (N=269) reported no differences in function (RDQ scores) between tramadol (MED 40 mg) plus diclofenac 100 mg and tizanidine 2 mg at 1, 3 or 4 weeks,[169] (Table 25).

Pain Intensity

Two trials (N=406) reported pain intensity for opioids versus nonopioids (Table 25). One trial (N=137) found single-dose IV morphine 0.1 mg/kg (MED 21 mg) associated with moderate improvement in pain intensity (0-10 VAS) versus IV dexketoprofen 50 mg at 30 minutes (MD -1.21, 95% CI -1.99 to -0.43); the difference versus acetaminophen 1,000 mg was similar but imprecise (MD -0.35, 95% CI -1.18 to 0.48).[77] Another trial (N=269) reported no differences in pain at rest or during activity between tramadol (MED 40 mg) plus diclofenac 100 mg and tizanidine 2 mg at 1, 3 or 4 weeks.[169]

Neither trial reported likelihood of experiencing a positive pain outcome (e.g., >30% improvement in pain)..Return to Work

Two trials (N=344) reported no difference in return to work outcomes for opioids versus nonopioids (Appendix E). One trial (n's not reported by group [N=75 total across 3 arms])[138] compared codeine 36 mg MED versus acetaminophen (dosages NR) at 2 weeks and another (N=269) tramadol 40 mg MED plus diclofenac 100 mg versus tizanidine 2 mg at 7 or 28 days.[169]

Harms

Serious Adverse Events

Serious adverse events and study withdrawal due to adverse events were not reported.

Any Adverse Event

One trial (N=137) found that intravenous morphine 0.1 mg/kg (MED 21 mg) was associated with similar likelihood of adverse events compared to dexketoprofen 50 mg (15.5% v. 8.7%; RR 1.79, 95% CI 0.56 to 5.69) or acetaminophen 1,000 mg (15.5% vs. 8.7%; RR 1.79, 95% CI 0.56 to 5.69) at 30 minutes. Specific adverse events were not reported.[77] Another trial (N=269)[169] found an increased risk of some, but not all, adverse events with tramadol 40 mg MED (plus diclofenac 100 mg) versus tizanidine 2 mg: dizziness (51.1 vs. 30.4%, RR 1.68, 95% CI 1.16 to 2.43), nausea (37.5% vs. 17.4%, RR 2.16, 95% CI 1.28 to 3.63), and vomiting (15.9% vs. 2.2%, RR 7.32, 95% CI 1.71 to 31.28) (Table 25).

Table 25. Opioid versus nonopioid: function, pain and adverse events

Intervention and Comparison	Author Year (Patients)	Outcome	Results Opioid vs. Nonopioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
Opioid vs. NSAID *IV Morphine vs. IV dexketoprofen*	Eken, 2014 (N=91)	Pain, VAS scores (0-10)	30 mins.: N=88, 1.55 (1.6) vs. 2.76 (2.04), MD -1.21, 95% CI -1.99 to -0.43
		Any AE	30 mins.:15.5% (4/45) vs. 8.7% (4/46); RR 1.79, 95% CI 0.56 to 5.69
Opioid vs. Acetaminophen *IV Morphine vs. IV Acetaminophen*	Eken, 2014 (N=91)	Pain, VAS scores (0-10)	30 mins.: N=88, 1.55 (1.6) vs. 1.90 (2.24), MD -0.35, 95% CI -1.18 to 0.48
		Any AE	30 mins.: 15.5% (7/45) vs. 8.7% (4/46); RR 1.79, 95% CI 0.56 to 5.69

Intervention and Comparison	Author Year (Patients)	Outcome	Results Opioid vs. Nonopioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
Opioid vs. Muscle relaxant *Tramadol + diclofenac vs. tizanidine**	Hung, 2024 (N=192)	Function, RDQ scores (0-24)	1 week: 7.2 (5.8) vs. 8.4 (6.4), MD -1.20, 95% CI -2.94 to 0.54 3 weeks: 4.6 (6.5) vs. 4.0 (4.8), MD 0.60, 95% CI -1.02 to 2.22 4 weeks: 4.6 (5.7) vs. 3.8 (4.7), MD 0.80, 95% CI -0.68 to 2.28
		Pain at rest, NRS scores (0-10)	1 week: 2.1 (2.0) vs. 2.0 (2.4), MD 0.10, 95% CI -0.53 to 0.73 3 weeks: 1.1 (1.4) vs. 1.6 (1.9), MD -0.50, 95% CI -0.98 to -0.02 4 weeks: 1.2 (1.8) vs. 1.0 (1.8), MD 0.20, 95% CI -0.31 to 0.71
		Pain during activity, NRS scores (0-10)	1 week: 3.8 (2.6) vs. 3.5 (2.6), MD 0.30, 95% CI -0.44 to 1.04 3 weeks: 2.2 (2.5) vs. 2.3 (2.1), MD -0.10, 95% CI -0.76 to 0.56 4 weeks: 1.9 (2.5) vs. 2.0 (2.3), MD -0.10, 95% CI -0.78 to 0.58
		Any AE	4 weeks Dizziness: 51.1 vs. 30.4%, RR 1.68, 95% CI 1.16 to 2.43 Nausea: 37.5% vs. 17.4%, RR 2.16, 95% CI 1.28 to 3.63 Vomiting: 15.9% vs. 2.2%, RR 7.32, 95% CI 1.71 to 31.28 Sleepiness: 52.3% vs. 57.6%, RR 0.91, 95% CI 0.70 to 1.18 Stomachache: 13.6% vs. 15.2, RR 0.90, 95% CI 0.44 to 1.83 Indigestion: 13.6% vs. 10.9%, RR 1.25, 95% CI 0.57 to 2.76

AE = adverse events; CI = confidence interval; MD = mean difference; NRS = numerical rating scale; NSAID = non-steroid anti-inflammatory drug; RDQ = Roland Morris Disability Questionnaire; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

Key Question 3b, e, h: Systemic Opioid Plus Nonopioid

Detailed Synthesis

Systemic Opioid Plus Nonopioid Versus Nonopioid

Four head-to-head trials (N=461) compared oral opioids in combination with a nonopioid (codeine plus acetaminophen;[105] oxycodone plus acetaminophen;[89,113] or oxycodone plus aspirin[138]) versus a nonopioid medication (diflunisal; ketorolac; acetaminophen; or cyclobenzaprine).[89,105,113,138] Two trials were rated low,[89,113] one moderate[105] and one high[138] risk of bias. The MED ranged from 22.5 mg to 45 mg across three trials[89,105,113] (MED not calculable in one trial[138]) and treatment duration ranged from 1 to 12 weeks.

Across trials, mean patient age ranged from 23 to 39 years, and most were male (proportion female: 0% in 1 trial[138] and a range of 21% to 55% across 3 trials[89,105,113]). Mean baseline duration of LBP ranged from 2.3 to 3.4 days across two trials,[105,113] and was not reported in the other two.[89,138] Baseline pain intensity was only reported in one trial: mean 6.7 on a 0 to 10 scale.[89] Radiculopathy was excluded in two trials,[113,138] and not reported in the other two.[89,105] Substance use disorders and psychiatric comorbidities were not reported in any trials.

Function

At 1 week, one head-to-head trial (N=216) found oxycodone 5 mg (MED 22.5 mg) plus acetaminophen 325 mg every 12 hours associated with similar improvement on the RDQ compared with cyclobenzaprine 5 mg every 8 hours (mean 7.8 vs. 8.2; difference -0.4, 95% CI -2.7 to 1.1) (Table 26).[113]

No trial reported likelihood of experiencing a positive function outcome (e.g., >30% improvement in RDQ). Pain Intensity

Two head-to-head trials (N=170) compared codeine plus acetaminophen versus another medication and reported pain using different measures (Table 26). One trial (N=40) reported no differences in pain intensity (0-2 scale) between an initial dose of codeine 60 mg (MED 45 mg) plus acetaminophen 600 mg (titrated to 30 mg + 300 mg, respectively, every 4 hours) versus an initial dose of diflunisal 1,000 mg (titrated to 500 mg every 4 hours) at 5, 12 and 15 days; data was limited and was estimated from a figure,[105] The other trial (N=123) reported a very imprecise estimate for codeine 60 mg (MED 45 mg) plus acetaminophen 600 mg every 4 hours versus ketorolac 10 mg every 4 to 6 hours in pain intensity at 6 hours (VAS; MD -0.5, 95% CI -9.02 to 8.02).[89]

No trial reported likelihood of experiencing a positive pain outcome (e.g., >30% improvement in pain).

Quality of Life and Pain Relief Satisfaction

One trial (N=216) found no difference in pain relief satisfaction between oxycodone 5 mg (MED 22.5 mg) in combination with acetaminophen 325 mg and cyclobenzaprine 5 mg[113] every 8 hours at 1 week or 3 months. Another trial (N=47) found no difference in satisfaction between codeine 60 mg (MED 45 mg) in combination with acetaminophen 600 (titrated to 30 mg and 300 mg, respectively) and diflunisal 1,000 mg (titrated to 500 mg) every 4 hours (very good or excellent acceptance: 43% vs. 47%; RR 0.90, 95% CI 0.46 to 1.79)[105] (Appendix E).

Return to Work and Medication Use

One trial (N=216) compared oxycodone 5 mg (MED 22.5 mg) plus acetaminophen 325 mg versus cyclobenzaprine 5 mg and found no differences between treatment groups in return-to-work outcomes at 1 week (Appendix E) or in medication use at 1 week (difference 3%, 95% CI -10 to 16%) or 3 months (difference 1%, 95% CI -2 to 4).[113] Another trial (n's not reported by group [N=75 total across 3 arms]) compared oxycodone in combination with aspirin versus acetaminophen (dosages NR) and found no differences between groups in return-to-work outcomes at 2 weeks (Appendix E).[138]

Harms

Severe Adverse Events and Withdrawal Due To Adverse Events

One trial (N=121) found codeine 60 mg (MED 45 mg) plus acetaminophen 600 mg every 4 hours associated with a substantially increased likelihood of severe adverse events (also referred to as serious adverse events by the authors), defined as incapacitating symptoms that precluded work or normal activities, that were prolonged, that led to discontinuation of study drugs, or that required medical evaluation or treatment, (17% vs. 3%; RR 5.25, 95% CI 1.20 to 23.0) and withdrawals due to adverse events (0% vs. 11.9%; p=0.005) versus ketorolac 10 mg every 4 to 6 hours at 1 week, though the estimates were imprecise[89] (Table 26).The trial was not designed to assess opioid use disorder, overdose, or long-term harms.

Any Adverse Event

Three trials reported adverse events (Table 26).[89,105,113] Two trials found opioid treatment associated with a similar likelihood of adverse events: one trial (N=216) compared oxycodone 5 mg (MED 22.5 mg) plus acetaminophen 325 mg versus cyclobenzaprine 5 mg at 3 months (40% vs. 33%; RR 1.19, 95% CI 0.84 to 1.70)[113] and the other trial (N=47) compared codeine 60 mg (MED 45 mg) plus acetaminophen 600 (titrated to 30 mg and 300 mg, respectively, every 4 hours) and diflunisal 1,000 mg (titrated to 500 mg every 4 hours) at 15 days (23.8% vs. 15.8%; RR 1.51, 95% CI 0.42 to 4.58).[105] One trial (N=121) of codeine 60 mg (MED 45 mg) plus acetaminophen 600 mg every 4 hours reported moderately increased likelihood of adverse events with opioid treatment versus ketorolac 10 mg every 4-6 hours at 1 week (64% vs. 34%; RR 1.90, 95% CI 1.28 to 2.83).[89]

Specific Adverse Events

One study (N=216) found oxycodone 5 mg (MED 22.5 mg) in combination with acetaminophen 325 mg every 8 hours associated with a higher likelihood of drowsiness (15% vs. 7%; difference 8%, 95% CI 0 to 16%), dizziness (15% vs. 3%; difference 12%, 95% CI 5 to 19), and nausea or vomiting (18% vs. 4%; difference 14%, 95% CI 6 to 22) compared with cyclobenzaprine 5 mg every 8 hours.[113] A trial (N=47) of codeine 60 mg (MED 45 mg) plus acetaminophen 600 (titrated to 30 mg and 300 mg, respectively, every 4 hours) reported greater likelihood of dizziness versus diflunisal 1,000 mg (titrated to 500 mg every 4 hours), though the estimate was highly imprecise and did not reach significance (9.5% vs. 0%; RR 4.55, 95% CI 0.23 to 89.08); no differences were observed for drowsiness or nausea.[105]

Table 26. Opioid plus nonopioid versus nonopioid: function, pain and adverse events

Intervention and Comparison	Author Year (Patients) Treatments	Outcome	Results Opioid + Nonopioid vs. Nonopioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
Opioid + nonopioid vs. NSAID	Innes, 1998 (N=123) *Codeine + acetaminophen vs. ketorolac*	Pain, VAS/NRS scores (0-10 scale)	6 hours: 6.21 (2.33) vs. 6.16 (2.31), MD 0.05, 95% CI -0.78 to 0.88
		Any AE	7-9 days: 64% (38/59) vs. 34% (21/62); RR 1.90, 95% CI 1.28 to 2.83
		Serious AEs	7-9 days: 16.9% (10/59) vs. 3.2% (2/62); RR 5.25, 95% CI 1.20 to 23.0
		Withdrawal due to AEs	7-9 days: 11.9% (7/59) vs. 0% (0/62); p=0.005

Intervention and Comparison	Author Year (Patients) Treatments	Outcome	Results Opioid + Nonopioid vs. Nonopioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
	Brown, 1986 (N=40) *Codeine + acetaminophen vs. diflunisal*	Pain (0-2 scale; 0=no pain, 1=mild, 2=moderate)	5 days: 0.8 (NR) vs. 1.1 (NR) 12 days: 0.4 (NR) vs. 1.0 (NR) 15 days: 1.1 (NR) vs. 0.5 (NR) Estimated from graph, p=NR
		Any AE	2 weeks: 23.8% (5/21) vs. 15.8% (3/19); RR 1.51, 95% CI 0.42 to 4.58
Opioid + nonopioid vs. muscle relaxant	Friedman, 2015 (N=216) *Oxycodone + acetaminophen vs. cyclobenzaprine*	Function, RDQ scores (0-24 scale)	1 week: 7.8 (8.5) vs. 8.2 (8.5); MD -0.4, 95% CI -2.7 to 1.2
		Any AE	40% (43/108) vs. 33% (36/108); RR 1.19, 95% CI 0.84 to 1.70

AE = adverse events; CI = confidence interval; MD = mean difference; NR = not reported; NRS = numerical rating scale; NSAID = non-steroid anti-inflammatory drug; RDQ = Roland Morris Disability Questionnaire; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

Systemic Opioid Plus Nonopioid Versus Opioid

Two trials (N=537) compared an opioid in combination with a nonopioid versus the same opioid alone. One trial (N=119) compared tramadol 37.5 mg [15 mg MED] plus acetaminophen 325 mg versus tramadol 50 mg [20 mg MED] alone.[162] Patients in both groups received four daily doses for the first 2 days (mean doses were 134.4 mg tramadol plus 1,165 mg acetaminophen versus 177.7 mg tramadol alone) after which they could increase as needed up to a maximum of 8 doses per day (mean doses over days 3-10 were 172.5 mg tramadol plus 1,495 mg acetaminophen versus 257.4 mg tramadol alone). The MED for each group in this trial for the first 2 days was 15 mg MED and 20 MED respectively, followed by 17.25 mg MED and 25.74 MED for days 3-10. The other trial compared tramadol 75 mg (15 mg MED) plus dexketoprofen 25 mg versus tramadol 100 mg (20 mg MED) alone.[168] This trial measured outcomes at 8 hours after only a single dose of medication (single-dose phase) and after 8 hours up to 5 days after multiple doses (i.e., every 8 hours; multiple-dose phase).

Mean patient age ranged from 43 to 55 years and most were female (52% to 58%). Mean baseline pain severity was 7.0 (on a 0-10 scale) in one trial;[168] the other trial did not report this. One trial reported a mean pain duration of 22 days,[162] while the other only reported ≤48 hours of pain as inclusion criterion. One trial[168] reported that 51% of patients had radiculopathy, while the other reported 0%.[162] Neither trial reported on history of substance abuse disorders or psychiatric comorbidities.

Function

At 5 days, one trial (N=304) found greater improvement in RMD scores (0-24) with tramadol (MED 15 mg) plus dexketoprofen 25 mg versus tramadol (MED 20 mg) alone every 8 hours, however the difference was not statistically significant (adjusted MD -5.70, 95% CI -13.2 to 1.8)[168] (Table 27).

Pain Intensity and Satisfaction

The combination of tramadol 37.5 mg (MED 15 mg) and acetaminophen 325 mg was associated with very similar pain intensity (mean VAS, 2.79 vs. 2.48, p=0.46) and self-reported adequate pain relief (82% vs. 83%; RR 0.98, 95% CI 0.81 to 1.18) compared to tramadol 50 mg (MED 20 mg) alone, respectively, at day 10.[162] Patient satisfaction with treatment was also

similar between groups in this trial (73% vs. 73%; RR 1.00, 95% CI 0.78 to 1.27). Another trial reported greater improvement in TOTPAR scores for tramadol (MED 15 mg) plus dexketoprofen 25 mg versus tramadol (MED 20 mg) alone at 8 hours after a single dose (MD 1.59, 95% CI 0.34 to 2.84) and after 8 hours up to 3 days after multiple doses (MD 11.99, 95% CI 0.72 to 23.26); however, the effect size is unclear and estimates were very imprecise[168] (Table 27). Neither trials reported quality of life outcomes.

Serious Adverse Events

Tramadol (MED 15 mg) plus dexketoprofen 25 mg was associated with a similar risk of serious adverse events compared with tramadol alone (MED 20 mg) at 8 hours after a single dose (N=418; 0.5% [1/211] vs. 0%) and after 8 hours up to 5 days after multiple doses (N=538; 0% in both groups);[168] the one severe adverse event which occurred in the combination group was urinary calculus and was deemed unrelated to treatment (Table 27).

Withdrawals Due To Adverse Events and Any Adverse Events

Tramadol plus acetaminophen was associated with a substantially decreased risk of withdrawals due to adverse events (8.5% vs. 21.7%; RR 0.39, 95% CI 0.15 to 1.03) and a moderately decreased risk of any adverse events (N=119; 50.8% vs. 73.3%; RR 0.69, 95% CI 0.52 to 0.93) compared with tramadol alone at 10 days in one trial.[162] In another trial, the risk of withdrawal due to adverse events was similar for tramadol plus dexketoprofen versus tramadol alone at 8 hours (N=418; 2.8% vs. 3.4%; RR 0.84, 95% CI 0.29 to 2.46) and between 9 hours and 5 days (N=538; 6.7% vs. 7.5%; RR 0.89, 95% CI 0.48 to 1.65), as was the risk of any adverse event at 8 hours (N=418; 13.3% vs. 15.0%; RR 0.89, 95% CI 0.55 to 1.42) and between 9 hours and 5 days (N=538; 25.2% vs. 27.2%; RR 0.92, 95% CI 0.70 to 1.23)[168] (Table 27). In the latter trial, patients received a single dose up to 8 hours and multiple doses after 8 hours.

Specific Adverse Events

The combination of tramadol and acetaminophen was associated with substantially lower likelihood of nausea (13.6% vs. 35%; RR 0.38, 95% CI 0.19 to 0.80) and dizziness (5.1% vs. 25%; RR 0.20, 95% CI 0.06 to 0.67) than tramadol alone.[162]

Table 27. Opioid plus nonopioid versus opioid: function, pain and adverse events

Intervention and Comparison	Author, Year (Patients)	Outcome	Results Opioid + Nonopioid vs. Opioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
Opioid + nonopioid vs. opioid *Tramadol + dexketoprofen vs. tramadol*	Varrassi, 2024* Single dose phase (N=418) Multiple dose phase (N=538)	Function, RDQ scores (0-24 scale)	5 days: N=304, adjusted MD –5.70, 95% CI –13.2 to 1.8
		Pain scores, TOTPAR (scale unknown)	8 hours: N=408, 10.1 (NR) vs. 8.5 (NR); MD 1.59; 95% CI 0.34 to 2.84 3 days: N=408, means NR; MD 11.99, 95% CI 0.72 to 23.26
		Any AE	8 hours: 13.3% (28/211) vs. 15.0% (31/207); RR 0.89, 95% CI 0.55 to 1.42
			9 hours to 5 days: 25.2% (68/270) vs. 27.2% (73/268); RR 0.92, 95% CI 0.70 to 1.23
		Serious AE	8 hours: 0.5% (1/211) vs. 0% (0/207)
			9 hours to 5 days: 0% (0/270) vs. 0.4% (1/268)
		Withdrawal due to AEs	8 hours: 2.8% (6/211) vs. 3.4% (7/207); RR 0.84, 95% CI 0.29 to 2.46

Intervention and Comparison	Author, Year (Patients)	Outcome	Results Opioid + Nonopioid vs. Opioid Mean (SD) and MD (95% CI) or % (n/N) and RR (95% CI)
			9 hours to 5 days: 6.7% (18/270) vs. 7.5% (20/268); RR 0.89, 95% CI 0.48 to 1.65
Opioid + nonopioid vs. opioid *Tramadol + acetaminophen vs. tramadol*	Perrot, 2006 (N=119)	Pain VAS scores (0-10 scale)	10 days: 2.79 (NR) vs. 2.48 (NR), p=0.46
		Any AE	10 days: 50.8% (30/59) vs. 73.3% (44/60); RR 0.69, 95% CI 0.52 to 0.93
		Withdrawal due to AEs	10 days: 8.5% (5/59) vs. 21.7% (13/60); RR 0.39, 95% CI 0.15 to 1.03

AE = adverse events; CI = confidence interval; MD = mean difference; NR = not reported; RDQ = Roland Morris Disability Questionnaire; RR = risk ratio; VAS = visual analogue scale.
* Varrassi, et al. included a third arm which received placebo. Patients randomized to tramadol/dexketoprofen fixed combination or tramadol alone during the single-dose phase (8 hours) continued to receive the same treatment during the multiple-dose phase (>8 hours to 5 days); however, patients randomized to placebo during the single-dose phase received either the tramadol/dexketoprofen fixed combination or tramadol alone during the multiple-dose phase.

Key Question 3c. What is the comparative effectiveness of systemic opioid therapy combined with a nonpharmacologic intervention versus (1) placebo or sham; (2) the same opioid alone; (3) the same nonpharmacologic intervention alone; or (4) a different nonpharmacologic therapy on health outcomes at short term and long term?

No trials were identified that met inclusion criteria.

Key Question 3d. What are the effects of prescribing opioid therapy versus not prescribing opioid therapy for ALBP on continued need for prescription pain relief, such as need for opioid refills, short term and long-term opioid use?

Description of Included Studies

Two NRSI, both retrospective claims database studies conducted in the United States, provided data regarding long-term opioid use and disability.[172,176] The studies compared workers' compensation claims for ALBP that involving early opioid prescribing versus claims without early opioid prescribing and evaluated long-term opioid prescribing and disability duration as outcomes. Both were rated as moderate risk of bias due to baseline imbalance and lack of attrition reporting.

Detailed Synthesis

Both studies reported an association between opioid therapy for ALBP and continued opioid use. One study (N=8,443) found early opioid use (defined as within 15 days following LBP onset) associated with an increased likelihood of late opioid use (defined as 5 or more opioid prescriptions from 30 to 730 days following onset) compared to nonuse.[176] This study found a dose-response relationship between higher early opioid exposure and increased likelihood of subsequent long-term use, after adjusting for age, sex, job tenure, and injury severity, compared to no early opioid use, the adjusted OR was 2.08 (95% CI 1.55 to 2.78) for 1 mg to 140 mg morphine equivalents and increased to 6.14 (95% CI 4.92 to 7.66) for ≥450 mg. Another study (N=2,887) of ALBP evaluated in the ED reported that receiving any opioid prescription within 2 days of the ED visit was associated with higher likelihood of long-term use (defined as ≥3 opioid

prescriptions filled between 4 days and 12 months from injury onset), after adjusting for injury severity, early MRI, age, sex, and job tenure (adjusted RR 1.29, 95% CI 1.05 to 1.58).[172] This study did not evaluate the association between dose of early opioid use and likelihood of long-term use.

Key Question 3f. Does effectiveness of a systemic opioid (alone or in combination with a nonopioid) vary based on prescribing strategy (dose, frequency, dosing interval, duration of therapy, quantity dispensed, type of opioid dispensed [e.g., long-acting, sustained release]), and what is the impact of different prescribing strategies on refill requests and quantity of pills?

No trials were identified that met inclusion criteria.

Key Question 3g. Do the harms of a systemic opioid (alone or in combination with a nonopioid) vary based on prescribing strategy (dose, frequency, dosing interval, duration of therapy, quantity dispensed, type of opioid dispensed [e.g., long-acting, sustained release]), and what is the impact of different prescribing strategies?

No trials were identified that met inclusion criteria.

Key Question 4. In adults with ALBP (<6 weeks duration), including back pain with radiculopathy, being considered for opioid therapy or who have been prescribed opioids for ALBP, what are the impacts of (a) policy and patient factors on the decision to prescribe opioids; (b) factors on patient health outcomes; (c) shared decision making on opioid prescription strategy, health outcomes, continued opioid use or harms; (d) accuracy of instruments for predicting risk of opioid misuse, withdrawal, opioid use disorder, or overdose; and (e) effectiveness of instruments for predicting risk of long-term use of opioids, opioid misuse, withdrawal, opioid use disorder, or overdose in patients with ALBP versus usual care (i.e., not using a formal instrument)?

No studies in patients with ALBP were identified to address any part of Key Question 4.

Two retrospective studies[234,235] that evaluated the impact of prescription drug monitoring programs (PDMP) in ED did not specifically address patients with ALBP (<6 weeks duration).

One retrospective study in adults presenting to a community emergency ED with back pain evaluated the impact of a state legislative mandate requiring providers to review data from a PDMP prior to prescribing opioids.[234] The population's mean pain score at baseline was 8 (0-10 scale) and 39% of patients were not opioid-naïve. Opioid prescriptions prior to and after the enactment of the law were compared. There was no difference in pre-PDMP and post-PDMP rates of opioid prescriptions at discharge from the ED (46% vs. 48%) or in the median MME prescribed (mean difference 8 mg, 95% CI -9 to 24 mg). Back pain duration was not specified, and the study did not meet inclusion criteria.

Another retrospective study in Medicaid beneficiaries presenting with pain to 86 EDs in Washington State evaluated the impact of an automated PDMP.[235] There was no difference in pre- and post-implementation rates of opioid prescriptions or in amounts prescribed. Only 7% of

patients presented with back pain and symptom duration was not specified; this study did not meet inclusion criteria.

Key Question 5. Effectiveness and Harms of Nonopioid Pharmacologic Therapy for ALBP

Forty-seven RCTs evaluated nonopioid pharmacologic therapies for the treatment of ALBP.[57,58,67-79,83,85-87,91,93,95,97,99-101,104,108,113,125,126,128,132,137,140-144,151,153-156,160,161,163]

Single Nonopioid Therapy (Key Questions 5a, 5d, 5e)

Thirty-five RCTs evaluated a single nonopioid pharmacologic therapy for the treatment of ALBP:[67,68,70,73-79,83,85,91,93,95,97,99-101,104,108,125,132,137,140,143,144,151,153-156,160,161,163] 18 evaluated an NSAID,[67,68,75-77,83,91,93,95,99-101,108,143,144,153-155] nine evaluated a muscle relaxant,[70,74,75,97,104,132,151,161,163] six evaluated a systemic steroid,[79,85,137,140,156,160] two evaluated acetaminophen,[73,125] and one evaluated cannabidiol.[78]

Nonsteroidal Anti-Inflammatory Drugs

Eighteen RCTs evaluated an NSAID.[67,68,75-77,83,91,93,95,99-101,108,143,144,153-155]

NSAID Versus Placebo

Description of Included Studies

Ten RCTs (N=2,110) compared an NSAID versus placebo (Table 28, Appendix E).[68,75,83,91,93,95,100,144,154,155] Patient mean age ranged from 38 to 50 years and 41% to 56% were female. Four trials were restricted to patients with radiculopathy or sciatica,[68,95,154,155] three trials excluded patients with radiculopathy,[83,91,100] and three trials did not specify inclusion or exclusion of patients with radiculopathy.[75,93,144] The baseline duration of the LBP episode was <3 days in four trials,[91,93,100,154] 6 to 9 days in two trials,[68,83,144] and not described in four trials (inclusion criteria specified a duration of <3 days in two trials[68,75] and ≤3 weeks in two trials[95,155]). Mean baseline pain intensity ranged from 5.4 to 8.4 (0-10) in 7 RCTs.[68,83,93,95,100,144,154] In most trials (6 RCTs), a history of previous LBP episodes was common (range, 40% to 77%; one trial only reported mean number of prior episodes: 3.7).[68,83,91,93,100,154] History of substance abuse disorders and psychiatric comorbidities were not consistently reported but were an exclusion criterion when mentioned.

The type of NSAID varied as did the doses, regimen and treatment durations (most typically 5-10 days) (Table 28); all were nonselective NSAIDs, except for meloxicam which is partially selective, and were given orally. Three trials had multiple arms and evaluated two different NSAIDs[68,100] or different doses of the same NSAID.[154] All but one trial[75] mentioned concomitant treatments: four trials did not give or permit any rescue medication or additional treatment,[68,93,154,155] and five trials provided rescue medication (acetaminophen, with or without codeine in one trial[95]),[83,91,95,100,144] but two of these trials terminated patients who used any rescue medication.[95,100] Two older trials prescribed bed rest (up to 7 days) in all patients.[93,155]

Five trials were assessed as low risk of bias,[68,83,100,144,154] four moderate risk of bias[91,93,95,155] and one high risk of bias (Appendix F).[75] Primary methodological limitations in the moderate and high risk of bias trials included unclear randomization and concealments methods and

baseline differences between groups (or unclear if similar at baseline). The trial at high risk of bias was older and was poorly reported overall.

Table 28. Study characteristics of NSAID versus placebo

Study, Year	Dreiser, 2001	Dreiser, 2003	Herrmann, 2009	Basmajian, 1989	Goldie, 1968	Amilie, 1987	Szpalski, 1994	Predel, 2019	Hancock, 2007	Weber, 1993
N Randomized	532	372	171	89	50	282	73	379 (207)*	120	214
Mean Age (Years)	47	41	50	NR	NR	38	38	45	41	48
Female	56%	50%	56%	NR	48%	41%	36%	59%	44%	NR
Radiculopathy	100%	Excluded	100%	NR	100%	Excluded	NR	NR	Excluded	100% (sciatica)
Pain Duration Inclusion Criteria	≤3 days	≤2 days	<72 hours	≤2 days	≤3 weeks	≤2 days	<2 weeks	<21 days	<6 weeks	<14 days
Mean Pain Duration (days)	3 days: 93%	1 day: 60% 2 days: 98%	NR	NR	NR	26.4 hours	Median 2 days	6.4 days	9.1 days	NR
Mean/Median Baseline Pain Severity (0-10)	7.6	7.2	8.4	NR	NR	NR	7.3	6.8	6.6	5.4
Previous LBP episodes	51%	40%	77% (mean 4.9 previous episodes)	NR	NR	65%	44%	NR†	Mean 3.7 previous episodes	NR
History of substance abuse disorder	NR	Excluded	Excluded	NR	NR	NR	NR	NR	NR	Excluded
Psychiatric Comorbidities	NR	NR	Excluded	NR	NR	NR	NR	NR	NR	Excluded
NSAID	Meloxicam	(1) Ibuprofen (2) Diclofenac	(1) Lornoxicam (2) Diclofenac	Diflunisal	Indomethacin	Piroxicam	Tenoxicam	Ibuprofen	Diclofenac	Piroxicam
Dosage Frequency	(1) 7.5 or (2) 15 mg; 1x/day	(1) 400 mg, (2) 25 mg; 3x/day	(1) 8 mg, 2x/day (2) 100-150 mg, 2-3x/day	500 mg, 2x/day	25 mg, 3x/day	20 mg, 1x/day	20 mg, 1x/day	400 mg, 3x/day	50 mg, 2x/day	40 mg 1x/day, 2 days; 20 mg 1x/day, 12 days
Treatment duration	7 days	7 days	5 days	7-10 days	14 days	7 days	14 days	6 days	4 weeks	14 days

Study, Year	Dreiser, 2001	Dreiser, 2003	Herrmann, 2009	Basmajian, 1989	Goldie, 1968	Amilie, 1987	Szpalski, 1994	Predel, 2019	Hancock, 2007	Weber, 1993
Duration follow-up	7 days	7 days	5 days	7-10 days	14 days	2 weeks	15 days	6 days	12 weeks	12 months
Setting	NR	Primary care	Primary care	NR	Hospital	Primary care, Rheumatology	NR		Primary care, Physiotherapist	Primary care
Country	Various (Europe)	France	Germany	Canada	Sweden	Norway	Belgium	Germany	Australia	Norway
Quality	Low	Low	Low	High	Moderate	Moderate	Moderate	Low	Low	Moderate

NR = not reported; NSAID = nonsteroid anti-inflammatory drug.
* Both neck and back were included; 207 patients randomized had back pain only. Demographics were reported for the overall population only, not by sub-group.
† Excluded if history of ≥3 or more episodes of back or neck pain in the last 6 months.

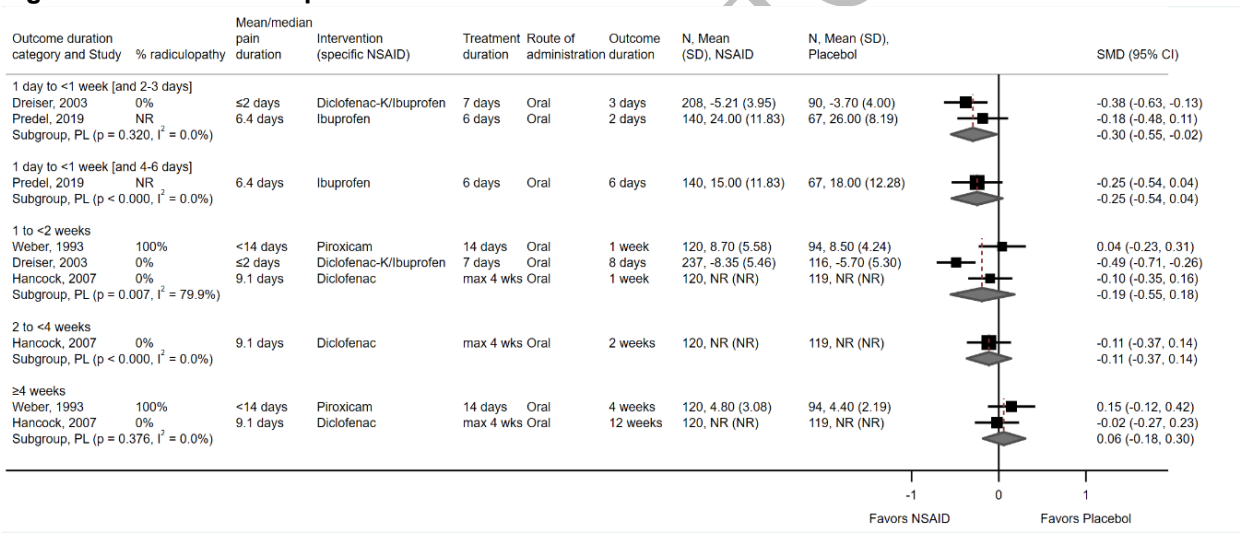
Detailed Synthesis

Function

Four RCTs(N=1,013) assessed function based on average changes in ODI (0-100 scale) or RDQ (0-24 scale) scores.[83,95,100,144] Only one or two trials were available for every timepoint except 1 to <2 weeks (3 trials). At <1 week, NSAIDs were associated with a small improvement in function, however this did not persist to later time frames up to 12 weeks across trials (Figure 4). Substantial heterogeneity (80%) is seen at 1 week to <2 weeks with one good quality (low risk of bias), large outlier trial[100] reporting a small effect with NSAID versus placebo at 8 days. Mean pain duration at baseline was shorter (≤48 hours) in this RCT than in the other two trials (9.1 and <14 days). Although two NSAIDs (diclofenac-K and ibuprofen) were combined into one intervention arm for the pooled analyses, each showed similar results compared with placebo. All trials evaluated nonselective NSAIDs. Across trials using the RDQ (0-24), differences were as follows: <1 week (1 RCT, MD -1.51, 95% CI to -2.49 to -0.52, small effect),[100] 1 to <2 weeks (2 RCTs, MD -1.61, 95% CI -4.18 to 1.02),[83,100] 2 to <4 weeks (1 RCT, MD -0.60, 95% CI -1.94 to 0.74)[83] and ≥4 weeks (1 RCT, MD -0.10, 95% CI -1.29 to 1.09).[83]

No trial reported the proportion of patients achieving a successful functional outcome.

Figure 4. NSAID versus placebo: ODI and RDQ function scores



CI = confidence interval; NR = not reported; NSAID = nonsteroid anti-inflammatory drug; ODI = Oswestry Disability Index; PL = profile likelihood; RDQ = SD = standard deviation.
For Dreiser 2003, the two NSAID groups (diclofenac-K and ibuprofen) were combined into one group.

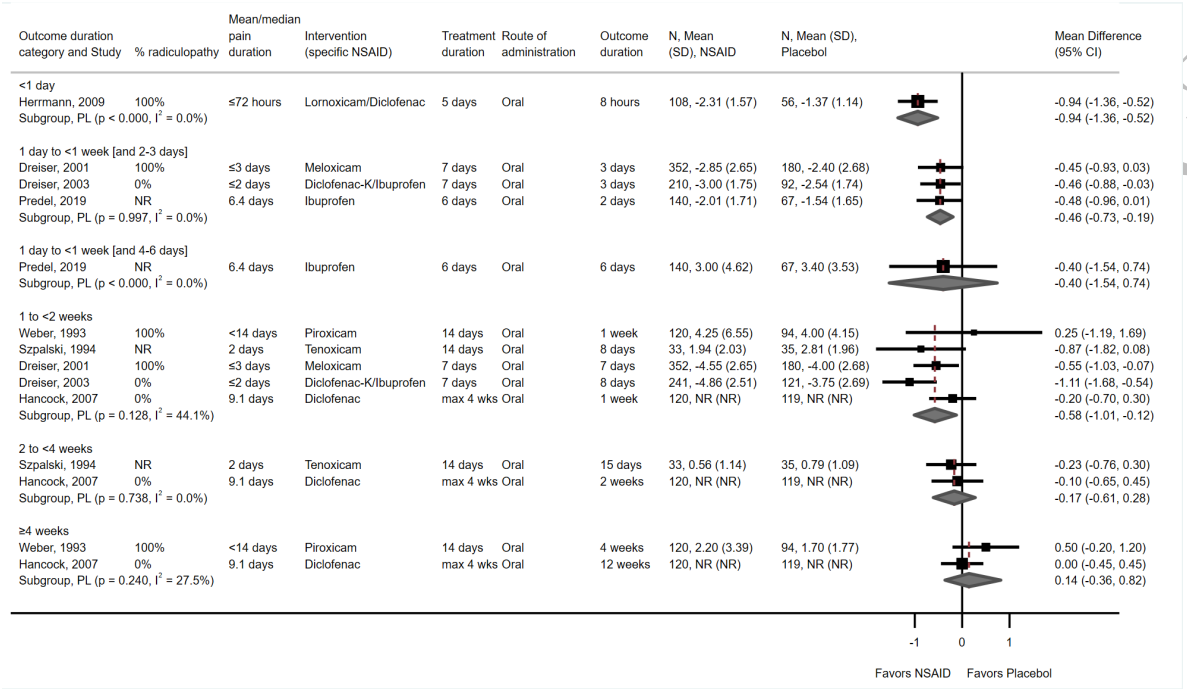
Pain

Seven RCTs (N=1,786) assessed pain based on average changes in VAS or NRS (0-10) pain scores (Figure 5).[68,83,93,95,100,144,154] At <1 day and 1 day to <2 weeks, NSAIDS were associated with small improvement in pain scores versus placebo (MD/pooled MD range -0.46 to -0.94, scale 0-10); at 1 day to <1 week, effects favored NSAIDs but were slightly below the threshold for a small effect. Heterogeneity (I^2=44.1%) was present at 1 to <2 weeks. When the trials were stratified by baseline pain duration, trials with pain duration ≤3 days found NSAIDs associated with a small improvement in pain compared with placebo[93,100,154] (3 RCTs, N=962, pooled MD -0.79, 95% CI -1.30 to -0.35, I^2=8.6%), but across trials with longer pain durations (mean 9.1 days and <14 days) no effect was seen[83,95] (2 RCTs, N=453, pooled MD 95% CI -0.15, 95% CI -

0.75 to 0.63 I^2=0%); heterogeneity was eliminated or significantly reduced. Benefits were not observed at 2 to <4 weeks and ≥4 weeks.

No trial reported the proportion of patients achieving a successful pain outcome.

Figure 5. NSAID versus placebo: VAS/NRS pain scores



CI = confidence interval; NR = not reported; NRS = Numeric Rating Scale; NSAID = nonsteroid anti-inflammatory drug; PL = profile likelihood; SD = standard deviation; VAS = visual analogue scale.
For Dreiser 2003, the two NAID groups (diclofenac-K and ibuprofen) were combined into one group.

Other Pain Outcomes

Three RCTs reported other measures of pain relief (Table 29).[100,154,155] One trial of diclofenac-K in patients without radiculopathy[100] and one trial of meloxicam in patients with radiculopathy[154] found NSAIDs associated with a moderate increase in the likelihood of pain relief at 6 hours versus placebo. In one trial, different doses of meloxicam showed similar results. Ibuprofen (1 trial)[100] and indomethacin (1 trial)[155] showed similar results compared with placebo at 6 hours and 1 and 2 weeks, respectively; one trial included and the other excluded patients with radiculopathy.

Table 29. NSAID versus placebo: pain relief

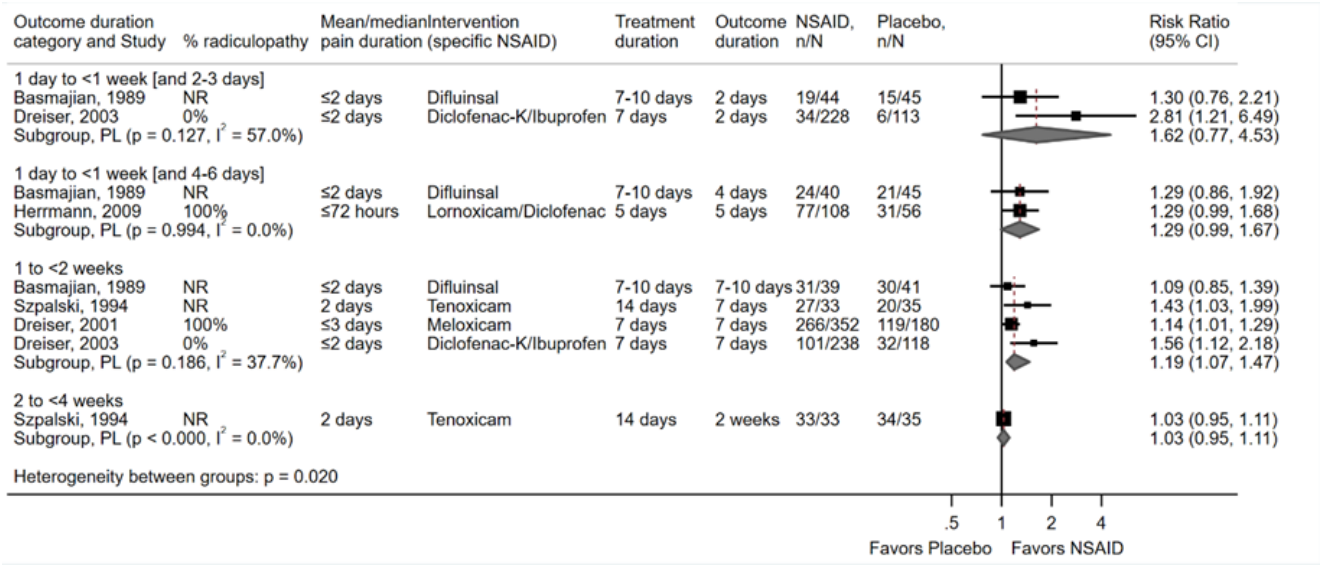
Author, year Symptoms NSAID Therapy	Outcome	Time	NSAID vs. Placebo* % (n/N) and RR (95% CI) or, MD (95% CI)
Dreiser, 2003 Nonradicular pain (1) Diclofenac-K 25 mg every 4-6 hours	First dose efficacy (pain relief)	6 hours	**17% (21/124) vs. 5% (6/121), RR 3.42 (95% CI 1.43 to 8.17)**
(2) Ibuprofen 400 mg every 4-6 hours	First dose efficacy (pain relief)	6 hours	5% (6/119) vs. 5% (6/121), RR 1.02 (0.34 to 3.06)
Dreiser, 2001 Radicular pain (1) Meloxicam 7.5 mg 1x/day	Pain relief (yes)	6 hours	**59% (101/171) vs. 41% (73/180), RR 1.46 (1.17 to 1.81)**
(2) Meloxicam 15 mg 1x/day	Pain relief (yes)	6 hours	**57% (103/181) vs. 41% (73/180), RR 1.4 (1.13 to 1.74)**
Goldie, 1968 Radicular pain Indomethacin 25 mg 3x/day	Pain relief: fair or complete (patient reported)	1 and 2 weeks	60% (15/25) vs. 72% (18/25), RR 0.83 (0.56 to 1.25)

CI = confidence interval; MD = mean difference; NSAID = nonsteroid anti-inflammatory drug; PL = profile likelihood; RR = risk ratio.
* Bold data indicates a statistically significant difference.

Global Efficacy

Five trials[68,75,93,100,154] reported global perceived effect defined as satisfactory, good or excellent improvement overall as rated by the patient (3 RCTs)[68,100,154] or moderate or marked improvement overall as rated by the clinician (2 trials).[75,93] At 4 to 5 days, two RCTs found NSAIDs associated with a small increase in the likelihood of a positive global efficacy rating compared with placebo (N=249, 68.2% vs. 51.5%, pooled RR 1.29, 95% CI 0.99 to 1.67, I^2=0%)[68,75]; while NSAIDs were favored at other timepoints, the differences were not statistically significant or slightly below the threshold for a small effect (Figure 6). When the trial assessed as high risk of bias[75] was excluded, a large effect was seen at 2 to 3 days in one trial (N=341, 14.9% vs. 5.3%, RR 2.81, 95% CI 1.21 to 6.49) and a small effect at 1 to <2 weeks across three trials (N=956, 63.2% vs. 51.2%, pooled RR 1.25, 95% CI 1.05 to 1.70, I^2=52.3%)[93,100,154] favoring NSAIDs; results at 4 to 5 days did not change (small effect, 1 RCT).[68] Although two NSAIDs (diclofenac-K and ibuprofen) were combined into one intervention arm for the pooled analyses, each favored the NSAID compared with placebo with diclofenac-K showing larger effects.[100]A sixth trial (N=239) reported patient-rated global effect on a continuous scale (-5 to 5 [best]) and found similar results for diclofenac versus placebo over 1 to 12 weeks (adjusted difference ranged from -0.3 to 0.1).[83]

Figure 6. NSAID versus placebo: global perceived effect



CI = confidence interval; NR = not reported; NSAID = nonsteroid anti-inflammatory drug; PL = profile likelihood; SD = standard deviation.

Quality of Life, Sleep Quality

Quality of Life or sleep quality were not reported.

Psychological Distress

Psychological distress was not reported.

Return to Work

One trial (N=266) found piroxicam associated with a similar likelihood of returning to work by day 3 (16% vs. 14%, RR 1.15, 95% CI 0.64 to 2.06) and a small increase in the likelihood of returning by day 7 (31% vs. 21%, RR 1.48, 95% CI 0.98 to 2.23) compared with placebo.[91]

Medication Use

Two trials (N=798) found NSAIDs associated with less additional analgesic use versus placebo at day 7: piroxicam 20 mg (mean 3.2 vs. 5.9 acetaminophen 500 mg tablets, p<0.05)[91] and meloxicam 15 mg (mean daily acetaminophen consumption: 869 mg vs. 1110 mg, p=0.03; proportion of patients taking acetaminophen: 58% vs. 71%, RR 0.82, 95% CI 0.70 to 0.95.[154] In the latter trial, fewer patients in the meloxicam 7.5 mg group were taking acetaminophen at 7 days compared with placebo patients (62% vs. 71%, RR 0.87, 95% CI 0.75 to 1.01), however, the mean daily consumption did not differ between groups (939 mg vs. 1110 mg, p>0.05). One trial (N=208) found piroxicam 20-40 mg and placebo associated with similar additional analgesics use (acetaminophen with or without codeine and/or levomepromazine 5 mg) at 4 weeks.[95] (Appendix E).

Adverse Events

Serious Adverse Events

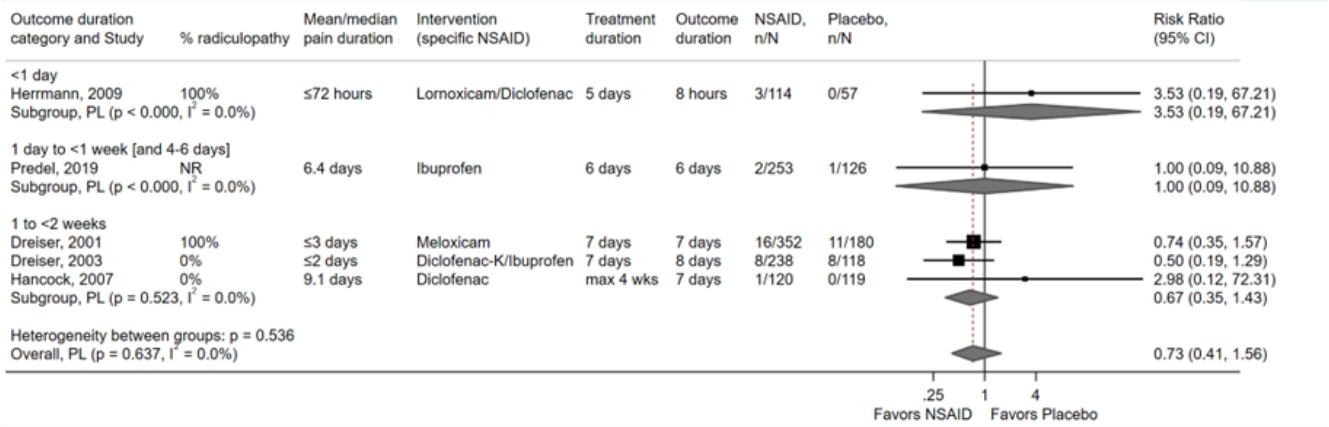
Across five RCTs (N=1,350), SAEs were rare and ranged from 0% to 1.8% with NSAIDs (ibuprofen, diclofenac, tenoxicam, lornoxicam, meloxicam) and 0% to 0.6% with placebo over follow-up periods ranging from 5 to 15 days.[68,93,100,144,154] Three RCTs (N=647) reported no

serious adverse effects in either group[93,100,144] In one RCT,[68] one patient (1.8%, 1/57) who received diclofenac (100 mg on day 1 and 5, 150 mg on days 2-4) experienced two cases of severe abdominal pain and nausea resulting in withdrawal from the trial; there were no SAEs in the lornoxicam and placebo arms (n=57 each). In a second trial,[154] there was one case of anaphylactic shock requiring steroid therapy in a patient (0.6%, 1/171) who received meloxicam 7.5 mg and one patient (0.6%, 1/180) who received placebo experienced marked deterioration of back pain; there were no serious adverse events in the meloxicam 15 mg group (n=181). The trials did not report GI bleed or cardiovascular events (but were also not designed to evaluate them).

Withdrawal due to Adverse Events

Five trials (N=1,677) found NSAIDs were not associated with an increased risk of withdrawal due to adverse events compared with placebo over follow-up periods ranging from 8 hours to 7 days (2.8% vs. 3.3% pooled RR 0.73, 95% CI 0.41 to 1.56, I^2=0%) (Figure 7).[68,93,100,144,154] The specific adverse events leading to withdrawal were not described in any trials. One of these trials enrolled patients with both back (n=207) and neck pain (n=172) and did not report adverse events specifically for the back pain group.[144] Findings from this trial were consistent with other trials.

Figure 7. NSAID versus placebo: withdrawal due to adverse events

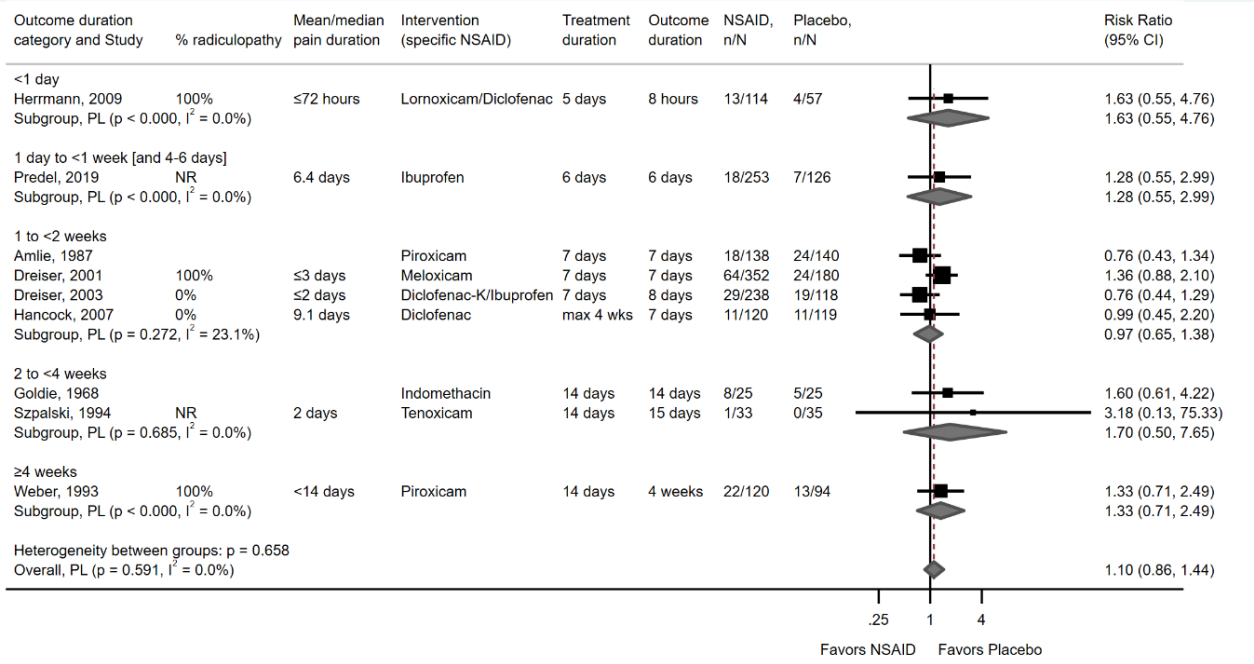


CI = confidence interval; NR = not reported; NSAID = nonsteroid anti-inflammatory drug; PL = profile likelihood.
For Dreiser 2003, the two NAID groups (diclofenac-K and ibuprofen) were combined into one group.

Any Adverse Event

Nine RCTs (N=2,287) found NSAIDs associated with similar risk of any adverse event compared with placebo over follow-up periods ranging from 8 hours to 4 weeks (pooled RR 1.10, 95% CI 0.86 to 1.44, I^2=0%) (Figure 8).[68,83,91,93,95,100,144,154,155] One trial enrolled patients with both back (n=207) and neck pain (n=172) and did not report adverse events specifically for the back pain group.[144] Findings from this trial were consistent with other trials.

Figure 8. NSAID versus placebo: any adverse event



CI = confidence interval; NR = not reported; NSAID = nonsteroid anti-inflammatory drug; PL = profile likelihood.
For Dreiser 2003, the two NAID groups (diclofenac-K and ibuprofen) were combined into one group.

NSAID Versus Acetaminophen

Three RCTs (N=474) compared NSAIDs versus acetaminophen for the treatment of ALBP (Table 30, Appendix E).[67,77,108] Patient mean age ranged from 32 to 66 years and 39% to 66% were female. Race and ethnicity were not reported. The mean duration of the LBP episode was 11 days in one trial[67] and not reported in the other two trials (one trial[77] specified a duration of ≤1 week for inclusion). Mean baseline pain severity (on a 0-10 scale) was 8.3 in two trials[77,108] and 4.8 in one trial.[67] All trials excluded pregnant or breastfeeding woman and patients with radiculopathy. One trial[67] excluded participants with a psychiatric comorbidity. History of substance abuse disorder or treatment for opioid use disorder, prior opioid use, and history of prior or recurrent LBP were not reported. All trials excluded patients with specific medical comorbidities (e.g., cancer, inflammatory conditions, renal dysfunction, cardiopulmonary disorders).

Two trials conducted in the ED evaluated a single IV dose of dexketoprofen 50 mg (both trials[77,108]) or ibuprofen 400 mg (1 trial[108]) compared with a single IV dose of acetaminophen 1000 mg. The third trial compared loxoprofen 60 mg three times daily with acetaminophen 600 mg four times daily taken orally for 4 weeks.[67] Concomitant treatments included rescue fentanyl in two trials[77,108] and lumbar corset (if needed) in one trial.[67]

Two trials[77,108] were considered low risk of bias, and one[67] was moderate risk of bias. Limitations in the moderate risk trial included lack of blinding, high attrition, and lack of intention-to-treat analyses (Appendix F).

Table 30. Study characteristics of NSAID versus acetaminophen

Study, Year	Dogan, 2022	Eken, 2014	Miki, 2018
N Randomized	210	137	127
Mean Age (Years)	NR (range, 18 to <65)	32	65
% Female	47%	39%	66%
% Radiculopathy	0% (Excluded)*	0% (Excluded)	0% (Excluded)
Pain Duration Inclusion Criteria	NR	≤1 week	<4 weeks
Mean Pain Duration (days)	NR	NR	11
Mean/Median Baseline Pain Severity (0-10)	8.3	8.2	4.8
History substance abuse disorder	NR	NR	NR
Psychiatric Comorbidities	NR	NR	0% (Excluded)
NSAID, dose† and frequency	(1) Dexketoprofen 50 mg, (2) Ibuprofen 400 mg, Single dose	Dexketoprofen 50 mg, Single dose	Loxoprofen 60 mg, 3x/day
Acetaminophen, dose† and frequency	Acetaminophen 1000 mg, Single dose	Acetaminophen 1000 mg, Single dose	Acetaminophen 600 mg, 4x/day
Route of Administration	IV	IV	Oral
Treatment duration	15 minutes	NR	4 weeks
Duration follow-up	60 minutes	30 minutes	4 weeks
Setting	ED	ED	Outpatient hospital
Country	Turkey	Turkey	Japan
Quality	Low	Low	Moderate

IV = intravenous; NR = not reported; NSAID = nonsteroid anti-inflammatory drug.
* Patients with neurological deficits excluded.
† In 100 ml normal saline.

Detailed Synthesis

Function

One RCT (N=70) found oral loxoprofen associated with similar improvement in Pain Disability Assessment Scale scores (0-60 scale) versus oral acetaminophen at 2 weeks (MD in change scores -2.98, 95% CI -8.72 to 2.78) and 4 weeks (MD in change scores -1.75, 95% CI -8.84 to 5.34).[67] Attrition was high (45%) in this trial.

Pain

Three RCTs (N=474) reported mean differences in back pain using a 0 to 10 VAS or NRS scale.[67,77,108] Results were inconsistent at <1 day across two trials of single IV doses of dexketoprofen 50 mg and acetaminophen 1000 mg. One trial[108] found dexketoprofen associated with a small improvement in pain at 15, 30, and 60 minutes while the second trial[77] found dexketoprofen associated with similar improvement at 15 and 30 minutes, though the estimate favored acetaminophen at 30 minutes. In the trial that found a small effect for dexketoprofen, patients in this group had significantly lower pain scores at baseline than patients in the acetaminophen group (7.88 vs. 8.46) and when the differences were analyzed in terms of change scores there was no effect. Authors of this trial concluded that all treatments were equally effective (based on change scores). Differences in duration of pain and presence of radiculopathy may also explain some of the inconsistency. One of these trials[108] also compared IV ibuprofen 400 mg with IV acetaminophen 1000 mg and reported similar reduction in mean NRS scores (as

well as change scores) with both treatments over 60 minutes. A third trial[67] found no difference between oral loxoprofen and oral acetaminophen at 2 or 4 weeks. Estimates were imprecise for all outcomes (Table 31).

Table 31. NSAID versus acetaminophen: VAS or NRS pain scores

Author, Year Route, Duration	NSAID	Acetaminophen	Time	VAS/NRS pain scores MD (95% CI)
Eken, 2014 (N=92) Single IV dose Radiculopathy excluded Pain duration ≤1 week	Dexketoprofen 50 mg (n=46)	Acetaminophen 1000 mg (n=46)	15 mins.	0.49 (-0.41 to 1.39)
			30 mins.	0.86 (-0.02 to 1.74)
Dogan, 2022 Single IV dose Radiculopathy NR Pain duration NR ("acute")	Dexketoprofen 50 mg (n=70)	Acetaminophen 1000 mg (n=71)	15 mins.	-0.74 (-1.09 to -0.39)*
			30 mins.	-0.59 (-1.08 to -0.10)*
			60 mins	-0.83 (-1.42 to -0.24)*
	Ibuprofen 400 mg (n=69)	Acetaminophen 1000 mg (n=71)	15 mins.	-0.01 (-0.43 to 0.41)
			30 mins.	0.08 (-0.52 to 0.68)
			60 mins	-0.22 (-0.97 to 0.53)
Miki, 2018 Oral, 3-4x/day Radiculopathy NR Pain duration, mean 11 days	Loxoprofen 60 mg (n=35)	Acetaminophen 600 mg (n=35)	2 wks.	MD in change scores 0.51 (-0.65 to 1.67)
			4 wks.	MD in change scores 0.80 (-0.46 to 2.06)

AE = adverse event; CI = confidence interval; GI = gastrointestinal; MD = mean difference; NR = not reported; NRS = Numeric Rating Scale; NSAID = nonsteroidal anti-inflammatory drug; RR = risk ratio; VAS = visual analogue scale.
*When analyzed in terms of change scores from baseline the differences between groups were not statistically significant: 15 minutes (MD 0.15, 95% CI -0.11 to 0.42), 30 minutes (MD 0.01, 95% CI -0.50 to 0.51), and 60 minutes (MD 0.25, 95% CI -0.38 to 0.89).

Psychological Distress and Quality of Life

One RCT (N=70) found similar improvement in Hospital Anxiety and Depression Scale (HADS) Anxiety and Depression scores and EQ-5D quality of life scores with oral loxoprofen versus acetaminophen at 2 and 4 weeks except for HADS depression scores at 2 weeks (moderately less improvement with loxoprofen, MD in change scores 2.43, 95% CI 0.45 to 4.41, [0-21 scale]) (Appendix E). Attrition was high (45%) in this trial.[67]

Medication Use

One RCT (N=92) found a single IV dose of dexketoprofen associated with a similar likelihood of using rescue medication (fentanyl) compared with a single IV dose of acetaminophen after 30 minutes (15.2% vs. 17.4%, RR 0.88, 95% CI 0.35 to 2.21).[77] In a second trial (N=120), no patient required rescue medication in any group (IV dexketoprofen, IV ibuprofen or IV acetaminophen).[108]

Adverse Events

Serious adverse events and withdrawal due to adverse events were not reported. One trial (N=92) found a similar incidence of any adverse event or nausea or vomiting with IV dexketoprofen compared with IV acetaminophen over 30 minutes (Table 32).[77] In a second trial (N=70) although adverse effects were observed more frequent in the oral loxoprofen group, the overall incidence showed no significant difference between the two medications (Table 28).[67] The third trial[108] reported that no adverse effects occurred in any patient following a single IV infusion of dexketoprofen or acetaminophen.

Table 32. NSAID versus acetaminophen: adverse events

Author, Year Therapies	Time	Outcome	NSAID	Acetaminophen	RR (95% CI)
Eken, 2014 Dexketoprofen 50 mg, Acetaminophen 1000 mg, Single IV dose	30 mins.	Any AE	8.7% (4/46)	8.7% (4/46)	1.00 (0.27 to 3.76)
		Nausea/Vomiting	4.3% (2/46)	4.3% (2/46)	1.00 (0.15 to 6.80)
		Allergic Reaction	0% (0/46)	4.3% (2/46)	NA
		Dry Mouth	4.3% (2/46)	0% (0/46)	NA
Miki, 2018 Loxoprofen 60 mg, 3x/day Acetaminophen 600 mg, 4x/day Oral	4 weeks	GI Issues	9.6% (3/35)	2.9% (1/35)	3.00 (0.33 to 27.46)
		Drowsiness	2.9% (1/35)	0% (0/35)	NA
		Leg Edema	2.9% (1/35)	0% (0/35)	NA

AE = adverse event; CI = confidence interval; GI = gastrointestinal; IV = intravenous; NA = not applicable; NSAID = nonsteroidal anti-inflammatory drug; RR = risk ratio.

NSAID Versus Muscle Relaxant

Description of Included Studies

One RCT (N=87) compared an NSAID (diflunisal 500 mg) versus a muscle relaxant (cyclobenzaprine 5 mg) taken twice daily for 7 to 10 days for the treatment of ALBP of 1 to 2 days duration (Appendix E).[75] Patient characteristics (e.g., age, race, sex, history of prior LBP episodes, history of substance abuse) and comorbidities (medical and psychiatric) were not reported. Pain severity at baseline was not reported.

The trial was conducted in Canada and was assessed as high risk of bias due to insufficient information, other than patient and provider blinding, to assess the study quality (Appendix F).

Detailed Synthesis

Function and Pain

Function and pain were not reported.

Global Efficacy and Adverse Events

Diflunisal was associated with a similar likelihood of achieving moderate or marked improvement on a patient-rated global improvement measure at all timepoints versus cyclobenzaprine: 2 days (N=87, 43% vs. 30%, RR 1.43, 95% CI 0.81 to 2.52), 4 days (N=84, 60% vs. 59.1%, RR 1.02, 95% CI 0.71 to 1.45), and 7 to 10 days (N=83, 79% vs. 84%, RR 0.95, 95% CI 0.77 to 1.16).[75] It is unclear how global improvement was measured but appeared to be a composite of the following: 1) presence of local pain 2) presence of muscle spasm 3) presence of muscle tenderness on palpation 4) limitation of range of motion, and limitation of activities of daily living. This trial did not report other clinical outcomes. Authors report that no significant adverse effects were attributable to the therapy.

NSAID Versus Lidocaine

One RCT (N=44) compared a single IV dose of ketorolac 30 mg versus lidocaine 100 mg for the treatment of acute radicular LBP in patients presenting to the ED in the United States (Appendix E).[99] Rescue medication was provided after 60 minutes if pain relief was not adequate. Mean age of patients was 38 years and 43% were female. Mean baseline pain intensity was 8.1 on a 0-10 scale. The duration of the current LBP episodes was not reported. Pregnant or breastfeeding women were excluded. Information on race or ethnicity, comorbidities (medical

and psychiatric), history of back pain, history or opioid use, and history or treatment of opioid or substance abuse disorder were not reported. The trial was considered moderate risk of bias due to unclear allocation concealment methods, baseline difference between group (sex, pain severity), and concerns regarding selective outcomes reporting (Appendix F).

Detailed Synthesis

Pain

IV ketorolac was associated with a greater reduction in VAS pain scores (0-10 scale) compared with IV lidocaine over 60 minutes, but the difference was not statistically significant (median change scores -1.4, 95% CI 0 to 28 vs. -0.8, 95% CI 0 to 23; p=0.84). Patient-reported pain relief (Pain Relief Scale scores, 0-5 scale, higher means more relief) was similar between groups at 60 minutes (median 2 vs. 2) and the 1-week follow-up (0 vs. 0).[99]

Medication Use

Fewer patients who received IV ketorolac compared with IV lidocaine required rescue medication, but the difference was not significant: 50% (10/20) vs. 67% (14/21) (RR 0.75,95% CI 0.44 to 1.28).[99]

Other Outcomes

Function, psychological distress, quality of life, medication use, and return to work was not reported.

Adverse Events

Although the authors stated that neurologic adverse events were recorded throughout treatment, they do not report on this outcome.[99]

NSAID Versus Manual Therapy

Three RCTs (N=269) compared NSAIDs versus manual therapy for the treatment of ALBP (Appendix E).[76,83,143] Patient mean or median age ranged from 35 to 41 years and 35% to 44% were female. Race or ethnicity was not reported. Two trials[76,143] excluded pregnant or breastfeeding women and the third[83] did not report on this. One trial[83] excluded patients with radiculopathy, one trial[143] enrolled a mixed population, and one trial[76] did not specify inclusion or exclusion of patients with radiculopathy. Mean duration of pain was ≤48 hours in one trial,[76] 9.1 days in one trial[83] and not reported in one trial (inclusion criteria was <1 month).[143] Baseline pain severity ranged from 5.5 to 6.9 on a 0-10 scale. Two trials included patients with prior LBP episodes;[83,143] in the third trial, history of previous LBP was unclear, but patients were excluded if they had a LBP episode within 1 month. One trial[76] excluded participants with a history of substance abuse disorder. History of opioid abuse disorder or treatment for opioid use disorder and prior opioid use were not reported. All trials excluded patients with specific medical comorbidities (e.g., cancer, inflammatory conditions, renal dysfunction, cardiopulmonary disorders).

Two trials[76,83] evaluated oral diclofenac 50 mg but total daily dosages varied (150 mg vs. 100 mg) as did treatment duration (3 days vs. 4 weeks) and one trial evaluated oral diflunisal (1000 mg per day) for 10 days.[143] Two trials[76,83] specified the use of high-velocity, low amplitude spinal manipulation and the third did not provide treatment details[143] though all trials allowed additional mobilization per the therapist's discretion. The frequency of therapy varied as did the

maximum number of spinal manipulation sessions (range, 2 to 12). Concomitant treatments included acetaminophen in two trials[76,83] and advice in two trial.[83,143]

One trial[83] was assessed as low risk of bias, and two[76,143] as moderate risk of bias. Methodological limitations in the low risk of bias trial included differential attrition between groups. Limitations in the moderate risk of bias trials included unclear randomization and allocation concealment methods and lack of blinding (Table 33; Appendix F).

Table 33. Study characteristics of NSAID versus manual therapy

Study, Year	Von Heymann, 2013	Hancock, 2007	Waterworth, 1985
N Randomized	75	120	74*
Mean Age (Years)	Median 36	41	35
% Female	37%	44%	38%
% Radiculopathy	NR	Excluded	Mixed (% NR)
Pain Duration Inclusion Criteria	<48 hours	<6 weeks	<1 month
Mean Pain Duration (days)	NR	9.1 days	NR
Mean Baseline Pain Severity (0-10)	5.5	6.6	2.1 (0-3 scale)
History of Prior LBP	NR†	Mean 3.7 previous episodes†	68%
History Substance Abuse Disorder	0% (Excluded)	NR	NR
Psychiatric Comorbidities	NR	NR	NR
NSAID, Dosage And Frequency	Diclofenac 50 mg 3x per day (plus sham manipulation)	Diclofenac 50 mg 2x per day (plus sham manipulation)	Diflunisal 1000 mg day 1, then 500 mg 2x per day
Comparator, Dosage And Frequency	HVLA Manipulation and/or mobilization; 1-2 sessions (plus placebo)	HVLA Manipulation and/or mobilization; 30–40-minute sessions, 2-3x/week (max. 12 treatment) (plus placebo)	Manipulation and/or mobilization; 5, 45-minute sessions weekly
Treatment Duration	3 days	4 weeks	10 days
Duration Follow-up	7-9 days	12 weeks	10-12 days
Setting	Orthopedic or primary care	Primary care	Primary care
Country	Germany	Australia	New Zealand
Quality	Moderate	Low	Moderate

HVLA = high velocity, low amplitude; IM = intramuscular; NR = not reported; NSAID = nonsteroid anti-inflammatory drug.
* Reflects patients who completed. Patients were not reported by arm as randomized. This trial included 3 arms: NSAID, manipulation and physiotherapy.
† von Heymann: Excluded patients with LBP episodes within the last 3 months; Hancock: excluded patients with LBP episodes with the last 1 month.

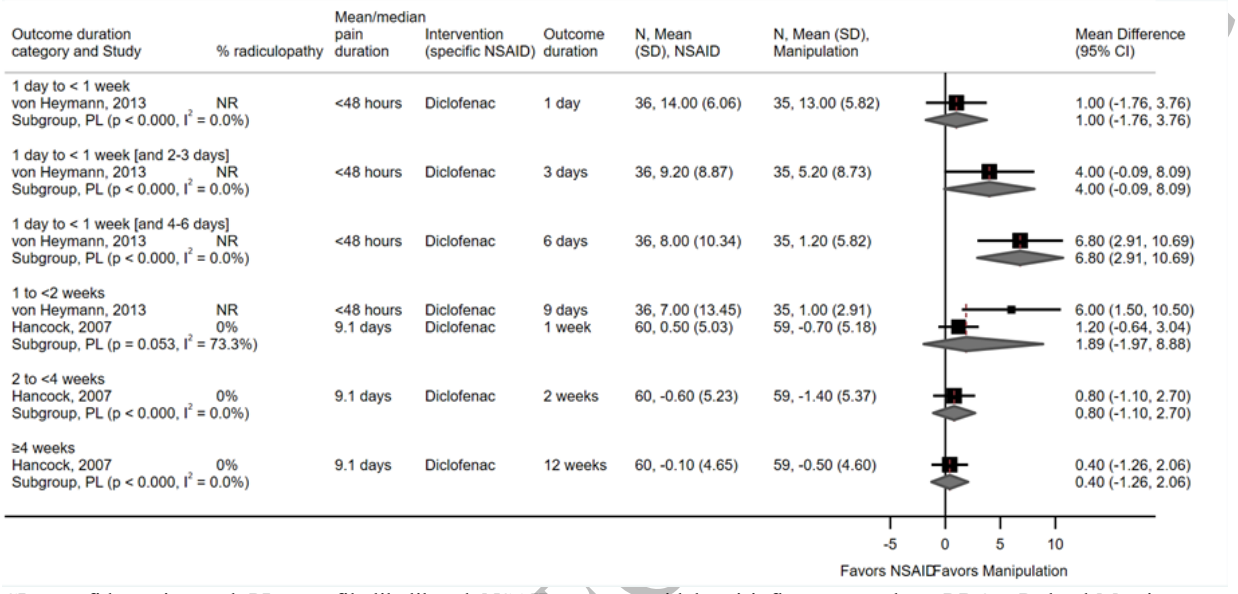
Detailed Synthesis

Function

Two RCTs (N=190) assessed function based on average changes in RDQ (0-24 scale) scores (Figure 9).[76,83] At 1 to <2 weeks, estimates from both trials favored manual therapy but the magnitudes of effect varied, and the pooled estimate was characterized by substantial heterogeneity (pooled MD 1.89, 95% CI -1.97 to 8.88, I^2=73.3%); baseline pain duration may explain some of the heterogeneity. One trial[76] in patients with pain duration of ≤48 hours found manual therapy associated with a large improvement at 6 days (MD 6.80, 95% CI 2.91 to 10.69) and 9 days (MD 6.00, 95% CI 1.50 to 10.50) compared with diclofenac; there was no effect between groups at 1 and 3 days, though the estimate at 3 days favored manual therapy (MD 4.00,

small effect). The second trial[83] in patients with nonradicular LBP of 9.1 days duration found similar improvement in scores between diclofenac and manual therapy at 1, 2, and 12 weeks (range of difference 0.40 to 1.20), though manual therapy tended to be favored at 1 week. All estimates were imprecise. In addition to baseline pain duration, differences in patient populations (presence or absence of radiculopathy), treatment regimens and length of treatment may have impacted outcomes across trials.

Figure 9. NSAID versus manual therapy: RDQ scores

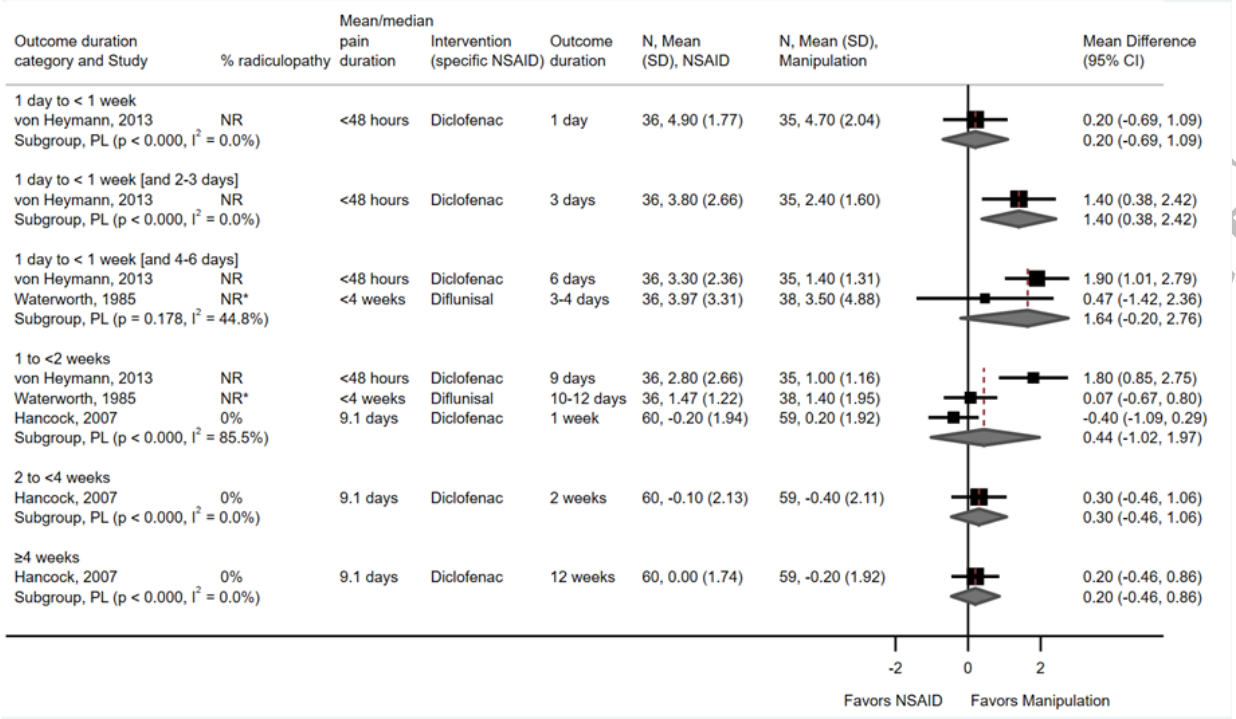


CI = confidence interval; PL = profile likelihood; NSAID = nonsteroidal anti-inflammatory drug; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation.

Pain

Three RCTs (N=264) assessed pain scores standardized to a 0-10 scale (Figure 10).[76,83,143] At 4 to 6 days estimates from both trials favored manual therapy but the magnitudes of effect varied, and the pooled estimate showed no effect (pooled MD 1.64, 95% CI -0.20 to 2.76, I^2=44.8%). At 1 to <2 weeks, the pooled estimate showed no effect but was marked by substantial statistical heterogeneity (pooled MD 0.44, 95% CI -1.02 to 1.97, I^2=85.5%); two trials showed no effect while the third showed a moderate effect favoring manual therapy. Baseline pain duration may explain some of the heterogeneity. One trial[76] in patients with pain duration of ≤48 hours found manual therapy associated with a moderate improvement in pain at 3, 6, and 9 days compared with diclofenac; results were similar between groups at 1 day. The other two trials included patients with longer baseline pain duration (9.1 days and <4 weeks [mean not reported])[83,143] and found no effect on pain scores between manual therapy versus diclofenac at 1, 2, and 12 weeks[83] and versus diflunisal at 3-4 and 10-12 days.[143] All estimates were imprecise. In addition to baseline pain duration, differences in patient populations (presence or absence of radiculopathy), treatment regimens and length of treatment may have impacted outcomes across trials.

Figure 10. NSAID versus manual therapy: pain scores



CI = confidence interval; PL = profile likelihood; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; SD = standard deviation.

Quality of Life

One trial (N=71) found manual therapy associated with a small improvement in 12-item Short Form Questionnaire physical component scale (SF-12 PCS) scores compared with diclofenac at 6 days (MD -9.2, 95% CI -14.53 to -3.87) and 9 days (MD -8.7, 95% CI -13.31 to -4.09), but results between groups were similar at earlier timepoints (1 and 3 days) and for SF-12 mental component scale (MCS) scores at all timepoints (Appendix E).[76]

Global Perceived Effect

One trial (N=120) found diclofenac associated with similar improvement in Global Perceived Effect scores (-5 to 5 scale, 5=completely recovered) compared with spinal manual therapy over 1 to 12 weeks (range of adjusted MDs in change scores, respectively, -0.3 to 0.1 vs. -0.1 to 0.4) (Appendix E).[83] In a second trial (N=74), similar proportions of patients who received diflunisal and manual therapy reported good or excellent overall improvement (77.8% vs. 73.7%).[143]

Return to Work and Medication Use

One trial (N=71) reported no statistical difference between NSAID and manual therapy in time off work (mean 1.8 [SD 2.1] vs. 1.2 [SD 1.7] days) or use of rescue medication (mean 6.4 [10.7] vs. 2.2 [3.7] tablets of acetaminophen 500mg) (Appendix E).[76]

Adverse Events

Serious Adverse Events and Withdrawal Due to Adverse Events

In two trials, 8.6% (3/35) vs. 2.9% (1/35) of patients[76] and 2.8% (1/36) vs. 2.6% (1/38) of patients[143] in the NSAID versus manual therapy groups, respectively, withdrew due to treatment

failure (not further defined). In the third trial (N=120), one patient (0.8%) had a suspected hypersensitivity reaction to diclofenac and discontinued treatment; this trial reported no serious adverse events with spinal manual therapy (Appendix E).[83]

Any Adverse Events

Adverse events were reported for NSAIDs only. In one trial, 9.2% (11/120) of patients experienced an adverse event which included GI disturbances, dizziness, and heart palpitations.[83] In the second trial, 2.8% (1/36) each had indigestion with the loading dose and nausea.[143] The third trial reported that no adverse events occurred (Appendix E).[76]

NSAID Versus Acupuncture

Two RCTs (N=102) compared NSAIDs versus acupuncture for the treatment of ALBP (Table 34, Appendix E).[101,153] Mean patient age ranged from 36 to 38 years, 35% to 44% were female, the mean duration of pain was ≤3 days (range,1.8 to 3.0 days) and baseline pain severity ranged from 6.1 to 8.2 on a 0-10 scale. One trial[101] required severe functional impairment (but without severe neurological symptoms or deficit) for study entry (ODI score ≥60%, dependent of wheelchairs or walkers). Both trials included patients with or without radiculopathy (proportions not provided).[101,153] One trial[101] excluded pregnant or breastfeeding women and the other trial[153] did not report whether they were included or excluded. No trial reported on prior history of LBP, but one trial[153] excluded individuals with back pain within 4 weeks of study entry. One trial[101] excluded participants requiring psychiatric medication. Race or ethnicity, history of or treatment of substance abuse disorder, including opioid use disorder, and prior opioid use was not reported. Both trials excluded patients with specific medical comorbidities (e.g., cancer, inflammatory conditions, renal dysfunction, cardiopulmonary disorders).

One trial[101] compared a single IM injection of diclofenac 75 mg versus a single, 20-minute session of motion style acupuncture and the other trial[153] compared 5 days of oral diclofenac (50 mg twice a day) versus 30-minute, daily sessions of traditional Chinese acupuncture. In one trial,[101] after the single injection, patients received additional treatment in the form of integrative therapy (herbal medicine, Chuna manipulation, bee venom pharmaco-acupuncture, and acupuncture) from physicians outside of the trial (but not blinded to previous treatment) as either inpatients (average of 5 session per week) or outpatients (1-2 times per week).

One trial[101] was assessed as moderate risk of bias and one trial[153] as high risk of bias. Methodological limitations in the trial at moderate risk of bias included baseline differences between groups and lack of blinding; additional concerns in the trial at high risk of bias included inadequate randomization method, unclear allocation concealment methods, and no reporting of attrition (Appendix F).

Table 34. Study characteristics of NSAID versus manual therapy, physiotherapy, or acupuncture

Study, Year	Shin, 2013	Liu, 2010
N Randomized	58	44
Mean Age (Years)	38	37
% Female	41%	43%
% Radiculopathy	Mixed (% NR)	Mixed (% NR)
Pain Duration Inclusion Criteria	<4 weeks	<2 weeks
Mean Pain Duration (days)	1.8	3.1
Mean/Median Baseline Pain Severity (0-10)	8.2	6.1
History of LBP	NR	NR (no episodes within 4 weeks of trial entry)
History substance abuse disorder	NR	NR
Psychiatric Comorbidities	Excluded	NR
NSAID, dosage and frequency	Diclofenac 75 mg, one IM injection	Diclofenac 50 mg, orally, 2x/day
Comparator, dosage and frequency	Motion-style Acupuncture* (traditional Chinese medicine points): one 20-minute session	Traditional Chinese Acupuncture: 30-minute sessions, once daily; needles manipulated every 10 mins.
Treatment duration	1 day	5 days
Duration follow-up	24 weeks	5 days
Country	Korea	China
Quality	Moderate	High

IM = intramuscular; NR = not reported; NSAID = nonsteroid anti-inflammatory drug.
* Similar to traditional acupuncture in that needles are inserted at specific acupuncture points and occasionally manually manipulated, but unique in that it requires passive or active movement of the patient's body while acupuncture needles are retained (e.g., patient walks with assistance while needles are retained).

Detailed Synthesis

Pain and Function

One RCT (N=58) found a single 20-minute session of motion style acupuncture associated with a large improvement in NRS back pain scores (0-10 scale) at 30 minutes (MD in change scores 3.12, 95% CI 2.26 to 3.98) and a moderate improvement at 2 weeks (MD in change scores 1.66, 95% CI 0.16 to 3.15) and 4 weeks (MD in change scores 1.50, 95% CI 0.08 to 2.92) compared with a single IM injection of diclofenac for patients with severe disabling LBP.[101] At 24 weeks, back pain scores similar between groups (MD -0.20, 95% CI -1.37 to 0.95).[101] Likewise, acupuncture was associated with a large improvement in ODI scores (0 to 100 scale) at 30 minutes (MD in change scores 32.95, 95% CI 26.88 to 39.03) and a moderate improvement at 2 weeks (MD in change scores 20.07, 95% CI 5.83 to 34.31) and 4 weeks (MD in change scores 16.88, 95% CI 3.19 to 30.57); this effect was not sustained at 24 weeks (MD in change scores -7.6, 95% CI -16.67 to 1.47).[101] Acupuncture was associated with a small reduction in NRS leg pain at 30 minutes compared with diclofenac (MD in change scores 0.97, 95% CI 0.22 to 1.71) but not at other timepoints and by 24 weeks the diclofenac injection group showed a greater reduction in pain (MD in change scores –1.85, 95% CI –3.47 to –0.22; moderate effect). During the follow-up period, 93% of patients who received diclofenac and 66% of patients who received acupuncture were hospitalized and received an average of 5 sessions a week of integrative medicine (versus those who remained outpatients and received this treatment once or twice a week).[101]

The second trial (N=44) found similar improvement in NRS back pain scores (0-10 scale, MD 0.37, 95% CI -0.47 to 1.21) and RDQ scores (0-24 scale, MD -0.20, 95% CI -1.85 to 1.45) between oral diclofenac and traditional acupuncture after 5 days of treatment.[153]

Other Outcomes

Quality of life, psychological distress, global efficacy, medication use, and return to work was not reported.

Adverse Events

One trial (N=58) stated there were no withdrawals or adverse events.[101]

NSAID Versus Physiotherapy

One RCT (N=70) compared oral diflunisal 500 mg twice daily (with an initial loading dose of 1000 mg) versus physiotherapy (local heat/shortwave diathermy and ultrasound for 15 to 20 minutes each and active flexion and extension exercises) for 10 days for the treatment of ALBP of <1 month duration (mean not reported) (Appendix E).[143] Patient mean age was 36 years, 35% were female and 69% had a history of prior LBP. Baseline pain severity was 2.1 on a 0-3 scale. Patient with and without radiculopathy were included but not further described. Pregnant or breastfeeding women were excluded. Race or ethnicity, comorbidities (medical and psychiatric) and history or treatment of opioid or substance abuse disorder were not reported. Patients were asked to refrain from bed rest, other nonopioid analgesics, or acupuncture. The trial was assessed as moderate risk of bias due to unclear randomization and allocation concealment methods and lack of blinding (Appendix F).

Detailed Synthesis

Pain and Global Efficacy

Diflunisal was associated with similar improvement in LBP (0-3 scale) at 3-4 days (mean 1.19 vs. 1.11) and 10-12 days (0.44 vs. 0.38) and leg pain (0-3 scale) at 3-4 days (mean 0.41 vs. 0.38) and 10-12 days (0.25 vs. 0.11) compared with physiotherapy.[143] Similar proportions of patients in both groups rated their overall improvement as good or excellent: 77.8% (28/36) vs. 70.6% (24/34), RR 1.10 (95% CI 0.83 to 1.46).[143]

Function outcomes were not reported.

Adverse Events

In the diflunisal group, one patient each (2.8%, 1/36) experienced indigestion with the loading dose, nausea and discontinued treatment due to lack of improvement (not further details provided).[143]

Muscle Relaxants

Nine trials evaluated a muscle relaxant.[70,74,75,97,104,132,151,161,163]

Muscle Relaxant Versus Placebo or Sham

Description of Included Studies

Eight RCTs (N=2,228) compared a muscle relaxant versus placebo or sham (Table 35, Appendix E).[70,74,75,97,104,132,161,163] In general, patients were in their early 40's and around half

were female (range, 43% to 55%). Inclusion or exclusion of patients with radiculopathy was not reported by most trials; two trials excluded patients with radiculopathy[97,163] and one[70] included a mixed population (53% with radiculopathy). Mean LBP episode duration was not well reported across the trials and inclusion criteria for pain duration varied widely (<72 hours to 2-6 weeks). One older trial[70] did not provide inclusion criteria or mean pain duration for its population but described patients as having ALBP; coupled with other signals from the population description, we gave the trial the benefit of the doubt that it met our inclusion criteria. Only one trial reported whether patients had a history of prior LBP episodes, which appeared to be rare (mean 0.47 episodes).[74] History of substance abuse disorders and psychiatric comorbidities was not consistently reported but was an exclusion criterion when mentioned.

The type of muscle relaxant varied as did the doses and regimen; the treatment duration was most commonly 5 to 10 days (Table 35). There were two trials of carisoprodol,[97,163] two trials of thiocolchicoside,[104,132] two trials of tizanidine,[70,104] and one trial each of baclofen,[161] cyclobenzaprine[75] and either cyclobenzaprine, carisoprodol or methocarbamol (not reported separately).[74] Except for one trial[132] that used intramuscular injections of thiocolchicoside for up to 5 days, all muscle relaxants were given orally. Three trials had multiple arms and evaluated two different muscle relaxants[74,104] or different doses of the same muscle relaxant.[163] In one trial, patients in both groups also received sham chiropractic manipulation (light pressure).[74] This trial included a third, true chiropractic manipulation arm (described in another section of this report) and the sham manipulation was done to preserve blinding. All but one trial[75] mentioned concomitant treatments: four trials did not give or permit any rescue medication or additional treatment[70,97,161,163] and three trials provided rescue medication (acetaminophen) to be taken as needed.[74,104,132] One older trial[161] prescribed best rest for the first 4 days; a small proportion of patients enrolled in this trial were using back brace/support (5%) and had received physical therapy recently (2.5%).

One trial was assessed at low risk of bias,[132] six at moderate risk of bias[70,74,97,104,161,163] and one at high risk of bias (Appendix F).[75] Primary methodological limitations included unclear randomization and concealments methods and higher than acceptable attrition in some trials. The trial at high risk of bias was older and was poorly reported overall.

Table 35. Study characteristics of muscle relaxant versus placebo or sham

Study, Year	Serfer, 2010	Ralph, 2008	Dapas, 1985	Basmajian, 1989	Tuzun, 2003	Berry, 1988	Ketenci, 2005	Hoiriis, 2004
N Randomized	828	562	200	88	149	112	97	192*
Mean Age (Yrs.)	41	40.4	42	NR	41	41	38	42
Female	55%	52%	52%	NR	54%	49%	48%	43%
Radiculopathy	Excluded	Excluded	NR	NR	NR	53%	NR	Excluded
Pain Duration Inclusion Criteria	≤3 days	≤3 days	<2 weeks	≤2 days	<3 days	NR†	<72 hours	2-6 weeks
Mean Pain Duration (days)	1.6	1.7	NR	NR	NR	NR†	NR	3.7 weeks
Mean/Median Baseline Pain Severity (0-10)	NR	Moderate: 74% Severe: 26%	4.2 (0-5 scale)	NR	7.2	5.2	5.7	4.2
Previous LBP episodes	NR	NR	NR	NR	NR	NR	NR	Mean 0.47 previous episodes
History substance abuse disorder	Excluded	Excluded	NR	NR	Excluded	NR	NR	NR
Psychiatric Comorbidities	NR	NR	NR	NR	Excluded	NR	NR	NR
Muscle Relaxant‡	Carisoprodol	Carisoprodol	Baclofen	Cyclobenzaprine	Thiocolchicoside (IM)	Tizanidine	(1) Thiocolchicoside; (2) Tizanidine	(1) Cyclobenzaprine, (2) Carisoprodol, (3) Methocarbamol
Dosage Frequency	(1) 250 mg, 4x per day (2) 350 mg, 4x per day	250 mg 4x per day	10-20 mg 3-4x per day	5 mg 2x per day	4mg 2x per day	4mg 3x per day	(1) 8 mg, (2) 6 mg; 2x per day	(1) 5 mg, (2) 350 mg, (3) 750 mg; Frequency NR§
Treatment duration	7 days	7 days	10 days, Taper, days 11-12	7-10 days	Up to 5 days	7 days	5-7 days	2 weeks
Duration of follow-up	7 days	7 days	14 days	7-10 days	5 days	7 days	5-7 days	4 weeks
Country	U.S.	U.S.	U.S.	Canada	Turkey	UK	Turkey	U.S.
Quality	Moderate	Moderate	Moderate	High	Low	Moderate	Moderate	Moderate

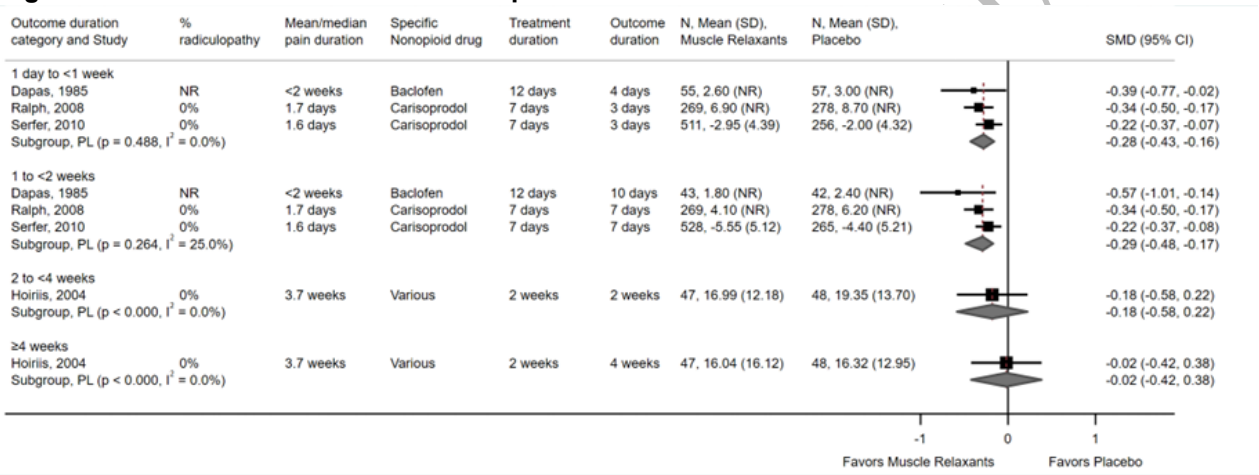
IM = intramuscular; LBP = low back pain; NR = not reported.
*Whole sample. Not reported by arm. The 3 trial arms included: chiropractic adjustments with placebo medicine, muscle relaxants with sham adjustments, or placebo medicine with sham adjustments.
†Older trial, described as acute pain; there were enough signals to give the trial the benefit of the doubt that this trial met our inclusion criteria for duration of pain.
‡Route of administration was oral unless otherwise stated.
§Medication usage instructions were chosen by the medical doctor based on his own clinical experience and were designed to mimic general medical care with a 2-week duration. No further information provided.

Function

Four RCTs assessed function using various measures (Figure 11).[74,97,161,163] Three RCTs found oral muscle relaxants associated with a small improvement in function scores compared with placebo at 1 day to <1 week (3-4 days, N=1,426) and 1 to <2 weeks (7-10 days, N=1,425);[97,161,163] two RCTs used the RDQ (0-24 scale, pooled MDs -1.24 to -1.48, I^2=51.1% to 55.6%)[97,163] and one used the Interference with ADLs Scale.[161] Results were similar for trials of carisoprodol[97,163] and a trial of baclofen.[161] The fourth trial (N=95) reported similar improvement in function scores (ODI scores) for muscle relaxants (cyclobenzaprine, carisoprodol, or methocarbamol) and placebo at 2 to <4 weeks and at ≥4 weeks.[74] This trial did not report the effect of each muscle relaxant separately.

No trial reported the proportion of patients achieving a successful functional outcome.

Figure 11. Oral muscle relaxants versus placebo: function scores



CI = confidence interval; NR = not reported; PL = profile likelihood; SD = standard deviation.

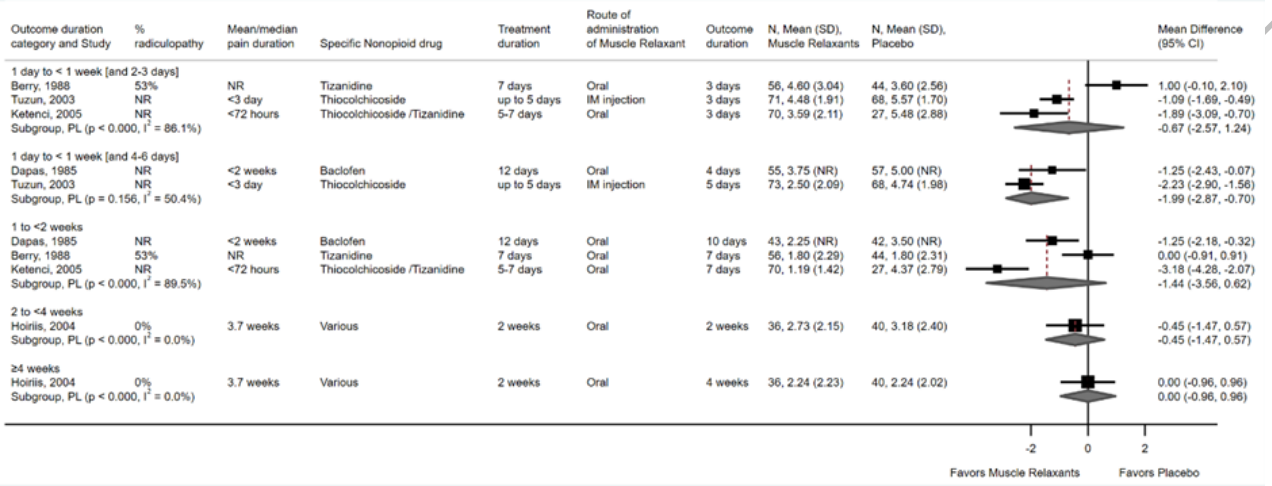
Pain

Compared with placebo, intramuscular injections of thiocolchicoside 4 mg twice daily for a maximum of 5 days were associated with a large increase in the likelihood of achieving a ≥50% improvement on the VAS pain scale at 5 days in one trial (N=143): 74.3% versus 30.4% (RR 2.44,95% CI 1.67 to 3.57).[132]

Five RCTs (N=526) reported mean improvement in VAS or NRS pain scores (0-10 scale).[70,74,104,132,161] Pooled estimates were characterized by substantial heterogeneity. At 3 days, one older RCT of tizanidine (mean pain duration not reported) reported pain improvement that tended to favor the placebo group.[70] In contrast, two other RCTs,[104,132] found moderate pain improvement with muscle relaxants (thiocolchicoside or tizanidine) versus placebo (N=236, pooled MD -1.25, 95% CI -2.31 to -0.55, I^2=28.0%). Similarly, at 7 to 10 days, the older RCT[70] found tizanidine and placebo associated with similar improvement in pain while two other RCTs of tizanidine or thiocolchicoside[104,161] reported moderate to large improvements with muscle relaxants. One trial[104] that reported pain at 3 and 7 days included two muscle relaxant groups, thiocolchicoside and tizanidine, which were combined into one group for the primary analyses. When each muscle relaxant was analyzed separately in the pooled analyses, conclusions were similar (data not shown). Individually, tizanidine was associated with a small effect at 3 days (MD -1.60) and a large effect at 7 days (MD -2.51) and thiocolchicoside was associated with large effects at both timepoints (MD -2.14 to 3.74) compared with placebo. At 4 to 5 days across

two trials, muscle relaxants were associated with a moderate improvement in pain; heterogeneity was less at this timepoint.[132,161] A single RCT found similar pain improvement for muscle relaxants (cyclobenzaprine, carisoprodol, or methocarbamol) and placebo at 2 and 4 weeks; this trial did not report the effect of each muscle relaxant separately.[74] Imprecision was present in all analyses (Figure 12).

Figure 12. Muscle relaxants versus placebo: pain scores

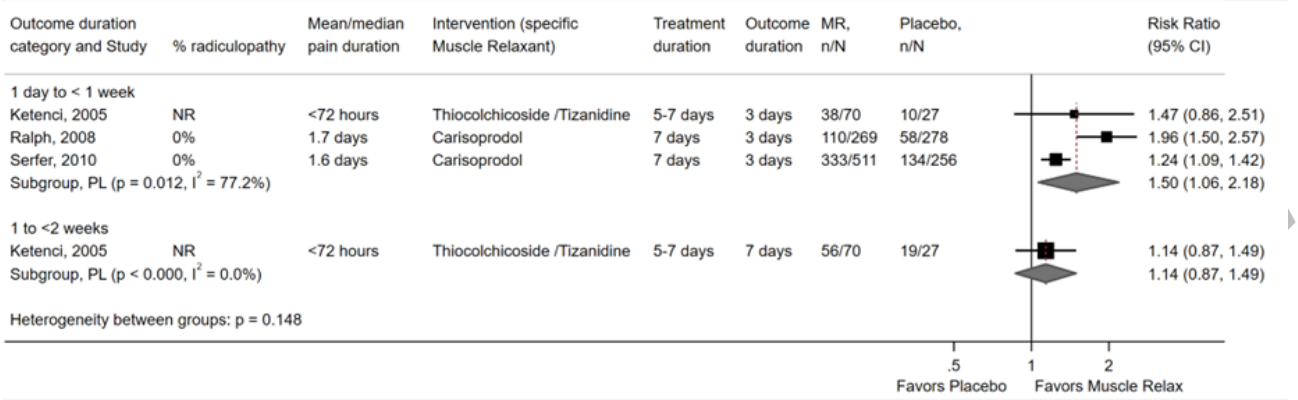


CI = confidence interval; NR = not reported; PL = profile likelihood; SD = standard deviation.

Other Pain Outcomes

Muscle relaxants were associated with a moderate increase in the likelihood of achieving patient-reported back pain relief (i.e., much or very much improved) compared with placebo at 1 day to <1 week (3 days, N=1,411) across three RCTs (Figure 13).[97,104,163] There was substantial heterogeneity (I^2=77.2%) in the pooled estimate but all trials favored muscle relaxants (RRs 1.24 to 1.96). The two large trials[97,163] with similar patient populations favored carisoprodol (though the magnitudes of effect differed, small vs. large), and the third trial finding similar improvement with thiocolchicoside or tizanidine compared with placebo. The latter, smaller trial (N=97) also reported similar improvement between groups at 1 to <2 weeks, though there is some imprecision in the estimate.[104] This trial included two different muscle relaxant groups (thiocolchicoside and tizanidine) which were combined into one group for our analyses; when considered individually, the results did not change.

Figure 13. Muscle relaxant versus placebo: patient-rated back pain relief

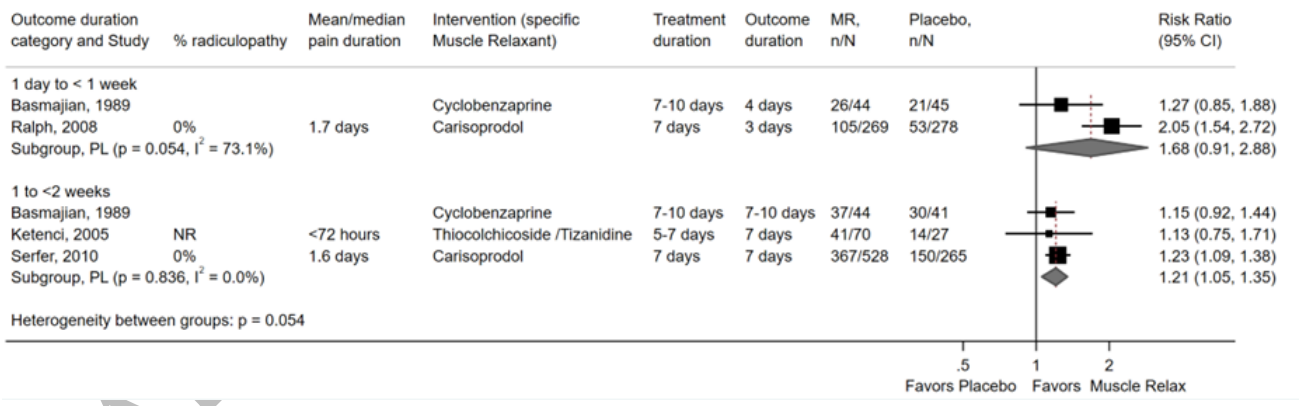


CI = confidence interval; MR = muscle relaxant; NR = not reported; PL = profile likelihood.

Global Efficacy

Four RCTs (N=1,526) reported patient-rated global impression of change defined as good or excellent improvement overall (Figure 14).[75,97,104,163] At 1 day to <1 week, the likelihood of a good or excellent rating was higher with muscle relaxants (cyclobenzaprine or carisoprodol) versus placebo across two trials (N=636) but the pooled estimate was not statistically significant (pooled RR 1.68, 95% CI 0.91 to 2.88, I^2=73.1%).[75,97] Individually, results were only significant for the moderate risk of bias trial of carisoprodol (large effect, N=547, RR 2.05, 95% CI 1.54 to 2.72). At 1 to <2 weeks, three trials (N=975) found muscle relaxants associated with a small increase in the likelihood of a good or excellent rating;[75,104,163] estimates were similar in the three trials even though they looked at different muscle relaxants.

Figure 14. Muscle relaxant versus placebo: global impression of change (good to excellent)



CI = confidence interval; MR = muscle relaxant; NR = not reported; PL = profile likelihood.

Quality of Life, Sleep Quality

Quality of life and sleep quality were not reported.

Psychological Distress

One trial (N=92) found muscle relaxants (cyclobenzaprine, carisoprodol or methocarbamol) associated with similar improvement in Modified Zung Self-Rating for Depression scores (scale, 20-80) compared with placebo at 2 weeks (MD -1.81, 95% CI -5.33 to 1.71) and 4 weeks (N=92, MD -0.08, 95% CI -3.48 to 3.32).[74] This trial did not report the effect of each muscle relaxant separately.

Return to Work

Return to work was not reported.

Medication Use

Four trials reported subsequent analgesic use during the study period.[70,74,104,132] Methods of reporting analgesic used varied (Table 36). The trials found muscle relaxants associated with less analgesic (i.e., acetaminophen, aspirin) consumption versus placebo, but differences were not always statistically significant.

Table 36. Muscle relaxant versus placebo: medication use

Author, Year Symptoms NSAID Therapy	Outcome	Time	NSAID vs. Placebo % (n/N) and RR (95% CI) or, MD (95% CI)
Tuzun, 2003 Thiocolchicoside 4 mg, twice daily IM injection	Total acetaminophen pills taken	5 days	**8 (n=73) vs. 10.5 (n=64) pills, p<0.001**
Berry, 1988 Tizanidine 4mg, three times daily	Mean number of aspirin tablets taken– All patients	3 days	N=108; MD -2.1 (-4.6 to 0.4)
		7 days	N=100; MD -1.3 (-6.4 to 3.8)
	Mean number of aspirin tablets taken – Subgroup without premedication	3 days	**N=NR; 3.4 (5.42) vs. 6.2 (6.95), p=0.037**
		7 days	N=NR, 7.9 (12.01) vs. 10.5 (12.88), p=0.094
Hoiriis, 2004 Cyclobenzaprine 5 mg, carisoprodol 350 mg or methocarbamol 750 mg, varied doses	Mean number of acetaminophen tablets used based on bottle counts	2 weeks	N=81; MD 2.5 (-8.0 to 12.9)
	Mean number of acetaminophen tablets used based on patient log counts	2 weeks	N=81; MD 0.7 (-7.0 to 8.3)
Ketenci, 2005 Thiocolchicoside 8 mg, twice daily	Acetaminophen consumption (any)	3 days	**45.7% (16/35) vs. 76.9% (20/26); RR 0.59 (0.39 to 0.90)**
		5-7 days	**28.5% (10/35) vs. 66.7% (16/26); RR 0.46 (0.25 to 0.85)**
Tizanidine 6 mg, twice daily	Acetaminophen consumption (any)	3 days	**45.1% (14/31) vs. 76.9% (20/26); RR 0.59 (0.38 to 0.91)**
		5-7 days	**25.8% (8/31) vs. 66.7% (16/26); RR 0.42 (0.21 to 0.82)**

CI = confidence interval; IM = intramuscular; MD = mean difference; NR = not reported; NSAID = nonsteroid anti-inflammatory drug; RR = risk ratio.

Adverse Events

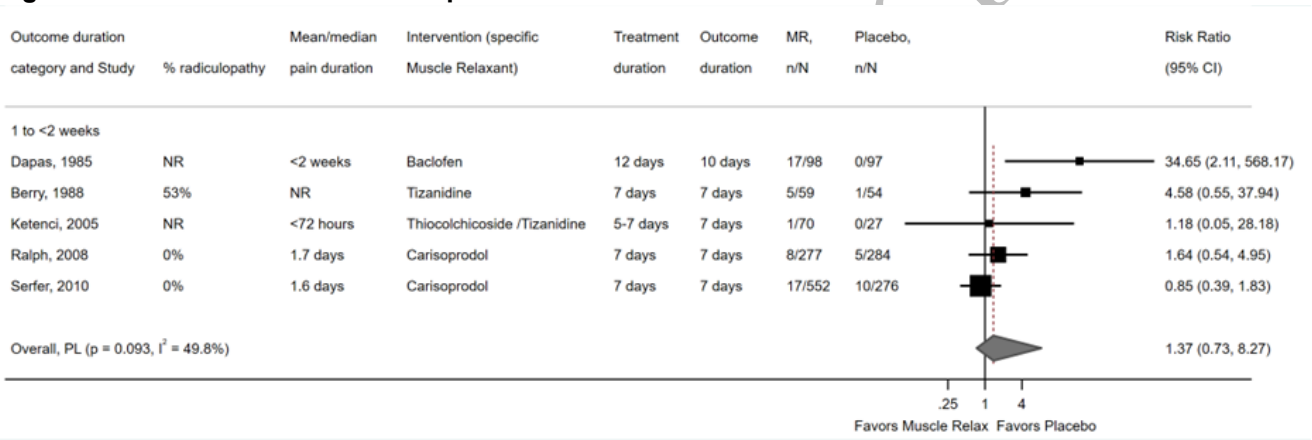
Serious Adverse Events

Serious adverse events were reported in two trials (N=1,023) over 7 to 10 days of follow-up and all occurred in patients who received a muscle relaxant (vs. 0% with placebo). One trial[163] reported that two patients (0.7%, 2/281) who received carisoprodol 350 mg four times daily for 7 days experienced severe drowsiness (there were no SAEs in the group randomized to carisoprodol 250 mg 4 times daily, n=271). In a second trial,[161] 6.1% (6/98) of patients reported "significant" adverse events with baclofen 10-20 mg 3-4 times daily over 12 days. Four trials (N=889) stated that no SAEs or "significant" adverse events occurred in any patient over treatment durations of 5 to 10 days.[75,97,104,132] Trials of carisoprodol did not report on or evaluate the addiction or abuse potential of this medication and were not designed to do so.

Withdrawal due to Adverse Events

Five RCTs (N=1,794) found muscle relaxants associated with a similar risk of withdrawal due to adverse events over treatment periods of 7 to 12 days compared with placebo (Figure 15).[70,97,104,161,163] Individually, only one, older trial[161] (moderate risk of bias) that evaluated baclofen (10 mg, 2 tablets, four times daily for 10-12 days) reported a significant difference between the two treatment groups: 17 patients in the muscle relaxant group versus none in the placebo group withdrew prematurely from the trial due to side effects from the medication, primarily nausea and sedation; however, the effect estimate was very imprecise (RR 34.65, 95% CI 2.11 to 568.17). Removal of this trial did not change conclusions but shifted the pooled estimate closer to the null, reduced imprecision and eliminated heterogeneity (N=1,599, pooled RR 1.19, 95% CI 0.64 to 3.21, I^2=0%). A sixth trial noted that no patient withdrew due to adverse events by day 5 but was not included in the meta-analysis because it reported no events.[132]

Figure 15. Muscle relaxants versus placebo: withdrawal due to adverse event

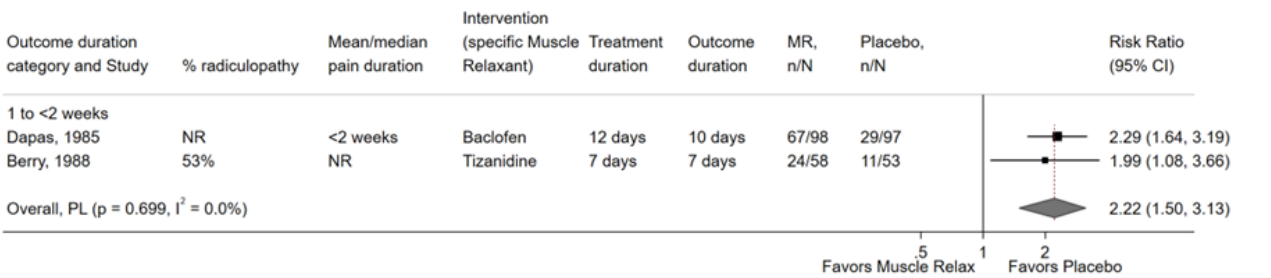


CI = confidence interval; MR = muscle relaxant; NR = not reported; PL = profile likelihood.

Any Adverse Event

Muscle relaxants (i.e., baclofen 10-20 mg, 3-4 times daily or tizanidine 4 mg, 3 times daily) were associated with a large increase in the risk of experiencing at least one adverse event compared with placebo over 7 to 10 days in two trials (N=360) (Figure 16).[70,161]

Figure 16. Muscle relaxants versus placebo: any adverse event



CI = confidence interval; MR = muscle relaxant; NR = not reported; PL = profile likelihood.

The frequency of various mild to moderate adverse events was higher in patients who received an oral muscle relaxant (range, 13.4% to 32.8%; n range, 58 to 277) versus a placebo (4.6% to 22%; n range, 27 to 284) over 7 days across four RCTs (N=1,597).[70,97,104,163] Excluding

the smallest trial,[104] the range for the placebo group is 4.6% to 9.4% (n range, 53 to 284; total N=1,500). The most common side effects reported were central nervous system effects (primarily sedation, drowsiness, dizziness, headache) followed by GI problems (diarrhea, dyspepsia). The trial (N=143) that compared IM injections of thiocolchicoside 4 mg versus placebo reported four cases of adverse events (nausea, heartburn, diarrhea, hypertension and dizziness) in both groups over 5 days.[132] See Appendix E for details.

Differential Efficacy and Safety

In one trial,[161] the patients were divided into two subgroups based on the severity of their symptoms at baseline: severe pain (N=123 randomized; 63 baclofen, 60 placebo) and moderate pain (N=77 randomized; 37 baclofen, 40 placebo). Authors state that patients with initially more severe or extremely severe symptoms benefited most from treatment with baclofen versus placebo, but data was not provided. Authors do not provide p-values for interaction.

Muscle Relaxant Versus Benzodiazepine

Description of Included Studies

One RCT (N=80) compared carisoprodol 350 mg versus diazepam 5 mg four times a day for 7 days for the treatment of acute low back strain or sprain (Appendix E).[151] Mean patient age was 39 years, 52% were female, and most were White (89%). Pregnant women, patients with a history of drug abuse, and patients with depression or taking MAOI inhibitors or tricyclic antidepressants were excluded. Pain duration was 7 days or less (mean duration not reported) and baseline pain intensity was at least moderate. Authors did not report if patients had a history of prior back pain episodes. The trial was conducted in the United States and was considered moderate risk of bias (Appendix F).

Detailed Synthesis

Function, Pain and Other Outcomes

Carisoprodol was associated with a moderate increase in the likelihood of achieving both function and pain response (i.e., slight to no activity impairment or pain) compared with diazepam at days 6 and 7 but the likelihood was similar at day 3 (Table 37). Patients who received carisoprodol showed greater improvement from baseline in VAS Activity Impairment scores at day 7 compared with diazepam; however, improvement was similar between groups at earlier timepoints and at all timepoints for VAS pain scores (Table 37). Effect sizes could not be determined for VAS activity and pain scores with the data provided. A similar pattern was observed for overall pain relief and sleep outcomes; carisoprodol was associated with moderate (for overall pain relief, i.e., marked to complete) and small (for sleep response, i.e., slight to no impairment) increases in the likelihood of achieving response at 6 and 7 days but not at 3 days, while improvement from baseline in VAS sleep scores was similar between groups at all timepoints (See Appendix E for data).[151]

Table 37. Carisoprodol versus benzodiazepine: function and pain outcomes (Boyles, 1983)[151]

Outcome Author, Year	Time	Carisoprodol	Diazepam	RR (95% CI) or p-value*
% Responders: Activity impairment slight to none (1 or 2 on a 1-5 scale)	2-3 days (day 3)	30.5% (11/36)	25.5% (9/35)	1.19 (0.56 to 2.51)
	4 days to <1 week (day 6)	65.0% (23/36)	40.0% (14/35)	1.60 (0.99 to 2.57)
	1 week to <2 weeks (day 7)	72.5% (26/36)	43.5% (15/35)	1.69 (1.09 to 2.60)
VAS Activity Impairment Scale (0-10), mean improvement from baseline (SD NR)	2-3 days (day 3)	-3.10 (n=36)	-2.75 (n=35)	p=NS
	4 days to <1 week (day 6)	-5.3 (n=36)	-4.0 (n=35)	p=NS
	1 week to <2 weeks (day 7)	-5.7 (n=36)	-4.0 (n=35)	p<0.05
% Responders: Pain intensity slight to none (1 or 2 on a 1-5 scale)	2-3 days (day 3)	44% (16/36)	29% (10/35)	1.56 (0.82 to 2.95)
	4 days to <1 week (day 6)	63% (23/36)	44% (15/35)	1.49 (0.95 to 2.35)
	1 week to <2 weeks (day 7)	67% (24/36)	44% (15/35)	1.56 (1.00 to 2.43)
VAS Pain Scale (0-10), mean improvement from baseline (SD NR)	2-3 days (day 3)	-3.65 (n=36)	-3.2 (n=35)	p=NS
	4 days to <1 week (day 6)	-5.45 (n=36)	-4.3 (n=35)	p=NS
	1 week to <2 weeks (day 7)	-5.65 (n=36)	-4.75 (n=35)	p=NS

CI = confidence interval; NS = not significant; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

Adverse Events

Serious adverse events were not reported. While the incidence of any adverse event was lower in the carisoprodol versus diazepam group, the difference was not statistically significant (N=80, 22.5% vs. 35.0%, RR 0.64, 95% CI 0.31 to 1.31).[151] The most common adverse events were central nervous system events (25.0% vs. 42.5%, respectively). One patient (2.5%) in the carisoprodol and two patients in the diazepam (5%) group withdrew due to side effects from the medication (not further described).

Muscle Relaxant Versus Manual Therapy

Description of Included Studies

One RCT (N=192) compared muscle relaxants versus chiropractic adjustments for uncomplicated ALBP of 2 to 6 weeks duration (mean 3.7 weeks) (Appendix E).[74] Mean patient age was 42 years and 43% were female. Pregnant women and patients with previous LBP within 18 months were excluded. Race, ethnicity, comorbidities (medical and psychiatric) and history of substance abuse were not reported. Patients randomized to muscle relaxants received a 2-week supply of one of three medications at the physician's discretion, cyclobenzaprine HCl 5 mg, carisoprodol 350 mg or methocarbamol 750 mg, as well as seven sham chiropractic adjustments (i.e., light pressure, attention control). Patients randomized to chiropractic adjustments received seven sessions involving high-velocity, low-amplitude thrusts in the lumbar, pelvic or sacral spinal region and self-administered placebo capsules over 2 weeks. All participants received acetaminophen as a rescue medication. The trial was conducted in the United States and was assessed as moderate risk of bias (Appendix F).

Detailed Synthesis

Function, Pain and Other Outcomes

Muscle relaxants were associated with similar improvement in ODI scores (0-100 scale) at 2 weeks (N=93, MD -0.03, 95% CI -5.38 to 5.32) and 4 weeks (N=93, MD 4.10, 95% CI -1.75 to 9.95) and VAS pain scores (0-10 scale) at 2 weeks (N=70, MD 0.29, 95% CI -0.75 to 1.33) and 4 weeks (N=70, MD 0.53, 95% CI -0.46 to 1.52) versus chiropractic adjustment.[74] Estimates were imprecise. Likewise, results between groups were similar at all timepoints for modified Zung Self-Rating for Depression scores and analgesic use (Appendix X).

Adverse Events

Adverse events were not reported.

Systemic Corticosteroids

Steroids Versus Placebo or Sham

Description of Included Studies

Six RCTs (N=640) compared a steroid with a placebo intervention (Table 38, Appendix E).[79,85,137,140,156,160] In general, patients were in their mid-30's to 40's and around half were female (range, 30% to 58%) with moderate to severe pain at baseline. Four trials were restricted to patients with radiculopathy or sciatica[79,137,140,160] and two trials excluded patients with radiculopathy.[85,156] One trial included patients with radiculopathy specifically due to disc herniation confirmed on MRI[140] and another included patients with LBP due to a bending or twisting injury specifically.[156] Four trials were conducted in ED settings and mean patient pain duration at baseline ranged from 19 to 53 hours;[79,85,156] one trial did not provide the duration of pain prior to ED admittance but appeared to be acute.[160] Pain duration in the remaining two trials was 15 to 30 days.[137,140] Though not well reported, it is likely that patients had a history of prior LBP episodes. Psychiatric comorbidities were not consistently reported, and no trial reported on history of substance abuse disorders.

Two trials each evaluated a single IM injection of methylprednisolone (same author group),[79,85] an IV bolus of dexamethasone or methylprednisolone,[137,160] and oral prednisone;[140,156] the doses and regimens varied across the two trials of oral prednisone (Table 38). All trials provided patients with standard care which included analgesics (including opioids), referral to physical therapy and education (see Table 39 for details).

Four trials were considered low risk of bias[79,85,137,140] and two moderate risk of bias.[156,160] Primary methodological limitations included baseline differences between groups on important variables and higher than acceptable attrition in some trials (Appendix F).

Table 38. Study characteristics of steroid versus placebo

Study, Year	Friedman, 2006	Friedman, 2008	Balakrishnamoorthy, 2015	Finckh, 2006	Goldberg, 2015	Eskin, 2014
N Randomized	87	82	58	65	269	79
Mean Age (Years)	36	38	38	47	46	40
% Female	58%	53%	52%	52%*	45%	30%
% Radiculopathy	Excluded	100%	100%	100% (sciatica)	100%	Excluded
Pain Duration Inclusion Criteria	<1 week	<7 days	NR (ED presentation)	>1 and <6 weeks	≤3 months	≤48 hours
Mean Pain Duration (days)	53 hours†	Median 48 hours	NR	Median 15	30.4	19 hours
Mean/Median Baseline Pain Severity (0-10)	8.8	9	7.6	5.1	6.7‡	8.0
Previous LBP episode	Mean 2.0 prior episodes	NR (episodes ≤1 month excluded)	NR	NR (recurrent episodes included)	NR	64% back problems
History substance abuse disorder	NR	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	Bipolar: 5.5%	NR	Excluded§	NR	NR
Steroid	Methylprednisolone (IM)	Methylprednisolone (IM)	Dexamethasone (IV)	Methylprednisolone (IV)	Prednisone (Oral)	Prednisone (Oral)
Steroid Dosage, Frequency	160 mg in 2 ml Single Dose	160 mg in 2 ml Single Dose	8 mg in 2 ml Single Dose	500 mg Single Dose	60 mg for 5 days, 40 mg for 5 days, 20 mg for 5 days	50 mg 1x/day
Placebo	NR (identical in look and volume)	NR (identical in look and volume)	Sodium Chloride 0.9% in 2 ml (IV)	Sodium Chloride 0.9% (IV)	NR (identical in look and size)	NR (inactive, identical in look and size)
Treatment duration	Once	Once	Once	Once	15 days	5 days
Duration follow-up	1 month	1 month	6 weeks	1 month	1 year	5-7 days
Country	U.S.	U.S.	Australia	U.S.	U.S.	U.S.
Quality	Low	Low	Moderate	Low	Low	Moderate

ED = emergency department; IM = intramuscular; IQR = interquartile range; IV = intravenous; NR = not reported.
* 45% of glucocorticoid patients were female compared to 59% of placebo patients.
† Steroid 44 hours vs. placebo 63 hours.
‡ Below-waist average pain on 0-10 scale.
§ Psychiatric comorbidities not defined beyond being contraindications to steroids.

Table 39. Concomitant therapies in trials of steroid versus placebo

Study, year	Concomitant Treatment Allowed or Provided to All Participants
Friedman, 2006	• Complementary 1-week supply of naproxen 500 mg tablets and oxycodone 5 mg/acetaminophen 325 tablets and a detailed, standardized low back pain instruction sheet
Friedman, 2008	
Balakrishnamoorthy, 2015	• Standardized regimen of regular analgesia (regular oral acetaminophen/codeine, ibuprofen and oral oxycodone as required), physiotherapy referral and education.
Finckh, 2006	• Standard care, including analgesics (acetaminophen and/or NSAID and/or tramadol) and physical therapy. • During the first 3 days, no concomitant steroid therapy or major analgesics were permitted. After day 3, the treating physician was allowed to use other therapeutic interventions, including glucocorticoids as needed. • Medication use, steroid vs. placebo: NSAID (26% vs. 24%), acetaminophen (77% vs. 79%), tramadol (55% vs. 62%)
Goldberg, 2015	• NSAIDs were not allowed for 3 weeks after randomization, but otherwise all patients in both treatment groups received usual care for their symptoms.
Eskin, 2014	• Rescue medications prescribed at discharge: Opioids 96%, muscle relaxants 82% and NSAIDs 11%

NSAID = nonsteroid anti-inflammatory drug.

Detailed Synthesis

Function

Two trials (N=345) of patients with radicular LBP found systemic corticosteroids associated with a greater likelihood of achieving function success (score of 0 on the 18-item Roland-Morris Disability Questionnaire (RDQ-18) or ≥30% or ≥50% improvement on ODI): the effect was moderate at 3 to 4 weeks and small at 52 weeks (Table 40).[79,140] In one of the trials, results were similar between groups at 1 week.[79] Conversely, two trials (N=153) of patients with nonradicular LBP, found systemic corticosteroids associated with similar likelihood of achieving function success (proportion of patients with a score of 0 on the RDQ-18 or able to return to normal activities) over 1 to 4 weeks (Table 40).[85,156]

Table 40. Proportions of patients achieving function "success" in trials evaluating steroid therapy versus placebo

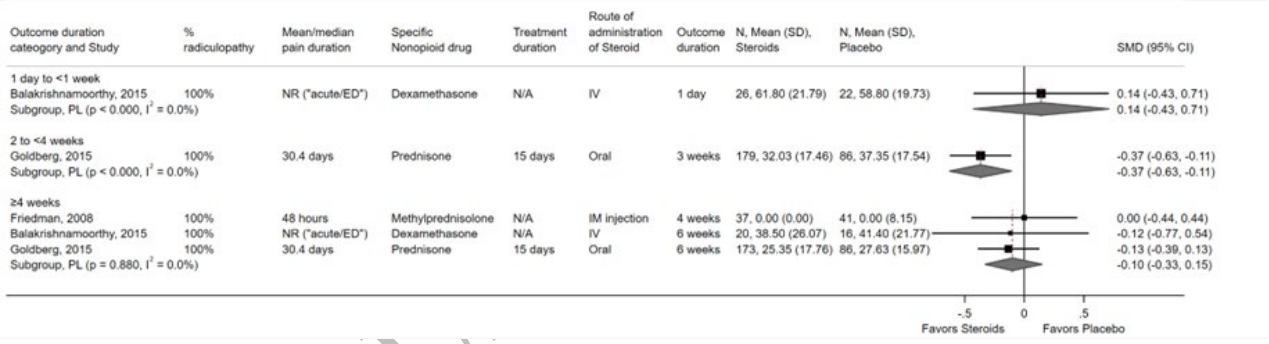
Author, Year Steroid Therapy Indication	Outcome	Time	Steroid vs. Placebo % (n/N) and RR (95% CI)
Friedman, 2006 Methylprednisolone 160 mg single IM injection Nonradicular LBP	RDQ score of 0 (no disability; 0-18 scale)	1 week	70.5% (31/44) vs. 73.8% (31/42), RR 0.95 (0.73 to 1.24)
		4 weeks	77.3% (34/44) vs. 73.8% (31/42), RR 1.05 (0.82 to 1.33)
Eskin, 2014 Prednisone 50 mg 1x/day Nonradicular LBP	Able to resume normal activities	5-7 days	68% (22/32) vs. 68% (24/35), RR 1.00 (0.73 to 1.39)
Friedman, 2008 Methylprednisolone 160 mg single IM injection Radicular LBP	RDQ score of 0 (no disability; 0-18 scale)*	1 week	58% (21/37) vs. 39% (16/41), RR 1.45 (0.90 to 2.34)
		4 weeks	81% (30/37) vs. 51% (21/41), RR 1.58 (1.13 to 2.22)
Goldberg, 2015 Prednisone 20 mg 3x/day for 5 days, 2x/day for 5 days, 1x/day for 5 days	≥30% improvement on ODI	3 weeks	26.8% (48/179) vs. 17.4% (15/88), adjusted RR 1.70 (1.1 to 2.9)
		52 weeks	70.7% (111/157) vs. 56.6% (43/77), adjusted RR 1.3 (1.0 to 1.6)

Author, Year Steroid Therapy Indication	Outcome	Time	Steroid vs. Placebo % (n/N) and RR (95% CI)
Radicular LBP (herniated disc)	≥50% improvement on ODI	3 weeks	33% (59/179) vs. 19.8% (17/88), adjusted RR 1.80 (1.10 to 2.90)
		52 weeks	86.6% (136/157) vs. 68.4% (52/77), adjusted RR 1.20 (1.10 to 1.50)

CI = confidence interval; IM = intramuscular; LBP = low back pain; ODI = Oswestry Disability Index; RDQ = Roland Morris Disability Questionnaire; RR = risk ratio;
* Took the inverse of any disability which was defined as an answer of yes (1 point) to any one of the 18 questions on the RDQ.

Three RCTs that enrolled patients with radiculopathy reported improvement in function based on the ODI [140,160] (0-100 scale) or the RDQ-18,[79] (0-18 scale) (Figure 17). One large RCT at low risk of bias[140] found that oral steroid was associated with a small functional improvement at 3 weeks (N=265, SMD -0.37, 95% CI -0.63 to -0.11) and 52 weeks (N=234, SMD -0.39, 95% CI -0.67 to -0.12, data not in figure) compared with placebo (corresponding MDs on the ODI: -6.40 and -7.11). The baseline duration of pain in this trial was a mean 30 days compared with 2 days or an acute presentation to the ED in the other trials. At other time frames, 1 day (1 RCT, N=48)[160] and 4-6 weeks (3 RCTs, N=373),[79,140,160] functional improvement was similar between steroid (various administration routes) and placebo.

Figure 17. Steroids versus placebo: function scores

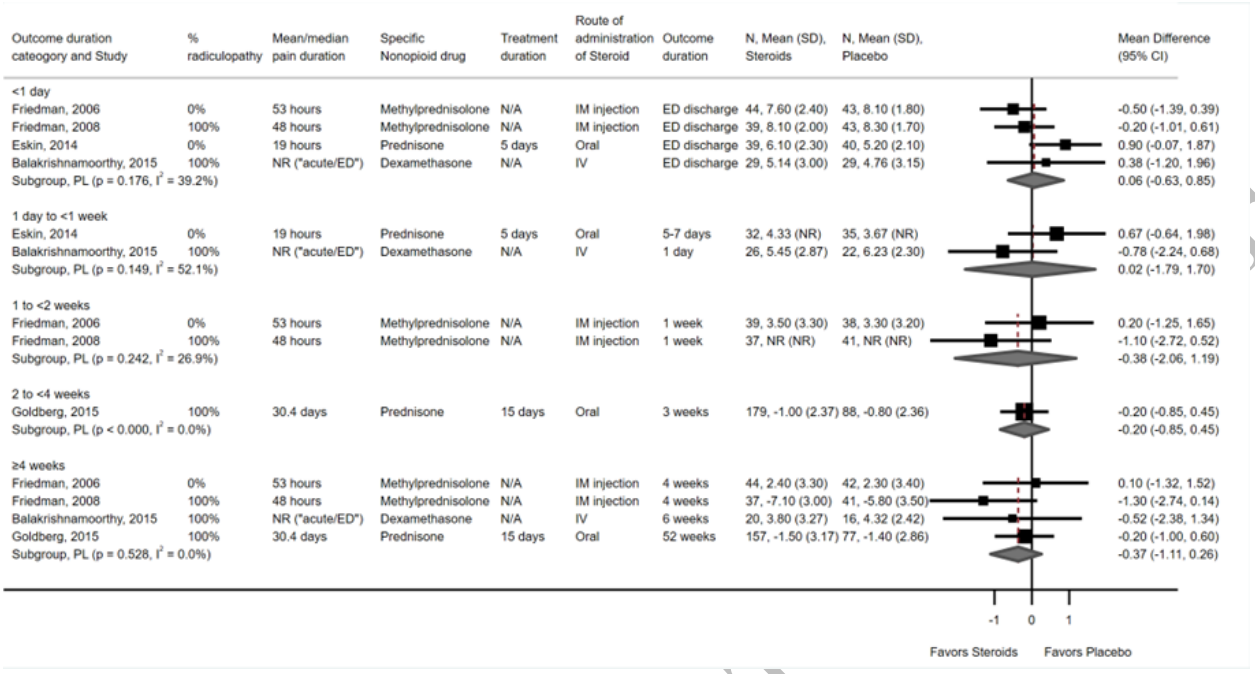


CI = confidence interval; ED = emergency department; IV = intravenous; NA = not applicable; NR = not reported; PL = profile likelihood; SD = standard deviation; SMD = standard mean difference.

Pain

Five RCTs (2 from the same author group) reported mean differences in back pain using a 0 to 10 VAS or NRS scale.[79,85,140,156,160] Overall, systemic corticosteroids were associated with similar improvement in pain compared with placebo at all time periods (Figure 18). Trials used different administration routes and three RCTs enrolled patients with radiculopathy[79,140,160] while two RCTs[85,156] excluded patients with radiculopathy. When trials were stratified based on the presence of radiculopathy there was no effect seen at any timepoint, however, effect estimates tended to favor steroids in patients with radiculopathy (MD range, -0.08 to -1.10) while there was no difference or a slight tendency to favor placebo in patients without radiculopathy (MD range, 0.10 to 0.67); however, estimates were imprecise. When just the two trials evaluating a single IM injection of methylprednisolone were pooled, there was no effect seen at any timepoint (pooed MD range, -0.34 to -0.59); results at 1 to <2 weeks and ≥4 weeks showed some heterogeneity.[79,85] A sixth trial that compared a single IV bolus of methylprednisolone (500 mg) versus placebo for acute sciatica also reported similar improvement between groups in back pain on VAS (p=0.73) over 10 days but did not provide data.[137]

Figure 18. Steroid versus placebo: back pain scores



CI = confidence interval; ED = emergency department; IM = intramuscular; IV = intravenous; NA = not applicable; NR = not reported; PL = profile likelihood; SD = standard deviation.

Two trials, one in patients with radicuopathy[137] and the other in patients with radicular LBP due to a herniated disc,[140] found systemic corticosteroid therapy associated with a similar likelihood of achieving a pain response (≥2 point improvement on 0-10 NRS for leg pain) compared with placebo at all timepoints measured (Table 41). The likelihood of pain response remained similar between groups when different cutoffs were considered (≥3 or ≥5 point improvement) in one of these trials.[140]

Likewise, across both trials, steroids were associated with similar improvement in NRS leg pain scores versus placebo at all timepoints, except for the earliest timepoint (2 days) in the trial evaluating a single IV bolus of methylprednisolone[137] at which time steroid therapy was associated with a small improvement in leg pain (Table 41).

Table 41. Leg pain outcomes in trials evaluating steroid therapy versus placebo for treatment of radicular pain

Author, Year Steroid Therapy Indication	Outcome	Time	Steroid vs. Placebo % (n/N) and RR (95% CI) or MD (95% CI
Finckh, 2006 Methylprednisolone 500 mg, single IV bolus Radicular LBP (sciatica)	≥2 point improvement on NRS leg pain	3 days	48% (15/31) vs. 28% (8/29), RR 1.75 (95% CI 0.88 to 3.51)
	NRS leg pain scores (0-10)	2 days	MD in change scores -0.57 (-1.09 to -0.03), N=60
		3-10 days	Similar improvement (data NR), p=0.22
Goldberg, 2015 Prednisone 20 mg 3x/day for 5 days, 2x/day for 5 days, 1x/day for 5 days Radicular LBP (herniated disc)	≥2 point improvement on NRS leg pain	3 weeks	67% (120/179) vs. 64.8% (57/88), adjusted RR 1.1 (0.9 to 1.3)
		52 weeks	89.8% (141/157) vs. 85.7% (66/77), adjusted RR 1.0 (0.9 to 1.2)
		3 weeks	51.4% (92/179) vs. 51.1% (45/88), adjusted RR 1.0 (0.8 to 1.3)

Author, Year Steroid Therapy Indication	Outcome	Time	Steroid vs. Placebo % (n/N) and RR (95% CI) or MD (95% CI
	≥3 point improvement on NRS leg pain	52 weeks	83.4% (131/157) vs. 77.9% (60/77), adjusted RR 1.1 (0.9 to 1.2)
	≥5 point improvement on NRS leg pain	3 weeks	28.4% (51/179) vs. 26.1% (23/88), adjusted RR 1.1 (0.8 to 1.7)
		52 weeks	67.5% (106/157) vs. 57.1% (44/77), adjusted RR 1.2 (0.9 to 1.4)
	NRS pain below the waist (0-10)*	3 weeks	adjusted MD -0.30 (95% CI -1.00 to 0.40), N=265
		52 weeks	adjusted MD -0.60 (95% CI -1.30 to 0.20), N=234

CI = confidence interval; IV = intravenous; LBP = low back pain; MD = mean difference; NR = not reported; NRS = Numeric Rating Scale; RR = risk ratio

* 6-, 12-, 24-week data was also reported and showed similar results. See detailed data abstraction in Appendix E.

Other Outcomes

One trial that compared a single IV bolus of methylprednisolone 160 mg versus placebo for patients with radicular pain found systemic steroid treatment associated with a substantial decrease in the likelihood of analgesic use at 1 month (22% vs. 43%, RR 0.49, 95% CI 0.24 to 0.99).[79] Across all other outcomes, steroids were associated with similar results compared with placebo, regardless of type of steroid, treatment regimen, follow-up duration, and presence of radiculopathy (Appendix E): 36-item Short Form Questionnaire (SF-36) PCS and MCS quality of life scores (1 RCT),[140] return to work/normal activities (3 RCTs),[79,85,156] recurrence of LBP (2 RCTs),[79,85] surgery within first month (1 RCT)[137] and medication use, to include opioids (3 RCTs).[85,156,160]

Adverse Events

Serious Adverse Events

Five SAEs occurred over a 52-week follow-up period in one trial (N=267) that enrolled patients with radiculopathy and none was considered related to the study medication: three SAEs occurred in the prednisone group (appendectomy, suicide attempt, and deep venous thrombosis) and two in the placebo group (upper gastrointestinal tract hemorrhage and partial nephrectomy for renal cell carcinoma).[140] A second trial (N=67) stated that there were no serious side effects from the study or prescribed medications through 1 week; this trial excluded patients with radiculopathy. Trials were not designed to address SAEs like fracture and serious infection, which are less likely with short-term therapy.[156]

Withdrawal due to Adverse Events

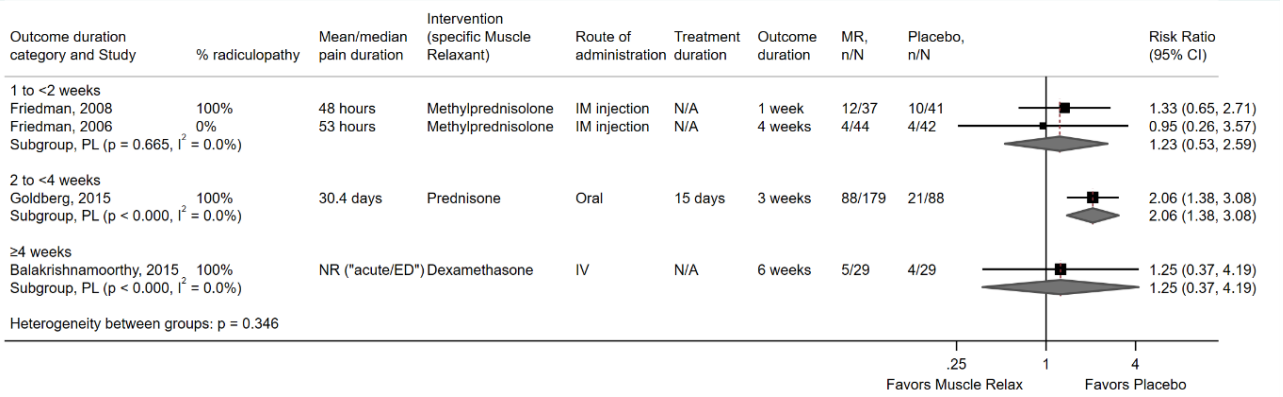
One trial (n=269) reported that one patient in both the steroid (methylprednisolone) and placebo groups discontinued therapy by 3 weeks due to GI symptoms (≤1% in both groups) (RR 0.49, 95% CI 0.03 to 7.77).[140] No other trial reported withdrawal due to adverse events.

Any Adverse Event

Four RCTs (N=489) reported the proportion of patients experiencing at least one adverse event (Figure 19).[79,85,140,160] Three small trials that evaluated IM or IV steroids for patients presenting to the ED with (2 RCTs)[79,160] and without (1 RCT)[85] radiculopathy, reported a similar risk of adverse events compared with placebo over follow-ups ranging from 1 to 6 weeks

(N=222, 19.1% vs. 16.1%, pooled RR 1.24, 95% CI 0.67 to 2.20, I^2=0%, data not shown in plot). In the fourth trial (N=267), oral steroid was associated with a large increase in the risk of adverse events over 3 weeks compared with placebo in patients with radiculopathy (49.2% vs. 23.9%, RR 2.06, 95% CI 1.38 to 3.08); symptom duration at baseline was much longer in this trial compared with the others.[140] A fifth trial (N=60) reported two cases of transient hyperglycemia and one case of facial flush attributed to corticosteroid IV infusion in patients with acute discogenic sciatica; there were no adverse events reported in placebo group.[137] Reporting was insufficient to determine whether all adverse events occurred in distinct participants.

Figure 19. Steroid versus placebo: any adverse event



CI = confidence interval; ED = emergency department; IM = intramuscular; IV = intravenous; MR = muscle relaxant; NA = not applicable; PL = profile likelihood; SD = standard deviation.

Differential Efficacy and Safety

Two trials provided limited information on differential efficacy; p-values for interactions and/or data or stratified analyses were not provided (Appendix E).[137,140] In one trial (N=267) that compared oral prednisone with placebo for radiculopathy due to disc herniation, subgroup analyses revealed no statistically significant interactions at either the 3-week or 52-week timepoints in the between-group changes in the ODI or pain NRS with baseline age, sex, race, ethnicity, elapsed time between onset of symptoms and randomization, lower extremity motor weakness, or presence of a positive straight-leg raise test result.[140] A second trial (N=60) stated that there was no effect modification by concomitant analgesics, NSAIDs, or the presence of neurological deficits for between-group changes in VAS leg pain scores over 10 days for patients with sciatica who received a single IV infusion of methylprednisolone versus placebo.[137]

Acetaminophen

Acetaminophen Versus Placebo

Description of Included Study

One large RCT (N=1,652) compared acetaminophen to placebo for the treatment of ALBP (with or without leg pain) of at least moderate intensity and <6 weeks duration (mean 9.9 days) at baseline (Appendix E).[125] The mean patient 45 years and 47% were female with a mean of 6.9 prior LBP episodes; however, patients were excluded if they had a previous episode within one month of enrollment. Pregnant women were excluded. Race, ethnicity, comorbidities (medical and psychiatric) and history of drug abuse were not reported.

Patients were randomized into three groups – regular doses of acetaminophen, as needed doses of acetaminophen, and placebo – and instructed to take two tablets every 6-8 hours from a regular box, in addition to two tablets from a second box as needed for a total of 4 weeks. Patients in the regular group received 665 mg of modified release acetaminophen tablets in the regular box and placebo in the as-needed box. Patients in the as-needed group received placebo in the regular box and 500 mg of modified release acetaminophen tablets in the as-needed box. Participants in the placebo group received placebo tablets in both boxes. Placebo tablets were identical in appearance to the active tablets.

The trial was conducted in Australia and was considered low risk of bias (Appendix F).

Detailed Synthesis

Function, Pain, and Opioid Use

Patients who took acetaminophen in regular doses or as needed showed similar improvement in RDQ scores and VAS pain scores over all timepoints measured up to 12 weeks compared with placebo.[125] The likelihood of recovery (i.e., 0 or 1 pain on NRS and sustained for 7 days) by 12 weeks was also similar between groups. Concomitant opioid use, both alone and in combination with an NSAID or acetaminophen, was rare during and after treatment (0% to 2.6%) and occurred with similar frequency between the groups (Table 42).

Table 42. Acetaminophen versus placebo: function, pain and opioid use outcomes (Williams, 2014)[125]

Outcome		Mean (SD) or % (n/N)			MD or RR (95% CI)	
Outcome	Time (weeks)	Acetaminophen Regular Dose	Acetaminophen as Needed	Placebo	Acetaminophen Regular Dose vs. Placebo	Acetaminophen as Needed vs. Placebo
Function: RDQ (0-24, worse)	1	7.7 (6.5) (n=513)	8.0 (6.5) (n=498)	8.3 (6.5) (n=500)	MD -0.6 (1.4 to 0.2)	MD -0.3 (-1.11 to 0.51)
	2	5.2 (6.1) (n=507)	5.4 (5.9) (n=496)	5.3 (6.1) (n=497)	MD -0.1 (-0.85 to 0.65)	MD 0.1 (-0.65 to 0.85)
	4	3.2 (5.2) (n=504)	3.5 (5.3) (n=506)	3.3 (5.1) (n=497)	MD -0.1 (-0.74 to 0.54)	MD 0.2 (-0.44 to 0.84)
	12	2.4 (4.7) (n=504)	2.6 (4.9) (n=514)	2.4 (4.5) (n=503)	MD 0 (-0.57 to 0.57)	MD 0.2 (-0.38 to 0.78)
Pain: NRS Pain Scale (0-10, worse)	1	3.7 (2.6) (n=517)	3.8 (2.7) (n=499)	3.6 (2.6) (n=504)	MD 0.1 (-0.22 to 0.42)	MD 0.2 (-0.13 to 0.53)
	2	2.6 (2.6) (n=509)	2.6 (2.5) (n=498)	2.5 (2.5) (n=497)	MD 0.1 (-0.22 to 0.42)	MD 0.1 (-0.21 to 0.41)
	4	1.7 (2.3) (n=509)	1.8 (2.4) (n=507)	1.7 (2.3) (n=499)	MD 0 (-0.28 to 0.28)	MD 0.1 (-0.19 to 0.39)
	12	1.2 (2.2) (n=506)	1.3 (2.2) (n=514)	1.3 (2.3) (n=505)	MD -0.1 (-0.38 to 0.18)	MD 0 (-0.28 to 0.28)
Recovery (pain)*	12	85% (466/550)	83% (452/549)	84% (461/553)	RR 1.01 (0.97 to 1.07)	RR 0.99 (0.94 to 1.04)
Acetaminophen with opioid	1-4	1.7% (9/515)	2.2% (11/505)	2.0% (10/508)	RR 0.89 (0.36 to 2.17)	RR 1.11 (0.47 to 2.58)
	5-12	1.6% (8/504)	2.6% (13/504)	1.8% (9/511)	RR 0.9 (0.35 to 2.32)	RR 1.46 (0.63 to 3.4)
NSAID with opioid	1-4	1.0% (5/515)	0.6% (3/505)	0.4% (2/508)	RR 2.47 (0.48 to 12.65)	RR 1.51 (0.25 to 8.99)
	5-12	0% (0/504)	0.2% (1/504)	0.2% (1/511)	RR NC	RR 1.01 (0.06 to 16.17)

Outcome		Mean (SD) or % (n/N)			MD or RR (95% CI)	
Outcome	Time (weeks)	Acetaminophen Regular Dose	Acetaminophen as Needed	Placebo	Acetaminophen Regular Dose vs. Placebo	Acetaminophen as Needed vs. Placebo
Opioid	1-4	1.4% (7/515)	0.8% (4/505)	0.8% (4/508)	RR 1.73 (0.51 to 5.86)	RR 1.01 (0.25 to 4)
	5-12	0.8% (4/504)	0.6% (3/504)	0.4% (2/511)	RR 2.03 (0.37 to 11.02)	1.52 (0.26 to 9.06)

CI = confidence interval; MD = mean difference; NC = not calculable; NRS = Numeric Rating Scale; RDQ = Roland-Morris Disability Questionnaire; RR = risk ratio; SD = standard deviation.
* The first day of 0 or 1 pain intensity, measured on a 0–10 pain scale, maintained for 7 consecutive days (sustained recovery).

Other Outcomes

Patients who took regular doses of acetaminophen, as needed doses of acetaminophen and placebo showed similar improvement over time across all other outcomes: Global Impression of Change, SF-12 PCS and MCS, sleep quality, hours absent from work, and any concomitant medication use or therapies (Appendix E).[125]

Harms and Adverse Events

Serious adverse events were rare (occurred in 1% of patients in all 3 groups) and considered unrelated to the study treatments. Compared with placebo, acetaminophen was associated with a similar risk of SAEs (any event resulting in hospitalization or death) when taken regularly (N=1,097, RR 0.99, 95% CI 0.29 to 3.42) or as needed (N=1,093, RR 0.8, 95% CI 0.22 to 2.97), and a similar risk of any adverse event when taken regularly (N=1,065, 19% vs. 18%, RR 1.0, 95% CI 0.78 to 1.29) or as needed (N=1,060, 19% vs. 18%, RR 1.01, 95% CI 0.79 to 1.3).[125] See Appendix E for further details regarding specific adverse events. Withdrawals due to adverse events were not reported.

Acetaminophen Versus Electroacupuncture

Description of Included Studies

One small RCT (N=41) compared acetaminophen versus electroacupuncture (Appendix E).[73] Patients ranged in age from 16 to 60 years, and the duration of pain at baseline was <3 days. Sex, race, ethnicity, comorbidities (medical and psychiatric) and history of substance abuse were not reported. Pregnant women were excluded. Patients were randomized to acetaminophen (dose unspecified) taken every 4 hours for 4 days plus two sham electroacupuncture treatments (no current) or to two 15-minute treatments of electroacupuncture and placebo tablets. The trial was conducted in United Kingdom, was poorly reported and considered to be at high risk of bias (Appendix F).

Detailed Synthesis

Function and Pain

Acetaminophen was associated with less improvement in function (1.58 vs. 0.19, 0-10 VAS mobility scores, p<0.05) and pain (1.37 vs. 0.33, 0-10 VAS pain scores, p<0.05) compared with electroacupuncture at 6 weeks; effects estimate were moderate, but no measure of variance was reported.[73] Results at earlier timepoints were similar between the treatment groups, respectively:

1 week (function scores: 2.52 vs. 2.65; pain scores: 2.34 vs. 2.32) and 2 weeks (function scores: 1.70 vs. 1.78; pain scores: 2.20 vs. 1.83).

Harms and Adverse Events

Adverse events were not reported.

Cannabidiol

Cannabidiol Versus Placebo

Description of Included Study

One RCT (N=100) compared cannabidiol (CBD) versus placebo for the treatment of ALBP of <30 days duration (median 1 day) in patients presenting to a hospital ED (Appendix E).[78] Mean Patient age was 47 years, 44% were female, and 21% had a prior history of LBP. Pregnant women were excluded, as were people who had used cannabis or CBD within 7 days of enrollment (no patients had previously used CBD). Race, ethnicity, comorbidities (medical and psychiatric) and history of substance abuse were not reported.

Patients in the CBD group received a single oral syringe with 400 mg CBD in 4 mL medium chain triglyceride (MCT) oil. The placebo was 4 mL of color matched MCT oil delivered in an identical manner. Most patients received standard drug treatment for LBP, acetaminophen 1000 mg (91%) and ibuprofen 400 mg (82%), which varied depending on the patient's past drug use and tolerance of NSAIDs. In addition, rescue analgesia – oxycodone 5 mg every 6 hours – was administered after 120 minutes if pain was not adequately managed and could be administered at any time the ED doctor deemed it clinically indicated.

The trial was conducted in Australia and was considered low risk of bias (Appendix F).

Detailed Synthesis

Pain and Opioid Use

A single dose of CBD was associated with similar improvement in NRS (0-10) pain scores as placebo 2 hours after administration (N=100, MD 0.40, 95% CI -0.64 to 1.44). Similar proportions of patients in both groups received oxycodone as a rescue medication within 4 hours (N=100, 62% vs. 54%, RR 1.15, 95% CI 0.82 to 1.61); the amounts of oxycodone received were also similar between groups (median 5 mg, IQR 0-10 for both groups; total mg per group: 230 vs. 215). Most patients had taken analgesic medications (most commonly acetaminophen and ibuprofen, some opioids) prior to arrival at the ED which may have affected pain scores and outcomes.

Harms and Adverse Events

No serious adverse events were reported, and no participants withdrew due to adverse events. CBD was associated with a similar risk of side effects compared with placebo; the most common side effects were sedation (3 hours: 16% vs. 14%, RR 1.14, 95% CI 0.45 to 2.91; 48 hours: 12% vs. 6%, RR 2.0, 95% CI 0.53 to 7.56), nausea (3 hours: 10% vs. 8%, RR 1.25, 95% CI 0.36 to 4.38; 48 hours: 10% vs. 20%, RR 0.50, 95% CI 0.18 to 1.36), and headache (48 hours: 8% in both groups).

Combination Nonopioid Therapy (Key Questions 5b, 5d, 5e)

Muscle Relaxant Plus NSAID

Muscle Relaxant Plus NSAID Versus Placebo

Description of Included Studies

One RCT (N=88) compared a muscle relaxant (cyclobenzaprine 5 mg) plus an NSAID (diflunisal 500 mg) versus placebo tablets taken twice daily for 7 to 10 days for the treatment of ALBP of 1 to 2 days duration (Appendix E).[75] Patient characteristics (e.g., age, race, sex, history of prior LBP episodes, history of substance abuse) and comorbidities (medical and psychiatric) were not reported. Pain severity at baseline was not reported.

The trial was conducted in Canada and was assessed as high risk of bias due to insufficient information, other than patient and provider blinding, to assess the study quality (Appendix F).

Function and Pain

Function and pain were not reported.

Global Efficacy and Adverse Events

The combination of cyclobenzaprine plus diflunisal was associated with a moderate increase in the likelihood of achieving moderate or marked improvement on a patient-rated global improvement measure versus placebo at 4 days (N=91, 69.6% vs. 46.7%, RR 1.49, 95% CI 1.03 to 2.15) but there was no difference at other timepoints (2 days: N=88, 46.5% vs. 33%, RR 1.40, 95% CI 0.83 to 2.35; 7 to 10 days: N=89, 77.1% vs.73.2%, RR 1.05, 95% CI 0.83 to 1.34).[75] It is unclear how global improvement was measured but appeared to be a composite of the following: 1) presence of local pain 2) presence of muscle spasm 3) presence of muscle tenderness on palpation 4) limitation of range of motion, and limitation of activities of daily living. This trial did not report other clinical outcomes. Authors reported that no significant adverse effects were attributable to the therapy.

Muscle Relaxant Plus NSAID Versus NSAID

Description of Included Studies

Eight RCTs (N=1,363) compared a muscle relaxant plus NSAID versus the same NSAID alone (Table 43; Appendix E).[57,58,69,75,87,113,126,128] Briefly, patient mean age ranged from 39 to 52 years and 39% to 55% were female. Patients with radiculopathy or sciatica were excluded from five RCTs[57,87,113,126,128] and one RCT[69] included a mixed population (moderate/severe (30%) and no/mild (70%) radiculopathy); two trials did not specific inclusion or exclusion of patients with radiculopathy.[58,75] Three trials reported mean pain duration at baseline (range 56 to 60 hours) and the proportion of patients with prior back pain (55% to 77%).[113,126,128] Mean pain intensity at baseline was ≥6 (on a 0 to 10 scale) in three trials.[57,58,87] In two trials, 2% to 4% of patients had a diagnosis of depression.[113,126,128] History of substance abuse disorders and current or prior opioid use was not reported by the trials.

The type of muscle relaxant and NSAID varied, as did the doses (Table 43). Three trials evaluated tizanidine,[69,87,126] two trials cyclobenzaprine,[75,113] two trials methocarbamol,[57,128] and one trial each orphenadrine,[128] baclofen,[126] metaxalone,[126] and thiocolchicoside.[58] The route of

administration was oral (taken 2 to 3 times per day for 7 to 10 days) in all but one trial which used a single, fixed-dose combination intramuscular (gluteal) injection (thiocolchicoside) and followed patients for 1 day.[58] The same NSAID, dose and regimen used in the combination therapy was used as the comparator. Two trials had multiple arms and evaluated two or three different muscle relaxants.[126,128] Concomitant therapies (e.g., other analgesics, steroids, etc.) were not allowed in five trials[58,69,113,126,128] and not reported in the remainder.[57,75,87] Three trials by the same author group provided a brief LBP educational session to all participants.[113,126,128] Three trials were conducted in ED settings[113,126,128] and the remainder in primary care or other outpatient settings.

Three trials (by the same author group) were assessed as low risk of bias,[113,126,128] four moderate risk of bias[57,58,69,87] and one high risk of bias.[75] Methodological limitations in the moderate and high risk of bias included unclear randomization and concealments methods and baseline differences between groups (Appendix F).

Table 43. Study characteristics of combination nonopioid therapy: NSAID plus muscle relaxant versus NSAID alone

Study, Year	Basmajian, 1989*	Friedman, 2015*	Friedman, 2018	Friedman, 2019	Samsamshariat, 2021	Iliopoulos, 2023	Pareek, 2009	Berry, 1988
N Randomized	88	215	240	320	64	134	197	105
Mean Age (yrs.)	NR	39	39	39	41	52	43	43
Female	NR	46%	45%	42%	41%†	55%†	39%	45%
Radiculopathy	NR	Excluded	Excluded	Excluded	Excluded	NR	Excluded	Mixed‡
Pain Duration Inclusion Criteria	≤2 days	NR	<2 weeks	<2 weeks	<6 weeks	≤7 days	≤30 days	NR
Mean Pain Duration (days)	NR	56 hours	56 hours	60 hours	NR	NR	NR	NR
Mean/Median Baseline Pain Severity (0-10)	NR	NR	NR	NR	7.1	6.9	≥6	NR
Prior LBP episodes	NR	55%	69%	77%	NR	NR	NR	NR
History substance abuse disorder	NR	NR	NR	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	NR	Depression: 2%	Depression: 4%	Excluded	NR	NR	NR
Muscle Relaxant (+ same NSAID)	Cyclobenzaprine 5 mg 2x/day	Cyclobenzaprine 5-10mg 3x/day	(1) Orphenadrine 100 mg 2x/day or (2) Methocarbamol 750 mg 3x/day	(1) Baclofen 10-20 mg (2) Metaxalone 400-800 mg or (3) Tizanidine 2-4 mg; 3x/day	Methocarbamol 550 mg 3x/day	Thiocolchicoside 4mg/4ml, FDC single IM injection	Tizanidine 2 mg 2x/day	Tizanidine 4 mg 3x/day
NSAID	Diflunisal 500 mg 2x/day	Naproxen 500mg 2x/day	Naproxen 500 mg 2x/day	Ibuprofen 600 mg every 3x/day	Indomethacin 25 mg 3x/day	Diclofenac 75 mg/3ml, single IM injection	Aceclofenac 100 mg 2x/day	Ibuprofen 400 mg 3x/day
Treatment duration	7-10 days	10 days	7 days	7 days	7 days	1 injection	7 days	7 days
Duration of follow-up	7-10 days	12 weeks	7 days	7 days	7 days	1 day	7 days	7 days
Country	Canada	U.S.	U.S.	U.S.	Iran	Greece	India	UK
Risk of Bias	High	Low	Low	Low	Moderate	Moderate	Moderate	Moderate

FDC = fixed dose combination; IM = intramuscular; LBP = low back pain; NR = not reported; NSAID = nonsteroid anti-inflammatory drug.

* Friedman 2015 had three arms: oxycodone/acetaminophen + naproxen vs. cyclobenzaprine + naproxen vs. naproxen + placebo. Only the combination of cyclobenzaprine plus naproxen vs. naproxen alone included here; Basmajain 1989 had three arms: combination Muscle Relaxant + NSAID vs. NSAID alone vs. placebo. Only the combination versus NSAID alone listed here.

† Samsamshariat, 2021: 50% of patients in the methocarbamol + NSAID group were female compared to 31% of patients in the NSAID alone group.

‡ Moderate/Severe (30%) and none/mild (70%) radiculopathy.

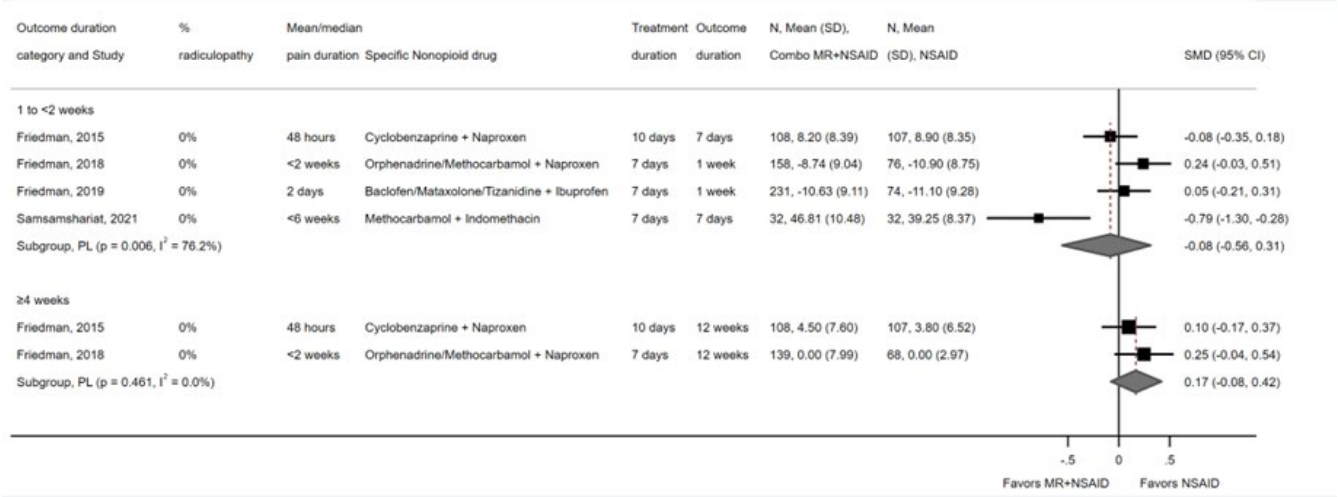
Detailed Synthesis

Function

Four RCTs (N=839) reported function using the RDQ (0-24 scale)[113,126,128] or Back Pain Function Scale (0-60 scale).[57] All trials excluded patients with radiculopathy. A muscle relaxant plus NSAID was associated with similar effects on function at 1 to <2 weeks (4 RCTs, N=818, pooled SMD -0.08, 95% CI -0.56 to 0.31, I^2=76.2%)[57,113,126,128] and at ≥4 weeks (2 RCTs, N=422, pooled SMD 0.17, 95% CI -0.08 to 0.42, I^2=0%)[113,128] compared with an NSAID alone (Figure 20). Statistical heterogeneity (I^2=76%) was present at 1 to <2 weeks. All trials were conducted in an ED setting except for one conducted in an outpatient clinic. Excluding this trial (that found a moderate effect for combination therapy on the Back Function Scale, MD 7.56),[57] which was also assessed as high risk of bias, did not change conclusions at 1 to <2 weeks but decreased statistical heterogeneity (3 RCTs, N=754, pooled SMD 0.07, 95% -0.14 to 0.28, I^2=27.6%).[113,126,128] The three ED trials[113,126,128] that evaluated function using the RDQ (0-24 scale) reported pooled differences of 0.57 (at 1 to <2 weeks) to 1.11 (≥4 weeks).

No trial reported the likelihood of achieving a successful functional outcome (e.g., >30% improvement in a functional scale).

Figure 20. Combination of muscle relaxant plus NSAID versus NSAID alone: function scores



CI = confidence interval; Combo = combination therapy; MR = muscle relaxant; NSAID = nonsteroidal anti-inflammatory drug; PL = profile likelihood; SD = standard deviation; SMD = standard mean difference.

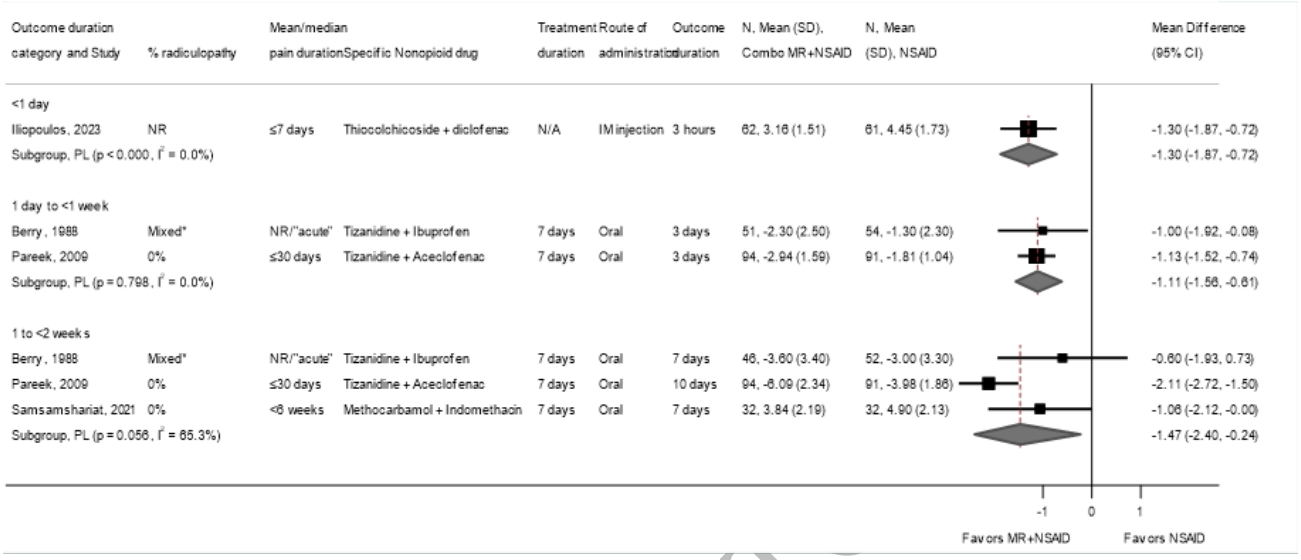
Pain

One trial (N=123) found a single intramuscular injection of thiocolchicoside plus diclofenac associated with a moderate increase in the likelihood of achieving a clinically meaningful improvement in pain (i.e., >30% improvement from baseline on a 0-10 VAS) compared with diclofenac injection alone at 1 hour (56.5% vs. 37.7%, RR 1.50, 95% CI 1.01 to 2.21) and 3 hours (91.9% vs. 57.4%, RR 1.60, 95% CI 1.28 to 2.01).[58]

Four RCTs (N=470) reported average differences in VAS pain scores (0-10 scale).[57,58,69,87] A muscle relaxant plus NSAID was associated with moderately reduced pain at all timepoints compared with NSAIDs alone (Figure 21): <1 day (1 RCT, N=123, IM steroid injection),[58] 1 day to <1 week (2 RCTs, N=290)[69,87] and 1 to <2 weeks (3 RCTs, N=347);[57,69,87] difference ranged from -1.11 to -1.47. At the latest timepoint, statistical heterogeneity was present (65.3%);

however, all trials reported results that favored combination therapy (differences ranged from -0.60 to -2.11 points).

Figure 21. Combination of muscle relaxant plus NSAID versus NSAID alone: VAS pain scores



CI = confidence interval; Combo = combination therapy; MR = muscle relaxant; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; PL = profile likelihood; SD = standard deviation; VAS = visual analogue scale.
*Moderate/severe (30%) and none/mild (70%) radiculopathy.

Other Pain Outcomes

Two RCTs (N=339) reported completeness of pain relief (Table 44).[69,128] One trial[69] in a mixed population with and without radiculopathy found tizanidine plus ibuprofen associated with a small increase in the likelihood of mild or no LBP during rest at 3 days (N=105, 90.2% vs. 72.2%, RR 1.25, 95% CI 1.03 to 1.51) and 7 days (N=98, 93.5% vs. 76.9%, RR 1.22, 95% CI 1.03 to 1.44) compared with ibuprofen alone, but a similar likelihood of mild or no sciatica. A second trial (N=234) that excluded patients with radiculopathy found orphenadrine and methocarbamol associated with a similar likelihood of mild or no LBP during the past 24 hours compared with naproxen at 1 and 12 weeks.[128]

Global Efficacy

One trial (N=185) in patients with nonradicular pain found tizanidine plus aceclofenac associated with a large increase in the likelihood of patients' rating their overall improvement as excellent or good at 7 days versus aceclofenac alone (75.5% vs. 34.1%, RR 2.22, 95% CI 1.63 to 3.02).[87] A second trial (N=87) found cyclobenzaprine plus diflunisal associated with a similar likelihood of moderate or marked improvement overall as rated by the clinician compared with diflunisal alone at all timepoints up to 10 days (Table 44).[75]

Table 44. Combination muscle relaxant plus NSAID versus NSAID alone: other pain outcomes and global efficacy

Author, Year Radicular Symptoms Combination Pharmacologic Therapy	Patient-Reported Outcome	Time	Combination Pharmacologic Therapy vs. Same NSAID Alone % (n/N) and RR (95% CI)
Friedman, 2018[128] Nonradicular LBP Orphenadrine 100 mg plus Naproxen 500 mg, 2x/day	**Mild or no LBP during past 24 hours**	1 week	66.7% (52/78) vs. 65.8% (50/76), RR 1.01 (0.81 to 1.27)
		12 weeks	84.1% (58/69) vs. 80.9% (55/68), RR 1.04 (0.89 to 1.21)
Friedman, 2018[128] Nonradicular LBP Methocarbamol 750-1500 mg plus Naproxen 500 mg, 2-3x/day		1 week	61.3% (49/80) vs. 65.8% (50/76), RR 0.93 (0.73 to 1.18)
		12 weeks	78.6% (55/70) vs. 80.9% (55/68), RR 0.97 (0.82 to 1.15)
Berry 1988[69] No/mild (70%) and moderate/severe (30%) radiculopathy Tizanidine 4 mg + Ibuprofen 400 mg, 3x/day	**Mild or no back pain at rest**	3 days	**90.2% (46/51) vs. 72.2% (39/54), RR 1.25 (1.03 to 1.51)**
		7 days	**93.5% (43/46) vs. 76.9% (40/52), RR 1.22 (1.03 to 1.44)**
	Mild or no sciatica pain at rest	3 days	88.2% (45/51) vs. 79.2% (42/53), RR 1.11 (0.94 to 1.32)
		7 days	84.8% (39/46) vs. 86.3% (44/51), RR 0.98 (0.83 to 1.16)
Pareek, 2009[87] Nonradicular LBP Tizanidine 2mg plus Aceclofenac 100 mg, 2x/day	**Global rating: excellent to good response (patient rated)**	7 days	**75.5% (71/94) vs. 34.1% (31/91), RR 2.22 (1.63 to 3.02)**
Basmajian, 1989[75] Radiculopathy NR Cyclobenzaprine 5 mg plus Diflunisal 500 mg, 2x/day	**Global Rating: moderate or marked improvement (clinician rated)**	2 days	46.5% (20/43) vs. 43% (19/44), RR 1.08 (0.68 to 1.72)
		4 days	69.6% (32/46) vs. 60% (24/40), RR 1.16 (0.84 to 1.59)
		7-10 days	77.1% (37/48) vs. 79% (31/39), RR 0.97 (0.78 to 1.21)

CI = confidence interval; LBP = low back pain; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; RR = risk ratio.

Quality of Life, Sleep Quality

Quality of life and sleep quality were not reported.

Psychological Distress

Psychological distress was not reported.

Return to Work

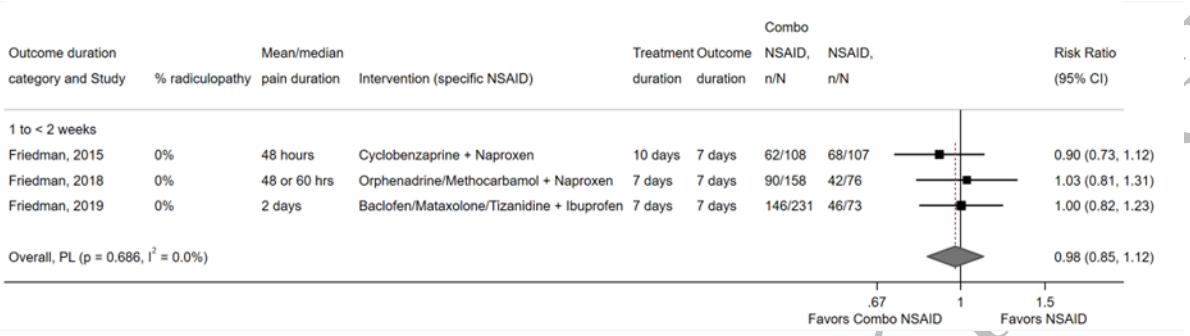
Two trials (N=527) reported no differences in return to work outcomes.[113,126] One trial (N=215) found no difference between cyclobenzaprine plus naproxen versus naproxen alone in time to return to work (median 3 days, IQR 2 to 4 days).[113] The other trial (N=312) found that, compared with ibuprofen alone, the addition of baclofen 10-20 mg (N=156, RR 1.08, 95% CI 0.78 to 1.50), metaxalone 400-800 mg (N=155, RR 0.88, 95% CI 0.61 to 1.25) or tizanidine 2-4 mg (N=155, RR 0.99, 95% CI 0.70 to 1.38) were each associated with similar likelihood of return to work/normal activities by 2 days.[126]

Medication Use

One trial (n=215) found cyclobenzaprine plus naproxen and naproxen alone associated with similar and low rates of opioid use at 3-month follow-up, 1% versus 3%, RR 0.33 (95% CI 0.03 to 3.13).[113]

Three RCTs (N=753) reported that the combination of a muscle relaxant plus NSAID was associated with similar likelihood at 7 days of additional medication use (unspecified) for back pain in the last 24 hours (Figure 22).[113,126,128]

Figure 22. Combination of muscle relaxant plus NSAID versus NSAID alone: medication use within 24 hours



CI = confidence interval; NSAID = nonsteroidal anti-inflammatory drug; PL = profile likelihood.

Adverse Events

Serious Adverse Events

Three RCTs (N=594) reported that no SAEs occurred in any patients over 7 days.[75,126,128] SAEs were not mentioned in any of the other trials.

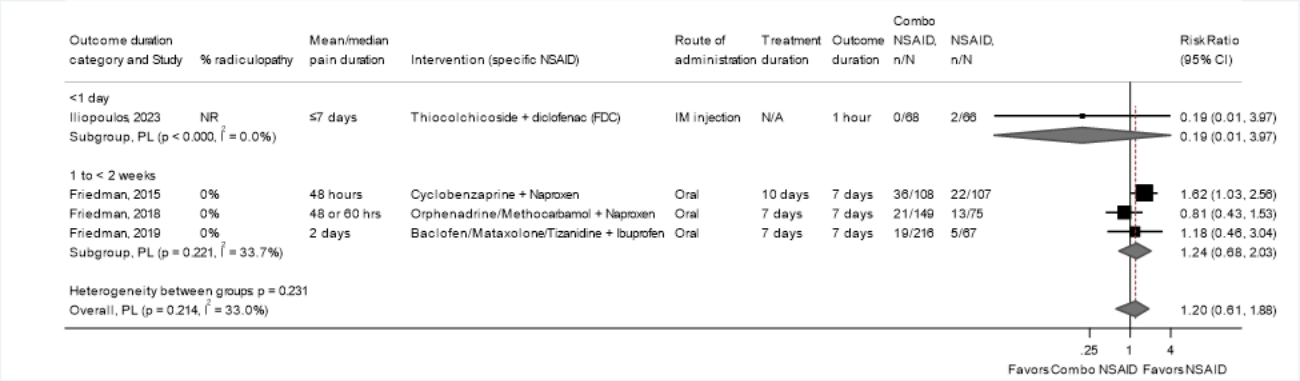
Withdrawals due to Adverse Events

One trial (N=105) found tizanidine plus ibuprofen associated with increased likelihood of study withdrawal due to adverse events versus ibuprofen alone; however, results were based on 6 events and the estimate was very imprecise (RR 5.29, 95% CI 0.64 to 43.78).[69] One RCT (N=134) reported that no patient withdrew from the study due to adverse events following a single IM injection of thiocolchicoside plus diclofenac.[58]

Any Adverse Event

Muscle relaxant plus NSAID therapy was associated with a similar risk of any adverse event (Figure 23) compared with the NSAID alone at 1 hour after IM injection (1 RCT, N=134, 0% vs. 3.0%, RR 0.19, 95% CI 0.01 to 3.97)[58] and over 7 days of oral therapy (3 RCTs, N=722, 16.1% [76/473] vs. 16.1% [40/249], pooled RR 1.24, 95% CI 0.68 to 2.03, I^2=33.7%).[113,126,128] All estimates were imprecise.

Figure 23. Combination of muscle relaxant plus NSAID versus NSAID alone: any adverse event

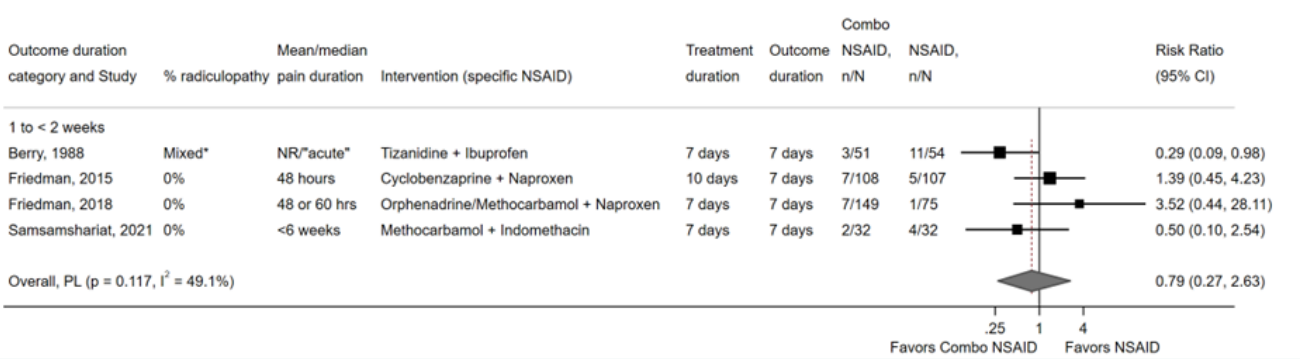


CI = confidence interval; FDC = ; NSAID = nonsteroidal anti-inflammatory drug; PL = profile likelihood.

Gastrointestinal-Related Adverse Events

Four RCTs (N=608) found combination muscle relaxant plus NSAID therapy was associated with a similar risk of GI problems compared with the NSAID alone at 7 days (5.6% vs. 7.8%, pooled RR 0.79, 95% CI 0.27 to 2.63, I^2=49.1%) (Figure 24).[57,69,113,128] The estimates were imprecise. Events in the muscle relaxant versus NSAID groups included "a lot" of stomach irritation (range, 4-6% vs. 1-5%) in two trials (N=264)[113] and nausea or vomiting (4% vs. 6%) in one trial (N=115).[113] The GI problems were unspecified in the other two trials (N=169).[57,69]

Figure 24. Combination of muscle relaxant plus NSAID versus NSAID alone: gastrointestinal-related adverse events



CI = confidence interval; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; PL = profile likelihood.
*Moderate/Severe (30%) and none/mild (70%) radiculopathy.

Central Nervous System-Related Adverse Events

One trial (N=105) reported a large increase in the likelihood of experiencing a central nervous system (CNS) adverse event (33.3% vs. 9.3%, RR 3.60, 95% CI 1.43 to 9.04) following combination drug treatment (tizanidine and ibuprofen) compared with ibuprofen alone at 7 days.[69] CNS events were usually mild and included drowsiness (n=10), dry mouth (n=3), tiredness and lightheadedness (n=2 each), and sedation and vertigo (n=1 each). The risk of other, unspecified adverse events in this trial was similar between treatment groups (5.9% vs. 2.0%, RR 3.18, 95% CI 0.34 to 29.56). Two trials (N=364) found methocarbamol, orphenadrine and cyclobenzaprine plus naproxen associated with a higher incidence of drowsiness compared with naproxen alone, respectively, but only the difference for methocarbamol was significant: 17% vs. 7%(RR 2.60,95% CI 0.98 to 6.93); 11% vs. 7%(RR 1.62,95% CI 0.56 to 4.73); 7% vs. 4% (RR 1.73,95% CI 0.53 to 5.75).[113,128] The incidence of dizziness (3-5% vs. 3%) was similar for all

muscle relaxant combinations versus an NSAID alone in these trials.[113,128] In a fourth trial, overall 12.2% (24/197) of patients experienced at least one adverse event possibly related to the study drugs (dizziness, drowsiness, vomiting, dyspepsia,) with no difference between treatment groups (data not provided by treatment group).[87]

Differential Efficacy and Safety

One small RCT (N=64) that compared methocarbamol plus indomethacin versus indomethacin alone provided 1-week pain and function outcomes stratified by patient sex (Table 45).[57] However, authors do not provide p-values for interaction and visual inspection of the confidence intervals indicates considerable overlap and likely no modification of treatment effect.

Table 45. Combination muscle relaxant plus NSAID versus NSAID alone: differential efficacy from RCTs

Author, Year Therapies	Follow-up	Outcome	Subgroup	Combination Therapy Mean (SD)	NSAID Alone Mean (SD)	MD (95% CI)*
Samsamshariat 2021[57] Methocarbamol 500 mg Indomethacin 25 mg 3x/day	1 week	**VAS pain** (0-10 (worst)	Males	3.31 (2.55) (n=16)	4.71 (2.14) (n=22)	-1.40 (-2.95 to 0.15)
			Females	4.37 (1.67) (n=16)	5.50 (2.14) (n=10)	-1.13 (-2.68 to 0.42)
		BPFS (0-60 (best)	Males	49.87 (10.82) (n=16)	39.67 (8.89) (n=22)	10.20 (3.91 to 16.69)
			Females	43.75 (9.47) (n=16)	38.00 (6.93) (n=10)	5.75 (-1.41 to 12.91)

BPFS = Back Pain Function Scale; CI = confidence interval; MD = mean difference; NSAID = nonsteroidal anti-inflammatory drug; RCT = randomized controlled trial; SD = standard deviation; VAS = visual analog scale.
* Calculated by AAI/OHSU to compare effect sizes and overlap of confidence intervals for male and female subgroups.

Acetaminophen Plus NSAID

Acetaminophen Plus NSAID Versus NSAID Alone

Description of Included Studies

Two RCTs (total N=280; range 120 to 160) compared the combination of acetaminophen (acetaminophen) and an NSAID versus an NSAID alone for the treatment of acute (mean 2 to 4 days), nonradicular LBP in an ED setting (Table 46, Appendix E).[72,141] The average study age of participants was 38 years and half were female. In one trial[72] most patients (90%) had a prior history of LBP; however, patients were excluded if their baseline LBP frequency was once (or more) per month. Pregnant women were excluded in both trials. Race, ethnicity, comorbidities (medical and psychiatric), and history of opioid or substance abuse disorder were not reported. Both trials included patients with functionally impairing back pain, defined as a score of >5 on the RDQ at baseline.

Both RCTs compared acetaminophen/acetaminophen combined with different NSAIDs versus the same NSAIDs alone: ibuprofen[72] or etodolac.[141] Patients in the NSAID only group also received a placebo capsule. The route of administration was oral in both trials. Individual doses and frequency of dosing varied across the trials and treatment duration ranged from 7 to 10 days (Table 46). Concomitant treatment included a 10-minute educational intervention on back

care in one trial.[72] In the other trial, additional medications, except for topical analgesics, were prohibited.[141]

One trial was considered low risk of bias[72] and the other moderate risk of bias.[141] Limitations in the moderate risk-of-bias trial included unclear allocation concealment methods and lack of masking of study team, care team, and/or patients (Appendix F).

Table 46. Study characteristics of combination nonopioid therapy: acetaminophen plus NSAID versus NSAID alone

Study, year	Bayram, 2021	Friedman, 2020
N Randomized	160	120
Mean Age (Years)	37	41
% Female	56%	48%
% Radiculopathy	0% (Excluded)	0% (Excluded)
Pain Duration Inclusion Criteria	<2 weeks	<2 weeks
Mean Pain Duration (days)	3.8	2
Mean Baseline Pain Severity (0-10)	6.9	NR*
Prior history of LBP	NR	90%
History substance abuse disorder	Chronic analgesic use: 0% (Excluded)	NR
Psychiatric Comorbidities	NR	NR
Nonopioid, NSAID combination	Acetaminophen + Etodolac	Acetaminophen + Ibuprofen
Comparator (NSAID alone)	Etodolac	Ibuprofen
Dosage Frequency	Acetaminophen: 500mg Etodolac: 400 mg 2x per day	Acetaminophen: 500-1000 mg Ibuprofen: 600 mg Every 6 hrs.; flexible dose
Treatment duration	10 days	7 days
Duration follow-up	10 days	7 days
Country	Turkey	U.S.
Quality	Moderate	Low

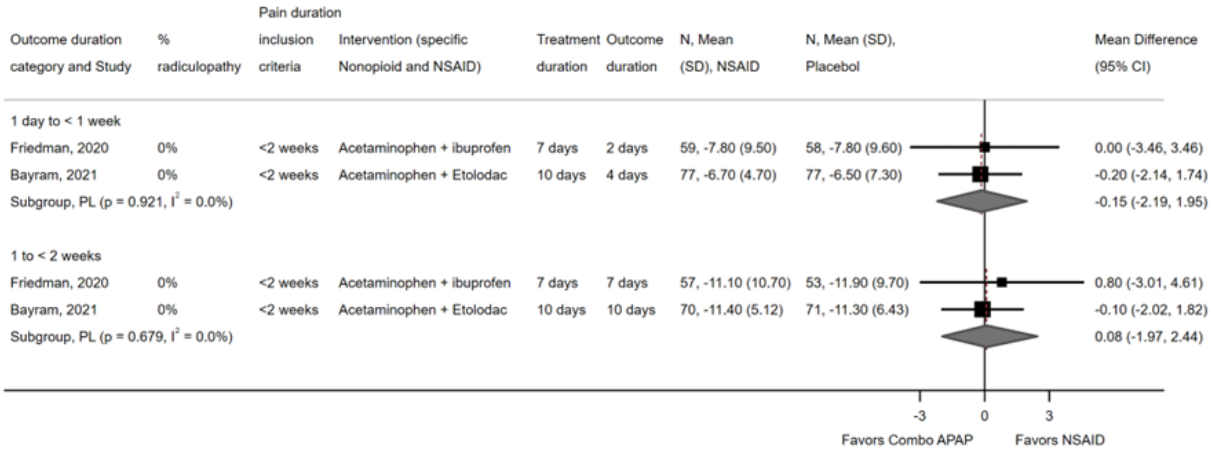
LBP = low back pain; NR = not reported; NSAID = nonsteroid anti-inflammatory drug.
* Only included patients who had functionally impairing back pain, defined as a score of >5 on the RDQ.

Detailed Synthesis

Function

Both RCTs found combination therapy with acetaminophen plus an NSAID (ibuprofen or etodolac) associated with similar improvement in RDQ scores (0-24 scale) at 2 to 4 days (N=271, pooled MD –0.15, 95% CI –2.19 to 1.95, I^2=0%) and 7 to 10 days (N=251, pooled MD 0.08, 95% CI –1.97 to 2.44, I^2=0%) of follow-up compared with NSAIDs alone (Figure 25).[72,141]

Figure 25. Combination acetaminophen plus NSAIDs versus NSAIDs alone: RDQ scores



CI = confidence interval; PL = profile likelihood; NSAID = nonsteroidal anti-inflammatory drug; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation.

Pain

One RCT found combination therapy with acetaminophen 500 mg plus etodolac 400 mg twice per day associated with similar improvement in VAS pain scores (0-10 scale) at 4 days (N=154, MD -0.01, 95% CI -0.58 to 0.56) and 10 days (N=141, MD -0.01, 95% CI -0.78 to 0.76) compared to etodolac alone.[141] In the second RCT, the likelihood of experiencing mild or no pain (self-reported worst pain during previous 24 hours) was similar for patients who received acetaminophen 500-1000mg plus ibuprofen 600 mg flexible dosing compared with ibuprofen alone at 2 days (N=117, 54% vs. 57%, RR 0.95, 95% CI 0.69 to 1.32) and 7 days (N=120, 72% vs. 72%, RR 1.08, 95% CI 0.83 to 1.40).[72]

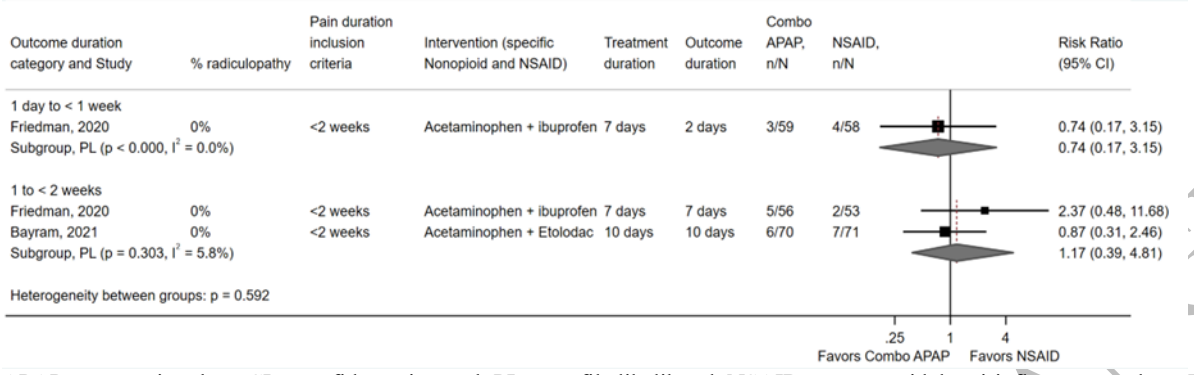
Return to Work and Recurrence of LBP

One trial (N=120) found that the number of days until return to usual activities was similar between the combination therapy (acetaminophen plus ibuprofen) and ibuprofen alone groups (MD 0.6 days, 95% CI -0.5 to 1.7).[72] The frequency of LBP during the previous 24 hours (frequent/always) was also similar between groups at 7 days in this trial (21% vs. 13%, 95% CI 0.72 to 4.05).

Adverse Events

One trial (N=141) reported that no serious adverse events (e.g., allergic reaction or shock) occurred.[141] Combination therapy (acetaminophen plus ibuprofen or etodolac) was associated with a similar risk of any adverse event compared with NSAID monotherapy at 2 days (1 RCT)[72] and 7 to 10 days (2 RCTs)[72,141] (Figure 26). Adverse events were minor and included primarily GI (e.g., abdominal pain, diarrhea, dyspepsia) and CNS (e.g., drowsiness, dizziness, nausea, blurry vision) symptoms.

Figure 26. Combination acetaminophen plus NSAIDs versus NSAIDs alone: Adverse Events



APAP = acetaminophen; CI = confidence interval; PL = profile likelihood; NSAID = nonsteroidal anti-inflammatory drug; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation.

Benzodiazepine Plus NSAID

Benzodiazepine Plus NSAID Versus NSAID Alone

Description of Included Studies

One RCT (N=114) compared the combination of a benzodiazepine plus an NSAID to an NSAID alone for the treatment of acute (median 60 hours), nonradicular LBP in an ED setting (Appendix E).[71] Patient mean age was 36 years, 45% were female, 56% had a prior history of LBP (participants with ≥1 episodes of back pain per month were excluded) and 4% screened positive for depression. Pregnant or breastfeeding woman were excluded. The trial did not report on medical comorbidities, history of substance abuse disorder, or history or treatment of opioid use disorder. Only patients with functionally impairing back pain, defined as a score of >5 on the RDQ, were enrolled. Patients in the combination therapy group received naproxen 500 mg plus diazepam 5 to 10 mg (1-2 tablets) every 12 hours and those in the monotherapy group received naproxen 500 mg every 12 hours plus 1-2 tablets of placebo. Treatment duration was 7 days. Concomitant treatment included a 10-minute educational intervention on back care. The trial was conducted in the United States and was considered at low risk of bias (Appendix F).

Detailed Synthesis

Function

Combination therapy with diazepam plus naproxen was associated with similar improvement in function scores on the RDQ (0-24) at 1 week (N=112, MD in change scores from baseline -0.3, 95% CI -3.5 to 2.8) and 3 months (N=103, MD in change scores from baseline -2.0, 95% CI -4.2 to 0.3) compared with naproxen alone.[71]

Pain

Combination therapy with diazepam plus naproxen was associated with a similar likelihood of experiencing mild or no pain (self-reported worst LBP) during the previous 24 hours at 1 week (N=112, 68% vs. 78%, RR 1.45, 95% CI 0.77 to 2.72) and during the previous 72 hours at 3 months (N=103, 88% vs. 91%, RR 0.97, 95% CI, 0.85 to 1.11).[71]

Return to Work and Recurrence of LBP

The number of days until return to usual activities was similar between patients who received combination therapy (diazepam plus naproxen) compared with naproxen alone (MD -0.4 days, 95% CI -0.6, 1.4.[71] The frequency of LBP (never or rarely) was also similar between treatment groups: 1 week (previous 24 hours, N=112, 49% vs. 56%, RR 0.90, 95% CI 0.63 to 1.29) and 3 months (previous 72 hours, N=103, 84% vs. 79%, RR 1.06, 95% CI 0.88 to 1.27).

Adverse Events

There were no serious or unexpected side effects. The risk of any adverse event was similar between the combination therapy and the naproxen only group: 21% versus 15% (N=109, RR 1.37, 95% CI 0.61 to 3.08).[71] Patients reported feeling tired, dizzy, and stomach irritation.

Pregabalin Plus NSAID

Pregabalin Plus NSAID Versus NSAID Alone

Description of Included Studies

One RCT (N=60) compared the combination of pregabalin and an NSAID versus an NSAID alone for the treatment of acute lumbar radicular pain due to lumbar disc herniation (see Appendix E).[142] Patient mean age was 51 years and 45% were female with moderate leg pain of 2 to 4 weeks duration that interfered with their activities of daily living. Pregnant or breastfeeding women and patients with psychiatric disorders were excluded. The trial did not report race, ethnicity, prior history of LBP, history or treatment of opioid or substance abuse disorder, or medical comorbidities. Treatments included pregabalin 150 mg per plus loxoprofen 180 mg per day (combination therapy) and loxoprofen 180 mg per day alone (monotherapy) taken orally for up to 4 weeks. Patients were instructed not to take gabapentin or sleeping pills. Prior to trial entry, patients had to have taken NSAIDs (loxoprofen 180 mg per day) for at least 2 weeks before the study without sufficient effect.

The trial was conducted in an outpatient clinic in Japan and was considered at moderate risk of bias.[142] Limitations included unclear allocation concealment methods and lack of patient, provider, and assessor blinding (Appendix F). No funding was reported.

Detailed Synthesis

Pain, Global Efficacy, and Sleep Quality

The combination therapy of pregabalin plus loxoprofen was associated with similar improvement in VAS (0-10 scale) leg pain scores compared with loxoprofen alone at 2 weeks (N=56, MD -0.84, 95% CI -2.16 to 0.48) and 4 weeks (N=56, MD of -0.72, 95% CI -1.54 to 0.10).[142] However, there was a moderate improvement in sleep disturbance (on a 0-10 VAS scale) at 2 weeks (N=56, MD -1.64, 95% CI -2.82 to -0.46) and 4 weeks (N=56, MD -1.47, 95% CI -1.89 to -1.05) and Patient Global Impression of Change (1-7 scale) at 4 weeks (N=56, MD -1.0, 95% CI -1.78 to -0.22) with combination therapy versus monotherapy.

Adverse Events

No serious adverse events occurred in either group. Four patients (13%) experienced pregabalin-related adverse events (severe dizziness, drowsiness, severe leg pain) that resulted in withdrawal from the study; no patient in the loxoprofen only group withdrew due to mediation

side-effects. Mild to moderate adverse events (nausea, edema, dizziness, headache) were more frequent with combination pregabalin and loxoprofen versus loxoprofen alone (N=60, 13.3% vs. 3.3%, RR 4.0, 95% CI 0.47 to 33.73) the difference was not statistically significant according to our calculations. The authors state that the incidence of adverse events overall was significantly greater (p=0.026) with combination therapy versus monotherapy but no details or data were provided.[142]

Combination Nonopioid Plus Nonpharmacologic Therapy (Key Questions 5c, 5d, 5e)

Description of Included Studies

Two studies compared a nonopioid pharmacologic therapy combined with a nonpharmacologic intervention versus the same nonopioid pharmacologic therapy alone (2 RCTs)[86,153] or versus the same nonpharmacologic intervention alone (1 RCT)[153] (Appendix E). Both trials had three treatment arms.

One trial (N=90) compared a 20-minute application of either a hot water bottle or ice combined with naproxen 500 mg (n=30 in both groups) versus naproxen 500 mg alone (n=30); all treatments were given twice daily for 1 week.[86] Mean patient age was 34 years and 52% were female with LBP of <4 weeks duration (mean 16 days). Patients with underlying diseases or trauma, those taking analgesics other than naproxen, or undergoing physiotherapy were excluded. No further information about patient demographics, substance abuse, or radiculopathy was provided. The trial was conducted in Iran and received university funding.

One trial (N=69) compared a combination of acupuncture plus diclofenac sodium versus diclofenac sodium alone (n=45) and versus acupuncture alone (n=49).[153] (See Key Question 5a, single nonopioid therapy versus single nonpharmacologic intervention, for results comparing diclofenac vs. acupuncture.) Mean patient age was 36 years and 42% were female with LBP with or without radiculopathy of <2 weeks duration (mean 3 days). Patients with underlying diseases or trauma or recent medical history of NSAIDs or anticoagulants or contraindications to treatment were excluded. No further information about patient demographics or substance abuse was provided. Acupuncture was given once a day (needles were retained for 30 minutes and manipulated every 10 minutes) and diclofenac sodium 50 mg was given orally twice a day for 5 days. The trial was conducted in China and the funding source was not reported.

Both trials were assessed as high risk of bias due to unclear randomization and allocation concealment methods and lack of blinding (Appendix F).

NSAID Plus Heat or Cold Versus NSAID Alone

Both combination therapies were associated with greater improvement in McGill Pain Questionnaire (MPQ) pain scores compared with naproxen alone at all timepoints except for cold therapy at 3 days (Table 47).[86] Patients who received the combination treatment with heat therapy showed greater improvement than those who received cold therapy. However, the scoring method for the MPQ and its scale were unclear so the clinical significance (i.e., effect sizes) of these differences could not be assessed. Other outcomes and adverse events were not reported.

Table 47. Pain outcomes for combination heat or cold therapy plus an NSAID versus NSAID alone

Time	MPQ Overall Pain Scores (scale unclear) Mean Difference (95% CI)	
-	Combination Heat Therapy plus Naproxen vs. Naproxen (n=29)	Combination Cold Therapy plus Naproxen vs. Naproxen (n=29)
Day 3	-2.58 (-4.03 to -1.13)	-0.59 (-1.89 to 0.71)
Day 8	-4.00 (-5.28 to -2.27)	-2.62 (-3.89 to -1.35)
Day 15	-4.83 (-5.59 to -4.07)	-3.39 (-4.48 to -2.30)

CI = confidence interval; MPQ = McGill Pain Questionnaire; NSAID = nonsteroid anti-inflammatory drug.

NSAID Plus Acupuncture Versus NSAID alone

The combination of acupuncture plus diclofenac sodium was associated with a moderate improvement in pain on the NRS (0-10 scale, N=45, MD -1.28, 95% CI -2.09 to -0.47) and a small improvement in function on the RDQ (0-24 scale, N=45, MD -1.81, 95% CI -3.48 to -0.14) compared with diclofenac sodium alone at 5 days in one trial.[153] Other outcomes and adverse events were not reported.

NSAID Plus Acupuncture Versus Acupuncture Alone

One trial compared an NSAID combined with a *nonpharmacologic* intervention (acupuncture) versus acupuncture alone (Appendix E).[153]

The combination of acupuncture plus diclofenac sodium was associated with small improvement in pain on the NRS (0-10 scale, N=49, MD -0.91, 95% CI -1.58 to -0.24) and a moderate improvement in function on the RDQ (0-24 scale, N=49, MD -2.01, 95% CI -3.45 to -0.59) at 5 days.[153] Other outcomes and adverse events were not reported.

Key Question 6. Noninvasive, Nonpharmacologic Therapy

Exercise

We employed the same general classification of exercise that was used in our prior review of noninvasive, nonpharmacologic pain as follows: Muscle performance, neuromuscular re-education, mobility/flexibility, cardiovascular/aerobic and combined exercise. (Details in Appendix I)

Exercise Versus Placebo/Sham

Description of Included Studies

One moderate-quality RCT (N=318) of mobility and flexibility exercises versus a placebo/attention control intervention[122] and a subsequent analysis that focused participants who were employed at the time of trial enrollment (N=196) was identified (Appendix E).[121] Most participants were 25 to 44 years of age (65%), male (56%) and had a history of prior LBP episodes (73%). Pain duration at baseline was 1-7 days in 66%, 8- 14 days in 26% and 15-21 days in 7% of patients and mean pain was 3.6 (VAS 1-10). The trial excluded patients with radiculopathy, psychiatric comorbidities, or history of substance use disorder. All patients were provided with acetaminophen for the first month. Exercise was performed for twice weekly in 20 minute sessions for 5 weeks. Adherence was reported as 82% at 2 months but 54% at 12 months.

The placebo/attention control consisted of low, almost zero dose (0.1 watt/cm^2), intermittent ultrasound without heat effect provided twice weekly for 5 weeks.[122] Methodologic limitations included unclear methods for randomization and concealment of treatments and inability to blind patients and providers (Appendix F).

Detailed Synthesis

Function, Pain, , Emotional and Sleep Problems, and Adverse Events

Improvement in pain and function from baseline was similar between those receiving exercise for mobility and flexibility and those receiving the placebo at all time frames in one RCT.[122]

Based on the Nottingham Health Profile (NHP) pain dimension (0-10), changes in pain from baseline were similar between those receiving exercise and those receiving the placebo at 4 weeks (-1.9 for both groups), 8 to 12 weeks (-2.4 vs. -2.2) and at 16 to 52 weeks (-2. 6 vs. -2.7). Similarly, there were no differences between groups on the NHP mobility problems (0-100 scale) at 4 weeks (-9 vs. -10), 8 to 12 weeks (-12 vs. -13) or 16 to 52 weeks (-14 vs. -15). (Appendix G.) Based on the NHP, the proportion of patients in each group who reported a reduction in emotional problems was similar at 4 weeks (38% vs. 32%), 8 to 12 weeks (33% vs. 31%) and at 16 to 52 weeks (37% vs. 36%) as was the proportion who reported a decrease in sleep problems at the respective times: 4 weeks and 8 to 12 weeks (23% vs. 28% for both times points), and at 16 to 52 weeks (28% vs. 33%). There were no differences at any time in the mean number of daily life influenced by LBP.[122] In analyses focused on participants who were employed at the time of trial enrollment, there were no differences between those receiving exercise and those who had placebo ultrasound in the proportions that reported at least one sick day at any of the follow-up times.[121]

Authors did not report on adverse events. Authors indicated that there was no effect modification for pain or function outcomes by various factors including age, sex, back pain history, prior treatment, mobility problems or radiation of pain at baseline, however, they do not provide data or results from tests of interaction to support this.

Exercise Versus Usual Care

Description of Included Studies

Ten RCTs (12 publications) (N=923) compared various forms of exercise with usual care (Appendix E).[59,88,94,96,107,111,119,121,122,127,131,149] Patient mean age ranged from 36 to 43 and 19% to 76% were female. Patients with radiculopathy or suspected neurologic symptoms such as leg pain were excluded from five RCTs,[88,96,111,122,127] and one RCT focused solely on those with radiculopathy.[149] Patients with or without radiculopathy were included in the remaining trials.[59,94,107,119] Mean pain duration ranged from 1 to 27 days and mean pain intensity ranged from 3.7 to 7.4 (VAS 0-10 scale) at baseline. The majority of patients in three RCTs had prior LBP episodes (66% to 73%).[59,107,122] One RCT in patients with radiculopathy reported that 24% of them had prior radicular symptoms.[149] Substance abuse disorder and psychiatric comorbidities were not reported in trials with the exception of one that excluded patients psychiatric comorbidities.[111] One trial (N=311) of mobility and flexibility exercises versus usual care[122] published a subsequent analysis focused on participants who were employed at the time of trial enrollment (N=244).[121]

Exercise focus was on mobility and flexibility in six trials[59,88,107,111,122,127] and neuromuscular re-education one trial.[96] Two trials employed a combination of exercises types,[94,119] and exercise type was not described in one RCT.[149] Exercise categorization is detailed in Appendix E. Frequency and duration of exercise were poorly reported and varied across trials. Treatment duration also varied across trials with most ranging from 2 to 6 weeks and frequency of exercise ranged from 2 or 3 times weekly to daily. One trial reported a single session with an option to return for additional session if pain persisted[88] and duration was unclear in another trial with patients instructed to perform the exercises every other hour during the day until pain subsided (Table 48 and Table 49).[127] Adherence to exercise was reported in three RCTs and ranged from 50% to 84%.[94,107,122]

Two trials indicated that no treatment was provided in the control groups, however their descriptions suggest that advice and use of medications as needed were provided, similar to usual care described in other trials.[94,127] Usual care and no treatment in trials generally consisted of advice to stay active and use of analgesics, NSAIDS and/or muscle relaxants. Medications at baseline were variably reported; specifics related to dose, frequency and duration were not provided for baseline or concurrent mediations in most trials. In one trial, all patients received oral acetaminophen for the first month[122] and another did describe suggested frequency or dosing for as needed analgesics and NSAIDS.[149]

All trials but one[119] (high) were at moderate risk of bias. Study limitations included unclear methods for randomization and/or intervention concealment and inability to blind participants to treatment. In addition, in the RCT rated at high risk of bias,[119] comparability between treatment groups at baseline was not reported and attrition was high (Appendix F).

Table 48. Patient and study characteristics: exercise versus usual care or no treatment

Study, Year	Luijsterburg, 2008	Faas, 1993 Faas, 1995*	Seferlis, 1998	Evans, 1987	Underwood, 1998
N Randomized	135†	Faas 1993: 311 Faas 1995: 244*	120	127‡	75
Mean Age (Years)	43	36	39§	41	41
% Female	48%**	44%	47%§	48%	40%
% Radiculopathy	100%	0%†† (Excluded)	Mixed (21% with positive Lasegue sign)	Mixed % NR	Unclear‡‡
% Previous Back Pain	24%§§	77%	NR***	0: 31% 1-4: 40% ≥5: 26%	NR
Pain Duration Inclusion Criteria	<6 weeks	≤3 weeks	NR	<30 days	<28 days
Mean Pain Duration (days)	13	1-7 days: 67% 8-14 days: 26% 15-21 days 8%	NR	0-5 days: 51% ≥6 days: 49%	7
Mean/Median Baseline Pain Severity (0-10)	5.7	3.6	5.1§†††	McGill pain score: 23.9	4.8
History substance abuse disorder	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	NR	Excluded	NR	NR
Exercise type	Unclear‡‡‡	Mobility, flexibility	Combo§§§	Mobility, flexibility	Mobility, flexibility
Comparator	Usual Care	Usual Care	Usual Care	Usual Care****	Usual Care

Study, Year	Luijsterburg, 2008	Faas, 1993 Faas, 1995*	Seferlis, 1998	Evans, 1987	Underwood, 1998
Dosage Frequency	NR	Exercise: 20 minutes twice weekly UC: NR	Exercise: 3 times per week UC: NR	Exercise: 20 minutes daily UC: NR****	One or two 1-hour sessions
Treatment duration	6 weeks	5 weeks	8 weeks	≥2 months	Exercise: Once or twice UC: NR
Duration follow-up	1 year	1 year	1 year	1 year	1 year
Country	Netherlands	Netherlands	Sweden	Canada	UK
Risk of bias	Moderate	Moderate	High	Moderate	Moderate

NR = not reported; UC = usual care.
* Faas 1995 is a subgroup analysis of Faas 1993, and included only patients with paid work.
† Four patients dropped out immediately after randomization because they no longer wished to participate.
‡ 14% of patients were workers' compensation cases.
§ Whole sample. Includes more than two interventions, but data not reported by arm.
** 57% of exercise patients were female compared to 40% of usual care patients.
†† 24% of patients had radiation into the upper leg. Patients with nerve root compression and neurological deficits were otherwise excluded.
‡‡ Excluded patients with neurological deficit due to nerve root issue not due to previous episode or another pathology, otherwise not specified.
§§ Proportion of patients that never have experienced lumbosacral radicular syndrome.
*** Patients were excluded if they had records of sick leave for previous low back pain.
††† Pain questionnaire was developed by Carlsson (0-10) as a modified Borg 10-point numerical scale (higher = greater pain). Mean represents all three groups, as pain scores were not reported by group.
‡‡‡ Type/content of the exercises were at the physical therapists' discretion but not described; Goal was gradual increase the activities to normal level in 6 weeks.
§§§ Combination of 2 or more exercise types - see appendix.

****** Patients in the usual care group received the same analgesic medications as all the other arms, and otherwise received no further instructions or care.**

Table 49. Patient and study characteristics: exercise versus usual care or no treatment – continued

Study, Year	Machado, 2010 Sheets, 2012	Ali, 2002	Buyuksireci, 2022	Malmivaara, 1995	Chok, 1999
N Randomized	146*	43	54	119	66
Mean Age (Years)	47	35	44	40	36
% Female	50%	33%	81%	71%	24%
% Radiculopathy	47%†	0% (Excluded)	0% (Excluded)	Unclear‡	Unclear§
% Previous Back Pain	71%	0%**	NR	Back pain >30 days in previous 12 months: 21%	57%
Pain Duration Inclusion Criteria	<6 weeks	<2 weeks	<3 months	≤3 weeks	7 days to 7 weeks
Mean Pain Duration (days)	<2 weeks: 67% 2-6 weeks: 34%	NR	Median 25	4.8	25
Mean/Median Baseline Pain Severity (0-10)	6.4	7.3	Median 7	5.9	Median 1.8††
History substance abuse disorder	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	NR	0% (exclusion)	NR	NR
Exercise type	Mobility, flexibility	Neuromuscular Re-Education	Mobility, flexibility	Mobility, flexibility	Combo‡‡
Comparator	Usual Care	Usual Care	Usual Care	No Treatment	No Treatment
Dosage Frequency	Exercise: Max 6 sessions	Exercise: 2-3x per weeks	Exercise: 3 days a week	NR	Exercise: 10 repetitions

Study, Year	Machado, 2010 Sheets, 2012	Ali, 2002	Buyuksireci, 2022	Malmivaara, 1995	Chok, 1999
	Frequency variable UC: NR	UC: NR	UC: NR		(exercise held for 10-20 secs) for 6 cycles, 30-45 minutes 3x/week
Treatment duration	3 weeks	4 weeks	2 weeks	NR	6 weeks
Duration follow-up	3 months	4 weeks	6 weeks	12 weeks	6 weeks
Country	Australia	U.S.	Turkey	Finland	Singapore
Risk of Bias	Moderate	Moderate	Moderate	Moderate	Moderate

NR = not reported; UC = usual care.
* 148 originally randomized. One patient in each group were misdiagnosed and excluded. ~62% already participating in moderate exercise.
† Pain radiating to the leg. Patients were excluded if they were diagnosed with a nerve root compromise.
‡ Patients were excluded if they had sciatic syndrome, but 10% of exercise and 9% of usual care patients had pain that radiated below the knee.
§ Patients could be included with or without radiating leg pain, but were excluded if they had signs of nerve root compromise.
** Inclusion criteria required patients to have been experiencing their first episode of acute nonspecific back injury.
†† Estimated from graphs. Authors report median pain as 1.60 and 1.95 for the exercise and usual care groups respectively, with minimum and maximum ranging from 0 to 8.30. Authors also report median pain in preceding 24 hours as 3.84 and 3.68 respectively, with minimum 0 and maximum 9.65.
‡‡ Combination of 2 or more exercise types - see Appendix E.

Detailed Synthesis

Function

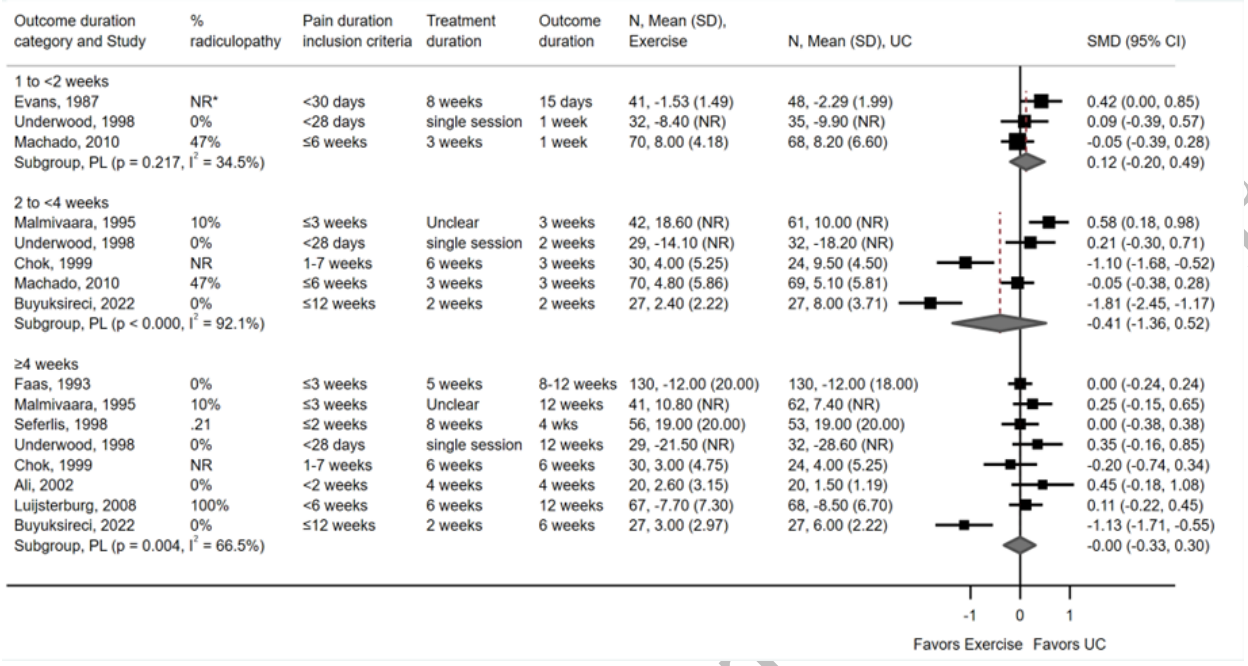
All 10 RCTs reported measures of function and contributed to pooled analyses at one or more times.[59,88,94,96,107,111,119,122,127,149] RDQ (0-24 scale) was reported in five trials[94,96,107,111,149] and ODI (0-100 scale) was reported in three trials[88,119,127] Other measures reported in single trials were the NHP mobility score (0-100 scale)[122] and Activities Discomfort Scale (0-78).[59]

Functional improvement was similar for exercise and usual care in pooled analyses at 1 to <2 weeks (3 RCTs, N=294)[59,88,107] but some heterogeneity is noted with one trial[59] suggesting a small functional improvement with usual care, however the effect estimate is imprecise (Figure 27). The reasons for this heterogeneity are unclear.

Functional improvement was also similar for exercise and usual care from 2 to <4 weeks, (5 RCTS, N=411)[88,94,111,127] with substantial heterogeneity noted with several outlier trials.[94,107,127] Exclusion of two trials, one which provided a single session[88] and which did not specify a number or duration sessions[127] had no effect on the heterogeneity but resulted in a larger, less precise effect size that was not statistically significant across the remaining trials (3 RCTs N=241, pooled MD -0.94, 95% CI -2.15 to 0.20, $I^2 = 93\%$).[94,107,111] (Figure 28; Appendix E). Two of the three trials[94,111] found an association between exercise and improved function but the third did not.[107] Mean difference in RDQ scores (0-24 scale), for these two trials were 5.50 (N=54, 95% CI -8.10 to -2.90)[94] and 5.60 (N=54, 95% CI -7.81 to -3.97)[111] consistent with large functional improvement. Reasons for the heterogeneity across these studies are unclear. Similar functional improvement with exercise and usual care was also seen at ≥4 weeks up to 12 weeks (8 RCTS, N=816, pooled MD 0.00, 95% CI -0.33 to 0.30, I^2=67%) (Figure 28). Exclusion of one trial for which the treatment duration was unclear[127] and another that required only a single session,[88] did not change the conclusion of no difference between exercise and usual care, and increased the heterogeneity of pooled estimates (Figure 28, Appendix E).

Two RCTs also found no differences between groups up to 52 weeks after intervention (Appendix E).[88,149]

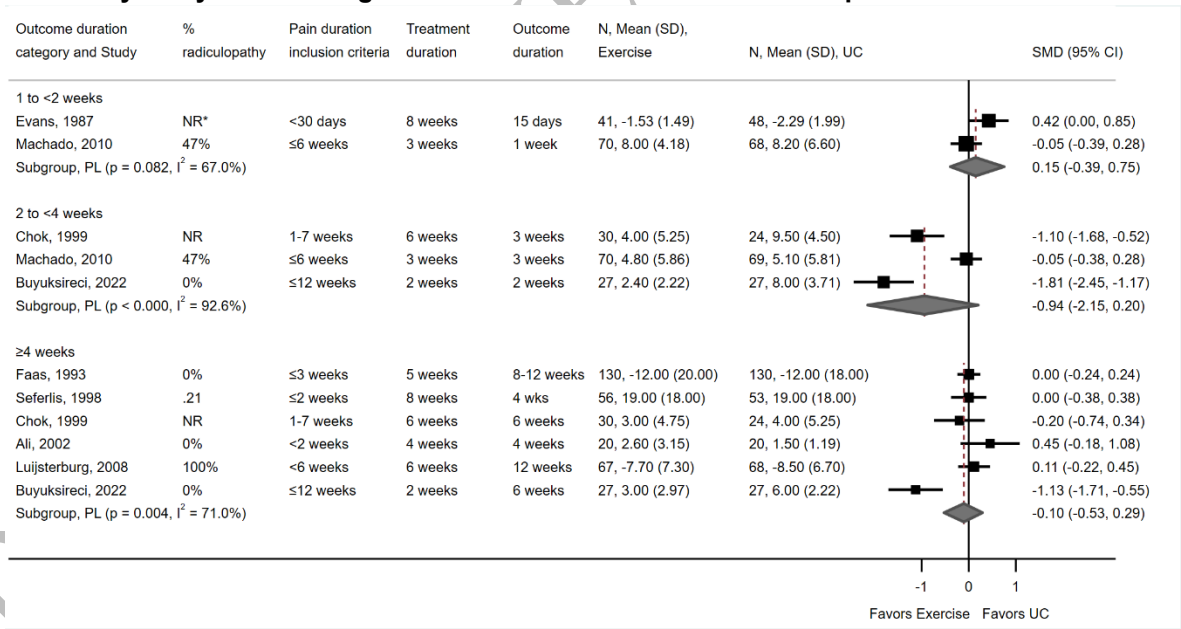
Figure 27. Exercise versus usual care: across function scores (RDQ/ODI/NHP Mobility/ADS)



ADS = Activities Discomfort Scale; CI = confidence interval; NHP = Nottingham Health Profile Mobility; NR = not reported; ODI = Oswestry Disability Index; PL = profile likelihood; RR = risk ratio; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation SMD = standardized mean difference.

* Population with and without radiculopathy but % with radiculopathy not reported

Figure 28. Exercise versus usual care: across function scores (RDQ/ODI/NHP Mobility/ADS) sensitivity analyses focusing on studies with defined intervention period of >1 session



ADS = Activities Discomfort Scale; CI = confidence interval; NHP = Nottingham Health Profile Mobility; NR= not reported; ODI = Oswestry Disability Index; PL = profile likelihood; RR = risk ratio; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation SMD = standardized mean difference.

* Population with and without radiculopathy but % with radiculopathy not reported

One RCT (N=68) reported on the likelihood of patients achieving a successful functional improvement (ODI <2.0 on 0-100 scale).[88] Within the first week, exercise was associated with a moderately higher likelihood of improved function versus usual care (71% vs. 47%) and a small likelihood of improvement at 2 weeks (88% vs. 70%), however likelihoods were similar at later times. Another RCT (n=113) reported no difference in the proportion of patients reporting no activity restriction due to back pain or restriction once or twice by 52 weeks[59] (Table 50).

Table 50. Exercise versus usual care: summary of other functional measures not in meta-analysis

Author, Year LPB Type Exercise Type	Outcome	Time	Exercise vs. UC or No Treatment % (n/N) and RR (95% CI) or, MD (95% CI)
Underwood, 1998 Nonradicular LBP Mobility, flexibility	ODI <20 (0-100 scale)	1 week	71% (23/32) vs. 47% (17/36) RR 1.52 (95% CI 1.01 to 2.29)
		2 weeks	88% (25/29) vs. 70% (23/33) RR 1.24 (95% CI 0.95 to 1.62)
		4 weeks	84% (27/32) vs. 78% (29/37) RR 1.08 (95% CI 0.86 to 1.35
		8 weeks	93% (27/29) vs. 93% (30/32) RR 0.99 (95% CI 0.87 to 1.13
		12 weeks	89% (26/29) vs. 91% (30/33) RR 0.99 (95% CI 0.84 to 1.16
		52 weeks	96% (22/23) vs. 86% (24/28) RR 1.12 (95% CI 0.94 to 1.33)*
Evans, 1987 Mixed (radicular and nonradicular) Mobility, flexibility	Daily activity restriction (none, once or twice vs. more often)	52 weeks	75% (44/59) vs. 78% (42/54) RR 0.96 (0.78 to 1.17)

CI = confidence interval; LBP = low back pain; MD = mean difference; NS = not statistically significant; ODI = Oswestry Disability Index; RR = risk ratio; SD = standard deviation; UC = usual care.
* Below the threshold for a small effect.

Pain

All 10 RCTs reported measures of pain and contributed to pooled analyses.[59,88,94,96,107,111,119,122,127,149] VAS or NRS (0-10) were reported in eight trials; the NHP Pain Dimension (0-10),[122] and the Borg numerical 1-11 scale, (rescaled to 0-10 for pooling)[119] were reported in the remaining RCTs.

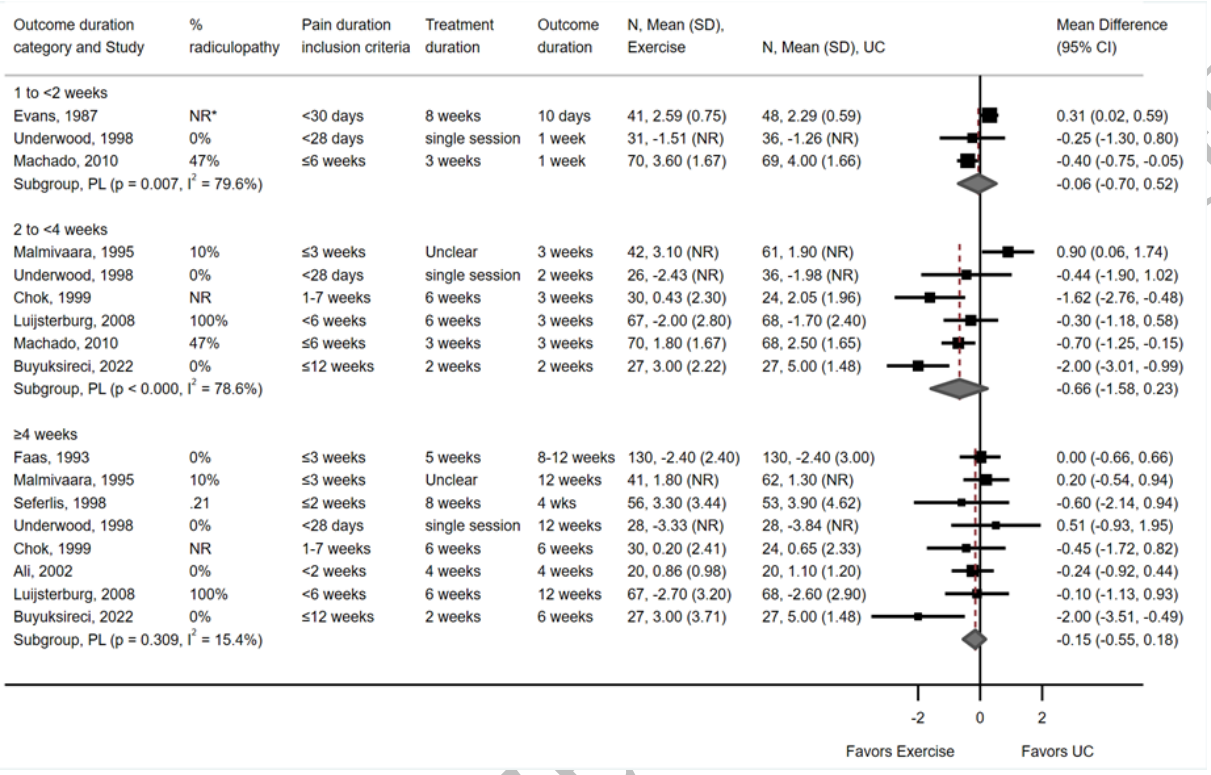
Three RCTS found similar pain for exercise and usual at 1 to <2 weeks (N=295, pooled MD –0.06, 95% CI –0.75 to 0.52, I^2=79.6), but some heterogeneity is noted with one trial favoring usual care[59] and another favoring exercise,[107] however effect estimates for all studies were below the threshold for a small effect and estimate are imprecise. The reasons for heterogeneity are unclear. Exclusion of one trial for which the treatment duration was unclear[127] did not change the conclusion of no difference between groups (Figure 29; Appendix E).

Across six trials, pain improvement was similar for exercise and usual care from 2 to 4 weeks (6 RCTs, N=546, pooled MD –0.66, 95% CI –1.58 to 0.23, I^2=78.6%) (Figure 29).[88,94,107,111,127,149] Substantial heterogeneity noted (I^2=79%) due to several outlier trials, one of which favored usual care although the effect was below the threshold for a small effect.[59] At 2 to <4 weeks, across the four trials that had defined intervention duration of >1 session,[94,107,111,149] exercise was associated with small pain improvement versus usual care and heterogeneity was reduced, however, estimates were imprecise. (4 RCTs, N=381, pooled MD -1.05, 95% CI -1.96 to -0.27, I^2= 64%, VAS scale 0-10)[94,107,111,149] (Figure 30, Appendix E).

Similar pain improvement was seen with exercise and usual care at ≥4 weeks up to 12 weeks (8 RCTs, N=811, pooled MD –0.29, 95% CI –0.84 to 0.09, I^2=17.5%) (Figure

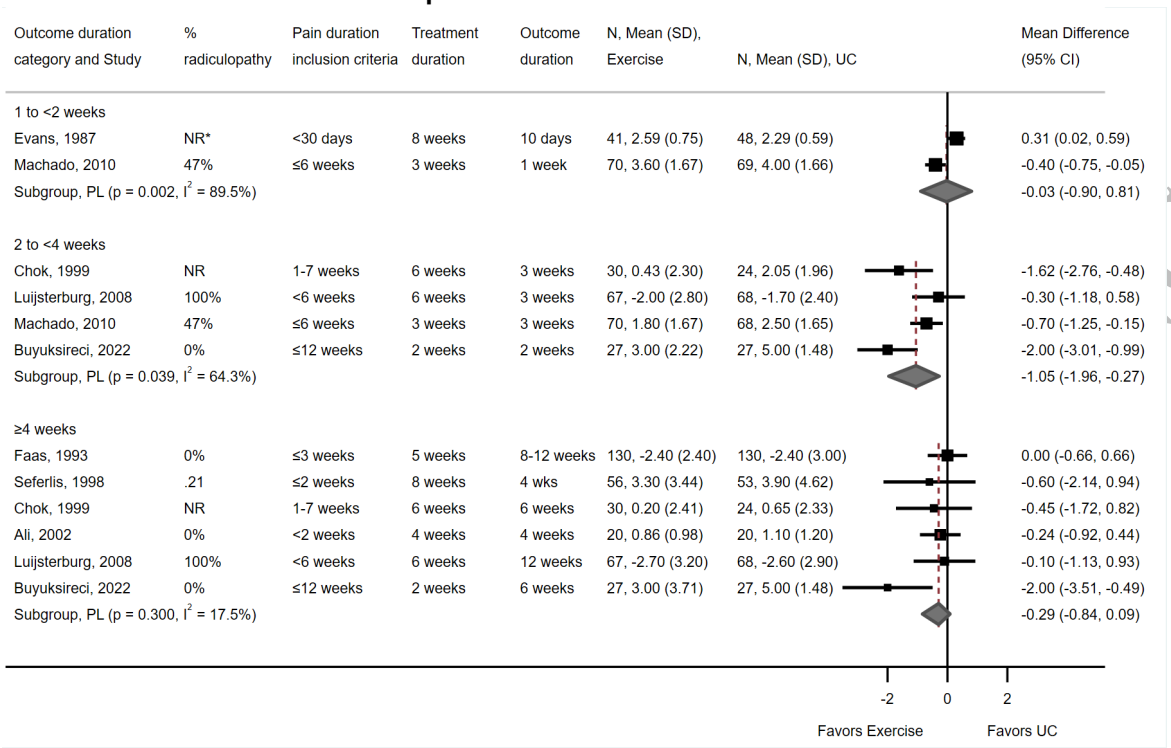
30).[88,94,96,111,119,122,127,149] Two RCTs also found no differences between groups 52 weeks after intervention[88,149] (Appendix E).

Figure 29. Exercise versus usual care: pain scores (0-10 scale)



CI = confidence interval; NR= not reported; NRS = Numeric Rating Scale; PL = profile likelihood; SD = standard deviation; MD = mean difference; VAS = visual analogue scale.
* Population with and without radiculopathy but % with radiculopathy not reported.

Figure 30. Exercise versus usual care: pain scores (0-10 scale) sensitivity analyses focusing on studies with defined intervention period of >1 session



CI = confidence interval; NR= not reported; NRS= Numeric Rating Scale; PL = profile likelihood; SD = standard deviation; MD = mean difference; VAS = visual analogue scale.
* Population with and without radiculopathy but % with radiculopathy not reported.

One RCT (N=67) reported on the likelihood of patients achieving a successful improvement in paint (VAS <2.0 on a 0-10 scale).[88] The likelihood of pain improvement was similar between exercise and usual care at all times (Table 50).

Other Pain Outcomes

One RCT (n=113) found no differences between the exercise and usual care groups at 52 weeks in the proportions of patients who had pain at that time, in the prior 5 months or int the proportion of patients reporting no or mild pain.[59] The same trial reported no difference between groups with regard the man number of daily life areas influenced by pain at any time. Another trial in patients with radiculopathy found no differences in VAS leg pain (0-10 scale) at any time from 3 to 52 weeks, however estimates are imprecise (Table 51).[149]

Table 51. Exercise versus usual care: summary of other reported pain measures

Author, Year LBP Type Exercise Type	Outcome	Time	Exercise vs. Usual Care or No Treatment % (n/N) and RR (95% CI) or, MD (95% CI)
Underwood, 1998 Nonradicular LBP Mobility, flexibility	VAS <2.0 (0-10 scale)	1 week	35% (11/31) vs. 28% (10/36) RR 1.28 (95% CI 0.63 to 2.60)
		2 weeks	50% (13/26) vs. 30% (10/33) RR 1.65 (95% CI 0.87 to 3.14)
		4 weeks	65% (20/31) vs. 48% (16/33) RR 1.33 (95% CI 0.86 to 2.06)
		8 weeks	76% (22/29) vs. 64% (18/28) RR 1.18 (95% CI 0.84 to 1.66)

Author, Year LBP Type Exercise Type	Outcome	Time	Exercise vs. Usual Care or No Treatment % (n/N) and RR (95% CI) or, MD (95% CI)
		12 weeks	68% (19/28) vs. 79% (22/28) RR 0.86 (95% CI 0.63 to 1.19)
		52 weeks	76% (16/21) vs. 69% (18/26) RR 1.10 (95% CI 0.78 to 1.56)
		>52 weeks	57% (12/21) vs. 44% (11/25) RR 1.30 (95% CI 0.73 to 2.31)
Luijsterburg, 2008 LPB with radiculopathy Exercise type unclear	Leg pain on NRS (0 to 10 scale (changes from baseline)	3 weeks	2.3 (2.4) (n=67) vs. -1.9 (2.4) (n=68), MD -0.4 (95% CI -1.2 to 0.4)
		6 weeks	-3.0 (2.7) (n=67) vs. -3.3 (2.8) (n=68), MD 0.3 (95% CI -0.6 to 1.2)
		12 weeks	-3.9 (2.8) (n=67) vs. -3.7 (3.1) (n=68), MD -0.2 (95% CI -1.2 to 10.8)
		52 weeks	-4.4 (2.7) (n=67) vs. -3.7 (2.7) (n=68), MD -0.7 (95% CI -1.7 to 0.2)
Evans, 1987 Mixed population Mobility, flexibility Exercise	Present pain (none or mild)	52 weeks	78% (46/59) vs. 85 (46/54) RR 0.91 (0.77 to 1.09)
	Pain in past 5 months (none or once/month)	52 weeks	61% (36/59) vs. 63% (34/54) RR 0.97 (0.73 to 1.29
	Pain intensity (none or mild)	52 weeks	0.47 (28/59) vs. 59% (32/54) 0.90 (0.57 to 1.13)

CI = confidence interval; LBP = low back pain; MD = mean difference; NRS = Numeric Rating Scale; NS = not statistically significant; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

Recurrence

Recurrence of LBP was common (29% to 80%) across three RCTs (N=444) at 52 months, however there were no differences between exercise and usual care and estimates are imprecise[88,119,122] (Table 52). In the largest trial (N=309), about one half of participants in each group reported no recurrence or only one recurrent episode (53% vs. 55%).[122]

Table 52. Exercise versus usual care: LBP recurrence

Author, Year LPB Type Exercise Type	Outcome	Time	Exercise vs. Usual Care or No Treatment % (n/N) and RR (95% CI) or, MD (95% CI)
Faas 1993 Nonradicular Mobility, flexibility	1 or more recurrence	52 weeks	69% (107/154) vs. 70% (108/155) RR 1.00 (0.86 to 1.16)
Seferlis, 1998 Mixed Combination exercise	1 or more recurrences during the study year	52 weeks	29% (12/41) vs. 40% (16/40) RR 0.73 (0.40 to 1.34)
Underwood, 1998 Nonradicular LBP Mobility, flexibility	1 or more days with LBP in prior 6 months	52 weeks	58% (14/24) vs. 80% (24/30) RR 0.73 (0.50 to 1.07)

CI = confidence interval; LBP = low back pain; MD = mean difference; NS = not statistically significant; RR = risk ratio; SD = standard deviation.

Quality of Life, Sleep Quality, Perception of Improvement

Two RCTs reported similar improvements on quality-of-life measures for exercise and usual care at all times reported (Table 53).[127,149] There was no difference between exercise and usual care in the proportion of patients reporting a decrease in *sleep problems* based on the Nottingham Health Profile at 4 weeks (N=291, 23% vs. 28%), 8 to 12 weeks (N=260, 23% vs. 23%) or 16 to 52 weeks (N=271, 28% vs. 33%) in one RCT[122] (Table 49).

One RCT (N=135) in patients with radiculopathy found similar likelihood of perceived improvement for exercise and usual care at all times across patients, however in analyses in patients with greater disability (RDQ ≥17) exercise was associated with a moderately greater likelihood of improvement or complete recovery versus usual care at 12 and 52 weeks but estimates are imprecise.[149] In contrast, a trial (n=138) in a mixed population of patients (i.e., patients who did and did not have radiculopathy) reported no difference on perceived global effect by group and 1 and 3 weeks[107] (Table 53).

At 52 weeks, one RCT (n=49) found substantially more patients in the exercise group reported that back pain had not been a problem in prior 6 months versus those who received usual care (50% vs. 14%, RR 3.25 95% CI 1.17 to 8.35), however estimates are imprecise[88] (Table 53).

Table 53. Exercise versus usual care: quality of life, sleep quality, perceived improvement

Author, year LBP Type Exercise Type	Outcome	Time	Exercise vs. Usual Care or No Treatment % (n/N) and RR (95% CI) or, MD (95% CI)
Quality of Life			
Luijsterburg, 2008 LPB with radiculopathy Exercise type unclear	SF-36 Total (0-100); general health (improvement from baseline)	6 weeks	2.2 (16.4) vs. -2.8 (13.9), MD 5.0 (95% CI -0.2 to 10.1)
		12weeks	-1.2 (18.4) vs. -4.7 (16.4) MD 3.5 (95% CI -2.4 to 9.5)
		52 weeks	3.1 (15.7) (n=67) vs. -4.1 (16.7) (n=68), MD 1.0 (95% CI -4.5 to 6.5)
Malmivaara, 1995, Mixed (patients with and without pain below knee) Mobility, flexibility	Health-related QoL index (measure NR; scale 0-1; higher = better	3 weeks	Mean (SD NR) 0.91 (n=42) vs. 0.94 (n=61) adj MD -0.02 (95% CI -0.04 to 0.001)
		12 weeks	0.95 (n=41) vs. 0.95 (n=62) at 12 adj MD -0.005 (95% CI -0.03 to 0.02)
Sleep Quality			
Faas 1993 Nonradicular Mobility, flexibility	Nottingham Health Profile: % of patients, decrease in sleeping problems	4 weeks	23% (33/145) vs. 28% (41/146) RR 0.81 (0.55 to 1.20)
		8 to 12 weeks	23% (30/130) vs. 23% (30/130)
		16 to 52 weeks	28% (38/137) vs. 33% (44/134) RR 0.85 (95% CI 0.59 to 1.21)
Perceived Improvement			
Luijsterburg, 2008 LPB with radiculopathy Exercise type unclear	Improved on global perceived effect (completely recovered or improved)	3 weeks;	45% (30/67) vs. 32% (22/68), RR 1.4 (95% CI 0.9 to 2.1)
		6 weeks	60% (38/67) vs. 44% (30/68), RR 1.3 (95% CI 0.9 to 1.8)
		12 weeks	70% (47/67) vs. 62% (42/68), RR 1.1 (95% CI 0.9 to 1.5)
		52 weeks	79% (53/67) vs. 56% (38/68), RR 1.4 (95% CI 1.1 to 1.8)
	Subgroup with severe disability (RDQ ≥17) at baseline: Improved on global perceived effect	12 weeks	78% (29/37) vs. 50% (15/30), RR 1.57 (95% CI 1.06 to 2.33)
		52 weeks	84% (31/37) vs. 53% (16/30) RR 1.57 (95% CI 1.09 to 2.26)
Machado, 2010 Mixed population (patients with and without radiculopathy) Mobility, flexibility	Global perceived effect (-5 to 5, <0 = worsening, >0 = improvement	1 week	Mean (SE) 2.6 (0.2) (n=70) vs. 2.1 (0.2) (n=68), adj MD 0.5 (95% CI -0.0 to 1.1)
		3 weeks	Mean (SE) 3.6 (0.1) (n=70) vs. 3.3 (0.2) (n=69), adj MD 0.3 (95 % CI -0.3 to 0.8)
Underwood, 1998 Nonradicular LBP Mobility, flexibility	Back pain not a problem in prior 6 months	52 weeks	50% (12/24) vs. 14% (4/25) RR 3.25 (1.17 to 8.35)

Adj = adjusted; CI = confidence interval; LBP = low back pain; MD = mean difference; NS = not statistically significant; RR = risk ratio; SD = standard deviation; SE = standard error; SF-36 = 36-item Short Form Questionnaire.

Psychological Distress

One RCT no difference between exercise and usual care in the proportion of patients reporting a decrease in emotional problems based on the Nottingham Health Profile at 4 weeks (38% vs. 32%), 8 to 12 weeks (33% vs. 34%) or 16 to 52 weeks (37% vs. 39%) (Appendix E).[122]

Return to Work and Work Absence

One RCT (N=103) found that return to work was similar in the exercise and usual care groups up to 12 weeks.[127] Analysis from another RCT which focused on participants who were employed at the time of trial enrollment (N=244) suggests that patients in the exercise group may have a slightly higher likelihood of reporting ≥1 sickness day versus those receiving usual care at 4 weeks (71.4% vs. 59.5% RR 1.19, 95% CI 0.99 to 1.44), but this did not persist to later times.[121] Similarly, another trial (N=135) in patients with radiculopathy found that at 12 weeks, slightly more patients in the exercise group reported absence from work (54% vs. 37% , RR 1.46, 95% CI 1.0 to 2.14) but there was no difference at 52 weeks; however mean days off work did not differ between groups at either time.[149] Similarly, another RCT (N=83) found no difference between exercise and usual care for days of due to LBP at 52 weeks.[119] All estimates were imprecise (Table 54).

Table 54. Exercise versus usual care: return to work and work absence

Author, Year LBP Type Exercise Type	Outcome	Time	Exercise vs. Usual Care or No Treatment % (n/N) and RR (95% CI) or, Mean (SD) and MD (95% CI)
Malmivaara, 1995 Mixed (patients with and without pain below knee) Mobility, flexibility	Return to work	3 weeks	94% (39/42) vs. 98% (60/61) RR 0.94 (95% CI 0.86 to 1.03);
		12 weeks	100% (41/41) vs. 100% (62/62)
Faas 1995 (Working participants) Nonradicular Mobility, flexibility	Sickness absence: % with ≥1 day:	4 weeks	71.4% (81/114) vs. 59.5% (69/116) RR 1.19 (0.99 to 1.44)
		8 to 12 weeks	29.8% (31/104) vs. 22.2% (22/97) RR 1.31 (95% CI 0.82 to 2.11
		16 to 52 weeks	33.3% (33/100) vs. 36% (35/97) RR 0.91 (95% CI 0.62 to 1.34)
Luijsterburg, 2008 LPB with radiculopathy Exercise type unclear	% Absence from work	12 weeks	54% (36/67) vs. 37% (25/68) RR 1.46 (95% CI 1.00 to 2.14)
		52 weeks	45% (30/67) vs. 37% (25/68) RR 1.22 (95% CI 0.81 to 1.83)
	Mean days off work	12 weeks	Mean 16.2 (21.3) vs. 13.1(24.1); NS
		52 weeks	Mean 29.2 (48.4) vs. 28.9 (72.3); NS
Seferlis, 1998 Mixed (radicular and nonradicular) Combination exercise	Days off work for low-back pain	52 weeks	49 (76) (n=42) vs. 52 (63) (n=41), MD 3 (95% CI -27.53 to 33.53)

CI = confidence interval; LBP = low back pain; MD = mean difference; NS = not significant; RR = risk ratio; SD = standard deviation.

Medication Use

Use of rescue medication was not described in studies of exercise versus usual care. Medication use was generally not reported across trials. One RCT reported similar analgesic use between exercise and usual care groups (N=311, 8.5% vs. 11.0%) at 52 weeks.[122]

Adverse Events and Withdrawal Due To Adverse Events

Adverse events were not reported in any trial. One RCT reported similar withdrawal due to increased pain for the exercise and usual care groups (N=54, 7.4% vs. 11.1%, RR 0.67, 95% CI 0.12 to 3.68).[111]

Heterogeneity of Treatment Effect

One RCT (N=309) of mobility and flexibility exercises versus usual care[122] reported that there was no effect modification for pain based on age, back pain duration at entry, sudden versus gradual development, or radiation of pain. Similarly, no modification for functional outcomes was seen based on mobility problems at entry, back pain history, or radiation of pain. P-values or data to evaluate this were not reported for effect modification analyses and the time frame for evaluation is unclear. Subanalysis of this RCT[121] in patients who were employed at study enrollment also found no effect modification by age, sex, level of education, duration of pain at entry, previous back pain, previous physiotherapy for back pain, or degree of pain or loss of mobility at the time of entrance into the trial for the outcome of work absence during the back pain episode. P-values or data to evaluate this were not reported for effect modification analyses and the time frame for evaluation is unclear in both publications.

Exercise Versus Bed Rest

Three RCTs (N=408)compared exercise to bed rest for the treatment of ALBP (Table 55; Appendix E).[59,127,139] Mean patient age (range, 38-41 years) and mean pain intensity (range, 4.8-6.3 on a 0-10 scale) and duration of pain (range, 3.8 to ≥6 days) at baseline were similar across the trials. The proportion of females ranged from 42% to 67%. In two trials, most patients had a history of prior back pain (69% to 70%)[59,139] compared with 22% in the third trial.[127] One RCT[139] was restricted to patients with radiculopathy and two[59,127] had mixed populations with and without radiculopathy but only one trial provided the proportion of patients with radicular pain (26%).[127] In two trials,[59,127] pregnant women were excluded. Medical and psychiatric comorbidities, history of substance use disorder, and opioid use were not reported.

Two RCTs[59,127] utilized mobility and flexibility exercises and one[139] utilized a combined mobility and flexibility and neuromuscular re-education approach. Exercise categorization is detailed in Appendix E. Exercise frequency ranged from every other day to twice weekly in two trials[127,139] and was not reported in the third trial.[59] Other exercise characteristics were not well reported (i.e., intensity, session duration, total number of sessions). Treatment duration ranged from 4 weeks to at least 2 months in two trials[59,139] and was unclear in one trial.[59] Bed rest ranged from 2-7 days. Concomitant treatment in all trials included education and advice and analgesic medication; in one trial, education consisted of a 30-minute video presentation.[59] One trial was conducted in an occupational health care setting,[127] one in a primary care setting[59] and one in an outpatient physiotherapy setting.[139]

Two RCTs[127,139] were conducted in Europe and one RCT[59] in Canada. All three trials[59,127,139] were assessed as moderate risk of bias. Common methodological concerns included unclear concealment allocation methods, baseline difference between treatment group, and inability to blind patients and providers (Appendix F).

Table 55. Study characteristics of exercise versus bed rest

Study, Year	Evans, 1987	Malmivaara, 1995	Hofstee, 2002
N Randomized	122	119	167
Mean Age (Years)	41	41	38
% Female	52%	67%	42%
% Radiculopathy	Mixed (% NR)	26%*	100%
% Previous Back Pain	Previous episodes 0: 27%; 1-4: 46%; ≥5: 23%	Previous 12 months: 22%	70%
Pain Duration Inclusion Criteria	NR	≤3 weeks	<1 month
Mean Pain Duration (days)	0-5 days: 48% ≥6 days: 62%	4.9	3.8†
Mean/Median Baseline Pain Severity (0-10)	4.8 (converted from McGill Pain Scale)	6	6.3
History substance abuse disorder	NR	NR	NR
Psychiatric Comorbidities	NR	NR	NR
Intervention	Mobility, Flexibility Exercise: 20-minute sessions, 5 exercises per session	Mobility, Flexibility Exercise: 10 repetitions per exercise (individualized), every other day	Combination Exercise: 2x/week
Comparator	Bed rest	Bed rest	Bed rest
Treatment duration	Exercise: ≥2 months Bed rest: ≥4 days	Exercise: NR Bed rest: 2 days	Exercise: 4-8 weeks Bed rest: 7 days
Duration follow-up	1 year	12 weeks	6 months
Country	Canada	Finland	Netherlands
Quality	Moderate	Moderate	Moderate

NR = not reported.
* Patients were excluded if they had sciatic syndrome.
† Whole sample. Includes more than two interventions, but data not reported by arm.

Detailed Synthesis

Function and Pain

Three RCTs (N=354) found exercise associated with similar improvement in function scores (Activity Discomfort Scale, ODI, and Quebec Back Pain Disability Scale) and pain scores (VAS pain or McGill Total Pain Score) compared with bed rest over follow-up periods ranging from 15 days to 6 months (Table 56).[59,127,139] Data was not well reported in general. At 1 year in one trial (N=112),[59] more patients who received exercise versus bed rest reported no restriction in activities (62.7% vs. 50.9%) but the difference was not statistically significant (RR 1.23, 95% CI 0.89 to 1.71) and similar proportions of patients in both groups reported that they had no pain presently (57.6% vs. 60.4%) and had no back pain episodes in the previous 5 months (27.1% vs. 26.4%) (Appendix E).

Table 56. Exercise versus bed rest: pain and function scores

Author, Year	Outcome	Time	N	Exercise vs. Bed Rest MD (95% CI)*
Pain				
Evans, 1987	McGill Total Pain Score (0-78)	15 days	88	MD in change scores 0.19 (-0.49 to 0.87)
Malmivaara, 1995	VAS Pain (0-10)	3 weeks	104	Mean 3.1 vs. 2.4, MD 0.7 (95% CI NC)
		12 weeks	100	Mean 1.8 vs. 2.1, MD -0.3 (95% CI NC)
Hofstee, 2002	VAS Pain (0-10)	4 weeks	162	Mean change scores -2.42 vs. -2.59, MD in change scores 0.17 (95% CI NC)
		8 weeks	158	Mean change scores -3.70 vs. -3.81, MD in change scores 0.11 (95% CI NC)
		6 months	150	Mean change scores -4.68 vs. -4.82, MD in change scores 0.14 (95% CI NC)
Function				
Evans, 1987	Activity Discomfort Scale (0-72)	15 days	88	MD in change scores 0.58 (-0.23 to 1.39)
Malmivaara, 1995	ODI (0-100)	3 weeks	104	Mean 18.6 vs. 16.0, MD 2.6 (95% CI NC)
		12 weeks	100	Mean 10.8 vs. 11.8, MD 1.0 (95% CI NC)
Hofstee, 2002	Quebec Disability Score (0-100)	4 weeks	162	Mean change scores -15.7 vs. -11.4, MD in change scores -4.3 (95% CI NC)
		8 weeks	158	Mean change scores -26.3 vs. -23.5, MD in change scores -2.8 (95% CI NC)
		6 months	150	Mean change scores -34.6 vs. -32.7, MD in change scores -4.3 (95% CI NC)

CI = confidence interval; MD = mean difference; NC = not calculable; ODI = Oswestry Disability Index; VAS = visual analogue scale

* All difference p>0.05

Other Outcomes

One RCT (N=109) found exercise associated with similar results compared with bed rest for the following: quality of life scores (scale 0-1; range, 0.91-0.93 vs. 0.92-0.95) and rate of return to work (range, 94%-100% vs. 95%-100%) at 3 and 12 weeks and analgesic or anti-inflammatory drug use at 12 weeks (91% vs. 93%) (Appendix E).[127] A second trial (N=167) found a similar incidence of treatment failure (intolerable pain and requested surgery) in both groups, respectively, at 1, 2, and 6 months (range 2% to 23% vs. 6% to 25%) (Appendix E).[139]

Adverse Events

One RCT reported one case each of cauda equina syndrome and pulmonary embolism in patients who received bed rest.[139]

Exercise Versus Manual Therapy

Two RCTs (N=180) compared exercise versus manual therapy for the treatment of ALBP (Appendix E).[84,119] Mean patient age ranged from 39 to 41 years and 45% to 47% were female. One trial[119] enrolled patients sick-listed due to LBP (severity 5.1 on a 0-10 scale) with (21%) or without (79%) sciatica of ≤2 weeks duration; patients with psychiatric comorbidities or a history of substance abuse and pregnant women were excluded. The second trial did not report any other patient characteristics.[84] History of back pain was not reported.

One trial[84] employed mobility and flexibility exercises (McKenzie method), 10 repetitions, twice daily for 4 weeks and the other thtrial[119] utilized combined mobility and flexibility and

neuromuscular re-education exercises conducted three times per week over 8 weeks. Exercise categorization is detailed in Appendix E. Manual therapy was comprised of manipulation and soft tissue mobilization both trials[84,119] with one trial[119] adding muscle energy techniques, autotraction, and stretches. One study[119] reported concomitant analgesic use in 67% of patients.

One study was conducted in Europe[119] and the other in Pakistan.[84] Both trials were assessed as high risk of bias due to unclear randomization and allocation concealment methods, baseline differences between groups or insufficient baseline information, and inability to blind patients and providers (Appendix F).

Detailed Synthesis

Function and Pain

In one trial,[119] mean function scores (ODI, 0-100 scale) were similar between patients who received combination exercise versus manual therapy (with muscle energy techniques) at 1 month (N=113, 19 vs. 22), 3 months (N=97, 14 vs. 12), and 12 months (N=82, 10 vs. 10); likewise, mean pain scores (modified Borg Pain Intensity, 1-11 scale) were similar between groups at all timepoints (1 month, 4.3 vs. 4.5; 3 months, 3.2 vs. 3.5; and 12 months, 2.5 vs. 2.5). The second trial (N=60) found McKenzie exercise associated with more functional impairment compared with manual therapy at 4 weeks as measured by combined ODI scores (scale NR, mean 388 vs. 244, respectively, p=0.0001) as well a substantially lower likelihood of having no to moderate pain (score 0-2 on the ODI pain scale) compared with manual therapy at 4 weeks (N=60, 46.7% vs. 96.7%, RR 0.48, 95% CI 0.33 to 0.71) (Appendix E).[84]

Other Outcomes

One RCT (N=82) found no effect of combination exercise compared with manual therapy for the following outcomes at 1 year: LBP recurrence (31% vs. 35%, RR 0.82, 95% CI 0.43 to 1.55), number of days off work for back pain (MD -8, 95% CI -41.85 to 25.85), or regular analgesic consumption (decrease from 67% at baseline to 21% for whole population, no difference between groups) (Appendix E).[119]

Adverse Events

Adverse events were not reported.

Exercise Versus Education (Back School)

One RCT (in two publications) (N=100) compared mobility and flexibility exercise (according to the McKenzie method) performed in 20-minute sessions for 2 weeks (frequency or total number of sessions not reported) versus one 45-minute "mini back school" lesson for the treatment of ALBP (Appendix E).[66,82] Mean patient age was 34 years, 23% were female and 55% had radiculopathy. Duration of symptoms at baseline was <2 weeks in 91% of patients and 2 to 4 weeks in the remainder. Mean baseline pain severity, medical or psychiatric comorbidities and history substance use disorder or opioid use were not reported. Pregnant women were excluded. Use of concomitant therapies was not reported. This trial[66,82] was performed in Sweden and was assessed as moderate risk of bias. Methodological limitations included unclear allocation concealment, baseline differences between groups and inability to blind patients and participants (Appendix F).

Detailed Synthesis

Function and Pain

The trial[82] did not provide data on function or pain, but noted significantly less pain ($p<0.001$) in the exercise group at both 3 weeks and 12 months.

Return to Work and LBP Recurrence

Exercise was associated with a large increase in the likelihood of returning to work at 2 weeks (80% vs. 40%, RR 2.00, 95% CI 1.39 to 2.89) and a small increase at 4 weeks (92% vs. 70%, RR 1.31, 95% CI 1.08 to 1.60) compared with education. At 6 weeks, the difference between groups was below the threshold for a small effect (100% vs. 87%, RR 1.14, 1.03 to 1.26) and by 11 weeks all patients in both groups had returned to work. Exercise was associated with a moderate decrease in likelihood of LBP recurrence at 12 months (44% vs. 74%, RR 0.59, 95% CI 0.42 to 0.85) and a small decrease at 60 months (63.8% vs. 88%, RR 0.72, 95% CI 0.57 to 0.92) versus education.

Adverse Events

Adverse events were not reported.

Exercise Versus Advice to Stay Active

One RCT compared combination exercise (mobility and flexibility and neuromuscular reeducation) performed twice a day for 4 to 8 weeks versus advice to stay active (continue normal ADLs) for the treatment of ALBP with radiculopathy (Appendix E).[139] Mean patient age was 39 years and 40% were female; at baseline, mean duration of pain was 3.8 weeks and mean pain severity was 6.2 (0-10 scale). Medical and psychiatric comorbidities, history of substance use disorder or opioid use, and inclusion or exclusion of pregnant women were not reported. Patients were allowed to use analgesic medication as needed. The trial[139] was conducted in the Netherlands and was assessed as moderate risk of bias. Methodological limitations included unclear allocation concealment and inability to blind patients and providers.(Appendix F).

Detailed Synthesis

Function and Pain

There was no effect of exercise on function or radicular pain outcomes compared with advice to stay active at any timepoint (Table 57).[139]

Table 57. Exercise versus advice to stay active: function and pain results from Hofstee 2002[139]

Outcome	Time	N	Exercise vs. Advice to Stay Active MD in Change Scores (95% CI)
Quebec Disability Score (0-100)	1 month	163	-0.5 (-6.3 to 5.3)
	2 months	156	0.0 (-7.2 to 7.3)
	6 months	147	-0.7 (-8.4 to 6.9)
VAS Radicular Pain (0-10)	1 month	163	0.08 (-0.82 to 0.98)
	2 months	156	-0.03 (-0.94 to 1.00)
	6 months	147	-0.10 (-1.00 to 0.80)

CI = confidence interval; MD = mean difference; VAS = visual analogue scale.

Treatment Failure

Similar proportions of patients in the exercise and advice to stay active groups failed treatment (i.e., intolerable pain and requested surgery) at 1 month (N=163, 2% vs. 7%, RR 0.35, 95% CI 0.07 to 1.66), 2 months (N=156, 13% vs. 12%, RR 1.13, 95% CI 0.51 to 2.50) and 6 months (N=147, 23% vs. 17%, RR 1.41, 95% CI 0.77 to 2.60).[139]

Adverse Events

Adverse events were not reported.

Exercise Versus Acupuncture

One RCT (N=54) compared myofascial meridian stretching exercise to traditional Chinese medicine acupuncture for the treatment of ALBP for a median of 25 days duration (Appendix E).[111] Mean patient age was 44 years, 76% were female and median pain at baseline was 8.0 (0-10 scale). Patients with radiculopathy and psychiatric comorbidities and pregnant women were excluded. History of substance use disorder or opioid use, and history of back pain were not reported. Exercises were administered by a physiotherapist (length of sessions unclear) and acupuncture was performed for 20 minutes per session; both treatments were given three times per week for 2 weeks. All groups received etodolac at 400 mg twice a day for 2 weeks. The trial[111] was performed in Turkey and considered moderate risk of bias. Methodological limitations included unclear allocation concealment methods, baseline differences between groups and inability in blind patients and care providers (Appendix F).

Detailed Synthesis

Function, Pain and Medication Use

Exercise was associated with similar function scores (RDQ, 0-24 scale) and pain scores (NRS, 0-10 scale) at 2 and 6 weeks compared with acupuncture, respectively: median 3 versus 4 for both outcomes and all timepoints (p range, 0.39 to 0.66) (Appendix E).[111] All patients reported using NSAIDs during the trial period.

Withdrawal due to Adverse Events

One RCT reported that two patients in the exercise group and one patient in the acupuncture group withdrew from the trial due to increased pain (7.4% vs. 3.7%, RR 2.0, 95% CI 0.19 to 20.77).[111]

Manual Therapy

Manual Therapy Versus Placebo or Sham: Description of Included Studies

Five RCTs (N=665) compared manual therapy with a sham or placebo intervention.[74,83,102,106,152] Patients were around 40 years of age and females comprised less than half of the study populations (13% to 45%). Patients with radiculopathy or suspected neurologic symptoms such as leg pain were excluded from two RCTs[83,152] and one only enrolled participants with radiating leg pain due to disc protrusion.[102] All trials enrolled patients with prior back pain but the proportion with prior back pain was not reliably reported across trials. Mean pain

duration at study entry ranged from a median of 9 days to a mean of around 26 days. Information on psychiatric comorbidities and history of substance use disorder were not provided by any trial. Patient and intervention characteristics are summarized in Table 58 and detailed in the data abstraction (Table 58 and Appendix E).

Manual therapy techniques described in this section included manipulation and/or mobilization and varied across trials. Three trials describe used of high velocity thrust procedures involving the lower spine pelvis or sacrum.[74,83,102] Two describe rotational manipulation of the trunk,[106,152] one of which also described mobilization of hypomobile segments.[106] Session length and numbers of sessions varied as did treatment duration (Table 58). Placebo treatments were short wave or detuned diathermy,[106,152] or detuned pulsed ultrasound,[83] and sham manipulation involving muscle pressure without thrusting were used in two trials.[74,102] In one RCT, all patients received detuned short-wave diathermy for 4 days.[152] Duration and frequency of placebo and sham procedures were the same as those for the intervention. Use of concurrent medications was poorly described. One RCT excluded patients using medications,[74] one prohibited opioid or steroid use but provided no further information,[102] another provided all patients with acetaminophen (1g, 4 times/day until recovery or for 4 weeks), and a subset of patients in both groups received diclofenac 50 mg twice per day,[83] and one allowed use of salicylates, acetaminophen or a combined analgesic.[106] Specific use of rescue medication was not reported.

Risk of bias was low in two RCTs,[83,102] moderate in two RCTs,[74,152] and high in one RCT[106] (Appendix F). In RCTs at moderate risk of bias, lack of information regarding randomization and/or concealment of treatment allocation and lack of similarity between groups at baseline were common limitations. In addition, attrition and differential loss to follow-up were not clearly reported in the high risk of bias trial.[106]

Table 58. Patient and study characteristics: manual therapy versus sham or placebo

Study, Year	Bergquist-Ullmann, 1977	Hancock, 2007	Hoiriis, 2004	Santilli, 2006	Glover, 1974
N Randomized	217*	239	192†	102	84
Mean Age (Years)	35†	41	42	<40: 39% 40-49: 29% ≥50: 31%	40
% Female	13%†	44%	45%‡	37%	13%
% Radiculopathy	Mixed (% NR)	NR (excluded)§	NR	100%	0% (Excluded)
% Previous Back Pain	1 episode: 32%† >1 episode: 25%†	% NR, mean of 3.7 prior episodes**	% NR††	% NR (needed to be pain free for 3 months)	59%‡‡
Pain Duration Inclusion Criteria	<3 months	<6 weeks	2 to 6 weeks	<10 days	NR
Mean Pain Duration (days)	<3 weeks: 83%† Median: 9 days†	9.1	3.8 weeks	25.5 days§§	<7 days: 52% ≥7 days: 48%
Mean/Median Baseline Pain Severity (0-10)	Pain index score: Median 42†***	6.5	4.4	6.4	NR
History Substance Abuse Disorder	NR	NR	NR	NR	NR
Psychiatric Comorbidities	NR	NR	NR	NR	NR

Study, Year	Bergquist-Ullmann, 1977	Hancock, 2007	Hoiriis, 2004	Santilli, 2006	Glover, 1974
Treatment	Mobilization of hypomobile segment with manual pressure and patient trunk rotation)	Mobilization or manipulation with high velocity thrust procedures)	Manipulation: specific high velocity, low amplitude thrusts to lumbar, pelvis or sacrum	Manual Therapy: Soft tissue manual therapy and brisk rotational thrusting away from greatest restriction	Manipulation (rotational manual therapy of lower spine and sacrum)
Comparator	Placebo (Short waves diathermy negligible heat)	Placebo: detuned pulsed ultrasound	Sham manipulation: hand on musculature, light pressure, no thrust, no instrumentation	Sham, simulated manipulation; soft muscle pressing; no thrust	Placebo Detuned diathermy
Dosage Frequency	MT: Max 10 sessions, mean 4 sessions Placebo: Max 10 sessions	Maximum of 12, 30–40-minute sessions, 2 to 3 times per week	7 visits	MT: 5 minutes, mean 12.8 (SD=4.8) sessions Sham: 5 minutes, mean 13.0 (SD=4.5) sessions	MT: One 15-minute manual therapy session + 4 15-minute sham diathermy sessions Sham: five 15-minute sessions
Treatment Duration	2 weeks	4 weeks	2 weeks	5 weeks	4 days
Duration Follow-up	1 year	12 weeks	4 weeks	180 days	7 days
Concurrent Medication	Salicylates, acetaminophen or combined analgesic if needed	All received acetaminophen, 1g, 4 x day until recovery or 4 weeks	Not specified; excluded patients with current mediation use	Opioid and steroid not allowed; NSAIDs used	NR
Country	Sweden	Australia	U.S.	Italy	Wales
Risk of Bias	High	Low	Moderate	Low	Moderate

HVLA = high-velocity low-amplitude; MT = manual therapy; MET = muscle energy technique; NR = not reported; OMT = osteopathic manipulative treatment; SD = standard deviation; UC = usual care.
* Patients were employees of the Automotive Division of AB Volvo.
† Whole sample. Includes more than two interventions, but data not reported by arm. For comparison of interest, Bergquist-Ullmann reports baseline outcome data for 146 patients, Hoiriis reports baseline data for 94 patients.
‡ 50% of MT patients and 39% of placebo patients were female respectively.
§ Exclusion criteria stipulated nerve root compromise with at least two of these signs: myotomal weakness, dermatomal sensory loss, or hyporeflexia of the lower limb reflexes.
** Patients reported mean 3.7 episodes.
†† Patients reported mean 0.47 episodes. 46% of patients reported previous chiropractic care.
‡‡ Patients were initially divided into those experiencing their first episode of pain and those with ≥2 episodes before being randomized, and then further divided into those with <7 days of pain and ≥7 days of pain, and then randomized into their respective intervention groups.
§§ Authors report a difference in the number of total days with pain ($p<0.005$), with MT patients experiencing 23.6 days versus 27.4 days in the placebo group.
*** Range of scores is 0 to 70 points.

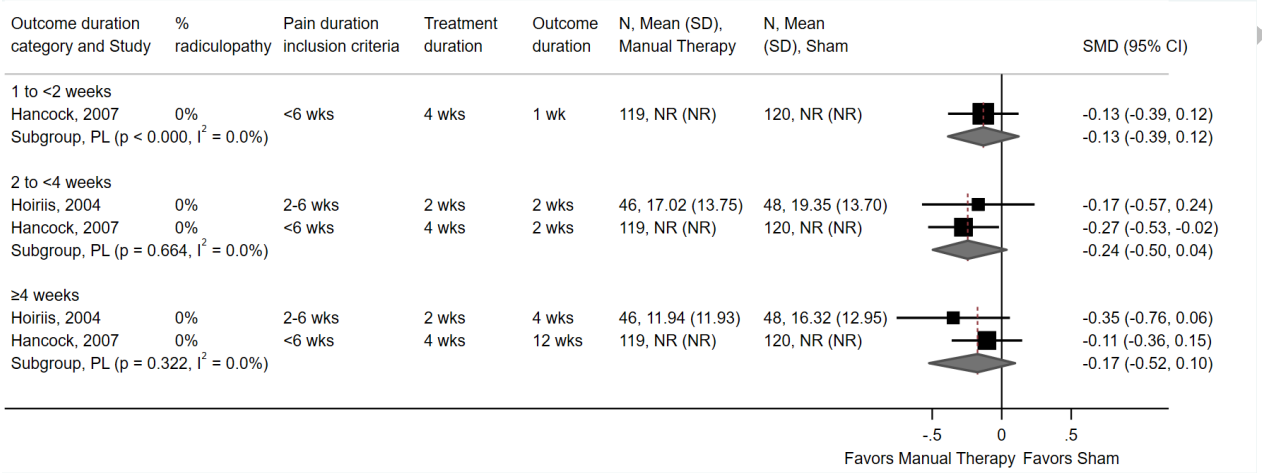
Detailed Synthesis

Function

Two RCTs (N=333) reported measures of function and contributed to pooled analyses. One reported ODI (0-100 scale)[74] and the other reported RDQ (0-24 scale).[83]

Functional improvement was similar between manual therapy and sham manual therapy or placebo 1 to <2 weeks (1 RCT, N=239). Manual therapy was associated with small functional improvement versus sham at 2 to <4 weeks (2 RCTs, N=333).[74,83] Improvement in function was similar across the same two trials at ≥4 weeks but tended to favor manual therapy. Neither trial included patients with radiculopathy (Figure 31).

Figure 31. Manual therapy versus sham or placebo: function scores



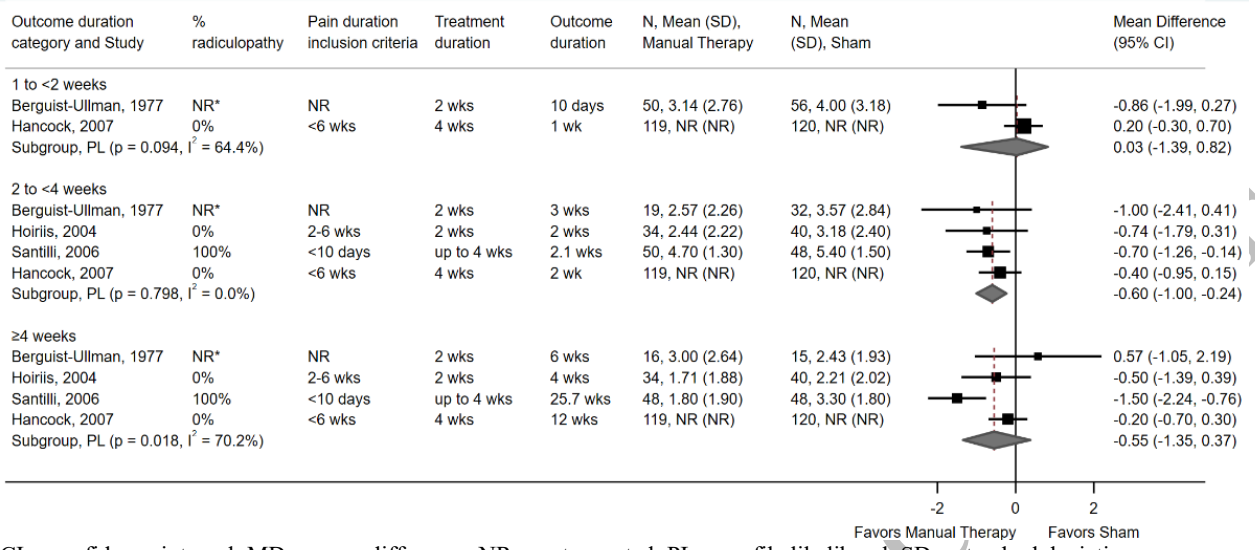
CI = confidence interval; NR = not reported; PL = profile likelihood; SD = standard deviation; SMD = standardized mean difference.

Pain

Three RCTs reported NRS or VAS (0-10 scale)[74,83,102] and a fourth reported a pain index (0 to 70 scale, rescaled to 0-10).[106]

Two trials reported similar pain improvement for manual therapy and placebo or sham at 1 to <2 weeks (2 RCTs, N=345).[83,106] Manual therapy was associated with a small improvement in pain between 2 and <4 weeks (4 RCTS, N=462, pooled MD -0.60, 95% CI -1.00 to -0.24, I^2= 0%),[74,83,102,106] but not at ≥4 weeks (4 RCTs, N=440) (Figure 32). Substantial heterogeneity was observed at ≥4 weeks. The mean difference between treatments was substantially higher in one outlier trial (MD -1.50) in patients with radiating leg pain due to disc protrusion[102] compared with the MD ranges from other trials[74,83,106] that did not include patients with radiculopathy (-0.2 to 0.57) and found no difference between treatments.

Figure 32. Manual therapy versus sham or placebo: pain scores (0-10 scale)



CI = confidence interval; MD = mean difference; NR = not reported; PL = profile likelihood; SD = standard deviation.
* Population with and without radiculopathy but % with radiculopathy not reported.

Other Pain Outcomes

One trial (N=98) enrolled participants with radiating leg pain due to disc protrusion.[102] Back pain improvement scores are reported in the meta-analysis above (Figure 32). Authors also report that on the proportions of patients who reported pain reduction and becoming free of local (back) pain. Manual therapy was associated with a slightly higher likelihood of pain reduction at 2 weeks versus placebo, however the likelihood was similar between groups at other times (effect estimates were below the threshold for a small effect) (Table 59). Manual therapy was also associated with substantially higher likelihood or becoming pain free beginning at 6.4 weeks, however estimates were imprecise. The same trial also reports VAS leg pain scores (0-10) and the proportions of patients who reported reduction in radiating pain and those becoming pain free (Table 59). Moderate improvement in radiating leg pain scores (VAS 0-10) at all but the earliest time (2 weeks) was seen with manual therapy versus placebo. Manual therapy was associated with a moderately higher likelihood of pain reduction at 2 weeks versus placebo but a small likelihood at subsequent times and with a larger likelihood of becoming pain free at all reported time frames. Effect sizes for these outcomes were imprecise, however (Table 59).

Table 59. Manual therapy versus placebo: other reported pain outcomes.

Author, Year LBP Type	Outcome	Time	Manual therapy vs. Sham or placebo % (n/N) and RR (95% CI) or, Mean (SD) and MD (95% CI)
Santilli, 2006 100% with leg pain	% of patients with reduction in local (back) pain	2.1 weeks (15 days)	**88.0% (44/50) vs. 62.5% (30/48), RR=1.41 (95% CI 1.11 to 1.79)**
		4.2 weeks (30 days)	94.0% (47/50) vs. 85.4% (41/48), RR=1.10 (95% CI 0.96 to 1.26)
		6.4 weeks	100% (48/48) vs. 91.7% (44/48), RR=1.09 (95% CI 1.00 to 1.19)
		12.9 weeks (90 days)	97.9% (47/48) vs. 89.6% (43/48), RR=1.09 (95% CI 0.98 to 1.21)
		25.7 weeks (180 days)	97.9% (47/48) vs. 93.8% (45/48) RR=1.04 (95% CI 0.96 to 1.14)

Author, Year LBP Type	Outcome	Time	Manual therapy vs. Sham or placebo % (n/N) and RR (95% CI) or, Mean (SD) and MD (95% CI)
Santilli, 2006 100% with leg pain	% of patients becoming pain free (back pain)	2.1 weeks (15 days)	0% vs. 0%
		4.2 weeks (30 days)	6% (/50) vs. 0% (0/48)
		6.4 weeks	17% (9/49) vs. 6% (3/48) RR=2.77 (95% CI 0.80 to 9.66)
		12.9 weeks (90 days)	**24% (13/48) vs. 6% (3/48)** **RR=4.01 (95% CI 1.21 to 13.22)**
		25.7 weeks (180 days)	**28% (15/48) vs. 6% (3/48)** **RR=4.62 (95% CI 1.42 to 15.00)**
Santilli, 2006 100% with leg pain	VAS Leg (radiating) Pain (0-10 scale)	2.1 weeks (15 days)	3.5 (1.9) (n=50) vs. 4.0 (1.7) (n=48) MD-0.50 (-1.22 to 0.22)
		4.2 weeks (30 days)	**2.2 (1.9) (n=49) vs. 3.5 (1.7) (n=48)** **MD -1.3 (-2.03 to -0.57)**
		6.4 weeks	**1.5 (1.9) (n=48) vs. 2.9 (1.8) (n=48)** **MD – 1.4 (-2.15 to -0.65)**
		12.9 weeks (90 days)	**1.2 (1.8) (n=48) vs. 2.9 (1.9) (n=48)** **MD -1.7 (-2.45 to -0.95)**
		25.7 weeks (180 days)	**1.2 (1.7) (n=48) vs. 2.4 (1.9) (n=48)** **MD =1.2 (-1.93 to -0.47)**
Santilli, 2006 100% with leg pain	% of patients with reduction in radiating (leg) pain	2.1 weeks (15 days)	**82% (42/50) vs. 53% (26/48)** **RR=1.55 (95% CI 1.16 to 2.08)**
		4.2 weeks (30 days)	**94% (47/49) vs. 77% (37/48)** **RR=1.22 (95% CI 1.03 to 1.44)**
		6.4 weeks	**100% (48/48) vs. 81% (39/48)** **RR=1.23 (95% CI 1.07 to 1.41)**
		12.9 weeks (90 days)	**98% (47/48) vs. 77% (37/48)** **RR=1.27 (95% CI 1.08 to 1.49)**
		25.7 weeks (180 days)	**100% (48/48) vs. 83% (40/48)** **RR=1.20 (95% CI 1.06 to 1.36)**
Santilli, 2006 100% had leg pain	% of patients becoming pain free (leg pain)	2.1 weeks (15 days)	**13% (7/50) vs. 4% (2/48)** **RR=3.24 (95% CI 0.71 to 14.84)**
		4.2 weeks (30 days)	**23% (12/49) vs. 12% (6/49)** **RR=1.85 (95% CI 0.75 to 4.55)**
		6.4 weeks	**41% (22/48) vs. 16% (8/48)** **RR=2.54 (95% CI 1.25 to 5.17)**
		12.9 weeks (90 days)	**55% (29/48) vs. 12% (6/48)** **RR=4.47 (95% CI 2.03 to 9.83)**
		25.7 weeks (180 days)	**55% (29/48) vs. 20% (10/48)** **RR=2.68 (95% CI 1.46 to 4.91)**

CI = confidence interval; LPB = low back pain, MD = mean difference; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

One RCT (N=84) of manual therapy versus placebo (detuned short-wave diathermy)[152] reported the patient's assessment of pain relief as a range from 0% (no relief) to 100% (complete relief); the reported mean pain relief was based on the percentage of relief for each patient. Across patients reporting one or more ABLP episodes, the mean percent relief from baseline was similar for manual therapy and placebo immediately post treatment (34% vs. 22%), at 3 days (50% vs. 56%) and at 7 days (75% vs. 80%). Authors report no differences between manual therapy and sham for sub-analyses of patients based on number of prior LPB episodes and current episode duration of <7 days or ≥7 days with one exception: In patients presenting with their first episode of ALBP and pain duration of <7 days, manual therapy was associated with a

higher percent pain improvement from baseline versus placebo immediately posttreatment (43% vs. 12%) but did not provide further information.

Recurrence

Recurrence of LBP was reported in one trial (N=142) and was similar at 52 weeks in patients receiving manual therapy and those receiving placebo: 60% vs. 71%, RR=0.85 (95% CI 0.67 to 1.07). The mean number of recurrences was 2.75 vs. 2.25.[106]

Quality of Life, Sleep Quality, Perception of Improvement

One RCT (N=102) reported no difference between manual therapy and sham manual therapy for any of the individual domain scores of the 36-item Short Form Questionnaire (SF-36) (0-100) at 6.4 weeks.[102] One trial (N=239) found that global perceived effect (-5 [much worse] to 5 [completely recovered] scale) was similar for patients receiving manual therapy and those receiving placebo at any time posttreatment (mean ranges -0.1 to 0.3 vs. -0.3 to 0.1).[83] A third trial (n=94) reported that manual therapy was associated with better Global Impression of Severity scores (0-31-scale) compared with sham at the end of the 2-week care phase (MD -2.2, 95% CI -3.93 to -0.47) (Appendix E).[74]

Psychological Distress

Two RCTs reported on measures of psychological distress.[74,102] No differences in scores were seen between manual therapy or sham/placebo on the Kellner Symptom Questionnaire (0-23 scale) for either depression (means 4.1 vs. 4.4) or anxiety (means 6.1 vs. 6.4) in one RCT[102] (N=102) or on the Modified Zung Self-Rating for Depression Scale (20-80 scale) at 2 weeks (means 12.90 vs. 13.83) or 4 weeks (means 11.91 vs. 11.81) in the other trial (N=90)[74] (Table 60).

Table 60. Manual therapy versus sham or placebo: psychological distress

Author, Year LBP Type	Outcome	Time	Manual therapy vs. sham or placebo MD (95% CI)
Santilli, 2006 100% had leg pain	Kellner Rating scale (0-23): Depression	6.4 weeks	4.1 (3.1) (n=53) vs. 4.4 (2.8) (n=49) MD -0.30 (-1.46 to 0.86)
	Kellner Rating scale (0-23): Anxiety	6.4 weeks	6.1 (3.6) (n=53) vs. 6.4 (3.8) (n=49) MD -0.30 (-1.75 to 1.15)
Hoiriis, 2004 LBP type NR	Modified Zung Self-Rating for Depression Score (20 to 80 scale)	2 weeks	12.49 (9.39) (n=43) vs. 13.83 (8.90) (n=47) MD -1.34 (95% CI -5.17 to 2.49)
		4 weeks	11.91 (10.53) (n=43) vs. 11.81 (7.37) (n=47); MD 0.10 (95% CI -3.62 to 3.82)

CI = confidence interval; LBP = low back pain, MD = mean difference RR = risk ratio; SD = standard deviation

Return to Work and Work Absence

Return to work was not reported for this comparison.

Medication Use

None of the trials reported on opioid use. One trial in patients with radiating leg pain due to disc protrusion reported nonsignificant differences between manual therapy and sham in NSAID use based on either the number of days or number of prescriptions.[102] Similarly, another trial reported similar use of acetaminophen over the 2-week treatment phase for manual therapy and sham.[74]

Adverse Events and Withdrawal Due To Adverse Events

No adverse events related to manual therapy were observed in two RCTs.[83,102] The other three RCTS did not report on adverse events.[74,106,152]

Heterogeneity of Treatment Effect

None of the included RCTs reported on modification of treatment effect.

Manual Therapy Versus Usual Care

Description of Included Studies

Five RCTs (N=473) compared various forms of manual therapy with usual care. Sample sizes ranged from 63 to 120.[115,116,119,148,166]

Patient and study characteristics are summarized in Table 61 and Appendix E. The mean participant age was 33 years and the average proportion of females was 40%. One trial[119] excluded individuals with psychiatric comorbidities or history of substance use disorder, but no other trials reported on these. All trials excluded pregnant people. Two trials in active duty military personnel[116,148] excluded individuals with radiculopathy, one trial[115] reported radiculopathy in 43% of individuals, one trial[119] enrolled patients with and without radiculopathy but did not report prevalence and one trial[166] did not report on radiculopathy specifically but noted that 22% of patients complained of "pins and needles" sensations. The mean pain duration at baseline was ≤7 days in one trial[116] and >7 days in two trials;[115,148] two trials did not report mean pain duration but enrolled participants with up to 2 weeks of pain[119] or less than 30 days[166] duration. Mean pain at baseline was moderate and similar across four trials (range, 5.1 to 5.8, 11-point scale).[115,116,119,148] One trial[166] reported previous LBP in 49% of participants in the prior 12 months and 54% of participants in the past 6 years.

Manual therapy consisted of both manipulation (predominately) and mobilization techniques in all trials with or without additional components such as massage and stretching; usual care related to medication appears to have been continued in the intervention groups (Table 61). Treatment sessions were once to twice weekly for 2 weeks to 3 months. Session frequency was not standardized in any trial. Usual care generally included advice to remain active, avoiding bed rest and use of medications across trials. Follow-up post-treatment ranged from 2 to 52 weeks.

Reporting of concomitant medication use was limited. One trial noted similar use between manual therapy and usual care of schedule 1 medications (88%, naproxen, ibuprofen, and acetaminophen) or schedule 2 medication (38%, cyclobenzaprine and acetaminophen with codeine)[148] while another reported less overall medication use (NSAID, muscle relaxants, narcotics) in the manual therapy group versus usual care (18% vs. 37%).[115] One trial[116] did not allow concomitant treatments and another[166] did not report their use. No trial described use of rescue medications. Two trials were conducted in the United States[115,148] and three in Europe.[116,119,166]

Three trials were rated as moderate risk of bias[115,116,148] and two were high risk of bias.[119,166] Common limitations included unclear randomization methods and inability to blind patients. Differences in baseline characteristics between treatment groups and unacceptable or unclear loss to follow-up were also observed in high risk of bias trials (Appendix F).

Table 61. Study characteristics of manual therapy versus usual care

Study, Year	MacDonald, 1990	Goertz, 2013	Juni, 2009	Seferlis, 1998	Cruser, 2012
N Randomized	95	91*	104	120	63*
Mean Age (Years)	NR	26	35	39†	27
% Female	55%	14%	36%	47%†	45%
% Radiculopathy	NR‡	43%	0% (Excluded)	Mixed % NR	0% (Excluded)
% Previous Back Pain	49% in prior ≤12 months; 54% in 1-6 prior years:	NR	NR	NR§	NR
Pain Duration Inclusion Criteria	<30 days	<4 weeks	<4 weeks	≤2 weeks	NR
Mean Pain Duration (days)	<14 days in 55%**	Mean 12 Median 9	6.5	NR	12
Mean Baseline Pain Severity (0-10 unless noted)	Pain analog (scale 0-75): 19.6††	5.8	7.0	5.1 (1-11 scale)†	5.4 (current) 8.3 (worst)
History substance abuse disorder	NR	NR	NR	0% (Excluded)	NR
Psychiatric Comorbidities	NR	NR	NR	0% (Excluded)	NR
Treatment	Manipulation, mobilization: direct pressure, stretching; low velocity, high amplitude oscillatory movement; HVLA thrust	Manipulation, HVLA thrust; may have also done mobilization, massage, mobility, flexibility exercise	Manipulation, spinal mobilization, MET, HVLA	Manipulation, general and passive mobilization, MET, stretching, auto traction	Manipulation, mobilization: direct pressure, stretching; myofascial release, counter-strain, isometric muscular contraction, sacroiliac articulation, HVLA
Comparator	Advice on posture and stress avoidance, graded resumption of activities, reassurance	Advice, medication, referral for imaging, PT exercise, stretching, modalities (e.g., heat, electrical stimulation)	Advice on rapid return to normal activity, avoid bedrest, medication	Rest, sick leave, drug prescription, advice about posture, information on the self-curing nature of the disease	Advice, muscle relaxants, passive modalities)
Dosage Frequency	MT: 2 x per week UC: As necessary	MT: Up to two times a week UC: NR	MT: Up to 5 sessions (median 3 sessions) UC: NR	Determined by therapist	MT: Once weekly, not closer than ≥7 or ≥10 days apart UC: NR
Treatment duration	Twice weekly for 3 weeks, then weekly until recovery; for 3 months if unrecovered	4 weeks	2 weeks	NR	4 weeks
Duration follow-up	3 months	4 weeks	2 weeks	12 months	4 weeks after last treatment
Concurrent medication	NR	NSAID, muscle relaxants, narcotics: 18% vs. 37% (timing NR)	Acetaminophen, diclofenac or dihydrocodeine; type and dose at patient discretion		Medications: Schedule 1: 88% vs. 90%; Schedule 2: 37% vs. 40% (timing NR)
Country	UK	U.S.	Switzerland	Sweden	U.S.

Study, Year	MacDonald, 1990	Goertz, 2013	Juni, 2009	Seferlis, 1998	Cruser, 2012
Risk of bias	High	Moderate	Moderate	High	Moderate

HVLA = high velocity low amplitude; MT = manual therapy: MET = muscle energy technique; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; PT = physical therapy; UC = usual care.
* All patients were U.S. active-duty military personnel.
† RCT Includes more than two interventions. Data across all treatment arms reported as data are not reported by arm.
‡ Patients were excluded if they had a neurological deficit in structures innervated by lumbar or sacral roots that could not be ascribed to a previous resolved episode or other pathology.
§ Patients were excluded if they had records of sick leave for previous low back pain.
** Present episode duration:<14 days pain as severe in last 24 hours: 55%; <30 days pain (most days each week): 71%
†† On 0 to 75 scale, higher score = worse pain

Detailed Synthesis

Function

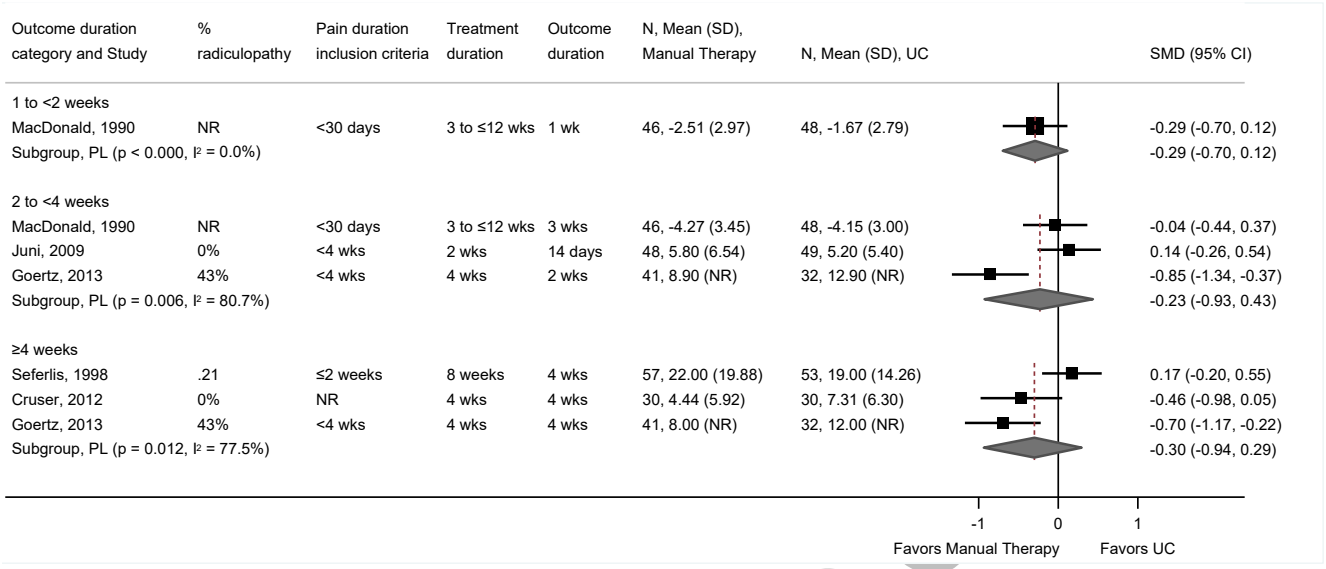
All five RCTs reported measures of function which included RDQ (0-24),[115,116,148] ODI (0-100 scale),[119] and one reported a Disability Index (DI).[166] The DI asks subjects to mark whether each of 12 everyday activities was comfortable with the score of 0 to 12 based on the number of activities left blank.

No trial reported the likelihood of achieving a successful functional outcome (e.g., >30% improvement on a functional scale).

One RCT (N=94) found no difference between manual therapy and usual care in functional scores at 1 to ≤2 weeks.[166] Pooled estimates at 2 to <4 weeks (3 RCTs, N=264)[115,116,166] also found similar improvement in function scores, however there is substantial heterogeneity (I^2=81%). One outlier RCT (N=73)[115] in active duty military personnel reported a moderate improvement (RDQ MD of 3.9 on 0-24 scale) in contrast to the other two trials that reported no difference between groups (Figure 33). Pooled estimates at ≥4 weeks also suggested no difference between manual therapy and usual care (3 RCTs, N=243, SMD -0.30),[115,119,148] and again substantial heterogeneity was seen (I^2=77%). The outlier trial[119] was at high risk of bias and data were not clearly reported. Following exclusion of it, manual therapy was associated with moderate improvement in function versus usual care across the two other RCTs, both in active duty military personnel (2 RCTs, pooled SMD -0.59, 95% CI -1.01 to -0.15, I^2=0%)[115,148] The pooled mean difference in RDQ (0-24 scale) across these two trials was -3.53 (95% CI -5.89 to -1.04) also corresponding to moderate improvement. One of these RCTs (N=73) also reported moderate improvement with manual therapy over usual care on the Back Pain Functional Scale (0-60 scale) at 2 and 4 weeks (Table 62).[115]

No trial reported the likelihood of achieving a successful functional outcome (e.g., >30% improvement on a functional scale). One RCT (N=94) reported that, in patients whose initial pain duration was 12 to 28 days, percentage improvement in DI scores (0-12 scale) was higher following manual therapy versus usual care (46% vs. 17%, p=0.04), but significant treatment effects were not seen in patients whose baseline pain duration was 1 to 13 days or >28 days; supporting data are not reported.[166]

Figure 33. Manual therapy versus usual care: function scores



CI = confidence interval; NR = not reported; PL = profile likelihood; SD = standard deviation; SMD = standardized mean difference; UC = usual care.

Pain

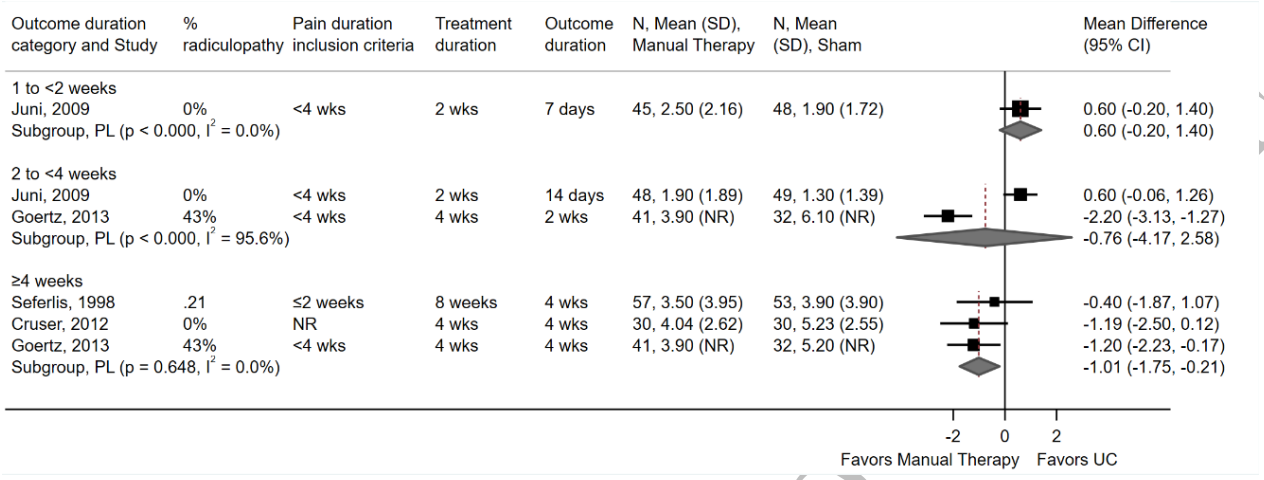
Four RCTs reported measures of pain (scale 0-10) and contributed to pooled analyses.[115,116,119,148]

One RCT (N=93) reported similar pain improvement between manual therapy and usual care at 1 to ≤2 weeks[116] (Figure 34). Results for two trials at 2 to <4 weeks were inconsistent: one RCT[115] (N=73) reported a large improvement in pain with manual therapy versus usual care (MD -2.20, 95% CI -3.13 to -1.27) while the other RCT[116] (N=97) showed similar improvement in pain between the two groups, and tended to favor usual care (MD 0.60, 95% CI -0.06 to 1.27). Estimates for both were imprecise (Figure 34). Population differences between the trials may partially explain the inconsistency: one trial (N=73) enrolled active-duty U.S. military personnel, 43% of whom reported pain radiation to extremities.[115] The other trial[116] (N=97) recruited slightly older patients (35 years vs. 26 years) and excluded patients with radiculopathy. Differences in manual therapy techniques may contribute to heterogeneity. At ≥4 weeks, manual therapy was associated with a small improvement in pain (3 RCTs, N=247).[115,119,148] One trial[148] reported current, typical, worst and best pain scores; worst pain is used in the meta-analysis, others are in Table 62. Sensitivity analysis using reported current pain showed a more moderate improvement (3 RCTs, N=247, pooled MD -1.28, 95% CI -2.02 to -0.36, I^2=13.7%),[115,119,148] but the estimate was less precise and more heterogeneity was present. Following exclusion of one high risk of bias trial,[119] manual therapy was associated with small improvement in pain versus usual care across the two remaining RCTs, both in active duty military personnel (2 RCTs, N=133, pooled MD -1.2, 95% CI -2.13 to -0.26, I^2=0%).[115,148]

No trial reported the likelihood of achieving a successful pain outcome (e.g., >30% improvement on a pain scale). Two trials reported on patient perceived pain relief. One RCT reported a substantially higher likelihood of patients reporting that their pain was gone, or much better or moderately better at 4 weeks (N=73, RR 5.27, 95% CI 2.05 to 12.53).[115] The other trial[116] (N=97) found no difference in the likelihood of being pain free at 2 weeks or 26 weeks (Table 62). As noted above, differences in patient population and manual therapy techniques

between these trials may partially explain the differences in findings. Estimates from both trials were imprecise.

Figure 34. Manual Therapy versus usual care: pain scores (0-10 scale)*



CI = confidence interval; MD = mean difference; NR = not reported; SD = standard deviation; MD = mean difference; UC = usual care.
* Cruser report of worst pain used in analysis; see text.

Table 62. Summary of function and pain outcomes not in meta-analyses

Author, Year LBP Type	Outcome	Time	Exercise vs. UC or no treatment % (n/N) and RR (95% CI) or, Mean (SD) and MD (95% CI)
Function			
Goertz, 2013 43% Radicular	Back Pain Functional Scale (0-60, higher = greater functional ability)	2 weeks	42.9 (NR) (n=41) vs. 32.9 (NR) (n=32), Adjusted MD 10.0 (95% CI 5.5 to 14.6)
		4 weeks	43.0 (NR) (n=NR) vs. 35.3 (NR) (n=NR), Adjusted MD 7.7 (95% CI 2.6 to 12.9)
Pain			
Cruser 2012 Nonradicular	VAS Pain - Current (0-10)	4 weeks	1.96 (1.47) (n=30) vs. 3.73 (2.39) (n=30) MD -1.77 (95% CI -2.77 to -0.77)
	VAS Pain - Typical (0-10)	4 weeks	2.15 (1.48) (n=30) vs. 3.65 (2.16) (n=30) MD -1.5 (95% CI -2.46 to 0.54)
	VAS Pain - Worst (0-10)	4 weeks	4.04 (2.62) (n=30) vs. 5.23 (2.55) (n=30) MD -1.19 (95% CI -2.50 to 0.12)
	VAS Pain - Best (0-10)	4 weeks	1.48 (0.96) (n=30) vs. 2.15 (1.65) (n=30) MD -0.67 (95% CI -1.37 to 0.03)
Goertz, 2013 43% Radicular	Pain gone, much or moderately better	4 weeks	65.9% vs. 12.5%, RR 5.27 (95% CI 2.05 to 12.53)
Juni, 2009 Nonradicular	Pain-free	2 weeks	31% (15/48) vs. 45% (22/49), RR 1.25 (95% CI 0.91 to 1.72)
		26 weeks	44% (22/50) vs. 59% (30/51), RR 0.75, 95% CI 0.51 to 1.10

CI = confidence interval; LBP = low back pain; MD = mean difference; NR = not reported; SD = standard deviation; MD = mean difference; RR = risk ratio; UC = usual care; VAS = visual analogue scale

Recurrence

One trial reported similar proportions of patients in manual therapy and usual care groups reporting no recurrence of LBP within 52 weeks (N=81, 65% vs. 60%, RR 1.11, 95% CI 0.79 to 1.57).[119]

Quality of Life, Sleep Quality, Perception of Improvement

None of the RCTs reported on QOL or sleep quality. Two RCTs in active duty military personnel reported greater patient satisfaction with manual therapy over usual care.[115,148] One trial[115] (N=73) reported greater patient satisfaction with their results with manual therapy versus usual care (0-10 scale) at 2 weeks (MD 4.4, 95% CI 3.42 to 5.38) and 4 weeks (MD 4.4, 95% CI 3.42 to 5.38) and the other RCT[148] (n=60) found greater patient satisfaction with treatment (0-10 scale) at 4 weeks (MD 2.93, 95% CI 1.75 to 4.11) following manual therapy (Appendix E).

One RCT[166] (N=94) found no statistically significant difference between manual therapy and usual care on the proportions of patients reporting recovery at any time period for subgroups based on initial pain duration 1 to 13 days, 12 to 28, or >28 days. Supporting data to evaluate effect estimates are not reported and estimates from graphs may not be accurate (Appendix E).

Psychological Distress

No trials reported on measures of psychological distress.

Return to Work and Work Absence

No trial reported on return to work. One trial (N=120) reported no difference between manipulation and usual care on the number of days off work for LBP at 52 weeks (MD, -5 days, 95% CI -36.32 to 26.32).[119]

Medication Use

Medication use during trials was poorly reported and is described in Table 62 above. Changes in medication use following manual therapy were not reported in any trial.

Adverse Events and Withdrawal Due To Adverse Events

Adverse events were poorly reported. Three trials do not report on adverse events[119,148,166] and no trial reported on withdrawal due to adverse events. No serious treatment-related adverse events were observed in the two trials reporting adverse events. One trial (N=97) reported that two participants in each treatment group experienced serious adverse events.[116] None were considered treatment related. One patient in the manual therapy group had acute loss of motor and sensory function of the left lumbar segment L5 due to a herniated disk identified prior to treatment. The other three events were acute pancreatitis, symptomatic cholelithiasis, and femoroacetabular impingement syndrome.

One RCT[115] (N=73) found no serious adverse events. They report two adverse events graded as mild. One patient reported sharp pain lower back considered to possibly be related to the CMT that prompted use of pain medication (not specified); it resolved within 48 hours. Another patient reported sharp pain in the right buttock that resolved within 24 hours; it was considered unrelated to the intervention.

Heterogeneity of Treatment Effect

One trial[116] (N=97) found no interaction on post-treatment differences in pain scores between treatment group and strata based on patient characteristics (age, sex, manual occupation) or by pain duration (<7 days, ≥7 days), pain intensity (< 7, ≥7 on 11 point scale), RMD score (< 14, ≥14 on 0-24 scale), dose of analgesics, or setting (emergency department, primary care) (Appendix E). The trial was underpowered to detect effect modification and confidence in these findings is very low.

One RCT[166] (N=95) reported Disability Index scores for subgroups of patients based on duration of the current LBP episode (1-13 days, 14-28 days and >28 days). They found no difference between manual therapy and usual care in mean functional improvement at 1, 2, or 3 weeks post treatment for any of the subgroups on DI scores. No test for interaction was done and the small number of patients in each subgroup precludes evaluation of effect modification (Appendix E).

Manual Therapy Versus Physiotherapy

Description of Included Trials

Two RCTs (N=160) compared manual therapy versus physiotherapy for the treatment of ALBP (Appendix E).[143,146] Patient age (range 36 to 43 years), proportion of females (range 35% to 38%), mean pain at baseline (mean 5.1 to 6.8 on a 0-10 scale), and duration of pain at baseline (≤3 weeks to <1 month) were similar across both trials. Most patients (69%) in one trial had previous LBP episodes.[143] One RCT[143] enrolled patients with and without radiculopathy but did not provide further details and the other trial did not report this.[146] Both trials excluded pregnant women. Medical and psychiatric comorbidities and history of substance use disorder or opioid use were not reported.

One RCT (N=112) compared spinal manipulation with or without mechanical therapy (McKenzie method) versus with local heat, ultrasound and active flexion/extension exercises; both groups received a maximum of five, 45-minute sessions per week for 2 weeks.[143] The second trial (N=48) compared passive mobilization and manipulation versus 15 minutes of microwave diathermy supplemented by isometric abdominal exercises and ergonomic instructions; sessions were conducted three times per week for up to 3 weeks.[146] Choices of manual therapy techniques were left to the discretion of the provider. Concomitant therapies were not reported by any study, but one trial[143] did not permit bed rest, nonopioid analgesics, or acupuncture during the study.

One study was conducted in Australia[146] and one in New Zealand.[143] Both trials were assessed as moderate risk of bias; methodological limitations included unclear randomization and allocation concealment methods and the inability to blind patients and providers as well as some differences between group on baseline characteristics in one trial[143] however the sample size was small (Appendix F).

Detailed Synthesis

Function

Function outcomes were not reported.

Pain

Both RCTs[143,146] reported similar pain scores in the manual therapy versus physiotherapy groups at all timepoints (Appendix E). One trial (N=48) reported VAS back pain scores (0-10 scale) immediately post treatment (mean 2.8 vs. 2.6) and at 3 weeks (mean 0.3 vs. 0.3).[146] The other trial[143] reported back pain and leg pain scores on a 0-3 Likert scale (none to severe) at 3 to 4 days and 10 to 12 days, respectively (mean 1.05 vs. 1.11 and 0.42 vs. 0.38 for back pain; mean 0.36 vs. 0.38 and 0.18 to 0.11 for leg pain). Effect estimates could not be calculated with the data provided.

Overall Efficacy and Return to Work

One RCT (N=72) found manual therapy associated with a similar likelihood of patient's reporting their overall improvement as good or excellent at 10 to 12 days (73.7% versus 70.6%, RR 1.04, 95% CI 0.78 to 1.39) and of being back at work with 2 weeks of treatment (90% vs. 89%, RR 1.01, 95% CI 0.86 to 1.20) compared with physiotherapy.[143] The second trial (N=48) found that patients in the manual therapy group achieved symptom-free status in fewer sessions than those in the physiotherapy group (MD -2.3 sessions, 95% CI -3.45 to -1.15).[146]

Adverse Events

One RCT (N=72) reported that one patient (2.6%) who received manual therapy withdrew due to lack of improvement by the second session; no patient in the physiotherapy group withdrew.[143]

Manual Therapy Versus Education (Back School)

Description of Included Trials

One RCT (N=142) compared manual therapy versus education for the treatment of ALBP (Appendix E).[106] Mean patient age was 35 years, 13% of participants were female and 57% had experienced one or more prior back pain episode. At baseline, median duration of pain was 9 days and median pain intensity was 42 (0-70 scale). Information on psychiatric comorbidities and history of substance use disorder was not provided. Manual therapy consisted of manipulation and/or mobilization (described as rotational manipulation of the trunk and/or mobilization of hypomobile segments) provided in two weekly sessions up to a total of 10 sessions; patients in this group also received postural and lifting education. The education group received one 45-minute "Back School" session. Patients were allowed to use salicylates, acetaminophen or a combined analgesic but specific use of rescue medication was not reported. The trial was conducted in Sweden and assessed as high risk of bias. Limitations include lack of information regarding randomization and/or concealment of treatment allocation, lack of similarity between groups at baseline and unclear reporting of attrition and differential loss to follow-up (Appendix F).

Detailed Synthesis

Function

Function outcomes were not reported.

Pain

One RCT[106] found manual therapy associated with similar pain scores (Pain Index, 0-70 scale) compared with education at 10 days (N=94, mean 22 vs. 20), 3 weeks (N=37, mean 18 vs. 19) and 6 weeks (N=43, mean 21 vs. 22). Both groups improved significantly from baseline. Effect estimates could not be calculated with the information provided.

Other Outcomes

One RCT (N=142) found manual therapy associated with a moderate reduction in the likelihood of requiring sick leave for more than 21 days compared with education (25.0% vs. 44.3%, RR 0.56, 95% CI 0.35 to 0.91).[106] The likelihood of sick leave less than 21 days (51.4%

vs. 42.9%, RR 1.20, 95% CI 0.84 to 1.70) and no sick leave (23.6 vs. 12.9%, RR 1.84, 95% CI 0.88 to 3.84) also favored manual therapy, but the differences between groups were not statistically significant. Recurrence of back pain was moderately increased with manual therapy versus education at 1 year (60% vs. 36%, RR 1.67, 95% CI 1.16 to 2.41), but the average number of recurrences was identical between groups at that time.

Adverse Events

Safety outcomes were not reported.

Manual Therapy Versus Massage

Description of Included Trials

One RCT (N=40) compared manual therapy versus massage for the treatment of ALBP (Appendix E).[109] Average patient age was 30 years and 43% were female. At baseline, duration of pain was <6 weeks (mean not reported) and mean pain severity was 4.6 (0-10 scale). Presence of radiculopathy, history of LBP, medical and psychiatric comorbidities, and history of substance use disorder were not reported. Manual therapy consisted of positional release techniques (commonly known as strain counterstain) in both prone and supine position. Massage consisted of deep transverse friction massage. Both groups received treatment three times per week for 4 weeks (overall treatment time was 30 minutes for both). Concomitant therapies were not reported. The trial was conducted in the Pakistan and assessed as moderate risk of bias. Limitations include unclear reporting of baseline characteristics, lack of blinding of patients and high attrition (Appendix F).

Detailed Synthesis

Function and Pain

One RCT (N=36) found manual therapy associated with similar improvement in function (ODI scores, 0-100 scale) at 2 weeks (MD -0.61, 95% CI -4.26 to 3.04) and 4 weeks (MD -0.70, 95% CI -4.36 to 2.96) and in pain (NRS pain score, 0-10) at 2 weeks (MD 0.45, 95% CI -0.13 to 1.03) and 4 weeks (MD 0.46, 95% CI -0.02 to 0.94).[109] Estimates were imprecise (Appendix E).

Other Outcomes and Adverse Events

Other outcomes and adverse events were not reported.

Manual Therapy (Manipulation) Versus Manual Therapy (Mobilization)

Description of Included Trials

One RCT compared manipulation versus mobilization for the treatment of ALBP (Appendix E).[147] Patient age ranged from 20 to 40 (mean age not reported) and 43% were female. Mean duration of pain was ≤4 weeks at baseline (<2 weeks in 48% and 2 to 4 weeks in 52%). Patients were excluded if they had an episode of LBP in the previous 6 months. Presence of radiculopathy, medical and psychiatric comorbidities and history substance use disorder or opioid use were not reported. This trial provided a single session of spinal manipulation using a high velocity thrust (described as including rotation of the facet and sacroiliac joints) versus

spinal mobilization without a manipulative thrust (described as without rotational forces and leverage required to move facet joints). Concomitant therapies were not reported. The trial was conducted in the United States and was assessed as moderate risk of bias. Limitations included unclear randomization methods and dissimilar groups at baseline.

Detailed Synthesis

Function

One RCT (N=54) found manipulation associated with similar improvement in function (RDQ scores, 0-24 scale) compared with mobilization for the overall population at all timepoints.[147] However, in the subgroup of patients with baseline pain duration of 2 to 4 weeks (n=26), manipulation was associated with a large improvement in RDQ scores at 3 days compared with mobilization (MD -5.2, 95% CI -8.78 to -1.62); there was no effect of treatment seen in the subgroup with pain duration of <2 weeks (n=28, mean 6.2 vs. 6.8). At other timepoints (6 to 12 days), there were no differences seen in RDQ scores between the manipulation and mobilization groups, respectively, within the different pain duration strata, though manipulation tended to be favored in the 2- to 4-week pain duration subgroup: baseline pain ≤2 weeks (range, mean 2.4-4.3 vs. 2-4) and baseline pain 2 to 4 weeks (range, 3.4-4.5 vs. 5-6) (Appendix E). Authors state that in multivariate analysis of variance there was a significant interaction (p=0.040) between treatment group and baseline pain duration.

Pain

Manipulation was associated with a similar likelihood of positive pain relief (much, somewhat, or a little better) compared with mobilization in this trial (N=54, 84.6% vs. 78.6%, RR 1.08, 95% CI 0.84 to 1.39).[147]

Other Outcomes and Adverse Events

Other outcomes and adverse events were not reported.

Acupuncture

Acupuncture Versus Sham or Placebo

Two RCTs compared acupuncture versus sham or placebo for the treatment of ALBP of 6 to 15 days duration (Table 63; Appendix E).[124,130] Mean patient age ranged from 43 to 46 years, 57% to 64% were female, and mean LBP severity at baseline ranged from 6.6 to 7.1 (on a 0-10 scale). In one trial,[130] 57% of patents had previous LBP episodes. One trial excluded patients with radiculopathy[124] and the other trial[130] included a mixed population (61% with radiculopathy). Neither trial reported on psychiatric comorbidities or history of substance use disorder or opioid use.

In both trials, patients received five, 20-minute sessions, and treatment periods ranged from 2 to 4 weeks. One trial[124] allowed 50 mg sodium diclofenac every 8 hours for lumbar pain if needed. In the second trial,[130] all patients received standard conservative care that included education and advice as well as possible pharmacologic treatment (cyclobenzaprine, diclofenac, ibuprofen, and acetaminophen).

One trial[124] was assessed as low risk of bias and the other trial[130] as moderate risk of bias. Limitations in the moderate trial included dissimilar groups at baseline and unacceptable loss to follow-up (Appendix F).

Table 63. Study characteristics of acupuncture versus sham, placebo, or usual care

Study, year	Hasegawa, 2014	Vas, 2012*
N Randomized	80	205†
Mean Age (Years)	46	43
% Female	64%	57%
% Radiculopathy	0%	61%
% Previous Back Pain	NR	Prior 2 years: 57%
Pain Duration Inclusion Criteria	<30 days	<2 days
Mean Pain Duration (days)	15	6
Mean/Median Baseline Pain Severity (0-10)	6.6	7.1
History Substance Abuse Disorder	NR	NR
Psychiatric Comorbidities	NR	NR
Treatment	Yamamoto's New Scalp Acupuncture	Traditional Chinese Acupuncture
Comparator	Sham (pressure at same acupoints, nonpenetrating)	Placebo (nonspecific acupuncture points, nonpenetrating); Sham (nonspecific acupuncture points, penetrating)†
Dosage Frequency	20 minutes, 5 sessions	20 minutes, 5 sessions
Treatment Duration	28 days	2 weeks
Duration Follow-up	28 days	48 weeks
Country	Brazil	Spain
Quality	Low	Moderate

NR = not reported.
* Trial included four randomized groups: (1) True acupuncture, (2) Sham acupuncture, (3) Placebo acupuncture, and (4) Usual care.
† The sham and placebo acupuncture groups are combined.

Detailed Synthesis

Function

One trial found acupuncture associated with a small increase in the likelihood of a successful functional outcome (≥35% improvement on the RDQ) compared with placebo acupuncture (nonspecific acupoints, nonpenetrating) at 3 months (N=102, 98% vs. 80.4%, RR 1.22, 95% CI 1.06 to 1.40).[130] For acupuncture versus sham acupuncture (nonspecific acupoints, penetrating), results at 3 months also favored acupuncture but the effect estimate was below the threshold for a small effect (N=109, RR 1.09, 95% CI 0.99 to 1.20).[130] At all other timepoints, acupuncture was associated with a similar likelihood of a successful functional outcome compared with sham acupuncture (3 weeks: 72.5% vs. 75.0%; 3 months: 98% vs. 89.7%; and 12 months: 86.3% vs. 88.7%) and compared with placebo acupuncture (3 weeks: 72.5% vs. 65.2%; 12 months: 86.3% vs. 83.7%) (Appendix E).[130]

Two RCTs[124,130] assessed function based on RDQ scores (0-24 scale) (Table 64). One trial (N=80) found scalp acupuncture associated with a moderate improvement in function at 3 weeks (MD -4.10, 95% CI -6.49 to -1.71) and 4 weeks (-4.40, 95% CI -6.79 to -2.01) compared with placebo but there was no effect seen at earlier timepoints (3 days and 1 week).[124] Similarly, the second trial found a significant improvement favoring traditional Chinese acupuncture versus

placebo at 3 weeks, but the effect did not persist to later timepoints.[130] There was no significant difference between acupuncture and sham acupuncture at any timepoint in this trial.

Table 64. Function outcomes

Author, Year Symptoms	Comparison	Time	Function, RDQ (0-24) Acupuncture vs. Placebo/Sham MD (95% CI)
Hasegawa, 2014 Scalp acupuncture	Sham acupuncture (pressure at same acupoints, nonpenetrating)	3 days	N=80, -2.10 (-4.31 to 0.11)
		1 week	N=80, -2.10 (-4.50 to 0.30)
		3 weeks	N=80, -4.10 (-6.49 to -1.71)
		4 weeks	N=80, -4.40 (-6.79 to -2.01)
Vas, 2012 Traditional Chinese acupuncture	Placebo acupuncture (pressure at points on the back, nonpenetrating)	Baseline	13.4 (4.6) (n=68) vs. 14.1 (5.2) (n=69)
		3 weeks	RDQ relative change* from baseline (SD) 63.9 (40.4) (n=64) vs. 43.1 (58.9) (n=64), p=0.020
		3 months	RDQ relative change* from baseline (SD); 91.6 (17.1) (n=51) vs. 76.3 (41.5) (n=51), p=0.251
		12 months	RDQ relative change* from baseline (SD); 67.4 (74.0) (n=51) vs. 70.7 (55.4) (n=49), p=0.994
	Sham acupuncture (nonspecific acupoints, penetrating as in true acupuncture)	Baseline	13.4 (4.6) (n=68) vs. 12.6 (6.3) (n=68)
		3 weeks	RDQ relative change* from baseline (SD) 63.9 (40.4) (n=64) vs. 65.0 (40.5) (n=65), p=0.707
		3 months	RDQ relative change* from baseline (SD) 91.6 (17.1) (n=51) vs. 84.7 (28.9) (n=58), p=0.792
		12 months	RDQ relative change* from baseline (SD) 67.4 (74.0) (n=51) vs. 75.6 (37.4) (n=53), p=0.447

CI = confidence interval; RDQ = Roland-Morris Disability Questionnaire; SD = standard deviation.
* Higher score means better function/less disability.

Pain

One RCT (N=80) found scalp acupuncture associated with a moderate improvement in VAS pain scores versus sham at 3 weeks (MD -1.69, 95% CI -2.79 to -0.59) and at 4 weeks (MD -1.40, 95% CI -2.38 to -0.42); there was no effect of acupuncture on pain at earlier timepoints: 3 days (MD -0.50, 95% CI -1.49 to 0.49) and 1 week (MD -0.57, 95% CI -1.62 to 0.48).[124]

Quality of Life

One RCT (N=80) reported on individual domains of the SF-36 questionnaire (0-100 scale) at 4 weeks. Scalp acupuncture was associated with significant improvements in functional capacity, pain, vitality (moderate effects, range of differences 10.8 to 13.1) and limitation in physical aspects scores (large effect, difference 23.0) versus sham; there were no effects for general health status, social aspects, emotional aspects and mental health domains (Appendix E).[124]

Recurrence of LBP and Return to Work

One RCT (N=160) found acupuncture associated with a similar likelihood of recurrence of the initial LBP episode at 3 and 12 months compared with sham acupuncture (5.9% vs. 10.3% and 51% vs. 43.4%, respectively) and with placebo acupuncture (5.9% vs. 5.9% and 51% vs. 51%, respectively) (Appendix E).[130] Similarly, there was no effect of acupuncture on the

proportion of patients not working compared with controls across timepoints (true acupuncture range, 0% to 10.3%; sham acupuncture range, 0% to 10%; and placebo acupuncture range, 2.8% to 11.8%).

Medication Use

One trial (N=80) found scalp acupuncture associated with less anti-inflammatory medication use compared with sham at 1 week (difference -2.3 pills, 95% CI -4.13 to -0.47) and 3 weeks (difference -2.5 pills, 95% CI -4.33 to -0.67) but not at 3 days or 4 weeks (Appendix E).[124] In the second trial (N=193), similar proportions of patients were taking the same, less or no medication (cyclobenzaprine, diclofenac, ibuprofen, acetaminophen) at 3 months in the true acupuncture (98.4%) compared with the sham (95.4%) and placebo (98.4%) groups (Appendix E).[130]

Adverse Events

Both trials (N=285) reported that no serious adverse events occurred and there were no withdrawals due to adverse events in any group.[124,130] One trial (N=205) reported a similar incidence of increased pain in the true (4%), sham (4%), and placebo (3%) acupuncture groups.[130] Possible adverse events to medication, including epigastralgia and nausea, were more common in the placebo acupuncture group (5.8%) versus the true or sham groups (1.5% each), but the difference was not statistically significant.

Acupuncture Versus Usual Care

Three RCTs compared acupuncture to usual care for the treatment of ALBP (Table 65; Appendix E).[111,112,130] Patient mean age ranged from 39 to 44 years and most were female (range 51% to 76%). In two RCTs[112,130] most patients had a history of previous LBP episodes (range 58% to 72%). The baseline duration of LBP varied across the trials and ranged from a mean of 6 days to a median of 25 days; one trial did not report mean duration, but inclusion criteria specified ≤14 days duration.[112] Mean or median pain at baseline ranged from 6.3 to 8.0 (on a 0-10 scale). Two trials[111,112] excluded patients with radiculopathy and one trial[130] included a mixed population (61% with radiculopathy). One trial[111] excluded individuals with psychiatric comorbidities and all three trials excluded pregnant or breastfeeding women. History of substance use disorder or opioid use was not reported.

In two trials,[111,130] the intervention consisted of five to six 20-minute sessions of traditional Chinese acupuncture over 2 weeks, and in the third trial patients received a single, 8- to 9-minute session of Western medical, motion-style acupuncture. In all three trials all patients received usual conservative care that included education and advice and possible pharmacologic treatment (cyclobenzaprine, diclofenac, etodolac, ibuprofen, and acetaminophen) as well as sick leave if needed in one trial.[112]

Two RCTs[112,130] were conducted in Europe and one[111] in Turkey. All trials were assessed as moderate risk of bias. Limitations across trials included inability to blind patients and providers, between group differences in baseline characteristics, and high attrition (Table 65; Appendix F).

Table 65. Study characteristics of acupuncture versus usual care

Study, Year	Vas, 2012*	Skonnord, 2020	Buyuksireci, 2022
N Randomized	138	185	54
Mean Age (Years)	42	39	44
% Female	64%	51%	76%
% Radiculopathy	61%	0% (Excluded)	0% (Excluded)

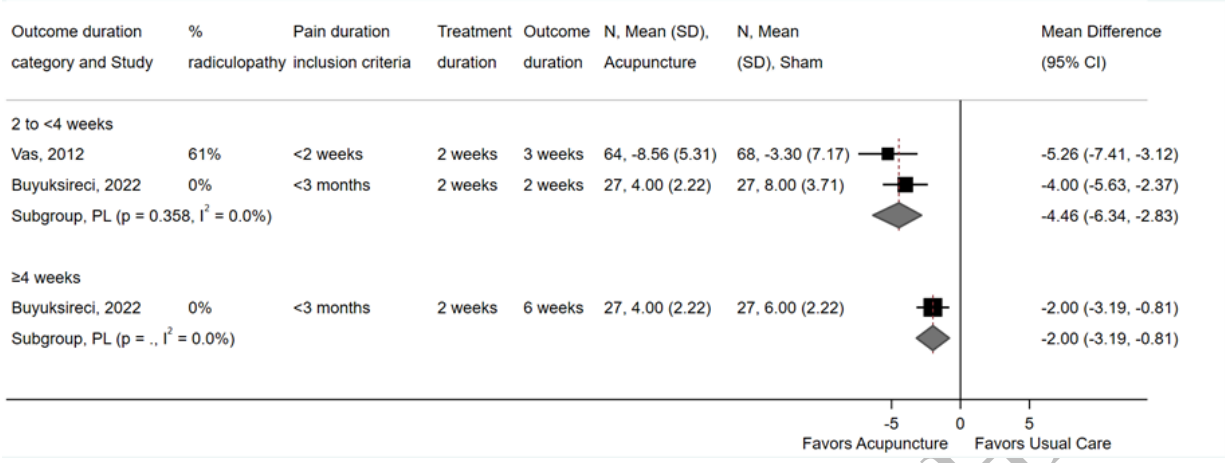
Study, Year	Vas, 2012*	Skonnord, 2020	Buyuksireci, 2022
% Previous Back Pain	Prior 2 years: 58%	72%	NR
Pain Duration Inclusion Criteria	<2 days	≤14 days	<3 months
Mean Pain Duration (days)	6†	NR	median 25
Mean/Median Baseline Pain Severity (0-10)	7.1	6.3	median 8
History Substance Abuse Disorder	NR	NR	NR
Psychiatric Comorbidities	NR	NR	0% (Excluded)
Treatment	Traditional Chinese Acupuncture: five 20-minute sessions	Western medical, motion-style acupuncture: one 8–9-minute session	Traditional Chinese Acupuncture: 20-minute sessions, 3x/week
Comparator	Usual Care: posture recommendations, analgesics, nonsteroidal anti-inflammatory drugs, myorelaxant drugs	Usual Care: advice about activity, prescription of analgesics (acetaminophen and/or ibuprofen) and sick leave if needed	Usual Care: advice, etodolac at 400 mg 2x/day)
Treatment Duration	2 weeks	Once	2 weeks
Duration Follow-up	48 weeks	1 year	6 weeks
Country	Spain	Norway	Turkey
Quality	Moderate	Moderate	Moderate

NR = not reported.
* Trial included four randomized groups: (1) True acupuncture, (2) Sham acupuncture, (3) Placebo acupuncture, and (4) Usual care.
† Whole sample. Includes more than two interventions, but data not reported by arm.

Function

Two RCTs[111,130] (N=186) found traditional Chinese acupuncture (5-6 sessions over 2 weeks) associated with a moderate improvement in RDQ scores (0-24 scale) compared with usual care at 2 to <4 weeks (pooled MD -4.46, 95% CI -6.34 to -2.83, I^2=0%) and with a small improvement at 6 weeks in one trial[111] (N=54, MD -2.00, 95% CI -3.19 to -0.81) (Figure 35). The third trial, that used a single 8- to 9-minute session of Western, motion-style acupuncture, found no effect on function (RDQ scores) compared with usual care at 2 weeks (N=141, MD 0, 95% CI -1.05 to 1.05) and 4 weeks (N=135, 0.50, 95% CI -0.55 to 1.55).[112]

Figure 35. Traditional Chinese acupuncture versus usual care: RDQ scores

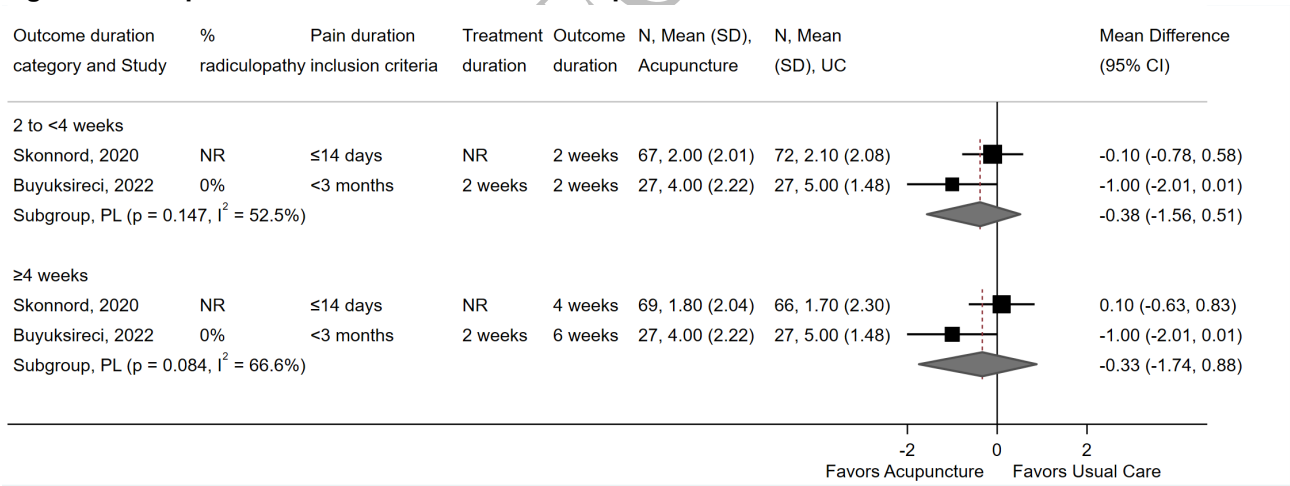


CI = confidence interval; NR = not reported; SD = standard deviation; UC = usual care.

Pain

One trial (N=57) found traditional Chinese acupuncture (6 sessions over 2 weeks) associated with a small improvement in NRS pain scores (0-10 scale) compared with usual care at 2 and 6 weeks (MD -1.00, 95% CI -2.01 to 0.01, at both timepoints).[111,130] One trial that used a single 8- to 9-minute session of Western, motion-style acupuncture, found no effect compared with usual care at 2 weeks (N=139, MD -0.10, 95% CI -0.78 to 0.58) and 4 weeks (N=135, 0.10, 95% CI -0.63 to 0.83) (Figure 36).[112]

Figure 36. Acupuncture versus usual care: NRS pain scores



CI = confidence interval, NR = not reported, NRS = Numeric Rating Scale; PL = profile likelihood; SD = standard deviation; UC = usual care.

Other Outcomes

One RCT[112] found a single session of Western, motion style acupuncture associated with a large increase in the likelihood of patients rating their global improvement as better or much better immediately post-treatment (N=148, 58.7% vs. 15.1%, RR 3.89, 95% CI 2.19 to 6.93) and with a small increase at 1 day (N=147, 67.9% vs. 51.3%, RR 1.24, 95% CI 1.03 to 1.49) versus usual care; over days 6 to 14, there was no longer an effect seen (range, 82.9% to 89.6% vs. 73.1% vs. 87.5%, respectively). There was no effect of acupuncture on the following outcomes

at any time over 52 weeks in this trial (Appendix E): quality of life measured by the EQ-5D-3L, recurrent LBP episodes, use of medication (opioid and nonopioid), and sick leave. One trial[130] found traditional Chinese acupuncture associated with a lower frequency (large effect) of occupational incapacity immediately post-treatment (6.4% vs. 30.9%, RR 0.21, 95% CI 0.06 to 0.66) as well as a large decrease in medication consumption at 3 weeks (1.6% vs. 11.8%, RR 0.13, 95% CI 0.02 to 0.97) versus usual care. There was no difference between groups at other timepoints or for other outcomes reported (Appendix E).

Serious Adverse Events and Withdrawal due to Adverse Events

Two RCTs[112,130] reported no serious adverse events. In one trial, one (3.7%) and three (11.1%) patients in the acupuncture and usual care groups, respectively, discontinued the trial due to increased pain. Another trial noted no withdrawals due to adverse events.[130]

Any Adverse Event

Two trials found acupuncture associated with a lower risk of any adverse events compared with usual care but the difference between groups was not statistically significant. In one trial[112] (N=167) events included pain, GI symptoms, tiredness, headache, and dyspnea (13.6% vs. 18.6%, RR 0.73, 95% CI 0.36 to 1.48), and in the other trial[130] (N=138) adverse events included epigastralgia and nausea, primarily (1.5% vs. 8.6%, RR 0.02 to 1.39). In one trial, 9.9% vs. 16.3% of patients reported GI symptoms specifically (RR 0.61, 95% CI 0.27 to 1.37).[112]

Acupuncture Versus Massage Therapy

One RCT (N=96) compared acupuncture versus massage therapy for the treatment of ALBP due to lumber strain of a mean 3.5 days duration (Appendix E).[136] Mean patient age was 44 years and 46% were female with baseline back pain severity of 6.8 (0-10 scale). Pregnant or breastfeeding women and patients with psychiatric comorbidities were excluded. Presence of radiculopathy and history substance use disorder and opioid use was not reported. Both acupuncture and massage were done along the same meridian lines. Acupuncture needles were left in for 25 minutes and occasionally manually twisted. Patients in the massage group received 30-minute sessions and were prescribed 300 mg aspirin once daily and 300 mg ibuprofen twice daily. Number and frequency of sessions was not reported. Treatment duration was 10 days. The trial[136] was conducted in China and assessed as moderate risk of bias. Methodological limitations included lack of blinding and unclear allocation concealment (Appendix F).

Detailed Synthesis

Function and Pain

One RCT (N=96) found acupuncture associated with a small improvement in function (RDQ scores, 0-24 scale, MD -1.23, 95% CI -1.44 to -1.02) and pain (VAS scores, 0-10 scale, MD -0.55, 95% CI -0.68 to -0.42) compared with massage at 10 days.[136]

Global Efficacy

More patients who received acupuncture were considered improved or cured compared with massage, but the difference was below the threshold for a small effect (N=96, 95.8% vs. 83.3%, RR 1.15, 95% CI 1.00 to 1.32).[136]

Adverse Events

Adverse events were not reported.

Physical Modalities Versus Sham or Placebo

Transcutaneous Electrical Nerve Stimulation (TENS)

Description of Included Studies

One RCT (N=63) compared active transcutaneous electrical nerve stimulation (TENS) with sham TENS (Appendix E).[65] All patients received emergency transportation during the intervention and were experiencing their first episode of ALBP (<2 hours duration). Patients in both groups were connected to same TENS device by two electrodes; the active TENS group received stimulation (100 Hz, pulse width 200 µs, voltage 2 milliampere) while the sham group did not. Treatment duration was 26 minutes. Mean patient age was 48 years, 45% were female, 100% were White, and mean baseline pain severity was 7.8 (0-10 scale). Patients with radiculopathy were excluded. No IV drugs were allowed and patients who took analgesics within 48 hours of study entry were excluded. Substance abuse disorder, opioid use and medical and psychiatric comorbidities were not reported. This trial was conducted in Austria and assessed as low risk of bias (Appendix F).

Detailed Synthesis

Pain and Anxiety

One trial (N=63) found TENS associated with a large improvement in VAS pain scores (0-10 scale, MD -2.80, 95% CI -3.28 to -2.32) and a moderate improvement in VAS anxiety scores (0-10, MD -1.50, 95% CI -2.02 to -0.98) compared with sham TENS upon arrival to the ED.[65]

Adverse Events

Adverse events were not reported.

Heat Therapy

Description of Included Studies

Two RCTs (N=151) compared heat therapy with placebo interventions for the treatment of ALBP (Appendix E).[63,114] Mean patient age ranged from 29 to 37 years and 33% to 51% were female. One trial[63] reported primarily moderate pain at baseline (mean 7.4, 0-10 scale). In one trial, 41% of patients were White and 43% were Asian.[114] Patients with radiculopathy were excluded from both trials. Psychiatric comorbidities and prior opioid use were not reported in either trial, but one trial noted that 16% of patients had previous medical conditions (not specified);[114] this trial also excluded patients with substance abuse disorder and pregnant women. In one trial (N=100), all patients received emergency transportation during the intervention and were experiencing their first episode of LBP (<6 hours duration).[63] Patients in the intervention group received active heat via an electric heating blanket while patients in the placebo group received only passive warming (nonheated blanket) during a mean of 26 minutes. No IV drugs were allowed and patients who took analgesics within 48 hours of study entry were excluded. The second trial (N=51) compared a back/hip heat wrap worn for 8 hours with an oral placebo.

The duration of back pain in this trial was <4 weeks (mean not reported). Patients were not permitted to lie down during the treatment period and were given acetaminophen 500 mg tablets as rescue medication. One trial was conducted in Austria[63] and one in the United States.[114]

Risk of bias was assessed as low in one trial[63] and moderate in the other trial.[114] Methodological limitations in the moderate risk of bias trial included unclear randomization and allocation concealment methods and lack of patients blinding and unclear attrition (though likely low given 8-hour treatment period) (Appendix F).

Detailed Synthesis

Pain

Both trials (N=141) found heat therapy associated with large improvements in pain across different measures compared with placebo interventions (Table 66).[63,114] In one trial (N=51), more patients who received a heat wrap reported that they experienced perceptible and meaningful pain relief compared with an oral placebo after 8 hours, but the difference was not statistically significant (53.8% vs. 28.0%, RR 1.92, 95% CI 0.93 to 3.96).[114]

Global Efficacy and Psychological Distress

Compared to oral placebo in one trial,[114] heat wrap therapy was associated with substantially better Global Assessment of Treatment scores (0-5 scale) and a large increase in the likelihood of patients rating their overall improvement as good, very good, or excellent after 8 hours (84% vs. 16%); however, estimates were imprecise (Table 66). One trial (N=90) found active heating (electric blanket) associated with a large improvement in VAS anxiety scores (0-10 scale) compared to placebo (passive heating via blanket) during emergency transport lasting a mean of 26 minutes (MD -3.45, 95% CI -4.12 to -2.78).[63]

Table 66. Heat therapy versus placebo: pain and global efficacy outcomes

Outcome	Author, year Treatments	Follow-up	Heat Therapy Mean (SD) or % (n/N)	Placebo Mean (SD) or % (n/N)	Effect Size (95% CI)
Pain					
VAS Pain Scores (0-10; lower = better)	**Nuhr, 2004** Electric blanket vs. wrap placebo	Mean 26 minutes*	4.19 (1.89) (n=47)	7.41 (1.20) (n=43)	MD -3.22 (-3.88 to -2.56)
Pain Relief Score (0-5 categorical scale; higher = better)	**Stark, 2014** Heat wrap vs. oral placebo	8 hours	3.20 (SD NR) (n=26)	1.70 (SD NR) (n=25)	MD 1.5, P<0.001
Total Pain Relief (sum of pain relief scores from 0-8 hours; 0-40 scale)			22.00 (SD NR) (n=26)	11.50 (SD NR) (n=25)	MD 10.50, P<0.001
Patient Perception of Improvement					
Global Assessment of Treatment (0-5; higher = better)	**Stark, 2014** Heat wrap vs. oral placebo	8 hours	3.40 (SD NR) (n=26)	1.70 (SD NR) (n=25)	MD 1.7, P<0.001
Global Assessment Good, Very Good, or Excellent			84% (22/26)	16% (4/25)	RR 5.29 (2.12 to 13.18)

CI = confidence interval; IP = immediately post-treatment; MD = mean difference; NC = not calculable; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale.

* Patients received intervention during emergency transport: duration of transport was an average of 26 minutes.

Adverse Events

Serious adverse events were not reported. In one trial (N=51), two patients (7.7%) who received heat wrap experienced mild headaches and one simultaneously experienced asthma (all considered unrelated to treatment), while no patients who received oral placebo experienced any events.[114]

Massage

Massage Versus Sham

One RCT (N=44) compared massage (compression at myofascial trigger points) to sham (compression at nontrigger points) for the treatment of ALBP episodes of 8.4 days duration (Appendix E).[61] Mean patient age was 38 years, 54% were female, and mean pain intensity at baseline was 5.1 (0-10 scale). Patients with radiculopathy or psychiatric comorbidities and women who were pregnant were excluded. History of prior LBP episodes, substance use disorder, or opioid use was not reported. Massage was performed for 15 minutes three times weekly over 2 weeks. Use of concomitant therapies was not permitted in either group. The trial was conducted in Japan and was assessed as moderate risk of bias (Appendix F).

Detailed Synthesis

Function and Pain

One RCT (N=39) found no effect of massage on reduction in RDQ scores (0-24 scale) at 1 week (MD in change scores -2.7, 95% CI -7.86 to 2.46) and 1 month (MD in change score -2.9, 95% CI -8.06 to 2.26) compared with sham massage.[61] Massage was associated with a large improvement in VAS pain with movement at 1 week (MD in change scores -4.10, 95% CI -7.38 to -0.83) and 1 month (MD in change scores -3.86, 95% CI -7.14 to -0.59) compared with sham massage, but there was no effect of massage immediately post-treatment (MD in change scores -2.22, 95% CI -5.50 to 1.06). Results between groups were almost identical for VAS pain at rest (Appendix E). All estimates were very imprecise.

Safety

One RCT (N=44) reported no adverse events or withdrawals due to adverse events in either group.[61]

Advice to Stay Active

Advice to Stay Active Versus Bed Rest

Three RCTs (N=487) compared advice to stay active with bed rest for the treatment of ALBP (Table 67; Appendix E).[129,135,139] Patient mean age ranged from 38 to 44 years, 32% to 53% were female, and most had a history of previous back pain episodes (47% to 69%). One trial restricted inclusion to patients with radiculopathy,[139] one trial excluded patients with radiculopathy,[129] and one trial included a mixed population (31% with radiculopathy). Mean LBP duration at baseline was ≤3 days in two trials[129,135] and 3.8 weeks in the third trial.[139] Mean baseline pain severity (0 to 10 scale) was lower in one trial[135] (3.8) compared with the other two trials[129,139] (6.3).

Psychiatric comorbidities, history of substance use disorder or opioid use, and inclusion or exclusion of pregnant or breastfeeding women were not reported.

Advice to stay active generally involved maintaining daily activities, adjusting intensity, frequency and duration of activities to pain levels, and limiting rest. The duration of bed rest ranged from 48 hours to 7 days. In one trial, patients were further stratified into bed rest at home or in the hospital for 7 days.[139] Concomitant treatments provided to all patients included prescription of 1000 mg acetaminophen (3 times per day) and 8 mg thiocolchicoside (2 times per day) in one trial,[129] and ibuprofen or co-proxamol in another trial.[135] The third trial[139] did not report on concomitant treatments.

All three RCTs[129,135,139] were conducted in Europe and were assessed as moderate risk of bias. Limitations across trials included difference between groups at baseline, inability to blind patients and investigators, and high attrition (Appendix F).

Table 67. Study characteristics of advice to stay active versus bed rest

Study, Year	Rozenberg, 2002	Wilkinson, 1995	Hofstee, 2002
N Randomized	278	42	167
Mean Age (Years)	44	38	40
% Female	53%	40%	32%
% Radiculopathy	0% (Excluded)	31%	100%
% Previous Back Pain	47%	Episodes: 0, 26%; 1, 38%; >1, 31%	Back pain: 68% Sciatica: 30%
Pain Duration Inclusion Criteria	≤72 hours	<1 week	<1 month
Mean Pain Duration	Days: 0, 12%; 1, 54% 2, 31%; 3, 2.7%	3 days	3.8 weeks*
Mean Baseline Pain Severity (0-10)	6.3	3.8	6.3
History Substance Abuse Disorder	NR	NR	NR
Psychiatric Comorbidities	NR	NR	NR
Treatment	Advice to stay active: continue ADLs as pain allows (bed rest <12 hours/day)	Advice to stay active: remain mobile, no daytime rest	Advice to stay active: continue ADLs as pain allows
Comparator	Bed Rest: ≥16 hours/day and no ADLs for 4 days	Bed Rest: 48 hours (no further information)	Bed Rest: only get up to use bathroom or shower for 7 days; in hospital 49% or at home 51%
Duration Follow-up	3 months	1 month	6 months
Country	France	UK	Netherlands
Quality	Moderate	Moderate	Moderate

ADL = activity of daily living; NR = not reported.
* Whole sample. Includes more than two interventions, but data is not reported by arm.

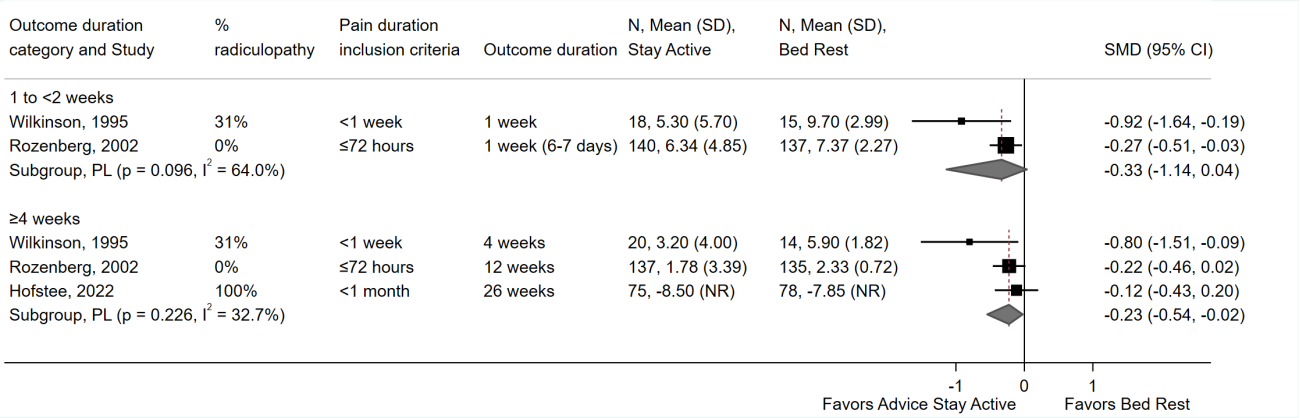
Detailed Synthesis

Function

Three RCTs[129,135,139] reported improvement in function using RDQ scores (0-24 scale)[129,135] and Quebec Disability scores[139] (Figure 37). Advice to stay active was associated with a small improvement in function scores at 1 week (2 RCTs, N=310, pooled SMD -0.33, 95% CI -1.14 to 0.04, I^2=64.0%) and 4 weeks (3 RCTs, N=459, pooled SMD -0.23, 95% CI -0.54 to -0.02, I^2=32.7%). At both timepoints, the two trials[129,135] in patients with a mean pain duration ≤3 days

favored advice to stay active but the magnitude of effect differed: the trial that included patients with (31%) and without (79%) radiculopathy showed a large effect (SMD -0.80 to -0.92)[135] and the trial that excluded patients with radiculopathy showed a small effect (SMD -0.22 to -0.33).[129] The latter trial also showed a small effect favoring advice to stay active at 12 weeks (N=272, SMD -0.22, 95% CI -0.46, to 0.02). The third trial, in patients with radiculopathy with a longer duration of pain (mean 3.8 weeks), found no effect of advice to stay active at any timepoint although the estimate favored advice to stay active at 4 weeks: 4 weeks (N=162, SMD -0.26, 95% CI -0.56 to 0.05) and 8 and 26 weeks (N=161 and 153, SMD -0.12, 95% CI -0.43 to 0.20).

Figure 37. Advice to stay active versus bed rest: function scores



CI = confidence interval; NR = not reported; PL = profile likelihood; SD = standard deviation; SMD = standardized mean difference.

Pain

Two RCTs[129,139] found advice to stay active associated with similar improvement in VAS back and radicular pain scores compared with bed rest across all timepoints. While advice to stay active tended to provide greater improvement in back pain in one trial,[129] all estimates were below the threshold for a small effect (Table 68).

Table 68. Advice to stay active versus bed rest: VAS pain scores

Author, Year	Outcome	Time	Advice to Stay Active vs. Bed Rest
Rozenberg, 2002	VAS back pain (0-10 scale)*	6-7 days	N=277, adjusted MD -0.41 (90% CI -0.81 to -0.04)
		1 month	N=279, adjusted MD -0.35 (97.5% CI, -0.26 to 0.05)
		3 months	N=272, adjusted MD -0.33 (97.5% CI -0.26 to 0.05)
Hofstee, 2002	VAS radicular pain (0-10 scale)	1 month	N=165, MD in change scores 0.25 (95% CI -0.64 to 1.14)
		2 months	N=161, MD in change scores 0.09 (95%CI -0.87 to 1.04)
		6 months	N=153, MD in change scores -0.27 (95% CI -1.02 to 0.48)

CI = confidence interval; MD = mean difference; VAS = visual analog scale.
* Adjusted for initial VAS scores and back pain history.

Other Outcomes

The trial that excluded patients with radiculopathy (N=278), found advice to stay active associated with a moderately lower likelihood of being on sick leave at 3 months compared with bed rest (N=171, 54.0% vs. 87.5%, RR 0.66, 95% CI 0.54 to 0.81).[129] All other outcomes in this trial were similar between the treatment groups, respectively: global efficacy (overall improved) at 1 week (85% vs. 78.8%), NSAID use at 1 week (5.0% vs. 1.5%), and recurrent back pain episode at 3 months (25% vs. 26.1%). The trial that was restricted to patients with radiculopathy

(and had a longer duration of baseline pain) found advice to stay active and bed rest associated with a similar likelihood of treatment failure (defined as intolerable pain and request for surgery for 2 months) at 1 month (7% vs. 6%), 2 months (12% vs. 19%), and 6 months (17% vs. 25%).[139] The trial with a mixed population with and without radiculopathy reported similar rates of time off work between the advice to stay active and bed rest groups (mean 11 vs. 10 days) (Appendix E).

Adverse Events

One RCT (N=42) reported no adverse events related to advice to stay active.[135] A second RCT (N=167) reported new sciatica episodes in 10 patients (unclear to which group they were randomized) and one patient each (1.2%) developed cauda equina syndrome and a pulmonary embolism during red rest at home.[139]

Education

Education Versus Attention Control and Versus Usual Care

Description of Included Studies

Three RCTs compared LBP education to usual care or attention control for treatment of ALBP (Table 69; Appendix E).[64,120,150] Mean patient age ranged from 39 to 45 years, 35% to 51% of patients were female, and most (73% to 81%) had experienced previous back pain episodes. Two trials[120,150] included a mixed population with and without radiculopathy (25% to 51% with radiculopathy) and the third[64] did not report this information. Mean baseline pain duration was 13 days in one trial,[150] <6 weeks in 87% of patients one trial,[120] and in the third trial patients had been off work for ≥3 days due to back pain.[64] Mean pain severity at baseline was 6.2 to 7.4 (0 to 10 scale) in two trials.[120,150] Psychiatric comorbidities,[150] history of substance abuse disorder[120] and pregnancy[64] were exclusion criteria in one trial each. All three trials reported up to 1-year follow-up.

Two trials evaluated the use of education leaflets[64] or booklets[120] compared with usual care in a primary care setting. One of these trials included two education groups: an education booklet plus 20-minutes with a nurse (provided the education, helped set goals and gave feedback and encouragement) and the education booklet alone.[120] Usual care was not described in these trials. The third trial evaluated an intensive LBP education session that included detailed explanations about the biopsychosocial nature of pain in the format of diagrams, metaphors, and stories tailored to individual patients.[150] Patients were also instructed on behavioral techniques, such as pacing, and advised to stay active and avoid bed rest. This group was compared with a placebo education group (i.e., attention control) in which the clinician listened, showed interest and attention, but did not provide education or advice. All patients received two 1-hour in-person sessions. Concomitant treatment was not reported but 51% of patients were using medication at baseline in this trial.

One trial was assessed as low risk of bias[150] and two as moderate risk of bias.[64,120] Methodological limitations in the moderate risk of bias trials included difference between groups at baseline, inability to blind patients and higher than acceptable attrition (Appendix F).

Table 69. Study characteristics of education versus placebo or usual care

Study, Year	Traeger, 2019	Cherkin, 1996	Roberts, 2002
N Randomized	202	299	64
Mean Age (Years)	45	43	39
% Female	51%	48%	35%
% Radiculopathy	51%	25%*	NR
% Previous Back Pain	81%	73%	73%†
Pain Duration Inclusion Criteria	≤3 months	<6 weeks	NR‡
Mean Pain Duration (days)	13	<6 weeks (87%)	NR
Mean/Median Baseline Pain Severity (0-10)	6.2	7.4	NR‡
History substance abuse disorder	NR	0% (Excluded)	NR
Psychiatric Comorbidities	Excluded§	NR	NR
Treatment	Education (intensive, "Explain Pain" book adapted)	Education (2 arms: nurse provided + booklet and booklet only)	Education (general practitioner-supported leaflets)
Comparator	Placebo/attention control	Usual care	Usual care
Dosage Frequency	Two 1-hour sessions	Education: Once (additional phone call after 3 days)	NR
Treatment duration	Twice	Once	Once
Duration follow-up	1 year	1 year	1 year
Country	Australia	U.S.	UK
Quality	Low	Moderate	Moderate

NR = not reported.
* Proportion of patients with a sciatica score ≥2 on a scale of 0-4.
† 89% of education patients and 57% of usual care patients had previous episodes of low back pain.
‡ Pain bad enough to require at least 3 days off work or equivalent.
§ Depressive symptoms: mean 4.6 on a 0-42 scale (higher = more depressive symptoms).

Detailed Synthesis

Function

Two trials (N=349) found education provided via leaflet or booklet (with or without the addition of a nurse consultation in the latter) associated with similar improvement in function scores (Aberdeen LBP Scale or modified 23-item Roland-Morris Disability Questionnaire [RDQ-23]) at all timepoints (Table 70).[64,120] Pain scores were similar between the nurse-led and booklet only education groups in one trial.[120] The trial (N=194) that compared an intensive education program with an attention control found education associated with a small improvement in RDQ scores (0-24) at 1 week (MD -1.6, 95% CI -3.1 to -0.1) and 3 months MD -1.7 (95% CI -3.2 to -0.2) but there was no effect seen at 6 or 12 months (Table 70).[150]

Pain

Two trials (N=488) found no association of patient education (provided via booklets without or without nurse consultation[120] or as an intensive education program[150]) compared with usual care or an attention control on pain improvement (Bothersomeness of Pain score, NRS pain scores) at any time measured (Table 70).[120,150] Pain scores were similar between the nurse-led and booklet only education groups in one trial.[120]

Table 70. Patient education versus usual care or attention control: function and pain scores

Outcome	Author, Year	Follow-up	Education* Mean (SD) or % (n/N)	Control* Mean (SD) or % (n/N)	Effect Size (95% CI)
Function					
Aberdeen LBP Scale (0-100; lower = better)	Roberts, 2002	2 days	42.70 (11.90) (n=35)	42.60 (13.60) (n=28)	MD 0.10 (-6.50 to 6.30)
		2 weeks	37.70 (14.80) (n=32)	35.60 (15.90) (n=26)	MD 2.00 (-6.10 to 10.10)
		3 months	14.60 (16.10) (n=32)	14.40 (17.60) (n=25)	MD 0.20 (-9.20 to 9.60)
		6 months	14.70 (16.10) (n=26)	8.60 (10.10) (n=23)	MD 6.10 (-1.80 to 13.90)
		12 months	11.00 (14.20) (n=25)	8.10 (9.60) (n=23)	MD 2.90 (-4.20 to 10.00)
Modified RDQ (0-23; lower = better)	Cherkin, 1996 Booklet	1 week	change score (SD NR): -5.4 (n=100)	change score (SD NR): -5.3 (n=93)	MD -0.1, p>0.05
	Cherkin, 1996 Nurse	1 week	change score: -5.2 (NR) (n=93)	change score: -5.3 (NR) (n=93)	MD 0.1, p>0.05
RDQ (0-24; lower = better)	Traeger, 2019	1 week	5.6 (5.2) (n=98)	7.1 (5.8) (n=96)	MD -1.6 (-3.1 to -0.1)
		3 months	3.5 (4.6) (n=97)	4.9 (6.0) (n=97)	MD -1.7 (-3.2 to -0.2)
		6 months	3.8 (5.2) (n=96)	4.3 (5.2) (n=96)	MD -0.8 (-2.4 to 0.7)
		12 months	3.0 (4.7) (n=94)	3.8 (5.1) (n=89)	MD -0.8 (-2.4 to 0.7)
Pain					
Bothersomeness of Pain Score (0-10; lower = better)	Cherkin, 1996 Booklet	1 week	change score (SD NR): -3.30 (n=100)	change score (SD NR): -3.60 (n=93)	MD 0.3, p>0.05
	Cherkin, 1996 Nurse	1 week	change score (SD NR): -3.50 (n=93)	change score (SD NR): -3.60 (n=93)	MD 0.1, p>0.05
Pain Intensity in Last Week, NRS (0-10; lower = better)	Traeger, 2019	1 week	3.2 (2.4) (n=101)	3.1 (2.2) ((n=101)	MD 0.10 (-0.50 to 0.80)
		3 months	2.1 (2.4) (n=101)	2.4 (2.2) (n=101)	MD -0.30 (-1.00 to 0.30)
		6 months	2.3 (2.6) (n=101)	2.5 (2.3) (n=101)	MD -0.20 (-0.80 to 0.50)
		12 months	1.8 (2.2) (n=101)	2.5 (2.4) (n=101)	MD -0.60 (-1.30 to 0.10)

CI = confidence interval; DASS = Depression, Anxiety, and Stress Scale; IP = immediately post-treatment; LBP = low back pain; MD = mean difference; NR = not reported; NRS = Numeric Rating Scale; OR = odds ratio; RDQ = Roland-Morris Disability Questionnaire.

* Roberts, 2002: Education Leaflet vs. Usual Care.

Cherkin, 1996: Education Booklet vs. Nurse Education vs. Usual Care.

Traeger, 2019: LBP Education vs. Placebo (no education).

Psychological Distress

One trial (N=194) reported similar scores on the Depression, Anxiety, Stress Scale (DASS) depression severity scale (0-42 scale, lower score means fewer symptoms) for the intensive patient education and attention control groups at 1 week (MD -0.70, 95% CI -1.80 to 0.50) or 3 months (MD -0.50, 95% CI -1.70 to 0.60).[150]

Days of Work Lost

There was no effect of nurse-led education or booklet only education compared with usual care on mean number of work-loss days at 7 weeks (range, 1.4 to 3.0 vs. 0.90) or 6 to 12 months (range, 0.20 to 0.30 vs. 0.50) or in the proportion of patients with any work-loss days at 7 weeks (range, 24%-36% vs. 29%) or 6 to 12 months (range, 6%-7% vs. 9%) in one trial (N=286).[120] Authors do not provide enough information to calculate effect estimates but report p-values >0.05.

LBP Recurrence

One trial (N=178) found intensive education associated with a moderate decrease in the likelihood of LBP recurrence compared to attention control by 12-month follow-up (28.6% vs. 47.1%, RR 0.61, 95% CI 0.41 to 0.90).[150]

Adverse Events

No trials reported any adverse events at any timepoint.

Self-Management

Standardized Self-Management (Back School) Versus Placebo or Usual Care

Description of Included Studies

Two trials (N=203) compared structured self-management in the form of Back School versus usual care[165] or versus a placebo intervention [106] for the treatment of ALBP (Appendix E).

One trial (N=145) compared four, 45-minute group sessions of Back School (provision of a physical therapy manual, lifting and posture education, home exercises, encouragement of physical activity) versus a mean of five sessions of placebo intervention (low intensity heating considered to have no meaningful clinical effect) over 2 weeks.[106] Additionally, patients attending the Back School were offered one session of supervised pool exercises. Mean patient age was 35 years, 13% were female, 57% had a history of previous LBP episodes, and 88% had LBP <4 weeks (range 1 to 8 weeks). Presence of radiculopathy, inclusion of pregnant women history of substance abuse or opioid use, and concomitant treatment was not reported. The trial was conducted in Sweden and was sponsored by the government. The trial was assessed as high risk of bias because randomization allocation concealment methods, blinding, and attrition, were not reported (Appendix F).

A second trial (N=56) compared a Back School program with usual care (advice and analgesics as needed).[165] Patients in the Back School program were given information about the function and anatomy of the spine and individualized advice and written instructions about positioning, posture and maintenance of activity levels and returned after a week for more

information and training. Patients could contact the physiotherapist for more information on a weekly basis for a total of 7 weeks. Mean patients age was 38 years, 57% were female and 68% had a prior history of back pain. Mean pain duration was 7 to 8 days at study entry. Patients with medical or psychiatric comorbidities were excluded. Presence of radiculopathy, substance abuse or opioid use was not reported. At baseline, 29% versus 50% of patents in the Back School and usual care groups, respectively were using analgesics. The trial was conducted in Sweden; no sponsor was reported. This trial was assessed as high risk of bias due to unclear randomization and allocation concealment methods, and inability to blind patients and providers (Appendix F).

Back School Versus Placebo

Pain

One trial (N=145) found Back School associated with a similar improvement in Pain Index scores (0-70 scale) at 10 days (mean change from baseline -15.10 vs. -12.52) and 3 weeks (mean change from baseline -15.21 vs. -17.53) compared with placebo intervention.[106] Effects estimates could not be calculated with the data provided.

Return to Work, LBP Recurrence, and Adverse Events

Compared with placebo, fewer patients who received Back School experienced a recurrence of LBP over 1-year follow-up, however the difference was not statistically significant (54% vs. 71%; RR 0.85, 95% CI 0.67 to 1.07).[106] The number of patients who did not miss work because of recurrence (51% vs. 46%) and the median number of sick days taken (7.5 vs. 8 days) due to a new LBP episode was similar between groups, respectively.

Adverse events were not reported.

Back School Versus Treatment as Usual

Pain

In one trial (N=56), there was no effect of Back School compared with usual care on the proportion of patients self-rated as pain-free at 1 week (21% vs. 16%, RR 1.33, 95% CI 0.43 to 4.09), 4 weeks (75% vs. 66%, RR 1.14, 95% CI 0.81 to 1.61), and 7 weeks (83% vs. 81%, RR 1.03, 95% CI 0.80 to 1.31).[165]

LBP Recurrence and Treatment Satisfaction

There was no effect of the Back School on recurrence of LBP necessitating sick leave (16% vs. 31%, RR 0.53, 95% CI 0.19 to 1.50) or on the risk of continued disability requiring further treatment (17% vs. 13%, RR 1.33, 95% CI 0.37 to 4.8) through 1 year compared with usual care.[165] Although none of the differences were statistically significant, there was a slight tendency for greater improvement in patients who received Back School and Back School was associated with substantially greater patient satisfaction at 1 year versus usual care (61% vs. 29%, RR 2.1, 95% CI 1.1 to 3.98).[165]

Adverse Events

Serious adverse events and withdrawals due to adverse events were not reported. One case (3%) of a herniated lumbar disc was reported in the usual care group.[165]

Minimal Psychological Intervention

Minimal Intervention Strategy Versus Usual Care

Description of Included Studies

One trial (N=314) compared a minimal intervention strategy with usual care for the treatment of ALBP (Appendix E).[81] This was a cluster randomized trial of 41 general practitioners; after the general practitioners were randomized, they referred consecutive patients to participate in the trial. Mean patient age was 43 years, 47% were female and 14% showed signs of radiculopathy. Median duration of current back pain episode was 13 days. All patient had a history of previous low back episodes. Patients were excluded if they had pathological conditions causing LBP, if they were currently being treated for LBP, or if they were pregnant. Medical or psychiatric comorbidities, history of substance use disorder, and opioid use were not reported. Patients in the minimal intervention group met with their general practitioners once (20-minute session) and explored psychosocial factors considered to be prognostic, set goals, and received an educational booklet. Patients were to refrain from seeing a physiotherapist for the first 6 weeks. Usual care was left to the discretion of the physician and generally followed accepted guidelines (wait and see, advice, analgesics, gradual increase in activities, home exercise). The trial was conducted in The Netherlands and was sponsored by the government. The trial was assessed as moderate risk of bias due to unclear allocation concealment methods and inability to blind patients and providers (Appendix F).

Results

One trial (N=314) found a minimal intervention strategy associated with similar improvement in function, pain and quality of life measured over 1 year of follow-up: RDQ scores (0-24 scale, adjusted MD 0.25, 95% CI -0.77 to 1.28), pain severity scores (0-10 scale, adjusted MD 0.015, 95% CI -0.41 to 0.44), and SF-36 perceived general health scores (1-5 scale, adjusted MD 0.056, 95% CI -0.07 to 0.17).[81] Similarly, both interventions were associated with similar odds of patients rating themselves as not recovered (i.e., slightly improved, no change or slightly/much/very much worse, OR 1.16, 95% CI 0.63 to 2.17) and sick leave due to pain (yes/no, OR 0.69, 95% CI 0.43 to 1.13). All analyses were adjusted for baseline values and estimated with multilevel analysis.

Differential Efficacy and Safety

One RCT (N=314) that compared minimal intervention strategy with usual care provided 1-year function and patient perceived recovery outcomes stratified by various psychosocial traits and by episodes of LBP (Table 71).[81] However, authors do not provide p-values for interaction and visual inspection of the confidence intervals indicates considerable overlap and likely no modification of treatment effect. Authors also state that the effect did not differ based on subgroup.

Table 71. Minimal intervention strategy versus usual care: differential efficacy from RCTs (Jellema, 2005)

Outcome, Timing	Subgroup*	MD or OR (95% CI)†
RDQ scores (0-24 (worst))	FABQ >15 (n=150)	MD −0.35 (−1.65 to 1.30)
	FABQ ≤15 (n=164)	MD 0.80 (−0.58 to 1.38)

Outcome, Timing	Subgroup*	MD or OR (95% CI)†
1 year	CSQ >11 (n=140)	MD 0.20 (−1.19 to 1.59)
	CSQ ≤11 (n=174)	MD 0.37 (−0.78 to 1.52)
	4DSQ >10 (n=107)	MD 0.35 (−1.55 to 2.26)
	4DSQ ≤10 (n=206)	MD 0.35 (−0.54 to 1.23)
	≥6 weeks or ≥3 episodes in past year (n=145)	MD −0.24 (−1.85 to 1.37)
	<6 weeks and <3 episodes in past year (n=169)	MD 0.54 (−0.27 to 1.34)
No recovery (%)‡ 1 year	FABQ >15 (n=150)	OR 0.89 (0.36 to 2.18)
	FABQ ≤15 (n=164)	OR 1.50 (0.64 to 3.53)
	CSQ >11 (n=140)	OR 0.72 (0.29 to 1.80)
	CSQ ≤11 (n=174)	OR 1.84 (0.80 to 4.22)
	4DSQ >10 (n=107)	OR 1.35 (0.51 to 3.60)
	4DSQ ≤10 (n=206)	OR 1.16 (0.52 to 2.56)
	≥6 weeks or ≥3 episodes in past year (n=145)	OR 1.37 (0.58 to 3.22)
	<6 weeks and <3 episodes in past year (n=169)	OR 0.88 (0.38 to 2.08)

4DSQ = four-dimensional symptom questionnaire; CSQ = coping strategies questionnaire; FABQ = fear-avoidance beliefs questionnaire; MD = mean difference; OR = odds ratio; RCT = randomized controlled trial; RDQ = Roland-Morris disability questionnaire.
*Higher scores mean more functional disability, more fear-avoidance, more catastrophizing thoughts, and more distress)
†Minimal intervention strategy minus usual care, adjusted for baseline values and estimated with multilevel analysis.
‡Slightly improved, no change or slightly/much/very much worse on a 1-7 Likert scale.

Key Question 7. Effectiveness and Harms of Selected Interventional Procedures (e.g., botulinum toxin, trigger point injection, epidural steroids/peri-radicular injections, SI joint/extra-articular injections, dry needling) compared with: (1) placebo or sham; (2) with each other; (3) systemic opioid therapy; (4) nonopioid pharmacologic treatments; or (5) nonpharmacologic therapy on Health Outcomes at Short Term and Long Term

Description of Included Studies

Two studies (N=127) evaluated selected interventional procedures treatment of ALBP(Appendix E).[103,134]

One trial (N=54) compared the effect of TPIs with an intravenously administered NSAID (dexketoprofen).[103] The TPI group (n=22) received TPIs of 2% lidocaine, 2.5 cc from 100 mg-5cc of ampoule, with 2.5 cc of saline mixture, injected to the trigger point, with the point needled several times. The NSAID group (n=32) received 50 mg of dexketoprofen in 100 cc isotonic solution intravenously, over a period of 5 minutes. Mean patient age was 43 years and 57% were female with LBP up to 48 hours duration (mean 12 hours). Patients with underlying diseases or

trauma as well as those taking NSAIDS or undergoing physiotherapy were excluded. No further information about patient demographics, substance use disorder, or radiculopathy was provided. The trial was conducted in Turkey and the source of funding was not reported. The trial was assessed as moderate risk of bias due to unclear allocation concealment methods and lack of blinding (Appendix F).

One trial (N=73) compared the effect of adding epidural steroid injection (ESI) s (n=37) to usual care with usual care alone (n=36) in patients with acute lumbosacral radicular syndrome (LRS) with profound sciatica.[134] The ESI group received an injection by an independent anesthesiologist within 48 hours of randomization of 80 mg triamcinolone in 10 ml normal saline administered using lumbar translaminar approach, one level above the presumed LRS. Usual care was not standardized and consisted of advice, analgesic medication, and/or referral, based on discretion of the attending clinician. Mean patient age was 44 years and 52% were female with LBP of up to 4 weeks duration. Patients with underlying diseases or trauma, those who were taking corticosteroids or anticoagulants, or those with body mass index (BMI) >35 were excluded. No further information about patient demographics or substance abuse was provided. The trial was conducted in The Netherlands and received funding from the University Medical Center Groningen. The trial was assessed as moderate risk of bias due to the inability to blind patients and providers (Appendix F).

TPIs Versus an Intravenously Administered NSAID

TPI was associated with a moderate increase in the likelihood of achieving pain response, defined as a score <4 on a 0-10 VAS scale, compared with NSAIDs at 60 minutes (N=54, 95% vs. 62%, RR 1.53, 95% CI 1.15 to 2.03) and a large improvement in VAS pain scores at all timepoints measured up to 60 minutes (MD -2.18, 95% CI -3.30 to -1.06); range of differences across 5 to 60 minutes (-3.77 to -2.18).[103] No side effects were reported during the study.

ESIs Plus Usual Care Versus Usual Care

After adjustment for baseline values, ESI was associated with a statistically significant improvement in function (RDQ scores, 0-24 scale) and back pain (NRS scores, 0-10 scale) compared with IV NSAIDs measured six times up to 52 weeks.[134] Over the entire course of the study, ESI was associated with a moderate improvement in both function (N=63, -2.50, 95% CI -4.55 to -0.46) and back pain (N=63, adjusted MD -1.12, 95% CI -1.98 to -0.26) versus IV NSAIDs using mixed models analysis. Both treatments resulted in similar improvement in leg pain (NRS, 0-10 scale) at all timepoints and over the entire course of the study (N=63, adjusted MD -0.67, 95% CI -1.67 to 0.33). Patients in the intervention group reported significantly greater satisfaction (Mean=9) than those in the control group (Mean=7.2) (NRS scores, 0-10 scale). There was one serious adverse event (1.6%); one patient (2.7%) in the ESI group died due to Burkitt lymphoma.

Key Question 8. In adults presenting with ALBP(<6 weeks duration), including back pain with radiculopathy:

Key Question 8a. What is the comparative effectiveness of specific management strategies (e.g., stratification by risk for poor prognosis, early referral to therapy [e.g., manipulation therapies, exercise, physical therapy modalities)], matching therapies through treatment-based classification of patients, stepped care

models, interdisciplinary models) versus (1) waitlist, usual care, or no treatment; or (2) other strategies on health outcomes short term and long term?

Description of Included Studies

Nine RCTs[60,80,90,98,110,118,123,145,157,164,167] (reported in 11 publications) and two comparative nonrandomized studies of interventions (NRSIs)[171,184] assessed the effectiveness of management strategies for the treatment of ALBP (Table 72; Appendix E).

The nine RCTs enrolled a median of 119 participants (range 70 to 226), apart from one large trial[60] that included 2,300 participants. The same study (the TARGET trial) included participants with LBP with a duration of up to 6 months, however the majority of the population (61%) reported a pain duration of less than 1 month. Among the other trials, baseline pain duration ranged from 24 hours to 90 days, and patients reported at least moderate pain (7 trials) and moderate to severe disability (8 trials) at baseline. Presence of radiculopathy was only reported in four of the trials, and ranged widely from 18% to 67%.[60,80,98,164]

In addition to the characteristics reported in Table 72, seven trials[98,118,123,157,164,167] excluded pregnant or breastfeeding woman and two[60,90] did not report whether they were included or excluded. Race and/or ethnicity was reported in five trials. White populations predominated, ranging from 62% to 82% of enrolled populations among the included studies, with fewer Black (2% to 18%) and Hispanic (6% to 14%) participants.[60,123,157,164,167] Medical comorbidities were infrequently reported, and the included trials did not report on history or treatment of opioid or substance abuse disorder.

Six trials[60,80,118,123,157,167] assessed early interventions to reduce the risk of progression to chronic pain. This included five trials[60,80,123,157,167] of early referral to physical therapy and one trial[118] of early interdisciplinary therapy. One trial[98] examined the effect of treatment-based classification that stratified patients into subgroups most likely to benefit from specific treatments (e.g., mobilization, exercise). One trial[90] assessed an interdisciplinary intervention that included psychology, physical therapy, occupational therapy, and case management, and the remaining trial[164] utilized a goal-directed treatment plan that included educational components for clinicians and patients. All trials had a usual care[60,80,98,123,157,164,167] or no intervention[90,118] control group. Detailed description of interventions appears in Appendix E.

Six trials were conducted in the United States,[60,90,98,123,157,167] and one trial each was conducted in China (Hong Kong),[80] New Zealand,[164] and the United Kingdom.[118] Six trials[90,118,123,157,164,167] reported receipt of government funding. Two trials were funded by a foundation[98] or other nonprofit organization,[60] and funding source was not reported in one trial.[80] One trial was considered low risk of bias,[164] and the remaining trials were rated moderate risk of bias. Common limitations in the moderate risk of bias trials included unclear group allocation concealment methods and lack of blinding for patients (Appendix F).

Two NRSIs of management strategies enrolled a total of 245 patients with ALBP (Table 72).[171,184] Study inclusion criteria differed between the studies. One study[184] enrolled patients presenting to the ED with a duration of pain of less than 2 weeks, although pain onset was within 3 days for 95% of the enrolled population. The other study[171] included patients with a duration of LBP of less than 90 days (median 16 days duration). Pain and functional limitations in both studies were moderate to severe based on VAS and ODI scores at baseline. Race and ethnicity were reported in one study. In that study, the population was 37% White, 36% Black, and 19% Hispanic.[184]

Although both NRSIs utilized treatment-based classification,[236,237] the studies differed in design and interventions (Appendix E). The ED-based study[184] prospectively enrolled and assigned patients to either early intervention or usual care. Those in the early intervention group underwent classification in the ED and were accordingly assigned to a multi-component, rehabilitation-matched intervention that started immediately. In the second study,[171] patients who had previously been randomly assigned to either manipulation, stabilization, or exercise were retrospectively classified using treatment-based classification. Patients whose randomized treatment matched their retrospective classification were compared with those patients whose randomized treatment did not match their retrospective classification.

Both studies were conducted in the United States; one was government-funded[184] and the other was foundation-funded.[171] Risk of bias was judged to be moderate in both studies (Appendix F).

Table 72. Study characteristics of management strategies versus usual care

Study, year	Brennan, 2006[171]	Darlow, 2019[164]	Delitto 2021[60] and Beneciuk 2023[110] TARGET Trial	Fritz 2003[98]	Fritz 2015[157] and Magel 2017[145]	Fritz 2021[123]	Gatchel 2003[90]	Kim 2021[184]	Lau 2008[80]	Rhon 2018[167]	Wand 2004[118]
N Randomized	NRSI; 123 included	226	2300	78	220	220	70	NRSI; 122 included	110	119	102
Mean Age (Years)	38	46	59	37	37	39	38	Median 40	50	27	35
% Female	45	46	59	38	52	49	35	59	61	17	50
% Radiculopathy	NR	67	32	18	NR	NR	NR	NR	26	NR	NR
Pain Duration Inclusion Criteria	<90 days	Initially <6 weeks; protocol revised to include any duration	LBP interfering with ability to do daily activities for ≤3 months or for <1/2 days for 6 months	<3 weeks	<16 days	≤90 days	<2 months	<2 weeks	≤24 hours	<90 days	≤6 weeks
Mean Pain Duration (days)	Median 16 days	88% <6 weeks	Mean duration NR; 61% <1 month; 29% 1-3 months; 10% >3 months	5.5	NR	36	3.8 weeks	Mean NR; 95% ≤3 days	NR	NR	NR
Mean/Median Baseline Pain Severity (0-10)	5.4	6.9	NR	6.7	5.2	5.0	NR	7.0	7.7	5.1	5.4
Function Inclusion Criteria	ODI ≥25	NR	LBP that interfered with regular daily activities for ≤3 months or for <1/2 the days during previous 6 months	Sufficient severity to necessitate modification of work duties	ODI ≥20	ODI ≥20	Constant daily pain when performing activities and decreased ability to perform normal job requirements	NR	NR	ODI >20	NR
Mean/Median Function	ODI: 43.3	RDQ: 12.64	ODI: 48.1	ODI: 42.9	ODI: 41.1	ODI: 37.4	NR	ODI: 40.0	RDQ: 15.5	ODI: 33.0	RDQ: 11.3

Study, year	Brennan, 2006[171]	Darlow, 2019[164]	Delitto 2021[60] and Beneciuk 2023[110] TARGET Trial	Fritz 2003[98]	Fritz 2015[157] and Magel 2017[145]	Fritz 2021[123]	Gatchel 2003[90]	Kim 2021[184]	Lau 2008[80]	Rhon 2018[167]	Wand 2004[118]
Psychiatric Comorbidities	NR	NR	Depression: 4% Anxiety: 4%	NR; mean CES-D depression score 13.1 (clinical cutoff 16.0)	Anxiety/ Depression: 28%	Anxiety: 21% Depression: 9%	Excluded (≥6 DSM-IV Axis 1 diagnoses, current psychosis, suicidal ideation)	NR	NR	NR	Excluded
Management strategy	Classification-based treatment	Goal-directed treatment	Early PT	Classification-based treatment	Early PT	Early PT	Interdisciplinary treatment	Early PT	Early PT	Early PT	Early interdisciplinary treatment
Duration follow-up	1 year	6 months	6 months	1 year	1 year	6 months	1 year	3 months	6 months	1 year	6 weeks
Country	U.S.	New Zealand	U.S.	U.S.	U.S.	U.S.	U.S.	U.S.	China	U.S.	UK
Risk of Bias	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate

CES-D = Center for Epidemiologic Studies Depression Scale; NR = not reported; NRSI = nonrandomized study of interventions; ODI = Oswestry Disability Index; PT = physical therapy; RDQ = Roland-Morris Disability Questionnaire.

Detailed Synthesis

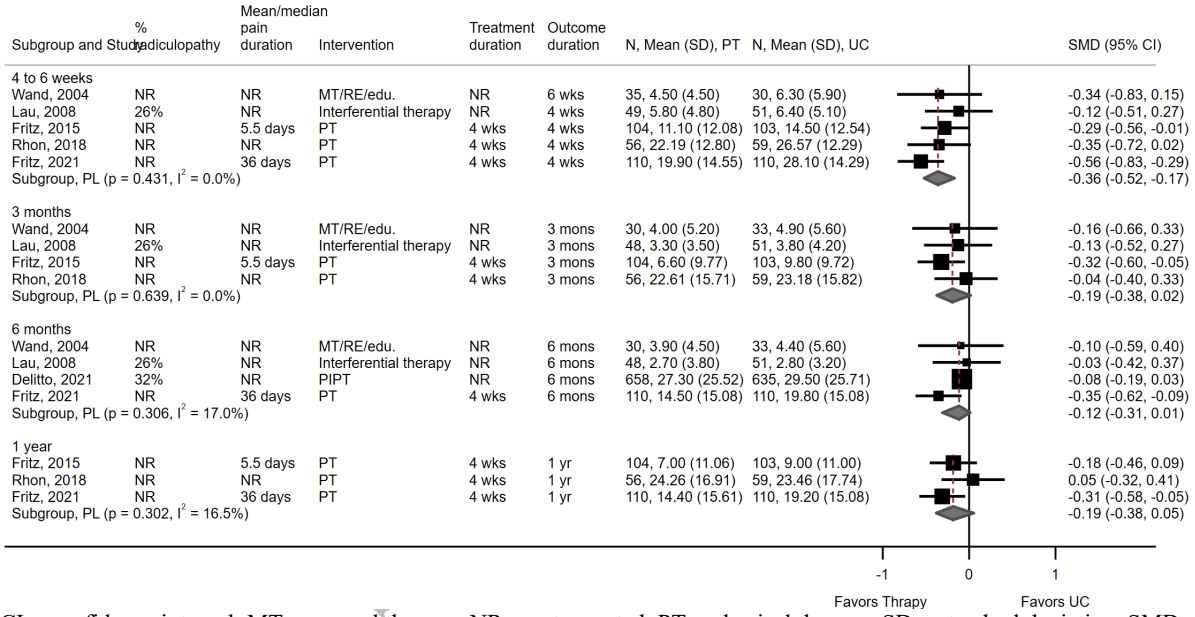
Early Interventions

Six trials[60,80,118,123,157,167] compared the effect of early interventions (physical therapy referral[60,80,123,157,167] or interdisciplinary treatment[118]) with usual care or no treatment. All trials had a moderate risk of bias (Appendix F).

Function

Six trials reported measures of function based on ODI[60,123,157,167] or RDQ[80,118] scores (Appendix E). Relative to usual care/no intervention, early interventions were associated with a small difference in function scores at 4 to 6 weeks (5 RCTs, N=707, SMD -0.36, 95% CI -0.52 to -0.17) (Figure 38; Table 73). There were slight but statistically insignificant differences between groups at 3 months, 6 months, or 1 year (Figure 38; Table 69).

Figure 38. Early interventions versus usual care: function scores



CI = confidence interval; MT = manual therapy; NR = not reported; PT = physical therapy; SD = standard deviation; SMD = standard mean difference; UC = usual care.

Table 73. Meta-analysis of RCTs of early interventions versus usual care: function

Number of RCTs N=	Time	Results SMD (95% CI or SD)
5 RCTs[80,118,123,157,167] N=707	4-6 weeks	**-0.36 (-0.52 to -0.17)**
4 RCTs[80,118,157,167] N=484	3 months	-0.19 (-0.38 to 0.02)
4 RCTs[60,80,118,123] N=167	6 months	-0.12 (-0.31 to 0.01)
3 RCTs[123,157,167] N=542	1 year	-0.19 (-0.38 to 0.05)

CI = confidence interval; RCT = randomized controlled trial; SD = standard deviation; SMD = standard mean difference.

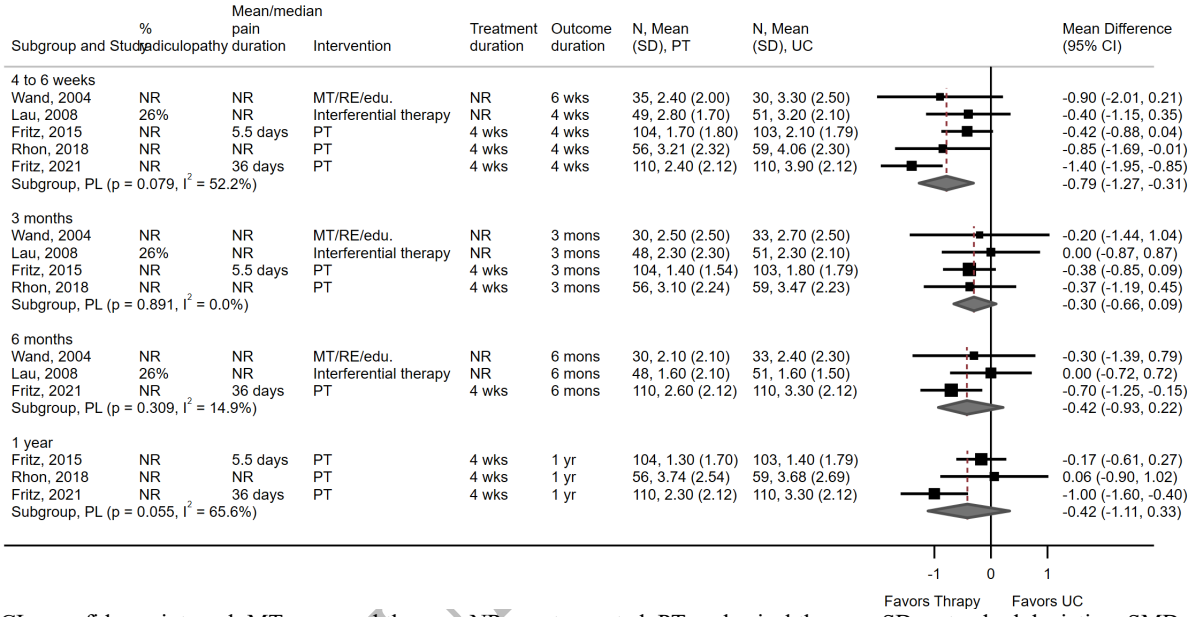
One NRSI not included in the meta-analysis compared the effect of physical therapy initiation in the ED for patients presenting with ALBP (Appendix E).[184] ED-initiated physical

therapy was associated with lower ODI scores at 2 months compared with usual care (adjusted MD -9.93, 95% CI -18.90 to -0.95) with no clear difference between groups at 1-week, 1-month, and 3-month follow-up (Appendix E).

Pain

The effect of early interventions on pain score was assessed in five RCTs, all moderate risk of bias (Appendix E).[80,118,123,157,167] Based on pooled analysis (N=707), early interventions were associated with moderately lower pain scores compared to usual care at 4 to 6 weeks follow-up (SMD -0.79, 95% CI -1.27 to -0.31; Figure 39; Table 74). There was no clear difference in pain scores between groups at 3 months, 6 months, or 1 year.

Figure 39. Early interventions versus usual care: pain scores



CI = confidence interval; MT = manual therapy; NR = not reported; PT = physical therapy; SD = standard deviation; SMD = standard mean difference; UC = usual care.

Table 74. Meta-analysis of RCTs of early interventions versus usual care: pain

Number of RCTs N=	Time	Results SMD (95% CI or SD)
5 RCTs[80,118,123,157,167] N=707	4-6 weeks	**-0.79 (-1.27 to -0.31)**
4 RCTs[80,118,157,167] N=484	3 months	-0.30 (-0.66 to 0.09)
3 RCTs[80,118,123] N=382	6 months	-0.42 (-0.93 to 0.22)
3 RCTs[123,157,167] N=542	1 year	-0.42 (-1.11 to 0.33)

CI = confidence interval; RCT = randomized controlled trial; SD = standard deviation; SMD = standard mean difference.

Patient Perception of Improvement

One trial (n=220) found patients who received physical therapy within 3 days of presentation more likely to describe their pain as a "great deal better" or a "very great deal better" compared to usual care patients at both 4-week (58% vs. 44%; RR 1.33, 95% CI 1.01 to 1.75) and 3-month 61% vs. 42%; RR 1.45, 95% CI 1.10 to 1.91) follow-up.[157]

Quality of Life

Quality of life, measured using the EuroQoL (EQ-5D) or 36-item Short Form Questionnaire (SF-36), was reported in four trials (Table 75).[80,118,123,157] The effect of early management strategies on quality of life is unclear based on these studies. There was a marginal benefit suggested based on EQ-5D scores (between-group differences ranging from 0.02 to 0.10) across timepoints from 4 weeks to 1 year in three of the trials,[118,123,157] while SF-36 mental and physical component scores were similar between groups at 1, 3, and 6 months in one trial.[80]

Table 75. Quality of life outcomes in trials of early intervention management strategies

Author, year N=	Scale	Time	Results Score (95% CI or SD)
Fritz 2015[157] N=220	**EQ-5D**	4 weeks	0.87 (0.85 to 0.89) vs. 0.84 (0.82 to 0.86) Mean between-group difference: 0.03 (0.0 to 0.07)
		3 months	0.91 (0.88 to 0.93) vs. 0.88 (0.86 to 0.90) Mean between-group difference: 0.03 (-0.01 to 0.06)
		1 year	0.92 (0.90 to 0.94) vs. 0.88 (0.86 to 0.90) **Mean between-group difference: 0.04 (0.01 to 0.07)**
Fritz 2021[123] N=220	**EQ-5D**	4 weeks	0.76 (0.73 to 0.8) vs. 0.7 (0.66 to 0.73) **Mean between-group difference: 0.07 (0.02 to 0.11)**
		6 months	0.80 (0.77 to 0.84) vs. 0.79 (0.75 to 0.82) Mean between-group difference: 0.02 (-0.03 to 0.07)
		1 year	0.82 (0.79 to 0.86) vs. 0.78 (0.74 to 0.81) Mean between-group difference: 0.04 (−0.01 to 0.09)
Lau 2008[80] N=110	**SF-36 PCS**	1 month	40 (9) vs. 42 (8) Mean between-group difference: -1 (-5 to 2)
		3 months	46 (10) vs. 47 (9) Mean between-group difference: -1 (-5 to 4)
		6 months	45 (11) vs. 46 (8) Mean between-group difference: -1 (-5 to 4)
	SF-36 MCS	1 month	46 (11) vs. 46 (9) Mean between-group difference: 1 (-4 to 5)
		3 months	51 (9) vs. 50 (9) Mean between-group difference: 1 (-3 to 5)
		6 months	54 (8) vs. 53 (7) Mean between-group difference: 1 (-3 to 4)
Wand 2004[118] N=102	**EQ-5D**	6 weeks	0.8 (0.1) vs. 0.7 (0.3) Mean between-group difference: 0.10 (-0.01 to 0.21)
		3 months	0.82 (0.22) vs. 0.73 (0.26) Mean between-group difference: 0.09 (-0.03 to 0.21)
		6 months	0.85 (0.19) vs. 0.75 (0.26) Mean between-group difference 0.10 (-0.01 to 0.21)

CI = confidence interval; EQ-5D = EuroQoL 5D; MCS = mental component scale; PCS = physical component scale; SD = standard deviation; SF-36 = 36-item Short Form Questionnaire.

Psychological Distress

Two trials (total N=440) of early physical therapy reported Fear-Avoidance Beliefs Questionnaire (FABQ) physical activity subscale scores (scale 0 to 24).[123,157] The studies were conducted by the same author, with similar designs, interventions, and patient populations. Baseline FABQ physical activity subscale scores were in the 14- to 15-point range, indicating elevated fear-avoidance beliefs (Appendix E). At timepoints ranging from 4 weeks to 1 year, FABQ physical activity subscale scores were similar between groups, and the mean between-group difference in scores was generally less than 1 point in both trials. Exceptions were found in one of the trials, which reported a between-group differences of -1.8 points (95 % CI -3.5 to -0.1) at 4 weeks and -3.6 points (95% CI -5.4 to -1.8) at 1 year.[123]

Work-Related Outcomes

Two trials of early physical therapy reported work-related outcomes (Appendix E).[123,167] At 1-year follow-up, one trial[123] reported no difference between groups in the proportion of patients reporting missing any days of work or in the mean number of days of work missed per month, and the second trial[167] found no difference in the proportion of patients requiring modified work duties or in the number of work days limited due to ALBP.

Opioid Use and Use of Rescue Medication

Use of opioids and other medications was reported in two trials,[60,167] including the large TARGET trial,[60] and one NRSI.[184] Study results are summarized in Table 76. Although absolute rates of opioid or other medication use were generally lower in patients who received early intervention, no risk estimates were statistically significant at any timepoint.

Table 76. Use of opioids and other medications in studies of early intervention management strategies versus usual care

Author, Year Study Design N=	Medication	Time	Risk of Opioid or Medication Use OR (95% CI)
Delitto 2021[60] RCT N=2261	Prescribed opioid	1 year	OR 0.92 (0.61 to 1.40)
	Benzodiazepine		OR 0.80 (0.55 to 1.16)
	Muscle Relaxants		OR 0.78 (0.54 to 1.12)
	Acetaminophen		OR 0.58 (0.16 to 2.12)
	NSAIDs		OR 0.72 (0.46 to 1.12)
Rhon 2018[167] RCT N=119	Prescribed opioid	1 year	OR 0.84 (0.41 to 1.73)
	Prescribed NSAID		OR 0.75 (0.30 to 1.90)
Kim 2021[184] NRSI N=122	High-risk medication*	1 week	OR 2.17 (0.83 to 5.64)
		1 month	OR 0.11 (0.02 to 0.58)
		2 months	OR 0.30 (0.06 to 1.56)
		3 months	OR 0.10 (0.01 to 1.11)

CI = confidence interval; NRSI = nonrandomized study of interventions; NSAID = nonsteroid anti-inflammatory drug; OR = odds ratio; RCT = randomized controlled trial.
* Opioids or other high-risk medications with the potential for central nervous system and respiratory depression

Treatment-Based Classification

One trial[98] (N=78) and one NRSI[171] (N=123) compared the effect of treatment-based classification versus usual care for patients with ALBP (Appendix E). Both studies classified patients according to the treatment-based classification system initially developed by Delitto, in which patients who undergo classification are assigned specific treatment plans based on clinical signs and symptoms at presentation. Both studies were judged to have moderate risk of bias (Appendix F).

Function

ODI scores were reported at 4 weeks and 1 year in both studies (Appendix E). At 4-week follow-up, the RCT found a mean difference of -11 points (95% CI -19.9 to -1.9) in ODI score in the group who underwent treatment-based classification compared with those who did not

receive treatment-based classification, indicating a moderate treatment effect.[98] In the NRSI, there was a slight difference between treatment-based classification and no classification (-5.60) that was not statistically significant (95% CI -11.69 to 0.48).[171] Findings were similar at 1 year, with the RCT reporting a small treatment effect favoring treatment-based classification (MD -9.9, 95% CI -17.7 to -3.0), with no difference reported in the NRSI (mean difference -3.10, 95% CI -9.33 to 3.13).

Pain

Neither study reported pain scale or pain response outcomes.

Other Outcomes

The RCT reported patient satisfaction with treatment, quality of life, and work outcomes (Appendix E).[98] At 4-week follow-up, patients who received treatment-based classification report slightly more satisfaction with treatment compared to those who did not receive classification, based on a patient satisfaction scale (scale 0 to 15; higher score=greater treatment satisfaction; mean difference 2 points, 95% CI 1 to 3 points). Quality of life, based on 36-item Short Form Questionnaire (SF-36) physical and mental component scores, also slightly favored treatment-based classification at 4 weeks (mean difference 5.6, 95% CI 0.6 to 10.7 and 5.7, 95% CI 1.8 to 9.5, respectively). The trial found a small association between treatment-based classification and the proportion of patients with no work restrictions at 4 weeks compared to no classification (RR 1.42, 95% CI, 95% CI 1.04 to 1.94). There were no clear differences between groups for any of these outcomes at 1-year follow-up.

Other Management Strategies

Two RCTs assessed the effect of other management strategies (Table 72; Appendix E).[90,164] One trial (N=226)[164] compared an educational, goal-directed treatment plan with usual care (Fear Reduction Exercised Early [FREE]) and the other trial[90] (N=70) compared interdisciplinary intervention program with no intervention. Both trials had moderate risk of bias.

The FREE trial found that, when compared to no intervention, there was no clear effect of interdisciplinary treatment on function (RDQ) or pain (Numeric Pain Rating Scale [NPRS]).[164] Mean between-group difference was less that one point in RDQ score and less than 0.35 points for NPRS and at follow-up ranging from 2 weeks to 6 months, with no statistical significance at any timepoint for either measure (Appendix E). Although there was no difference between intervention and no intervention in anxiety scores at 2 and 6 weeks, fear avoidance scores were significantly lower (indicating less fear avoidance) in the FREE group at 2 weeks (MD -0.65, 95% CI -1.16 to -0.14) but not at 6 weeks. There were also no differences between groups for work-related outcomes (time off or restricted work duties) or in use of opioids or other medications at final, 6-month follow-up.

The second trial[90] randomized high-risk patients to interdisciplinary treatment or no intervention; the trial also included a nonrandomized low-risk group. At 1-year follow-up, high-risk patients who received the interdisciplinary intervention reported lower pain scores over the previous 3 months compared to the high-risk no intervention group (2.68 [SD NR] vs. 4.31 [SD NR]; p=0.001) and had fewer days of self-reported disability (38.2 [SD NR] vs. 102.4 [SD NR]; p=0.001) (Appendix E). The interdisciplinary intervention was also associated with a small increase in the likelihood of being able to return to work at 1 year compared to no intervention (91% vs. 68%; RR 1.32, 95% CI 1.05 to 1.67). The study found no difference between groups in the use of rescue medications.

Key Question 8b. What are the comparative harms of management strategies versus waitlist, usual care, no treatment or other strategies including (1) harms specific to the strategy/pathway; (2) withdrawal due to adverse events; (3) psychological harms (e.g., depression, suicidal ideation, suicidal behavior); (4) overdose; (5) misuse, withdrawal, opioid use disorder/substance use disorder and related outcomes; (6) serious adverse events?

Detailed Synthesis

One trial (N=226) comparing the FREE management strategy with usual care included a narrative report of no serious adverse events in either group and one moderate adverse event (gluteal muscle tear) in the FREE group at 6 months.[164] Total adverse events were not different between groups (31% vs. 20%; RR 1.23, 95% CI 0.75 to 2.02). Harms associated with management strategies were otherwise not reported among the included studies.

Key Question 8c. Do effectiveness or harms (questions a, b) differ based on (1) patient demographics or characteristics including those related to risk factors related to of health; (2) patient medical or psychiatric comorbidities; (3) type or characteristic of back pain (e.g., onset, duration, severity, recurrent, radicular, nonradicular); (4) timing of strategy; (5) dose, frequency of strategy; (6) duration of strategy; (7) type or components of the strategy; (8) current treatment for opioid use disorder; (9) opioid use history; (10) substance use history; (11) use of concomitant therapies; (12) clinical setting, provider type?

Detailed Synthesis

Two trials[60,110,145,157] (in 4 publications) comparing management strategies with usual care conducted subgroup analyses according to patient characteristics. Both trials utilized an early intervention management strategy. Neither trial found any interaction for measures of function (2 trials) or pain (1 trial) when results were stratified according to patient characteristics (Table 77; Appendix E).

Table 77. Effectiveness of management strategies in subgroups

Author, year	Management Strategy*	Outcome	Time	Subgroup	Management Strategy vs. Usual Care
Delitto 2021[60] and Beneciuk 2023[110]	**Early intervention**	**ODI, mean score**	6 months	Age (≤44, 45-64, ≥65 years)	p for interaction=0.53
				Sex (male, female)	p for interaction=0.24
				Race (Black, White, Other)	p for interaction=0.61
				Ethnicity (Hispanic, non-Hispanic)	p for interaction=0.39
				BMI (<30, ≥30)	p for interaction=0.16
				Smoking (smoker, nonsmoker)	p for interaction=0.08
				Area Deprivation Index (quartiles 1 [low deprivation] to 4 [high deprivation])	p for interaction=0.19

Author, year	Management Strategy*	Outcome	Time	Subgroup	Management Strategy vs. Usual Care
				Insurance status (Medicaid, Medicare, private, self-pay, other)	p for interaction=0.49
				LBP location (axial LBP, leg pain + LBP)	p for interaction=0.24
				Pain medication use (no prescription, prescribed 1-2 pain medications, prescribed ≥3 pain medications)	p for interaction=0.10
				Ethnicity (Hispanic, non-Hispanic)	p for interaction=0.14
		ODI, ≥50% improvement in disability score	6 months	Sex (male, female)	p for interaction=0.17
				Area Deprivation Index (quartiles 1 [low deprivation] to 4 [high deprivation])	p for interaction>0.05
Fritz 2015[157] and Magel 2017[145]	**Early intervention**	**ODI, mean score**	4 weeks	Baseline risk (medium risk, high risk)	p for interaction=0.72
			3 months		p for interaction=0.25
			1 year		p for interaction=0.94
		NPRS, mean score	4 weeks	Baseline risk (medium risk, high risk)	p for interaction=0.89
			3 months		p for interaction=0.25
			1 year		p for interaction=0.94

BMI = body mass index; LBP = low back pain; NPRS = Numeric Pain Rating Scale; ODI = Oswestry Disability Index.
* See Appendix E for detailed descriptions of interventions

Subgroup analyses according to other patient characteristics (e.g., history of opioid use disorder or substance use disorder; use of concomitant therapies), management strategy characteristics, clinical setting, or provider type were not reported.

Key Question 9. What is the cost-effectiveness of pharmacologic, nonpharmacologic, interventional procedures and management strategies for treatment of ALBP (<6 weeks duration), including back pain with radiculopathy? (Include only full economic studies.)

Description of Included Studies

Four full economic studies, each addressing different patient populations and interventions, were identified. Two studies were in patients with acute radiculopathy (lumbosacral radicular syndrome, sciatica). One study[192] compared the cost-effectiveness of ESI with usual care alone.

The other study, a cost-utility analysis,[193] compared exercise with usual care. One cost-utility analysis[195] that compared a single acupuncture session with usual care was in patients with nonspecific ALBP. In one study[194] that evaluated the cost-effectiveness of a minimal psychosocial intervention strategy (MIS) with usual care, pain radiating below the knee was reported in 14% of patients at baseline. All studies excluded patients with signs or symptoms of serious pathology as a cause of LBP and used clinical data from RCTs for analysis of benefit. Interventions in all studies were in addition to usual care and compared with usual care alone. All used a 1-year time horizon for economic analyses. Three studies were conducted in the Netherlands[192,194,195] and one in Norway.[193] None were industry funded. Details regarding economic study quality can be found in Appendix F.

Detailed Synthesis

Epidural Steroid Injection Versus Usual Care in Patients with Radiculopathy

Study Overview

One study[192]from the Netherlands compared the cost-effectiveness of ESI with usual care alone in patients with acute radiculopathy (< 6 weeks) or 48% women with a mean age of 44 years and mean baseline pain intensity of 5.4 (0-10 scale) enrolled in a RCT. A societal perspective and 1 year time-horizon are reported (no discounting reported). Direct costs related to medical care components (pharmacy, imaging, other testing, provider visits, physical therapy, intervention costs), as well as indirect costs such as out-of-pocket travel expenses and absence from work were included. Medication costs from the Dutch health insurance board and information on other costs based on published guidelines for cost-effectiveness studies were used. Costing year was not described. Clinical data from the small RCT (N=63) were used to estimate benefit based on improvement in pain (Numeric Rating Scale [NRS] 0-10 scale) over 1 year. No funding was reported.

Base Case and Sensitivity Analyses

Total costs were lower for ESI versus usual care. The cost difference at year between ESI and usual care was -$942 (95% CI $3248 to $1530). Authors cite loss of productivity (absence from work) as the major contributor to the difference. Authors do not report differences in NRS scores at follow-up, noting only that symptom severity remained significantly greater in the usual care group versus the ESI group (p=0.0115), thus, an incremental benefit difference is not apparent. Authors reported a base case incremental cost-effectiveness ratio (ICER) of $990 saved with ESI per 1 point improvement in NRS back pain in one patient during 1 year compared with usual care. Bootstrapping estimated cost-effectiveness of ESI versus usual care (lower cost, greater improvement) in 81% of simulations.

Limitations

Analyses did not include adverse events and the basis for the reported incremental benefit is unclear. The study sample size was small. Sensitivity analyses around assumptions and potential cost drivers were not reported however uncertainty was addressed by bootstrap analyses.

Exercise Versus Usual Care in Patients with Radiculopathy

Study Overview

One cost-effectiveness analysis[193] compared exercise with usual care in patients with lumbosacral radicular syndrome (acute radiculopathy). Exercise type was not described. Patients were enrolled in an RCT (N=135). The mean age was 43 years, 49% were female, and reported mean pain duration of 13 days at baseline; pain intensity was not reported. A societal perspective and 1 year time-horizon are reported. Direct health care costs included those for primary care, physical therapy, manual therapy, Cesar or Mensendieck therapy, specialist consultation, hospitalization and surgery (laminectomy). Direct nonhealthcare costs included devices, out-of-pocket costs and housekeeping assistance. Cited cost sources were Dutch guidelines and professional association data for 2005. Indirect costs were lost productivity due to work absence based on mean income of the Dutch population. Patient information on some direct and indirect costs appears to have been used; authors cite this as a potential source of recall bias. Clinical outcomes data from the RCT were used to determine quality-adjusted life years (QALYs) based on EuroQol 5-dimension EQ-5D scores (0-1 scale); other health benefits were based on the Global Perceived Effect (GPE, 1 to 7 scale) over 1 year. GPE was dichotomized to delineate the percentage of patients that improved (completely or much improved) versus not improved (slight improvement, no improvement slightly worse, much worse and worse than ever) and authors express the primary health effect as "patient improved gain". Foundation funds supported this study.

Base Case and Sensitivity Analyses

Authors did not report cost-utility analyses citing no difference in quality of life between exercise and usual care. They report that total costs were primarily comprised of productivity loss but no statistically significant differences between groups on mean number of days absent from work were seen. The study reports that there were significant differences in total costs (direct and indirect) in favor of usual care. At 1 year, there was a slightly higher likelihood of perceived improvement in the exercise group improved versus usual care (79% vs. 56%, RR 1.4, 95% CI 1.1 to 1.8). ICERs for perceived recovery based on the GPE dichotomization were 837 € (95% CI -732 to 3186€) per patient improved gained using direct costs only and 6224€ (95% CI -10,419 to 27551€) per patient improved gained using direct and indirect costs at 1 year. Confidence interval width for both estimates indicates imprecision. Bootstrap sensitivity analyses suggests that if 600 € per patient improved is an acceptable threshold for direct costs the ICER is acceptable with 35% certainty, but this would increase to 69% at a threshold of 1200 € per patient improved gain. Thresholds for total costs of 4000 € and 1200 € resulted in 37% and 68% certainty of having an acceptable ICER.

Limitations

Sensitivity analyses around assumptions and potential cost drivers were not reported however uncertainty was addressed by bootstrap analyses. Study sample size may have impacted finding of no effect for the EQ-5D. Discussion of possible sources of bias and their potential impact was limited.

Minimal Psychosocial Intervention Strategy Versus Usual Care in Patients with Nonspecific ALBP

Study Overview

One study[194] evaluated the cost-effectiveness of a minimal intervention strategy (MIS) with usual care. MIS consisted of a 20-minute consultation as part of the patient's general practitioner visit for back pain intended to help identify, explore and discuss psychosocial prognostic factors, provide relevant information on treatment and course of ALBP and help the patient set specific goals for resuming work and activities. Patients were enrolled in a cluster RCT. The mean age was 43 years, 47% were female, 14% reported radiation of pain below the knee. Mean baseline back pain intensity was 4.9 (0 to 10 scale) and pain duration was a median of 13 days. A societal perspective and 1-year time-horizon are reported. Direct costs related to medical care included a range of provider services (e.g., primary care, manual therapy, psychology, physical therapy, podiatry) as well as imaging, medications, clinic visits, and hospitalization. Direct costs also included integrative/complementary care methods (e.g., acupuncture, massage). Indirect costs included absenteeism from both paid and unpaid work.

Sources of cost included the Dutch Central organization for Healthcare Charges, the Royal Ducht Society for Pharmacy, and Dutch guidelines for cost-effectiveness studies as well as patient prospective cost diaries (e.g., for informal care, aids, over-the-counter medications). Costs were in 2002 Euros. Clinical data from the cluster RCT (N=314) were used to estimate benefit based on the RDQ (0-24 scale), EuroQol (0-1 scale) and a perceived recovery score (scale NR) for 1 year. The study was government funded.

Base Case and Sensitivity Analyses

The study found that total costs were lower for MIS versus usual care (MD -490 €, 95% CI -987 to 92 €, p>0.05), however the groups had similar improvement in function based on the RDQ (MD -0.74, 95% CI -2.31 to 0.83, 0 to 24 scale), quality of life based on EuroQol (EQ-5D) (MD 0.004, 95% CI -0.04 to 0.03, 0-1 scale) and recovery rate (-2%, 95% CI -14 to 10, scale not provided). Base case ICERs reflect the finding that MIS resulted in less improvement versus usual care but indicate a 690 € saving per RDQ point, 239 € saving per 1% recovery rate and 47,348 € saving per QALY at 1 year. For RDQ improvement, 78% of bootstrapped ICERs indicate that MIS is less expensive, but may be less effective than usual care even though 96% of the cost-effectiveness pairs were close to the origin of the cost-effectiveness plane which may suggest similar effectiveness and cost. Authors performed sensitivity analyses using imputation of missing cost data for total costs and for days of absenteeism; approximately 20% of data were missing. Imputation resulted in statistically significant cost differences between MIS and usual care favoring MIS for total costs (MD -628€, 95% CI -1123 to -81€) and for absenteeism (MD -545 €, 95% CI -1031 to -40€). Authors do not report ICERs accounting for these differences but state that sensitivity analyses results are inconsistent with the primary analyses.

Limitations

Sensitivity analyses around assumptions and potential cost drivers were not reported, however uncertainty was addressed by bootstrap analyses. Authors evaluated the impact of missing data on costs, however results should be interpreted cautiously given the imprecision of many estimates.

Acupuncture Versus Usual Care in Patients with Nonspecific ALBP

Study Overview

One cost-utility analysis[195] compared acupuncture with usual care in patients with nonspecific ALBP (<14 days duration) enrolled in an RCT. Patient mean age was 40 years, 51% were female and baseline back pain intensity was 6.3 (0-10 scale). Patients received a single Western-style medical acupuncture treatment. Healthcare and societal perspectives are reported at 28 days and 1 year. Direct costs included primary care providers, physiotherapists, other therapists (not specified), acupuncture equipment, day back surgery and medications. Medication costs were determined from Norwegian Medicines Agency and Pharmacy Selling Prices documentation. Sources of some specific cost data are not clearly reported; authors indicate that information from official professional organization websites was used and cite Norwegian Medical association. Costs in 2018 U.S. dollars and Norwegian krone and were reported. Indirect costs for productivity loss were based on Statistics Norway report of average wages by sex and age. Data on the EuroQol 5-dimensions 3-level (EQ-5D-3L) from the RCT (N=167 with baseline data) were used to determine QALY. The study was funded by the University of Oslo and private endowments.

Base Case and Sensitivity Analyses

Authors report a small incremental difference in costs from a healthcare perspective between acupuncture and usual care at both 28 days ($10) and 1 year ($3) and no difference in QALYs at 28 days (incremental difference 0.00056) but a gain in QALYs with acupuncture at 1 year (incremental difference 0.0487), resulting in base case ICERs of $17,857/QALY gained at 28 days and $62/QALY at 1 year. Authors report that there were no statistical differences for QALYs between acupuncture and usual care at any timepoint, however. Data were missing from 24% of participants at 1 year and imprecision of cost estimates is noted. The ICER from sensitivity analyses using imputation of missing cost and health-related quality of life (HRQOL) data at 1 year was -568/QALY, suggesting cost savings with acupuncture. Other sensitivity analyses varying costs and use of healthcare services were not reported for the healthcare perspective.

Base case ICERs from a societal perspective suggest cost-savings with acupuncture versus usual care at 28 days (ICER -$1,055,357/QALY) and at 1 year (ICER -$91,877) however 95% confidence intervals for costs are wide, calling into question the precision and stability of these estimates. Sensitivity analyses using imputation of missing cost and HRQOL data at 1 year suggest a much lower cost savings (ICER -$15,389/QALY). Authors additional sensitivity analyses based on imputation method and for imputation of both cost and QALY data together and for each separately, resulted in a range of ICERS at 1 year (-$15,389/QALY to -$63,203/QALY). ICERs for low use of healthcare services and high use of services were -$14,846/QALY and $16,590/QALY respectively based on imputed data.

Limitations

Sensitivity analyses suggest that ICER estimates are imprecise; a broad range of possible ICERS, particularly for the societal perspective at 1 year, were reported and confidence intervals for costs are wide. Sensitivity analyses varying costs and use of frequency of healthcare services were not reported for the healthcare perspective. Authors evaluated the impact of missing data on costs and HRQOL, however results should be interpreted cautiously.

Discussion

Key Findings and Strength of Evidence

This review synthesized evidence on the assessment of acute ALBP and the comparative effectiveness of management options including pharmacologic and nonpharmacologic treatments as well as special interventional procedures and specific management strategies. Key findings are presented below.

Evidence was available from five systematic reviews[11,53-56] and 15 nonrandomized studies on factors or tools for assessing patients with back pain. Few patient or clinical exam factors for serious spine pathology (red flags) were considered highly informative for identifying or ruling out such pathologies. We relied on systematic reviews that reported the accuracy of such factors for specific pathologies. Therefore, the findings depend on how well the reviews were conducted and the quality of the studies they included. The reviews were reasonably well conducted. However, our confidence in findings is low for most conditions due to methodologic limitations of the underlying evidence and the small numbers of studies for many factors for serious pathologies. For factors predicting future chronic LBP or disability (yellow flags), we updated a prior systematic review conducted by a core team member.[11] Incorporation of the new evidence reaffirms that the presence of high levels of maladaptive pain coping behaviors, presence of nonorganic signs, high levels of baseline functional impairment, low general health status, and psychiatric comorbidities can increase the likelihood of correctly predicting the development of persistent disabling LBP through 1 year. Limitations of the evidence include substantial heterogeneity in how factors and outcomes were defined and evaluated in the studies, as well as methodological limitations of the studies. Risk prediction instruments could be more helpful than individual yellow flags for predicting outcomes and for matching care with individual patient needs. However, no instrument has been extensively validated. Validation was most robust for the Acute Low Back Pain Screening Questionnaire and related modifications, and most instruments include individual items that were not predictive (such as sex, education level, or previous LBP episodes). No trials directly evaluating the effectiveness of screening tools versus not using them to inform clinical decision making in patients with ALBP (<6 weeks duration) specifically were identified.

Evidence on early imaging of patients with ALBP was available from three RCTs,[62,158,159] three nonrandomized studies and one cost-effectiveness study. RCTs in patients with ALBP that compared early or immediate imaging with usual care without immediate imaging found no difference in function or pain at ≥4 weeks. However, our confidence in these findings is low due to study limitations and imprecision. Studies in workers compensation populations suggest an association between early MRI and the likelihood of continuing to receive disability benefits and use of additional medical services, however these results are based on observational studies and may not be generalizable and should be interpreted cautiously.

Evidence from 109 RCTs informed Key Questions regarding the comparative effectiveness evaluation of management options in patients with ALBP. Key findings related to the comparative effectiveness of management option are summarized in Tables 78 to 80 below. The tables focus on evidence for function, pain and adverse events for medications stratified for which there is at least low evidence. An expanded version showing comparisons with no or insufficient evidence and evidence at <1 day is found with the detailed strength of evidence tables (Appendix G).

Function was inconsistently reported across treatments and studies reported a variety of measures. Pain was the main reported outcome across trials. However, there was variability in measures used to assess pain. Where possible, we reported the likelihood of patients achieving clinically important differences in pain and function (e.g., ≥30% improvement or similar) as well mean improvement in instrument scores. In general, although some interventions were associated with beneficial effects on pain or function, effect sizes for most interventions were small for these outcomes, based on mean differences. Few trials reported on the likelihood of patients meeting clinically important differences. Results for efficacy outcomes other than pain or function were limited and are detailed in the report.

Table 78. Acute low back pain: effects on function from RCTs evaluating treatment options

Key Question	Intervention and Comparison	Function *1 day to <1 week* Effect Size/SOE	Function *1 week to <2 weeks* Effect Size/SOE	Function *2 to <4 weeks* Effect Size/SOE	Function *≥4 weeks* Effect Size/SOE
KQ 3a: Single Opioid	**Opioid vs. placebo**	no evidence	none +	no evidence	Small favoring placebo +
KQ 3b: Opioid + Nonopioid	**Opioid + nonopioid vs. muscle relaxant**	no evidence	none +	no evidence	no evidence
KQ 5a: Single Nonopioid	**NSAID vs. placebo**	small ++	none +	none +	none +
	NSAID vs. acetaminophen	no evidence	no evidence	none +	no evidence
	NSAID vs. manual therapy	Insufficient evidence	none +	none +	none +
	Muscle relaxant vs. placebo	small ++	small ++	none +	none +
	Muscle relaxant vs. benzodiazepines	**Responders:** moderate + **Scores:** Insufficient evidence	**Responders:** moderate + **Scores:** Insufficient evidence	no evidence	no evidence
	Muscle relaxant vs. manual therapy	no evidence	no evidence	none +	none +

Key Question	Intervention and Comparison	Function *1 day to <1 week* Effect Size/SOE	Function *1 week to <2 weeks* Effect Size/SOE	Function *2 to <4 weeks* Effect Size/SOE	Function *≥4 weeks* Effect Size/SOE
	Steroid vs. placebo	**Radiculopathy:** *Scores* none +	**With or without radiculopathy:** *Responders* None ++	**Radiculopathy:** *Responders* moderate + *Scores* small +	**Radiculopathy:** *Responders* small to moderate ++ *Scores* small + **No radiculopathy:** *Responders* none +
	Acetaminophen vs. placebo	no evidence	none ++	none ++	none ++
KQ 5b: Combination 2 Nonopioids	**Muscle relaxant + NSAID vs. NSAID**	no evidence	none ++	no evidence	none ++
	Acetaminophen + NSAID vs. NSAID	none +	none +	no evidence	no evidence
	Benzodiazepine + NSAID vs. NSAID	no evidence	none +	no evidence	none +
KQ6: Nonpharmacologic - Exercise	**Exercise vs. placebo**	no evidence	no evidence	no evidence	none +
	Exercise vs. UC	no evidence	**Responders:** moderate + **Scores:** none +	**Responders:** small + **Scores:** none +	**Responders:** none + **Scores:** none ++
	Exercise vs. bed rest	no evidence	no evidence	none +	none +
	Exercise vs. advice to stay active	no evidence	no evidence	no evidence	none +
	Exercise vs. acupuncture	no evidence	no evidence	none +	none +
KQ6: Nonpharmacologic – Manual Therapy	**Manual therapy vs. sham or placebo**	no evidence	none +	small +	none +
	Manual therapy vs. UC	no evidence	Insufficient evidence	none +	moderate +
	Manual therapy vs. massage	no evidence	no evidence	none +	none +

Key Question	Intervention and Comparison	**Function** *1 day to <1 week* Effect Size/SOE	**Function** *1 week to <2 weeks* Effect Size/SOE	**Function** *2 to <4 weeks* Effect Size/SOE	**Function** *≥4 weeks* Effect Size/SOE
	Manipulation vs. mobilization	Subgroup pain duration < 2 wks: none + Subgroup pain duration 2-4 wks: large +	no evidence	no evidence	no evidence
KQ6: Nonpharmacologic – Acupuncture	**Acupuncture vs. placebo**	**Scores:** none +	**Scores:** none +	**Responders:** none + **Scores:** moderate ++	**Responders:** 3 months small + 12 months none + **Scores:** *Scalp acupuncture* Moderate + *Traditional Chinese acupuncture* None +
	Acupuncture vs. sham	no evidence	no evidence	**Responders:** none + **Scores:** none +	**Responders:** 3, 12 months none + **Scores:** *Traditional Chinese acupuncture* None +
	Acupuncture (traditional Chinese) vs. UC	no evidence	no evidence	moderate ++	4-12 weeks: small + 52 weeks: none +
	Acupuncture (motion-style) vs. UC	no evidence	no evidence	insufficient evidence	none +
	Acupuncture vs. Massage Therapy	no evidence	small +	no evidence	no evidence

Key Question	Intervention and Comparison	Function *1 day to <1 week* Effect Size/SOE	Function *1 week to <2 weeks* Effect Size/SOE	Function *2 to <4 weeks* Effect Size/SOE	Function *≥4 weeks* Effect Size/SOE
KQ6: Nonpharmacologic – Advise to Stay Active	**Advice to stay active vs. bed rest**	no evidence	small +	no evidence	small ++
KQ6: Nonpharmacologic – Education	**Education vs. UC/placebo**	insufficient evidence	small +	insufficient evidence	3 months: small + 6-12 months: none +
KQ6: Nonpharmacologic – Minimal Psychosocial Intervention	**Minimal psychosocial intervention vs. UC**	no evidence	no evidence	no evidence	none +
KQ7: Select Interventional Procedures	**Epidural steroid injections vs. UC**	no evidence	no evidence	no evidence	small +
KQ8: Management Strategies or Pathways	**Early intervention vs. UC**	no evidence	no evidence	no evidence	4-12 weeks: small ++ 6 months: similar + 12 months: small +
	Treatment-based classification vs. UC	no evidence	no evidence	no evidence	4 weeks: moderate + 12 months: small +
	Educational, goal-directed treatment plan vs. UC	no evidence	no evidence	none +	none +

NSAID = nonsteroidal anti-inflammatory drug; SOE = strength of evidence; UC = usual care.
Effect Size: none, small, moderate, or large improvement; none = no effect/no statistically significant effect.
Strength of Evidence: + = low, ++ = moderate, +++ = high
* Moderate effect at 4 weeks, small effect at 52 weeks.

Table 79. Acute low back pain: effects on pain from RCTs evaluating treatment options

Key Question	Intervention and Comparison	Pain *1 day to <1 week* Effect Size/SOE	Pain *1 week to <2 weeks* Effect Size/SOE	Pain *2 to <4 weeks* Effect Size/SOE	Pain *≥4 weeks* Effect Size/SOE
KQ3a: Single Opioid	**Opioid vs. placebo**	small +	no evidence	none +	4, 52 weeks: none + 6, 12, 26 weeks: small favoring placebo +
KQ5a: Single Nonopioids	**NSAID vs. placebo**	small ++	small +	none ++	none +
	NSAID vs. acetaminophen	no evidence	no evidence	Loxoprofen: none +	no evidence
	NSAID vs. manual therapy	none +	none +	none +	none +
	NSAID vs. acupuncture	Insufficient evidence	Insufficient evidence	Insufficient evidence	Motion style acupuncture: none +
	Muscle relaxant vs. placebo	**Responders** (IM thiocolchicoside) large + **Scores:** moderate ++	Insufficient evidence	none +	none +
	Muscle relaxant vs. benzodiazepines	**Responders:** moderate + **Scores:** Insufficient evidence	**Responders:** moderate + **Scores:** Insufficient evidence	no evidence	no evidence
	Muscle relaxant vs. manual therapy	no evidence	no evidence	none +	none +
	Steroid vs. placebo	**Radiculopathy** *Responders* none +	**Radiculopathy** *Responders* none ++	**Radiculopathy** *Responders* none ++	**Radiculopathy** none ++

Key Question	Intervention and Comparison	**Pain** *1 day to <1 week* Effect Size/SOE	**Pain** *1 week to <2 weeks* Effect Size/SOE	**Pain** *2 to <4 weeks* Effect Size/SOE	**Pain** *≥4 weeks* Effect Size/SOE
		With or without radiculopathy *Scores* Insufficient evidence	**With or without radiculopathy** *Scores* none +	*Scores* none +	
	Acetaminophen vs. placebo	no evidence	none ++	none ++	none ++
KQ5b: Combination 2 Nonopioids	**Muscle relaxant + NSAID vs. NSAID**	no evidence	moderate ++	no evidence	no evidence
	Acetaminophen + NSAID vs. NSAID	none +	none +	no evidence	no evidence
	Pregabalin + NSAID vs. NSAID*	no evidence	no evidence	none +	none +
KQ6: Nonpharmacologic - Exercise	**Exercise vs. placebo**	no evidence	no evidence	no evidence	none +
	Exercise vs. UC	no evidence	none +	none† +	none ++
	Exercise vs. bed rest	no evidence	no evidence	none +	none +
	Exercise vs. advice to stay active	no evidence	no evidence	no evidence	none +
	Exercise vs. acupuncture	no evidence	no evidence	none +	none +
KQ6: Nonpharmacologic – Manual Therapy	**Manual therapy vs. sham or placebo**	no evidence	none +	small ++	none +
	Manual therapy vs. UC	no evidence	none +	Insufficient evidence	small +
	Manual therapy vs. physiotherapy	none +	none +	none +	no evidence
	Manual therapy vs. massage	no evidence	no evidence	none +	none +
KQ6: Nonpharmacologic – Acupuncture	**Acupuncture vs. placebo**	none +	none +	moderate +	small +
	Acupuncture (traditional Chinese) vs. UC	no evidence	no evidence	small +	small +
	Acupuncture (motion-style) vs. UC	no evidence	no evidence	none +	none +

Key Question	Intervention and Comparison	Pain *1 day to <1 week* Effect Size/SOE	Pain *1 week to <2 weeks* Effect Size/SOE	Pain *2 to <4 weeks* Effect Size/SOE	Pain *≥4 weeks* Effect Size/SOE
	Acupuncture vs. massage therapy	no evidence	small +	no evidence	no evidence
KQ6: Nonpharmacologic – Advise to Stay Active	**Advice to stay active vs. bed rest**	no evidence	small +	no evidence	back pain: small + radicular pain: none +
KQ6: Nonpharmacologic – Education	**Education vs. UC/placebo**	insufficient evidence	none +	insufficient evidence	none +
KQ6: Nonpharmacologic – Minimal Psychosocial Intervention	**Minimal psychosocial intervention vs. UC**	no evidence	no evidence	no evidence	none +
KQ7: Select Interventional	**Epidural steroid injections vs. UC**	no evidence	no evidence	no evidence	back pain: moderate + leg pain: none +
KQ8: Management Strategies or Pathways	**Early intervention vs. UC**	no evidence	no evidence	no evidence	4-6 weeks: moderate + 3-12 months: none +
	Educational, goal-directed treatment plan vs. UC	no evidence	no evidence	none +	none +

IM = intramuscular; IV = intravenous; NSAID = nonsteroidal anti-inflammatory drug; TENS = transcutaneous electrical nerve stimulation; SOE = strength of evidence; UC = usual care.
Effect Size: none, small, moderate, or large improvement; none = no effect/no statistically significant effect.
Strength of Evidence: + = low, ++ = moderate, +++ = high
* Leg pain (100% had radiculopathy); authors did not report back pain.
† In subgroup analysis across 4 RCTs with 2-6 weeks treatment, exercise was associated with a small improvement in pain: N=381, pooled MD -1.05, 95% CI -1.96 to -0.27, I^2 = 64%.

Table 80. Acute low back pain: adverse events from RCTs evaluating treatment options

Key Question	Intervention and Comparison	Serious AEs Effect Size/SOE	Withdrawal due to AEs Effect Size/SOE	Any AE Effect Size/SOE
KQ3a: Single Opioid	**Opioid vs. placebo**	insufficient evidence	moderate (tramadol) +	**Oxycodone (+ naloxone):** none + **Oxycodone (+ acetaminophen):** moderate + **Tramadol (+ acetaminophen):** large +
	Opioid vs. NSAID (IV)	no evidence	no evidence	none +
	Opioid vs. acetaminophen (IV)	no evidence	no evidence	none +
KQ3b: Opioid + Nonopioid	**Opioid + nonopioid vs. NSAID**	large +	large +	moderate +
	Opioid + nonopioid vs. muscle relaxant	no evidence	no evidence	none +
KQ5a: Single Nonopioids	**NSAID vs. placebo**	none +	none +	none ++
	NSAID vs. manual therapy	insufficient evidence	none +	GI side effects occur with NSAIDs* +
	Muscle relaxant vs. placebo	none +	none +	higher risk* +
	Muscle relaxant vs. benzodiazepine	no evidence	no evidence	none +
	Steroid vs. placebo	insufficient evidence	**Oral steroid** none +	**IM or IV steroid** none ++ **Oral steroid** large +
	Acetaminophen vs. placebo	none ++	no evidence	none ++
	Cannabidiol vs. placebo	insufficient evidence	insufficient evidence	none +
KQ5b: Combination 2 Nonopioids	**Muscle relaxant + NSAID vs. NSAID**	none +	insufficient evidence	**Any** none + **CNS-events** moderate to large +

Key Question	Intervention and Comparison	Serious AEs Effect Size/SOE	Withdrawal due to AEs Effect Size/SOE	Any AE Effect Size/SOE
	Acetaminophen + NSAID vs. NSAID	no evidence	no evidence	none +
	Benzodiazepine + NSAID vs. NSAID	insufficient evidence	no evidence	none +
	Pregabalin + NSAID vs. NSAID	insufficient evidence	higher risk* +	large +
KQ6: Nonpharmacologic - Exercise	**Exercise vs. UC**	no evidence	none +	no evidence
	Exercise vs. acupuncture	no evidence	none +	no evidence
KQ6: Nonpharmacologic – Manual Therapy	**Manual Therapy vs. physiotherapy**	no evidence	none +	no evidence
KQ6: Nonpharmacologic – Acupuncture	**Acupuncture vs. sham/placebo**	insufficient evidence	insufficient evidence	none +
	Acupuncture vs. UC	insufficient evidence	insufficient evidence	none +
KQ8: Management Strategies or Pathways	**Educational, goal-directed treatment plan vs. UC**	nsufficient evidence	no evidence	none +

AE = adverse event; CNS = central nervous system; GI = gastrointestinal; IM = intramuscular; IV = intravenous; NSAID = nonsteroidal anti-inflammatory drugs; TENS = transcutaneous electrical nerve stimulation; SOE = strength of evidence; UC = usual care.
Effect Size: none, small, moderate, or large improvement; none = no effect/no statistically significant effect.
Strength of Evidence: + = low, ++ = moderate, +++ = high
* Evidence for medication groups only; comparative data not available or no events in the control group so an effect size is not possible.

There were no high strength of evidence ratings for any treatment outcome due to methodological limitations, inconsistency across trials and/or imprecision. The strength of evidence ratings for most treatments was low based on these limitations. There was moderate strength of evidence for several interventions at one or more timepoints.

The most robust evidence base for pharmacologic treatments was for the comparisons of nonsteroidal anti-inflammatory drugs (NSAIDS) versus placebo[68,75,83,91,93,95,100,144,154,155] and for muscle relaxants versus placebo[70,74,75,97,104,132,161,163] or combination of muscle relaxants with NSAID versus NSAID.[57,58,69,75,87,113,126,128] Oral NSAIDs were associated with small functional and pain improvements within the first week versus placebo with similar risk of adverse events between groups. Similarly, muscle relaxants improved function slightly within the first 2 weeks and moderate pain improvement was seen versus placebo. The combination of muscle relaxants with NSAID moderately improved pain versus an NSAID alone but did not improve function. The strength of evidence was moderate for the above comparisons. Similar improvements in pain and function were seen for NSAIDs versus acetaminophen, but evidence was limited to one trial and strength of evidence was low.

There was moderate strength of evidence for additional nonopioid therapies. In patients with radiculopathy, systemic steroids were associated with moderately higher likelihood of functional improvement at 4 weeks versus placebo, however pain improvement was similar between groups as was risk of any adverse event for intramuscular or intravenous steroid administration versus placebo. Acetaminophen did not improve pain or function versus placebo, however (Tables 78 to 80).

Evidence comparing different non-opioid medications to each other, alone or in combination, was sparse and insufficient for most comparisons. In general, there were no differences between treatment groups in functional improvement or pain with one exception. Muscle relaxants were associated with a moderately higher likelihood of achieving a clinically important improvement in pain and function versus benzodiazepines in one trial; the strength of evidence was low. The addition of an NSAID to acetaminophen, a benzodiazepine or pregabalin did not improve patient outcomes versus NSAID alone, however findings are confined to single studies.

Evidence on opioids for management of ALBP was particularly sparse. Evidence from the few available trials of opioid versus placebo suggest there may be limited benefit over placebo.[28,113,133] The strength of evidence was low due to imprecision and unknown consistency. Functional improvement was similar within the first 2 weeks but favored placebo >4 weeks. Pain improvement with opioid was small compared to placebo within the first 2 weeks (tramadol plus acetaminophen) in one trial,[133] however, pain improvement was similar for oxycodone versus placebo between 2 days up to 6 weeks, after which placebo was favored in another trial.[28] There were no differences in the risk of adverse events for comparisons of opioids versus NSAID or acetaminophen. The risk of any adverse event was moderately to substantially higher for opioid (with acetaminophen or acetaminophen) versus placebo and for combinations of codeine with acetaminophen versus NSAID (ketorolac). Retrospective studies found no differences in rates or amounts of opioids prescribed before and after implementation of PDMP in EDs but they did not specifically address patients with ALBP (<6 weeks duration) and did not meet inclusion criteria (Tables 78 to 80).

The largest evidence base for nonpharmacologic treatments was for exercise (usual care was the most common comparator)[59,88,94,96,107,111,119,121,122,127,131,149] followed by studies of manual therapy verses placebo or sham[74,83,102,106,152] or usual care[115,116,119,148,166] and of acupuncture.[111,124,130] Exercise was associated with a small to moderate likelihood of improving function compared with usual care in single trial although mean improvement on instrument scores was similar between groups for pain and function in pooled estimates across trials. Exercise was associated with a small improvement in pain versus usual care across four trials that had employed >1 exercise session. There were no differences in function or pain when exercise was compared to acupuncture, bed rest or advice to stay active. However, evidence was from a small number of trials and the strength of evidence was generally low for these comparisons. In contrast, advice to stay active was associated with small improvements in function and pain versus bed rest. Manual therapy was associated with small improvements in function and pain compared with sham or placebo intervention, and with moderate functional improvement and small pain improvement versus usual care at isolated timepoints. The strength of evidence was moderate for small pain improvement and 2 to 4 weeks with manual therapy versus sham/placebo, but low for other times and comparators. Some forms of acupuncture were associated with moderate functional improvement and small to moderate pain improvement versus placebo. Traditional Chinese acupuncture was associated with moderate to small

functional and pain improvement versus usual care. No differences in function or pain between acupuncture and sham acupuncture were seen, however.

Evidence from eight RCTs (in 10 publications) formed the primary evidence base to inform evaluation of specific management strategies.[60,80,98,110,118,123,145,157,164,167] Early intervention was associated with small functional improvement (moderate strength of evidence) and moderate improvement in pain at 4 to 12 weeks versus usual care (low strength of evidence). Similarly, treatment based on stratification of patients into subgroups most likely to benefit from specific treatments were associated with moderately improved function at 4 weeks; strength of evidence was low for both.

Evidence for select interventional procedures was limited to two small RCTs[103,134] and the strength of evidence was low. Epidural steroid injections were associated with small functional improvement and moderate back pain improvement but no leg pain improvement at >4 weeks in one trial of patients with radiculopathy. Trigger point injections (TPIs) may be more effective than an intravenously administered NSAID in reducing short-term pain.

Few trials reported on outcomes <1 day (Appendix E); six trials[63,65,77,78,89,103] were conducted in an ED setting and two in primary care settings.[101,114] Over timepoints ranging from 30 minutes to 8 hours, four trials found a large improvement in pain with motion-style acupuncture versus NSAID (1 RCT),[101] transcutaneous electrical nerve stimulation (TENS) versus sham (1 RCT)[65] and heat versus placebo (2 RCTs)[63,114] and two trials found moderate improvement in pain with intravenous (IV) morphine versus IV NSAID (1 RCT)[77] and trigger point injections versus IV NSAID (1 RCT).[103] There was no effect on pain up to 6 hours for IV morphine versus IV acetaminophen in one trial,[77] codeine plus acetaminophen versus NSAID in one trial,[89] and cannabidiol versus placebo in one trial.[78] Strength of evidence was low for all.

Adverse events were poorly reported across pharmacologic and nonpharmacologic trials, and the strength of evidence was low or insufficient for most treatment/comparison pairs. Serious adverse events appeared to be rare; however, they were rarely reported, and evidence from studies was insufficient due to study limitations, imprecision and unknown consistency (Table 80 and Appendix E). Mild to moderate adverse events were more common with muscle relaxants compared with placebo but fewer compared with benzodiazepines although neither were statically significant. Adverse events were substantially more common with oral steroids compared with placebo. Gastrointestinal-related adverse events occurred with NSAIDS (versus manual therapy). Such events were similar between a combination of muscle relaxants with NSAID versus NSAID alone, however the risk of central nervous system adverse events was greater with the combination therapy. Most trials of nonpharmacologic treatment did not report on adverse events.

Four studies evaluated the cost-effectiveness of different interventions versus usual care all from a societal perspective. These studies provided limited information on the cost-effectiveness of steroid injections and exercise in patients with radiculopathy and for a minimal psychosocial intervention and acupuncture in patients with nonspecific ALBP versus usual care. They are snapshots based on specific assumptions, clinical scenarios and modeling strategies that may not be widely generalizable. Their inclusion is for informational purposes only.

Limitations Of The Evidence Base

There was limited evidence for some interventions in patients with ALBP as we defined it (<6 weeks) including opioids and some nonpharmacologic management options. Few trials comparing opioids to other pharmacologic options were identified and no studies on dosing

strategies, decision making, or the impact of prescribing opioids or not on long-term patient outcomes were identified. Trials of opioid therapy were not designed to assess long-term opioid use or outcomes such as misuse or substance use disorder. Evidence from nonrandomized studies of interventions (NRSIs) in workers' compensation populations suggest an association between opioid therapy for ALBP and continued opioid use and increased duration of disability, however these findings should be considered within the limitations of such studies (e.g., selection bias, residual confounding). No studies that evaluated the impact of policy and patient factors or of shared decision-making regarding prescription of opioids on patient health outcomes were identified in patients with ALBP. Further, we found no studies evaluating the accuracy or application of instruments for predicting risk of opioid misuse, withdrawal, opioid use disorder or overdose specific to patients with ALBP. No study evaluated the impact of discontinuation of opioids. No studies addressing the effectiveness or harms of prescribing strategies or factors related to clinical decision making around their prescription were identified.

Trials of pharmacologic interventions generally reported on single doses or treatment sessions, so the impact of multiple doses, various dosing strategies and duration of treatment are unclear. Evidence on cannabidiol was limited to a single trial compared with placebo and only single trials for a number of nonpharmacologic interventions were identified. Evidence comparing pharmacologic and nonpharmacologic therapies is sparse. We did not identify any trials that compared opioids with nonpharmacologic treatments and evidence comparing nonopioids with nonpharmacologic treatments was insufficient. Many of the included trials were older and may not reflect current practice. It is likely that patients may employ a combination of pharmacologic and nonpharmacologic treatments for pain management, however concomitant treatments were usually not described in included trials. The variety of potential treatment combinations cannot be captured in studies. Evidence for many nonpharmacologic treatments and for management pathways such as early physical therapy report only on times >4 weeks.

In studies evaluating the accuracy of patient and clinical exam factors for identifying serious spinal pathology, evidence for some factors was confined to a small number of studies or single studies. While likelihood ratios provide useful information on such factors, they are used in conjunction with patient symptoms, presentation and other clinical findings. These may differ by clinical setting (primary care versus ED or tertiary care) as might the reported prevalence of the condition being evaluated. Systematic reviews generally did not pool studies due to substantial heterogeneity in populations, assessment methods, reporting and use of reference standards. The systematic reviews did not confine their inclusion to studies of patients with ALBP. Studies evaluating yellow flag factors and use of screening instruments as predictors for chronic LPB and disability differed in how they defined or analyzed risk factors in different domains and in how they defined worse LBP outcomes. Several risk factors were only evaluated in a relatively small number of studies, making it difficult to reach strong conclusions. None of the studies we reviewed evaluated different potential etiologies of nonradicular pain (e.g., piriformis syndrome, facet joint pain, myofascial pain, sacroiliac pain or discogenic pain) so we do not know whether outcomes from these more specific diagnoses might be associated with disability. However, we are aware of no reliable methods for identifying pain caused by these different potential etiologies, and many experts would lump all of these conditions and others as nonspecific LBP. Finally, we focused on likelihood ratios as the measure of utility for predicting worse outcomes. Likelihood ratios are informative because findings for a specific risk factor or risk prediction instrument can be directly applied to a patient. However, likelihood ratios assess each predictive factor in isolation and do not assess the independent contributions of other potential predictors.

Analyses of the impact of early imaging are limited by the small number of available trials in patients with ALBP, all of which are older. This limited the ability to pool across studies and precluded evaluation of any potential impact of imaging modality on results or modification of effect by patient factors. Multivariate models would provide information regarding predictive utility while accounting for potential confounders and can address collinearity between predictors. However, multivariate models are usually not designed and impractical for clinical use. Clinicians often evaluate patients on the basis of individual predictors, so likelihood ratios may be more applicable clinically. Because individual risk factors are relatively weak, risk prediction instruments could be more helpful than individual yellow flags for predicting outcomes. No instrument has been extensively validated, some validation studies show likelihood ratios similar to estimates for individual predictors, and most instruments include individual items that aren't predictive (such as sex, education level, or previous LBP episodes). No RCTs explicitly comparing the use of screening tools versus not using them or on their use to inform clinical decision making in patients with ALBP (<6 weeks duration) were identified. The trial directly comparing use of the STarT Back instrument to inform a stratified model of care based on a patient's prognosis (low, medium, or high) to not using it that we identified was not in patients with ALBP.[238] Two United States-based pragmatic trials did not directly compare use of STarT Back to not using it.[60,239] One trial randomized clinicians to receive training on its use or not.[239] Clinicians who received the training used the tool for about half of their patients and did not use it to change treatments they recommended. The other cluster randomized trial[60] used the tool to identify high risk patients to receive earlier intensive psychologically informed physical therapy or usual care. It is included in Key Question 8.

Limitations Of The Systematic Review Process

Some limitations of this systematic review should be noted. Most are linked to the limitations of the underlying evidence base. We were unable to perform meta-analysis for many interventions due to the paucity of trials, small sample sizes in many trials, and heterogeneity across trials in study design, populations, study designs and outcomes and methodologic limitations of studies.

Our definition of ALBP of <6 weeks duration may have limited the number of available studies for some interventions. In addition, the duration of back pain was not clearly reported in a number of trials and we may have excluded some that were truly in our population of interest. Variability in pain duration between trials was noted and may have impacted outcomes, but there was insufficient evidence to evaluate the impact of baseline pain duration or severity. The combination of these factors may have resulted in clinical and population heterogeneity across included trials. There was substantial heterogeneity in how and whether trials reported prior LBP episodes and the presence of radiculopathy. Therefore, evidence was insufficient to evaluate the impact of such factors on outcomes.

Across treatments, trials were not designed or powered to evaluate how benefits or harms may be impacted by patient factors, clinical factors or social risk factors. Studies rarely reported on patient social risk factors for health and were not designed to evaluate such factors or the impact of them on management of ALBP and patient outcomes. One study on prediction of chronic LBP found that Black race increased the likelihood of high-impact LBP at 6 months and Hispanic ethnicity slightly increased the likelihood while White race was associated with slightly decreased likelihood.[185] Such findings provide limited information to evaluate or understand the impact of demographic and other factors. There was substantial heterogeneity in how patient

back pain presentation was reported (e.g., prior LBP episodes and the presence or absence of radiculopathy) and evidence was insufficient to evaluate the impact of such factors on outcomes. Trials that did evaluate modification were likely underpowered. In addition, there were too few trials to do stratified analyses across trials on any factor.

Applicability

The applicability of our findings continues to be impacted by several factors. Symptom duration, clinical characteristics, comorbid medical and psychiatric conditions, the presence of radiculopathy and concomitant treatment use were rarely reported. Many trials excluded patients with psychiatric comorbidities or history of substance use disorder. All trials excluded pregnant persons. Thus, the applicability of our findings to clinical populations presenting in primary care, urgent care or emergency settings where there is likely substantial heterogeneity in patient populations is unclear. Heterogeneity in interventions, comparators and co-interventions may impact applicability as well. There is likely substantial variability in how pharmacologic and nonpharmacologic treatments are prescribed in clinical settings and patients likely will employ a broad range of treatments simultaneously to enhance function and relieve pain.

Strengths

This review has some notable strengths. First, we believe that this review provides one of the most comprehensive syntheses of evidence specific to the evaluation and management of ALBP (<6 weeks) to inform development of a clinical guideline. Second, input on Key Questions and populations, interventions, comparators, outcomes, timing, and settings (PICOTS) during topic refinement was obtained from an independent advisory board and Guideline Development Group, both of which include patient advocates or persons with lived experience, in addition to clinicians from a broad range of disciplines, research areas and other perspectives including policy. In addition, input from the U.S. Food and Drug Administration (FDA) and clinical experts from American Academy of Pain Medicine was used to help hone the questions, PICOTS, and Scope. This input complemented perspectives from the clinical and methodological experts of the Pacific Northwest Evidence-based Practice Center (EPC) systematic review team. This review includes the evaluation of medications and various treatments versus placebo or sham treatments, allowing consideration of the impact of nonspecific factors. Lastly, 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 reviews.[18-20,39,40] Using this approach, beneficial effects identified were generally in the small or, less commonly, the moderate range. Effect sizes below the threshold for a small effect were considered to be no effect. 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 for some measures. 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. 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.

Research Gaps

Many research gaps exist precluding strong conclusions regarding the comparative effectiveness and harms of pharmacologic and nonpharmacologic treatment options for ALBP. In particular, there are many gaps related to the role and appropriate prescription of opioids in patients with ALBP. Substantial variation between providers in the rate of prescribing opioids for ALBP has been described in a variety of settings.[239-242] Patient, provider and policy factors likely impact decision making and contribute to the variation. Evidence is lacking for factors related to shared decision-making around opioid use, risk assessment instruments and potential for long-term opioid use for all medications prescribed for ALBP. Similarly, there is a need for evidence on the management of special patient populations with psychiatric comorbidities and those with a history of substance use disorder and how outcomes in such populations may or may not differ from populations without such factors. Further research into patient, provider, and policy factors that impact decision making are needed and should include evaluation of how patients value the risk and harms of therapy options, as well as how to best tailor treatments to patient needs. Long-term evidence on the impact of prescribing opioids for ALBP on potential harms (e.g., opioid use disorder, overdose, disability) is needed. A better understanding of how to best assess risk for problematic long-term opioid use and development of validated instruments to do so is an important area for further research. Well-designed prospective and registry studies may provide opportunities to address these gaps. Adequately powered treatment trials are needed to explore differential effectiveness and harms by patient and clinical factors. Studies designed to evaluate patient demographic and social risk factors and their impact on decision making, management of ALBP, and patient outcomes are also needed. More research is also needed to understand the clinical usefulness of risk prediction instruments for identifying high-risk patients and the optimal strategies to decrease the likelihood of chronic disabling back pain.[243]

Conclusions

There was moderate evidence that oral NSAIDs and muscle relaxants improved function and pain and the combination of muscle relaxants with NSAID moderately improved pain versus NSAID alone. Acetaminophen did not improve pain or function versus placebo. In patients with radiculopathy, systemic steroids improved function and pain versus placebo. Evidence on opioids was sparse; Although one trial reported a small pain improvement for a limited time, others did not. Additionally, no improvement in function was reported and placebo was favored at later time frames. Nonpharmacologic treatments which may improve function and/or pain include manipulation, acupuncture, exercise, early physical therapy, and advice to stay active. In patients with radiculopathy, systemic steroids improved function and pain versus placebo and ESIs were associated with small pain improvements versus usual care, however evidence for these treatments was limited. Adverse events were poorly reported across all treatment trials and serious adverse events appear to be rare. The strength of evidence was low for most outcomes.

Assessment of patient and clinical factors for identification of serious spine pathology and for factors that may predict chronic LBP or disability is important. The evidence for many factors is sparse, however. Psychosocial factors were most predictive of developing persistent LBP. Early imaging in patients with ALBP without symptoms or signs of serious spine pathology does not improve patient outcomes.

Further research in patients with ALBP is needed to evaluate the comparative effectiveness and risks of opioids versus nonopioids and nonpharmacologic management options as well to

evaluate the effects of therapies on longer-term outcomes including long-term opioid use. Adequately powered studies are needed to explore how benefits and harms of therapies may vary by subgroups. Research into patient, provider and policy factors that impact decision making, management and patient outcomes is also needed.

Conclusion Section FROM SR MANUSCRIPT

Assessment of patient and clinical factors to identify serious spine pathology is important; however, few factors were considered highly informative for ruling such pathologies in or out, and the quality of studies for these factors was generally low. Psychosocial factors were the strongest predictors of persistent LBP. Early imaging in patients with ALBP without symptoms or signs of serious spine pathology did not improve patient outcomes but SOE was low.

There was moderate evidence that oral NSAIDs and muscle relaxants improved function and pain within the first few weeks. Acetaminophen did not improve pain or function. Opioids did not improve function and while one trial found small, short-term pain improvement, others did not and the SOE was low. Moderate evidence for benefit at one or more times was seen for nonpharmacologic treatments that may improve function and/or pain including manipulation (mobilization and manipulation), acupuncture, exercise, early PT referral, and advice to stay active. In patients with radiculopathy, systemic steroids improved function and pain versus placebo and ESIs were associated with small pain improvements versus usual care, however evidence for these treatments was limited. AEs across treatments were poorly reported and SAEs were rare.

Further research in patients with ALBP is needed to further delineate the role, benefits, and harms of pharmacologic and nonpharmacologic treatments and to assess longer-term outcomes, including longer term opioid use. Trials are needed to explore subgroup differences in benefits and harms.

References

1. GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. The Lancet. Rheumatology. 2023 Jun;5(6):e316-e29. doi: 10.1016/S2665-9913(23)00098-X. PMID: 37273833.
2. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) results. https://www.healthdata.org/research-analysis/gbd. Accessed August 9 2023.
3. World Health Organization. Low back pain. 2023. https://www.who.int/news-room/fact-sheets/detail/low-back-pain. Accessed August 10 2023. 2023.
4. Dieleman JL, Cao J, Chapin A, et al. US Health Care Spending by Payer and Health Condition, 1996-2016. JAMA. 2020 Mar 3;323(9):863-84. doi: 10.1001/jama.2020.0734. PMID: 32125402.
5. Kim LH, Vail D, Azad TD, et al. Expenditures and Health Care Utilization Among Adults With Newly Diagnosed Low Back and Lower Extremity Pain. JAMA Netw Open. 2019 May 3;2(5):e193676. doi: 10.1001/jamanetworkopen.2019.3676. PMID: 31074820.
6. Kent ML, Tighe PJ, Belfer I, et al. The ACTTION-APS-AAPM Pain Taxonomy (AAAPT) Multidimensional Approach to Classifying Acute Pain Conditions. Pain Med. 2017 May 1;18(5):947-58. doi: 10.1093/pm/pnx019. PMID: 28482098.
7. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ. 2006 Jun 17;332(7555):1430-4. doi: 10.1136/bmj.332.7555.1430. PMID: 16777886.
8. Pengel LH, Herbert RD, Maher CG, et al. Acute low back pain: systematic review of its prognosis. BMJ. 2003 Aug 9;327(7410):323. doi: 10.1136/bmj.327.7410.323. PMID: 12907487.
9. Itz CJ, Geurts JW, van Kleef M, et al. Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. Eur J Pain. 2013 Jan;17(1):5-15. doi: 10.1002/j.1532-2149.2012.00170.x. PMID: 22641374.
10. Mehling WE, Gopisetty V, Bartmess E, et al. The prognosis of acute low back pain in primary care in the United States: a 2-year prospective cohort study. Spine (Phila Pa 1976). 2012 Apr 15;37(8):678-84. doi: 10.1097/BRS.0b013e318230ab20. PMID: 22504516.
11. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? Jama. 2010 Apr 7;303(13):1295-302. doi: 10.1001/jama.2010.344. PMID: 20371789.
12. Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med. 2007 Oct 2;147(7):478-91. doi: 10.7326/0003-4819-147-7-200710020-00006. PMID: 17909209.
13. Spears CA, Hodges SE, Kiyani M, et al. Health Care Resource Utilization and Management of Chronic, Refractory Low Back Pain in the United States. Spine (Phila Pa 1976). 2020 Oct 15;45(20):E1333-E41. doi: 10.1097/BRS.0000000000003572. PMID: 32453242.
14. Chou R, Fu R, Carrino JA, et al. Imaging strategies for low-back pain: systematic review and meta-analysis. Lancet. 2009 Feb 7;373(9662):463-72. doi: 10.1016/s0140-6736(09)60172-0. PMID: 19200918.
15. Koes BW, van Tulder M, Lin CW, et al. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2010 Dec;19(12):2075-94. doi: 10.1007/s00586-010-1502-y. PMID: 20602122.

16. Koes BW, van Tulder MW, Ostelo R, et al. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine (Phila Pa 1976). 2001 Nov 15;26(22):2504-13; discussion 13-4. doi: 10.1097/00007632-200111150-00022. PMID: 11707719.

17. Oliveira CB, Maher CG, Pinto RZ, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2018 Nov;27(11):2791-803. doi: 10.1007/s00586-018-5673-2. PMID: 29971708.

18. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatments for Low Back Pain. Comparative Effectiveness Review No. 169. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 16-EHC004-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2016. https://effectivehealthcare.ahrq.gov/products/back-pain-treatment/research. PMID: 26985522.

19. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review Update. Comparative Effectiveness Review No. 227. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No. 20-EHC009. Rockville, MD: Agency for Healthcare Research and Quality; April 2020. https://effectivehealthcare.ahrq.gov/products/noninvasive-nonpharm-pain-update/research. PMID: 32338846.

20. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review. Comparative Effectiveness Review No. 209. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No 18-EHC013-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2018. https://effectivehealthcare.ahrq.gov/topics/nonpharma-treatment-pain/research-2018. PMID: 30179389.

21. Dowell D, Ragan KR, Jones CM, et al. CDC Clinical Practice Guideline for Prescribing Opioids for Pain - United States, 2022. MMWR Recomm Rep. 2022 Nov 4;71(3):1-95. doi: 10.15585/mmwr.rr7103a1. PMID: 36327391.

22. Qaseem A, Wilt TJ, McLean RM, et al. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med. 2017 Apr 4;166(7):514-30. doi: 10.7326/M16-2367. PMID: 28192789.

23. Framing Opioid Prescribing Guidelines for Acute Pain: Developing the Evidence National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Evidence-Based Clinical Practice Guidelines for Prescribing Opioids for Acute Pain. Framing Opioid Prescribing Guidelines for Acute Pain: Developing the Evidence. Washington (DC): National Academies Press (US); 2020 Dec 19. PMID: 32182005.

24. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Content last reviewed October 2022. Effective Health Care Program, Agency for Healthcare Research and Quality, Rockville, MD. Chapters available at: https://effectivehealthcare.ahrq.gov/products/collections/cer-methods-guide. .

25. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71. PMID: 33782057.

26. Treadwell JR, Singh S, Talati R, et al. A Framework for "Best Evidence" Approaches in Systematic Reviews. AHRQ Methods for Effective Health Care. (Prepared by the ECRI Institute Evidence-based Practice Center under Contract No. HHSA 290-2007-10063-I.) AHRQ Publication No. 11-EHC046-EF. Rockville, MD: Agency for Healthcare Research and Quality (US); 2011.

27. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008. doi: 10.1136/bmj.j4008. PMID: 28935701.

28. Jones CMP, Day RO, Koes BW, et al. Opioid analgesia for acute low back pain and neck pain (the OPAL trial): a randomised placebo-controlled trial. Lancet. 2023 Jul 22;402(10398):304-12. doi: 10.1016/s0140-6736(23)00404-x. PMID: 37392748.

29. Viswanathan M, Ansari MT, Berkman ND, et al. Chapter 9: Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2017. Chapters available at: https://effectivehealthcare.ahrq.gov/products/methods-bias-update/methods.

30. Higgins JP, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.handbook.cochrane.org.

31. 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.

32. Furlan AD, Pennick V, Bombardier C, et al. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976). 2009 Aug 15;34(18):1929-41. doi: 10.1097/BRS.0b013e3181b1c99f. PMID: 19680101.

33. Sterne J, Hernán M, McAleenan A, et al. Chapter 25: Assessing risk of bias in a non-randomized study. In: Cochrane Handbook for Systematic Reviews of Interventions version. 2019;60.

34. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012 Feb;65(2):163-78. doi: 10.1016/j.jclinepi.2011.05.008. PMID: 21959223.

35. U.S. Preventive Services Task Force. U.S. Preventive Services Task Force Procedure Manual. 2015. https://www.uspreventiveservicestaskforce.org/uspstf/procedure-manual. Accessed August 25 2020.

36. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009. PMID: 22007046.

37. Yang B, Mallett S, Takwoingi Y, et al. QUADAS-C: A Tool for Assessing Risk of Bias in Comparative Diagnostic Accuracy Studies. Ann Intern Med. 2021 Nov;174(11):1592-9. doi: 10.7326/M21-2234. PMID: 34698503.

38. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013 Feb 19;158(4):280-6. doi: 10.7326/0003-4819-158-4-201302190-00009. PMID: 23420236.

39. Chou R, Hartung D, Turner J, et al. Opioid Treatments for Chronic Pain. Comparative Effectiveness Review No. 229. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No. 20-EHC011. Rockville, MD: Agency for Healthcare Research and Quality. April 2020. DOI: https://doi.org/10.23970/AHRQEPCCER229. 2020. PMID: 32338848.

40. McDonagh MS, Selph SS, Buckley DI, et al. Nonopioid Pharmacologic Treatments for Chronic Pain. AHRQ Comparative Effectiveness Reviews No. 228. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No. 20-EHC010. Rockville, MD: Agency for Healthcare Research and Quality; April 2020. DOI: https://doi.org/10.23970/AHRQEPCCER228.

41. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1187-97. doi: 10.1016/j.jclinepi.2010.08.010. PMID: 21477993.

42. Morton SC, Murad MH, O'Connor E, et al. AHRQ Methods for Effective Health Care Quantitative Synthesis—An Update. (Prepared by the Scientific Resource Center under Contract No. Oregon Health & Science University–Pacific Northwest EPC Appendix 2, Page 4 290-2012-0004-C). AHRQ Publication No. 18-EHC007-EF. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality (US); 2018.

43. Morton SC, Murad MH, O'Connor E, et al. Quantitative Synthesis—An Update. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 18-EHC007-EF. Rockville, MD: Quality AfHRa; 2018.

44. 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.

45. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

46. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 11-EHC019-EF Rockville, MD: Agency for Healthcare Research and Quality (US); 2010.

47. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011 Jul 22;343:d4002. doi: 10.1136/bmj.d4002. PMID: 21784880.

48. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC medical research methodology. 2003 Nov 10;3:25. doi: 10.1186/1471-2288-3-25. PMID: 14606960.

49. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med. 2006 Mar 21;144(6):427-37. doi: 10.7326/0003-4819-144-6-200603210-00010. PMID: 16549855.

50. McGinn TG, Guyatt GH, Wyer PC, et al. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama. 2000 Jul 5;284(1):79-84. doi: 10.1001/jama.284.1.79. PMID: 10872017.

51. Sackett DL. The rational clinical examination. A primer on the precision and accuracy of the clinical examination. Jama. 1992 May 20;267(19):2638-44. PMID: 1573753.

52. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. Bmj. 2004 Jul 17;329(7458):168-9. doi: 10.1136/bmj.329.7458.168. PMID: 15258077.

53. Han CS, Hancock MJ, Downie A, et al. Red flags to screen for vertebral fracture in people presenting with low back pain. Cochrane Database of Systematic Reviews. 2023(8) PMID: 00075320-100000000-12637.

54. Galliker G, Scherer DE, Trippolini MA, et al. Low Back Pain in the Emergency Department: Prevalence of Serious Spinal Pathologies and Diagnostic Accuracy of Red Flags. The American journal of medicine. 2020 Jan;133(1):60-72 e14. doi: 10.1016/j.amjmed.2019.06.005. PMID: 31278933.

55. Henschke N, Maher CG, Ostelo RW, et al. Red flags to screen for malignancy in patients with low-back pain. The Cochrane database of systematic reviews. 2013 Feb 28;2013(2):CD008686. doi: 10.1002/14651858.CD008686.pub2. PMID: 23450586.

56. Dionne N, Adefolarin A, Kunzelman D, et al. What is the diagnostic accuracy of red flags related to cauda equina syndrome (CES), when compared to Magnetic Resonance Imaging (MRI)? A systematic review. Musculoskelet Sci Pract. 2019 Jul;42:125-33. doi: 10.1016/j.msksp.2019.05.004. PMID: 31132655.

57. Samsamshariat S, Sharifi-Sade M, Zoofaghari S, et al. Efficacy of the combination of indomethacin and methocarbamol versus indomethacin alone in patients with acute low back pain: a double-blind, randomized placebo-controlled clinical trial. Journal of research in pharmacy practice. 2021;Vol.10(2):96-101p. doi: https://doi.org/10.4103/jrpp.JRPP2131. PMID: 34527615.

58. Iliopoulos K, Koufaki P, Tsilikas S, et al. A randomized controlled trial evaluating the short-term efficacy of a single-administration intramuscular injection with the fixed combination of thiocolchicoside-diclofenac versus diclofenac monotherapy in patients with acute moderate-to-severe low back pain. BMC musculoskeletal disorders. 2023;Vol.24(1):476p. doi: https://doi.org/10.1186/s12891-023-06599-0. PMID: 37301824.

59. Evans C, Gilbert JR, Taylor W, et al. A randomized controlled trial of flexion exercises, education, and bed rest for patients with acute low back pain. Physiotherapy canada. 1987;Vol.39(2):96-101p. doi: https://doi.org/10.3138/ptc.39.2.096. PMID: CN-00175532.

60. Delitto A, Patterson CG, Stevans JM, et al. Stratified care to prevent chronic low back pain in high-risk patients: the TARGET trial. A multi-site pragmatic cluster randomized trial. EClinicalMedicine. 2021;34doi: https://doi.org/10.1016/j.eclinm.2021.100795. PMID: 33870150.

61. Takamoto K, Bito I, Urakawa S, et al. Effects of compression at myofascial trigger points in patients with acute low back pain: a randomized controlled trial. European journal of pain (London, England). 2015;Vol.19(8):1186-96p. doi: https://doi.org/10.1002/ejp.694. PMID: 25808188.

62. Ash LM, Modic MT, Obuchowski NA, et al. Effects of diagnostic information, per se, on patient outcomes in acute radiculopathy and low back pain. Ajnr. 2008;American journal of neuroradiology. Vol.29(6):1098-103p. doi: https://doi.org/10.3174/ajnr.A0999. PMID: 18467522.

63. Nuhr M, Hoerauf K, Bertalanffy A, et al. Active warming during emergency transport relieves acute low back pain. Spine. 2004;Vol.29(14):1499-503p. doi: https://doi.org/10.1097/01.brs.0000131439.87553.99. PMID: 15247569.

64. Roberts L, Little P, Chapman J, et al. The back home trial: general practitioner-supported leaflets may change back pain behavior. Spine. 2002;Vol.27(17):1821-8p. doi: https://doi.org/10.1097/00007632-200209010-00002. PMID: 12221342.

65. Bertalanffy A, Kober A, Bertalanffy P, et al. Transcutaneous electrical nerve stimulation reduces acute low back pain during emergency transport. Academic emergency medicine. 2005;Vol.12(7):607-11p. doi: https://doi.org/10.1197/j.aem.2005.01.013. PMID: 15995091.

66. Stankovic R, Johnell O. Conservative treatment of acute low back pain. A 5-year follow-up study of two methods of treatment. Spine. 1995;Vol.20(4):469-72p. doi: https://doi.org/10.1097/00007632-199502001-00010. PMID: 7747231.

67. Miki K, Ikemoto T, Hayashi K, et al. Randomized open-label [corrected] non-inferiority trial of acetaminophen or loxoprofen for patients with acute low back pain. Journal of orthopaedic science. 2018 Netherlands Elsevier (E-mail: orders@springer;Vol.23(3):483-7p. doi: https://doi.org/10.1016/j.jos.2018.02.007. PMID: 29503036

68. Herrmann WA, Geertsen MS. Efficacy and safety of lornoxicam compared with placebo and diclofenac in acute sciatica/lumbo-sciatica: an analysis from a randomised, double-blind, multicentre, parallel-group study. International journal of clinical practice. 2009;Vol.63(11):1613-21p. doi: https://doi.org/10.1111/j.1742-1241.2009.02187.x. PMID: 19832818.

69. Berry H, Hutchinson DR. Tizanidine and ibuprofen in acute low-back pain: results of a double-blind multicentre study in general practice. Journal of international medical research. 1988;Vol.16(2):83-91p. doi: https://doi.org/10.1177/030006058801600202. PMID: 2967781.

70. Berry H, Hutchinson DR. A multicentre placebo-controlled study in general practice to evaluate the efficacy and safety of tizanidine in acute low-back pain. Journal of international medical research. 1988;Vol.16(2):75-82p. doi: https://doi.org/10.1177/030006058801600201. PMID: 2967780.

71. Friedman BW, Irizarry E, Solorzano C, et al. Diazepam Is No Better Than Placebo When Added to Naproxen for Acute Low Back Pain. Annals of emergency medicine. 2017;Vol.70(2):169-76.e1p. doi: https://doi.org/10.1016/j.annemergmed.2016.10.002. PMID: 28187918.

72. Friedman BW, Irizarry E, Chertoff A, et al. Ibuprofen Plus Acetaminophen Versus Ibuprofen Alone for Acute Low Back Pain: an Emergency Department-based Randomized Study. Academic emergency medicine. 2020;Vol.27(3):229-35p. doi: https://doi.org/10.1111/acem.13898. PMID: 31811673.

73. Hackett GI, Seddon D, Kaminski D. Electroacupuncture compared with paracetamol for acute low back pain. Practitioner. 1988;Vol.232(1443):163-4p. PMID: 2973008.

74. Hoiriis KT, Pfleger B, McDuffie FC, et al. A randomized clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain. Journal of manipulative and physiological therapeutics. 2004;Vol.27(6):388-98p. doi: https://doi.org/10.1016/j.jmpt.2004.05.003. PMID: 15319761.

75. Basmajian JV. Acute back pain and spasm. A controlled multicenter trial of combined analgesic and antispasm agents. Spine. 1989;Vol.14(4):438-9p. PMID: 2524114.

76. von Heymann WJ, Schloemer P, Timm J, et al. Spinal high-velocity low amplitude manipulation in acute nonspecific low back pain: a double-blinded randomized controlled trial in comparison with diclofenac and placebo. Spine. 2013;Vol.38(7):540-8p. doi: https://doi.org/10.1097/BRS.0b013e318275d09c. PMID: 23026869.

77. Eken C, Serinken M, Elicabuk H, et al. Intravenous paracetamol versus dexketoprofen versus morphine in acute mechanical low back pain in the emergency department: a randomised double-blind controlled trial. Emergency medicine journal. 2014;Vol.31(3):177-81p. doi: https://doi.org/10.1136/emermed-2012-201670. PMID: 23407378.

78. Bebee B, Taylor DM, Bourke E, et al. The CANBACK trial: a randomised, controlled clinical trial of oral cannabidiol for people presenting to the emergency department with acute low back pain. Medical journal of Australia. 2021;Vol.214(8):370-5p. doi: https://doi.org/10.5694/mja2.51014. PMID: 33846971.

79. Friedman BW, Esses D, Solorzano C, et al. A randomized placebo-controlled trial of single-dose IM corticosteroid for radicular low back pain. Spine. 2008;Vol.33(18):E624-9p. doi: https://doi.org/10.1097/BRS.0b013e3181822711. PMID: 18665021.

80. Lau PM, Chow DH, Pope MH. Early physiotherapy intervention in an Accident and Emergency Department reduces pain and improves satisfaction for patients with acute low back pain: a randomised trial. Australian journal of physiotherapy. 2008;Vol.54(4):243-9p. doi: https://doi.org/10.1016/s0004-9514(08)70003-5. PMID: 19025504.

81. Jellema P, van der Windt DA, van der Horst HE, et al. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomised clinical trial in general practice. BMJ (Clinical research ed.). 2005;Vol.331(7508):84p. doi: https://doi.org/10.1136/bmj.38495.686736.E0. PMID: 15967762.

82. Stankovic R, Johnell O. Conservative treatment of acute low-back pain. A prospective randomized trial: mcKenzie method of treatment versus patient education in "mini back school". Spine. 1990;Vol.15(2):120-3p. PMID: 2139241.

83. Hancock MJ, Maher CG, Latimer J, et al. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. Lancet (london, england). 2007;Vol.370(9599):1638-43p. doi: https://doi.org/10.1016/S0140-6736(07)61686-9.

84. Hussain I, Hussain Shah SI, Amjad I. Efficacy of spinal manual therapy in non-specific acute low back pain. Rawal medical journal. 2013;Vol.38(4):358-60p.

85. Friedman BW, Holden L, Esses D, et al. Parenteral corticosteroids for Emergency Department patients with non-radicular low back pain. Journal of emergency medicine. 2006;Vol.31(4):365-70p. doi: https://doi.org/10.1016/j.jemermed.2005.09.023. PMID: 17046475.

86. Dehghan M, Farahbod F. The efficacy of thermotherapy and cryotherapy on pain relief in patients with acute low back pain, a clinical trial study. Journal of clinical and diagnostic research. 2014;Vol.8(9):LC01-LC4p. doi: https://doi.org/10.7860/JCDR/2014/7404.4818. PMID: 25386469.

87. Pareek A, Chandurkar N, Chandanwale AS, et al. Aceclofenac-tizanidine in the treatment of acute low back pain: a double-blind, double-dummy, randomized, multicentric, comparative study against aceclofenac alone. European spine journal. 2009;Vol.18(12):1836-42p. doi: https://doi.org/10.1007/s00586-009-1019-4. PMID: 19421791.

88. Underwood MR, Morgan J. The use of a back class teaching extension exercises in the treatment of acute low back pain in primary care. Family practice. 1998;Vol.15(1):9-15p. doi: https://doi.org/10.1093/fampra/15.1.9. PMID: 9527292.

89. Innes GD, Croskerry P, Worthington J, et al. Ketorolac versus acetaminophen-codeine in the emergency department treatment of acute low back pain. Journal of emergency medicine. 1998;Vol.16(4):549-56p. doi: https://doi.org/10.1016/s0736-4679(98)00044-4. PMID: 9696169.

90. Gatchel RJ, Polatin PB, Noe C, et al. Treatment- and cost-effectiveness of early intervention for acute low-back pain patients: a one-year prospective study. Journal of occupational rehabilitation. 2003;Vol.13(1):1-9p. doi: https://doi.org/10.1023/a:1021823505774. PMID: 12611026.

91. Amlie E, Weber H, Holme I. Treatment of acute low-back pain with piroxicam: results of a double-blind placebo-controlled trial. Spine. 1987;Vol.12(5):473-6p. doi: https://doi.org/10.1097/00007632-198706000-00010. PMID: 2957801.

92. Modic MT, Obuchowski NA, Ross JS, et al. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. Radiology. 2005;Vol.237(2):597-604p. doi: https://doi.org/10.1148/radiol.2372041509. PMID: 16244269.

93. Szpalski M, Hayez JP. Objective functional assessment of the efficacy of tenoxicam in the treatment of acute low back pain. A double-blind placebo-controlled study. British journal of rheumatology. 1994;Vol.33(1):74-8p. doi: https://doi.org/10.1093/rheumatology/33.1.74. PMID: 8162464.

94. Chok B, Lee R, Latimer J, et al. Endurance training of the trunk extensor muscles in people with subacute low back pain. Physical therapy. 1999;Vol.79(11):1032-42p. PMID: 10534796.

95. Weber H, Holme I, Amlie E. The natural course of acute sciatica with nerve root symptoms in a double-blind placebo-controlled trial evaluating the effect of piroxicam. Spine. 1993;Vol.18(11):1433-8p. PMID: 8235813.

96. Ali AA. Management of the first episode of acute low back pain: a comparison between two treatment protocols (PhD). Texas woman's university. 2002.

97. Ralph L, Look M, Wheeler W, et al. Double-blind, placebo-controlled trial of carisoprodol 250-mg tablets in the treatment of acute lower-back spasm. Current medical research and opinion. 2008;Vol.24(2):551-8p. doi: https://doi.org/10.1185/030079908x261014. PMID: 18194591.

98. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine. 2003;Vol.28(13):1363-71; discussion 72p. doi: https://doi.org/10.1097/01.BRS.0000067115.61673.FF. PMID: 12838091.

99. Tanen DA, Shimada M, Danish DC, et al. Intravenous lidocaine for the emergency department treatment of acute radicular low back pain, a randomized controlled trial. Journal of emergency medicine. 2014;Vol.47(1):119-24p. doi: https://doi.org/10.1016/j.jemermed.2012.12.014. PMID: 24768290.

100. Dreiser RL, Marty M, Ionescu E, et al. Relief of acute low back pain with diclofenac-K 12.5 mg tablets: a flexible dose, ibuprofen 200 mg and placebo-controlled clinical trial. International journal of clinical pharmacology and therapeutics. 2003;Vol.41(9):375-85p. doi: https://doi.org/10.5414/cpp41375. PMID: 14518597.

101. Shin JS, Ha IH, Lee J, et al. Effects of motion style acupuncture treatment in acute low back pain patients with severe disability: a multicenter, randomized, controlled, comparative effectiveness trial. Pain. 2013 Netherlands Elsevier (P;Vol.154(7):1030-7p. doi: https://doi.org/10.1016/j.pain.2013.03.013. PMID: 23639822.

102. Santilli V, Beghi E, Finucci S. Chiropractic manipulation in the treatment of acute back pain and sciatica with disc protrusion: a randomized double-blind clinical trial of active and simulated spinal manipulations. Spine journal. 2006;Vol.6(2):131-7p. doi: https://doi.org/10.1016/j.spinee.2005.08.001. PMID: 16517383.

103. Kocak AO, Ahiskalioglu A, Sengun E, et al. Comparison of intravenous NSAIDs and trigger point injection for low back pain in ED: a prospective randomized study. American journal of emergency medicine. 2019 W;Vol.37(10):1927-31p. doi: https://doi.org/10.1016/j.ajem.2019.01.015. PMID: 30660342.

104. Ketenci A, Ozcan E, Karamursel S. Assessment of efficacy and psychomotor performances of thiocolchicoside and tizanidine in patients with acute low back pain. International journal of clinical practice. 2005;Vol.59(7):764-70p. doi: https://doi.org/10.1111/j.1742-1241.2004.00454.x. PMID: 15963201.

105. Brown FL, Bodison S, Dixon J, et al. Comparison of diflunisal and acetaminophen with codeine in the treatment of initial or recurrent acute low back strain. Clinical therapeutics. 1986;9:52-8. PMID: 2951012.

106. Bergquist-Ullman M, Larsson U. Acute low back pain in industry. A controlled prospective study with special reference to therapy and confounding factors. Acta orthopaedica Scandinavica. 1977;Vol.48(SUPPL. 170):1-117p. PMID: 146394.

107. Machado LA, Maher CG, Herbert RD, et al. The effectiveness of the McKenzie method in addition to first-line care for acute low back pain: a randomized controlled trial. BMC medicine. 2010;8(10)doi: https://doi.org/10.1186/1741-7015-8-10. PMID: 20102596.

108. Dogan C, Yilmaz A, Ozen M, et al. Comparative evaluation of the effectiveness of intravenous paracetamol, dexketoprofen and ibuprofen in acute low back pain. American journal of emergency medicine. 2022;56:223-7. doi: https://doi.org/10.1016/j.ajem.2022.04.017. PMID: 35461026.

109. Ammara U, Shahzadi F, Nazir S, et al. Comparative Effects of Positional Release Technique Versus Deep Transverse Friction Massage on Pain and Disability in patients with Acute Mechanical Low Back Pain. Pakistan Journal of Medical and Health Sciences. 2023;17(4):200-4. doi: 10.53350/pjmhs2023174200.

110. Beneciuk JM, George SZ, Patterson CG, et al. Treatment effect modifiers for individuals with acute low back pain: Secondary analysis of the TARGET trial. Pain. 2023;164(1):171-9. doi: 10.1097/j.pain.0000000000002679. PMID: 35543647.

111. Büyükşireci DE, Demirsoy N, Mit S, et al. Comparison of the Effects of Myofascial Meridian Stretching Exercises and Acupuncture in Patients with Low Back Pain. JAMS Journal of Acupuncture and Meridian Studies. 2022;15(6):347-55. doi: 10.51507/j.jams.2022.15.6.347. PMID: 36537117.

112. Skonnord T, Skjeie H, Brekke M, et al. Acupuncture for acute non-specific low back pain: A randomised, controlled, multicentre intervention study in general practice - The Acuback study. BMJ Open. 2020;10(8)doi: 10.1136/bmjopen-2019-034157. PMID: 32764081.

113. Friedman BW, Dym AA, Davitt M, et al. Naproxen with cyclobenzaprine, oxycodone/acetaminophen, or placebo for treating acute low back pain: A randomized clinical trial. JAMA - Journal of the American Medical Association. 2015;314(15):1572-80. doi: 10.1001/jama.2015.13043. PMID: 26501533.

114. Stark J, Petrofsky J, Berk L, et al. Continuous low-level heatwrap therapy relieves low back pain and reduces muscle stiffness. Physician and Sportsmedicine. 2014;42(4):39-48. doi: 10.3810/psm.2014.11.2090. PMID: 25419887.

115. Goertz CM, Long CR, Hondras MA, et al. Adding chiropractic manipulative therapy to standard medical care for patients with acute low back pain: Results of a pragmatic randomized comparative effectiveness study. Spine. 2013;38(8):627-34. doi: 10.1097/BRS.0b013e31827733e7. PMID: 23060056.

116. Jüni P, Battaglia M, Nüesch E, et al. A randomised controlled trial of spinal manipulative therapy in acute low back pain. Annals of the Rheumatic Diseases. 2009;68(9):1420-7. doi: 10.1136/ard.2008.093757. PMID: 18775942.

117. Jellema P, Van Der Horst HE, Vlaeyen JWS, et al. Predictors of outcome in patients with (sub)acute low back pain differ across treatment groups. Spine. 2006;31(15):1699-705. doi: 10.1097/01.brs.0000224179.04964.aa. PMID: 16816766.

118. Wand BM, Bird C, McAuley JH, et al. Early intervention for the management of acute low back pain: A single-blind randomized controlled trial of biopsychosocial education, manual therapy, and exercise. Spine. 2004;29(21):2350-6. doi: 10.1097/01.brs.0000143619.34308.b4. PMID: 15507794.

119. Seferlis T, Németh G, Carlsson AM, et al. Conservative treatment in patients sick-listed for acute low-back pain: A prospective randomised study with 12 months' follow-up. European Spine Journal. 1998;7(6):461-70. doi: 10.1007/s005860050109. PMID: 9883955.

120. Cherkin DC, Deyo RA, Street JH, et al. Pitfalls of patient education: Limited success of a program for back pain in primary care. Spine. 1996;21(3):345-55. doi: 10.1097/00007632-199602010-00019.

121. Faas A, Van Eijk JTM, Chavannes AW, et al. A randomized trial of exercise therapy in patients with acute low back pain: Efficacy on sickness absence. Spine. 1995;20(8):941-7. PMID: 7644960.

122. Faas A, Chavannes AW, Van Eijk JTM, et al. A randomized, placebo-controlled trial of exercise therapy in patients with acute low back pain. Spine. 1993;18(11):1388-95. PMID: 8235809.

123. Fritz JM, Lane E, McFadden M, et al. Physical Therapy Referral From Primary Care for Acute Back Pain With Sciatica : A Randomized Controlled Trial. Ann Intern Med. 2021 Jan;174(1):8-17. doi: 10.7326/m20-4187. PMID: 33017565.

124. Hasegawa TM, Baptista AS, de Souza MC, et al. Acupuncture for acute non-specific low back pain: a randomised, controlled, double-blind, placebo trial. Acupuncture in medicine : journal of the British Medical Acupuncture Society. 2014 Apr;32(2):109-15. doi: 10.1136/acupmed-2013-010333. PMID: 24316509.

125. Williams CM, Maher CG, Latimer J, et al. Efficacy of paracetamol for acute low-back pain: a double-blind, randomised controlled trial. Lancet. 2014 Nov 1;384(9954):1586-96. doi: 10.1016/s0140-6736(14)60805-9. PMID: 25064594.

126. Friedman BW, Irizarry E, Solorzano C, et al. A Randomized, Placebo-Controlled Trial of Ibuprofen Plus Metaxalone, Tizanidine, or Baclofen for Acute Low Back Pain. Ann Emerg Med. 2019 Oct;74(4):512-20. doi: 10.1016/j.annemergmed.2019.02.017. PMID: 30955985.

127. Malmivaara A, Häkkinen U, Aro T, et al. The treatment of acute low back pain--bed rest, exercises, or ordinary activity? The New England journal of medicine. 1995 Feb 9;332(6):351-5. doi: 10.1056/nejm199502093320602. PMID: 7823996.

128. Friedman BW, Cisewski D, Irizarry E, et al. A Randomized, Double-Blind, Placebo-Controlled Trial of Naproxen With or Without Orphenadrine or Methocarbamol for Acute Low Back Pain. Ann Emerg Med. 2018 Mar;71(3):348-56 e5. doi: 10.1016/j.annemergmed.2017.09.031. PMID: 29089169.

129. Rozenberg S, Delval C, Rezvani Y, et al. Bed rest or normal activity for patients with acute low back pain: a randomized controlled trial. Spine (Phila Pa 1976). 2002 Jul 15;27(14):1487-93. doi: 10.1097/00007632-200207150-00002. PMID: 12131705.

130. Vas J, Aranda JM, Modesto M, et al. Acupuncture in patients with acute low back pain: a multicentre randomised controlled clinical trial. Pain. 2012 Sep;153(9):1883-9. doi: 10.1016/j.pain.2012.05.033. PMID: 22770838.

131. Sheets C, Machado LA, Hancock M, et al. Can we predict response to the McKenzie method in patients with acute low back pain? A secondary analysis of a randomized controlled trial. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2012 Jul;21(7):1250-6. doi: 10.1007/s00586-011-2082-1. PMID: 22109566.

132. Tüzün F, Unalan H, Oner N, et al. Multicenter, randomized, double-blinded, placebo-controlled trial of thiocolchicoside in acute low back pain. Joint Bone Spine. 2003 Sep;70(5):356-61. doi: 10.1016/s1297-319x(03)00075-7. PMID: 14563464.

133. Lasko B, Levitt RJ, Rainsford KD, et al. Extended-release tramadol/paracetamol in moderate-to-severe pain: a randomized, placebo-controlled study in patients with acute low back pain. Curr Med Res Opin. 2012 May;28(5):847-57. doi: 10.1185/03007995.2012.681035. PMID: 22458917.

134. Spijker-Huiges A, Winters JC, van Wijhe M, et al. Steroid injections added to the usual treatment of lumbar radicular syndrome: a pragmatic randomized controlled trial in general practice. BMC Musculoskelet Disord. 2014 Oct 11;15:341. doi: 10.1186/1471-2474-15-341. PMID: 25304934.

135. Wilkinson MJ. Does 48 hours' bed rest influence the outcome of acute low back pain? The British journal of general practice : the journal of the Royal College of General Practitioners. 1995 Sep;45(398):481-4. PMID: 7546872.

136. Li Y, Chen X, Li L, et al. Effect of Acupuncture along Meridians on Pain Degree and Treatment of Acute Lumbar Sprain. Disease markers. 2022;2022:5497805. doi: 10.1155/2022/5497805. PMID: 35915733.

137. Finckh A, Zufferey P, Schurch MA, et al. Short-term efficacy of intravenous pulse glucocorticoids in acute discogenic sciatica. A randomized controlled trial. Spine (Phila Pa 1976). 2006 Feb 15;31(4):377-81. doi: 10.1097/01.brs.0000199917.04145.80. PMID: 16481946.

138. Wiesel SW, Cuckler JM, Deluca F, et al. Acute low-back pain. An objective analysis of conservative therapy. Spine (Phila Pa 1976). 1980 Jul-Aug;5(4):324-30. doi: 10.1097/00007632-198007000-00006. PMID: 6450448.

139. Hofstee DJ, Gijtenbeek JM, Hoogland PH, et al. Westeinde sciatica trial: randomized controlled study of bed rest and physiotherapy for acute sciatica. J Neurosurg. 2002 Jan;96(1 Suppl):45-9. doi: 10.3171/spi.2002.96.1.0045. PMID: 11797655.

140. Goldberg H, Firtch W, Tyburski M, et al. Oral steroids for acute radiculopathy due to a herniated lumbar disk: a randomized clinical trial. JAMA. 2015 May 19;313(19):1915-23. doi: 10.1001/jama.2015.4468. PMID: 25988461.

141. Bayram S, Şahin K, Anarat FB, et al. The effect of oral magnesium supplementation on acute non-specific low back pain: Prospective randomized clinical trial. The American journal of emergency medicine. 2021 Sep;47:125-30. doi: 10.1016/j.ajem.2021.03.078. PMID: 33812327.

142. Nakashima H, Kanemura T, Ando K, et al. Is Pregabalin Effective Against Acute Lumbar Radicular Pain ? Spine surgery and related research. 2019 Jan 25;3(1):61-6. doi: 10.22603/ssrr.2018-0003. PMID: 31435553.

143. Waterworth RF, Hunter IA. An open study of diflunisal, conservative and manipulative therapy in the management of acute mechanical low back pain. The New Zealand medical journal. 1985 May 22;98(779):372-5. PMID: 3157894.

144. Predel HG, Ebel-Bitoun C, Lange R, et al. A randomized, placebo- and active-controlled, multi-country, multi-center parallel group trial to evaluate the efficacy and safety of a fixed-dose combination of 400 mg ibuprofen and 100 mg caffeine compared with ibuprofen 400 mg and placebo in patients with acute lower back or neck pain. J Pain Res. 2019;12:2771-83. doi: 10.2147/jpr.S217045. PMID: 31576162.

145. Magel J, Fritz JM, Greene T, et al. Outcomes of Patients With Acute Low Back Pain Stratified by the STarT Back Screening Tool: Secondary Analysis of a Randomized Trial. Phys Ther. 2017 Mar 1;97(3):330-7. doi: 10.2522/ptj.20160298. PMID: 28204740.

146. Farrell JP, Twomey LT. Acute low back pain. Comparison of two conservative treatment approaches. The Medical journal of Australia. 1982 Feb 20;1(4):160-4. PMID: 6210835.

147. Hadler NM, Curtis P, Gillings DB, et al. A benefit of spinal manipulation as adjunctive therapy for acute low-back pain: a stratified controlled trial. Spine (Phila Pa 1976). 1987 Sep;12(7):702-6. PMID: 2961085.

148. Cruser d A, Maurer D, Hensel K, et al. A randomized, controlled trial of osteopathic manipulative treatment for acute low back pain in active duty military personnel. The Journal of manual & manipulative therapy. 2012 Feb;20(1):5-15. doi: 10.1179/2042618611y.0000000016. PMID: 23372389.

149. Luijsterburg PA, Verhagen AP, Ostelo RW, et al. Physical therapy plus general practitioners' care versus general practitioners' care alone for sciatica: a randomised clinical trial with a 12-month follow-up. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2008 Apr;17(4):509-17. doi: 10.1007/s00586-007-0569-6. PMID: 18172697.

150. Traeger AC, Lee H, Hubscher M, et al. Effect of intensive patient education vs placebo patient education on outcomes in patients with acute lowback pain a randomized clinical trial. JAMA Neurology. 2019 Feb;76(2):161-9. doi: https://dx.doi.org/10.1001/jamaneurol.2018.3376. PMID: 30398542.

151. Boyles WF, Glassmann JM, Soyka JP. Management of acute musculoskeletal conditions: Thoracolumbar strain or sprain. A double-blind evaluation comparing the efficacy and safety of carisoprodol with diazepam. Today's Therapeutic Trends. 1983;1(1):1-16.

152. Glover JR, Morris JG, Khosla T. Back pain: a randomized clinical trial of rotational manipulation of the trunk. Br J Ind Med. 1974 Jan;31(1):59-64. doi: 10.1136/oem.31.1.59. PMID: 4274488.

153. Liu J, Li N. Clinical observation of a combination of acupuncture and drug administration for non-specific acute lumbar sprain. J Acupunct Tuina Sci. 2010;8(1):47-9.

154. Dreiser RL, Le Parc JM, Velicitat P, et al. Oral meloxicam is effective in acute sciatica: two randomised, double-blind trials versus placebo or diclofenac. Inflammation research : official journal of the European Histamine Research Society ... [et al.]. 2001 Mar;50 Suppl 1:S17-23. doi: 10.1007/PL00022375. PMID: 11339516.

155. Goldie I. A clinical trial with indomethacin (indomee(R)) in low back pain and sciatica. Acta Orthop Scand. 1968;39(1):117-28. doi: 10.3109/17453676808989446. PMID: 4239771.

156. Eskin B, Shih RD, Fiesseler FW, et al. Prednisone for emergency department low back pain: a randomized controlled trial. The Journal of emergency medicine. 2014 Jul;47(1):65-70. doi: 10.1016/j.jemermed.2014.02.010. PMID: 24739318.

157. Fritz JM, Magel JS, McFadden M, et al. Early Physical Therapy vs Usual Care in Patients With Recent-Onset Low Back Pain: A Randomized Clinical Trial. JAMA. 2015 Oct 13;314(14):1459-67. doi: 10.1001/jama.2015.11648. PMID: 26461996.

158. Deyo RA, Diehl AK, Rosenthal M. Reducing roentgenography use. Can patient expectations be altered? Arch Intern Med. 1987 Jan;147(1):141-5. doi: 10.1001/archinte.147.1.141. PMID: 2948466.

159. Djais N, Kalim H. The role of lumbar spine radiography in the outcomes of patients with simple acute low back pain. APLAR Journal of Rheumatology. 2005;8:45-50.

160. Balakrishnamoorthy R, Horgan I, Perez S, et al. Does a single dose of intravenous dexamethasone reduce Symptoms in Emergency department patients with low Back pain and RAdiculopathy (SEBRA)? A double-blind randomised controlled trial. Emergency medicine journal : EMJ. 2015 Jul;32(7):525-30. doi: 10.1136/emermed-2013-203490. PMID: 25122642.

161. Dapas F, Hartman SF, Martinez L, et al. Baclofen for the treatment of acute low-back syndrome. A double-blind comparison with placebo. Spine (Phila Pa 1976). 1985 May;10(4):345-9. doi: 10.1097/00007632-198505000-00010. PMID: 2931831.

162. Perrot S, Krause D, Crozes P, et al. Efficacy and tolerability of paracetamol/tramadol (325 mg/37.5 mg) combination treatment compared with tramadol (50 mg) monotherapy in patients with subacute low back pain: a multicenter, randomized, double-blind, parallel-group, 10-day treatment study. Clin Ther. 2006 Oct;28(10):1592-606. doi: 10.1016/j.clinthera.2006.10.001. PMID: 17157115.

163. Serfer GT, Wheeler WJ, Sacks HJ. Randomized, double-blind trial of carisoprodol 250 mg compared with placebo and carisoprodol 350 mg for the treatment of low back spasm. Curr Med Res Opin. 2010 Jan;26(1):91-9. doi: 10.1185/03007990903382428. PMID: 19908948.

164. Darlow B, Stanley J, Dean S, et al. The Fear Reduction Exercised Early (FREE) approach to management of low back pain in general practice: A pragmatic cluster-randomised controlled trial. PLoS medicine. 2019 Sep;16(9):e1002897. doi: 10.1371/journal.pmed.1002897. PMID: 31498799.

165. Lindequist S, Lundberg B, Wikmark R, et al. Information and regime at low back pain. Scandinavian journal of rehabilitation medicine. 1984;16(3):113-6. PMID: 6238404.

166. MacDonald RS, Bell CM. An open controlled assessment of osteopathic manipulation in nonspecific low-back pain. Spine (Phila Pa 1976). 1990 May;15(5):364-70. doi: 10.1097/00007632-199005000-00005. PMID: 2141951.

167. Rhon DI, Miller RB, Fritz JM. Effectiveness and Downstream Healthcare Utilization for Patients That Received Early Physical Therapy Versus Usual Care for Low Back Pain: A Randomized Clinical Trial. Spine (Phila Pa 1976). 2018 Oct 1;43(19):1313-21. doi: 10.1097/BRS.0000000000002619. PMID: 29489568.

168. Varrassi G, Hanna M, Coaccioli S, et al. Dexketoprofen Trometamol and Tramadol Hydrochloride Fixed-Dose Combination in Moderate to Severe Acute Low Back Pain: a Phase IV, Randomized, Parallel Group, Placebo, Active-Controlled Study (DANTE). Pain and therapy. 2024;Vol.13(4):1007-22p. doi: https://doi.org/10.1007/s40122-024-00623-4. PMID: CN-02723612 NEW.

169. Hung KKC, Lam RPK, Lee HKH, et al. Comparison of diclofenac with tramadol, tizanidine or placebo in the treatment of acute low back pain and sciatica: multi-center randomized controlled trial. Postgraduate medical journal. 2024 Sep 22;100(1188):741-50. doi: 10.1093/postmj/qgae052. PMID: 38679808.

170. Hancock MJ, Maher CG, Latimer J, et al. Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. European journal of pain (London, England). 2009;Vol.13(1):51-5p. doi: https://doi.org/10.1016/j.ejpain.2008.03.007. PMID: 18448369.

171. Brennan GP, Fritz JM, Hunter SJ, et al. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. Spine. 2006;Vol.31(6):623-31p. doi: https://doi.org/10.1097/01.brs.0000202807.72292.a8. PMID: 16540864.

172. Lee SS, Choi Y, Pransky GS. Extent and Impact of Opioid Prescribing for Acute Occupational Low Back Pain in the Emergency Department. Journal of Emergency Medicine. 2016;50(3):376-84. doi: 10.1016/j.jemermed.2015.10.015. PMID: 26754559.

173. Mehling WE, Avins AL, Acree MC, et al. Can a back pain screening tool help classify patients with acute pain into risk levels for chronic pain? European Journal of Pain (United Kingdom). 2015a;19(3):439-46. doi: 10.1002/ejp.615. PMID: 25381748.

174. Mehling WE, Ebell MH, Avins AL, et al. Clinical decision rule for primary care patient with acute low back pain at risk of developing chronic pain. Spine Journal. 2015b;15(7):1577-86. doi: 10.1016/j.spinee.2015.03.003.

175. Webster BS, Bauer AZ, Choi Y, et al. Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. Spine. 2013;38(22):1939-46. doi: 10.1097/BRS.0b013e3182a42eb6. PMID: 23883826.

176. Webster BS, Verma SK, Gatchel RJ. Relationship between early opioid prescribing for acute occupational low back pain and disability duration, medical costs, subsequent surgery and late opioid use. Spine. 2007;32(19):2127-32. doi: 10.1097/BRS.0b013e318145a731. PMID: 17762815.

177. Tan CIC, Liaw JSC, Jiang B, et al. Predicting outcomes of acute low back pain patients in emergency department: A prospective observational cohort study. Medicine (Baltimore). 2018 Jun;97(26):e11247. doi: 10.1097/md.0000000000011247. PMID: 29952991.

178. Pozzobon D, Nogueira LAC, Ferreira PH, et al. Return to self-reported physical activity level after an event of acute low back pain. PLoS One. 2019;14(7):e0219556. doi: 10.1371/journal.pone.0219556. PMID: 31291377.

179. Friedman BW, Gensler S, Yoon A, et al. Predicting three-month functional outcomes after an ED visit for acute low back pain. The American journal of emergency medicine. 2017 Feb;35(2):299-305. doi: 10.1016/j.ajem.2016.11.014. PMID: 27856138.

180. Williams CM, Hancock MJ, Maher CG, et al. Predicting rapid recovery from acute low back pain based on the intensity, duration and history of pain: a validation study. Eur J Pain. 2014 Sep;18(8):1182-9. doi: 10.1002/j.1532-2149.2014.00467.x. PMID: 24648103.

181. Webster BS, Choi Y, Bauer AZ, et al. The cascade of medical services and associated longitudinal costs due to nonadherent magnetic resonance imaging for low back pain. Spine (Phila Pa 1976). 2014 Aug 1;39(17):1433-40. doi: 10.1097/brs.0000000000000408. PMID: 24831502.

182. Shaw B, Kinsella R, Henschke N, et al. Back pain "red flags": which are most predictive of serious pathology in the Emergency Department? European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2020 Aug;29(8):1870-8. doi: 10.1007/s00586-020-06452-1. PMID: 32495276.

183. Fuhro FF, Fagundes FRC, Manzoni ACT, et al. Discriminative and Predictive Analysis of the Brazilian Version of the Örebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ) Short-Form in Patients With Low Back Pain. J Chiropr Med. 2021 Dec;20(4):191-8. doi: 10.1016/j.jcm.2021.12.009. PMID: 35496720.

184. Kim HS, Ciolino JD, Lancki N, et al. A Prospective Observational Study of Emergency Department-Initiated Physical Therapy for Acute Low Back Pain. Phys Ther. 2021 Mar 3;101(3)doi: 10.1093/ptj/pzaa219. PMID: 33351942.

185. Roseen EJ, Smith CN, Essien UR, et al. Racial and Ethnic Disparities in the Incidence of High-Impact Chronic Pain Among Primary Care Patients with Acute Low Back Pain: A Cohort Study. Pain Med. 2023 Jun 1;24(6):633-43. doi: 10.1093/pm/pnac193. PMID: 36534910.

186. Law RK, Lee EW, Law SW, et al. The predictive validity of OMPQ on the rehabilitation outcomes for patients with acute and subacute non-specific LBP in a Chinese population. J Occup Rehabil. 2013 Sep;23(3):361-70. doi: 10.1007/s10926-012-9404-y. PMID: 23179745.

187. Grimmer-Somers K, Prior M, Robertson J. Yellow flag scores in a compensable New Zealand cohort suffering acute low back pain. J Pain Res. 2008 Dec 1;1:15-25. doi: 10.2147/jpr.s3710. PMID: 21197284.

188. Graves JM, Fulton-Kehoe D, Jarvik JG, et al. Early imaging for acute low back pain: one-year health and disability outcomes among Washington State workers. Spine (Phila Pa 1976). 2012 Aug 15;37(18):1617-27. doi: 10.1097/BRS.0b013e318251887b. PMID: 22415000.

189. Jellema P, van der Windt DA, van der Horst HE, et al. Prediction of an unfavourable course of low back pain in general practice: comparison of four instruments. The British journal of general practice : the journal of the Royal College of General Practitioners. 2007 Jan;57(534):15-22. PMID: 17244419.

190. Tsang Chi Chung R, Lau Sau Ying J, Kwong So Fong S, et al. Reliability, Construct and Predictive Validity of the Hong Kong Chinese Orebro Musculoskeletal Pain Screening Questionnaire. J Occup Rehabil. 2017 Dec;27(4):584-92. doi: 10.1007/s10926-016-9689-3. PMID: 28028688.

191. Tsang RCC, Lee EWC, Lau JSY, et al. Development of a short form of the Hong Kong Chinese orebro musculoskeletal pain screening questionnaire. Hong Kong Physiother J. 2019 Jun;39(1):57-66. doi: 10.1142/s1013702519500057. PMID: 31156317.

192. Spijker-Huiges A, Vermeulen K, Winters JC, et al. Costs and cost-effectiveness of epidural steroids for acute lumbosacral radicular syndrome in general practice: An economic evaluation alongside a pragmatic randomized control trial. Spine. 2014;39(24):2007-12. doi: 10.1097/BRS.0000000000000597. PMID: 25202937.

193. Luijsterburg PAJ, Lamers LM, Verhagen AP, et al. Cost-effectiveness of physical therapy and general practitioner care for sciatica. Spine. 2007;32(18):1942-8. doi: 10.1097/BRS.0b013e31813162f9. PMID: 17700438.

194. Jellema P, van der Roer N, van der Windt DA, et al. Low back pain in general practice: cost-effectiveness of a minimal psychosocial intervention versus usual care. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2007 Nov;16(11):1812-21. doi: 10.1007/s00586-007-0439-2. PMID: 17659363.

195. Skonnord T, Fetveit A, Skjeie H, et al. Cost-effectiveness analysis of acupuncture compared with usual care for acute non-specific low back pain: secondary analysis of a randomised controlled trial. Acupuncture in medicine : journal of the British Medical Acupuncture Society. 2022 Apr;40(2):123-32. doi: 10.1177/09645284211055747. PMID: 34847780.

196. Liang M, Komaroff AL. Roentgenograms in primary care patients with acute low back pain: a cost-effectiveness analysis. Arch Intern Med. 1982 Jun;142(6):1108-12. PMID: 6212032.

197. Koontz NA, Wiggins RH, 3rd, Mills MK, et al. Less Is More: Efficacy of Rapid 3D-T2 SPACE in ED Patients with Acute Atypical Low Back Pain. Academic radiology. 2017 Aug;24(8):988-94. doi: 10.1016/j.acra.2017.02.011. PMID: 28385420.

198. Reito A, Kyrölä K, Pekkanen L, et al. Specific spinal pathologies in adult patients with an acute or subacute atraumatic low back pain in the emergency department. Int Orthop. 2018 Dec;42(12):2843-9. doi: 10.1007/s00264-018-3983-y. PMID: 29845368.

199. Henschke N, Maher CG, Refshauge KM, et al. Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. Arthritis Rheum. 2009 Oct;60(10):3072-80. doi: 10.1002/art.24853. PMID: 19790051.

200. Ala A, Shams-vahdati S, Meshkini A, et al. Diagnostic value of Lasegue test in discopathy in patients with acute low back pain. Life Science Journal. 2012;9:4315–8.

201. Thiruganasambandamoorthy V, Turko E, Ansell D, et al. Risk factors for serious underlying pathology in adult emergency department nontraumatic low back pain patients. The Journal of emergency medicine. 2014 Jul;47(1):1-11. doi: 10.1016/j.jemermed.2013.08.140. PMID: 24725822.

202. Cherkin DC, Deyo RA, Street JH, et al. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. Spine (Phila Pa 1976). 1996 Dec 15;21(24):2900-7. doi: 10.1097/00007632-199612150-00023. PMID: 9112715.

203. Dionne CE, Bourbonnais R, Frémont P, et al. Determinants of "return to work in good health" among workers with back pain who consult in primary care settings: a 2-year prospective study. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2007 May;16(5):641-55. doi: 10.1007/s00586-006-0180-2. PMID: 16868783.

204. Dixon AN, Gatchel R. Gender and parental status as predictors of low back pain disability: a prospective study. J Occup Rehabil. 1999;9(3):195-200.

205. Fransen M, Woodward M, Norton R, et al. Risk factors associated with the transition from acute to chronic occupational back pain. Spine (Phila Pa 1976). 2002 Jan 1;27(1):92-8. doi: 10.1097/00007632-200201010-00022. PMID: 11805644.

206. Grotle M, Brox JI, Veierød MB, et al. Clinical course and prognostic factors in acute low back pain: patients consulting primary care for the first time. Spine (Phila Pa 1976). 2005 Apr 15;30(8):976-82. doi: 10.1097/01.brs.0000158972.34102.6f. PMID: 15834343.

207. Grotle M, Vøllestad NK, Brox JI. Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II. Spine (Phila Pa 1976). 2006 Apr 20;31(9):1038-46. doi: 10.1097/01.brs.0000214878.01709.0e. PMID: 16641782.

208. Heneweer H, Aufdemkampe G, Van Tulder MW, et al. Psychosocial variables in patients with (sub)acute low back pain: An inception cohort in primary care physical therapy in the Netherlands. Spine. 2007;32(5):586-92. doi: 10.1097/01.brs.0000256447.72623.56.

209. Henschke N, Maher CG, Refshauge KM, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. BMJ. 2008 Jul 7;337(7662):a171. doi: 10.1136/bmj.a171. PMID: 18614473.

210. Poiraudeau S, Rannou F, Le Henanff A, et al. Outcome of subacute low back pain: influence of patients' and rheumatologists' characteristics. Rheumatology (Oxford). 2006 Jun;45(6):718-23. doi: 10.1093/rheumatology/kei231. PMID: 16377729.

211. Schiøttz-Christensen B, Nielsen GL, Hansen VK, et al. Long-term prognosis of acute low back pain in patients seen in general practice: a 1-year prospective follow-up study. Fam Pract. 1999 Jun;16(3):223-32. doi: 10.1093/fampra/16.3.223. PMID: 10439974.

212. Singer J, Gilbert JR, Hutton T, et al. Predicting outcome in acute low-back pain. Canadian family physician Medecin de famille canadien. 1987 Mar;33:655-9. PMID: 21263854.

213. Swinkels-Meewisse IE, Roelofs J, Schouten EG, et al. Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. Spine (Phila Pa 1976). 2006 Mar 15;31(6):658-64. doi: 10.1097/01.brs.0000203709.65384.9d. PMID: 16540870.

214. Thomas E, Silman AJ, Croft PR, et al. Predicting who develops chronic low back pain in primary care: a prospective study. Bmj. 1999 Jun 19;318(7199):1662-7. doi: 10.1136/bmj.318.7199.1662. PMID: 10373170.

215. Truchon M, Côté D. Predictive validity of the Chronic Pain Coping Inventory in subacute low back pain. Pain. 2005 Aug;116(3):205-12. doi: 10.1016/j.pain.2005.04.003. PMID: 15927382.

216. Turner JA, Franklin G, Fulton-Kehoe D, et al. ISSLS prize winner: early predictors of chronic work disability: a prospective, population-based study of workers with back injuries. Spine (Phila Pa 1976). 2008 Dec 1;33(25):2809-18. doi: 10.1097/BRS.0b013e31817df7a7. PMID: 19050587.

217. Turner JA, Franklin G, Fulton-Kehoe D, et al. Worker recovery expectations and fear-avoidance predict work disability in a population-based workers' compensation back pain sample. Spine. 2006;31(6):682-9. doi: 10.1097/01.brs.0000202762.88787.af.

218. Hazard RG, Haugh LD, Reid S, et al. Early prediction of chronic disability after occupational low back injury. Spine (Phila Pa 1976). 1996 Apr 15;21(8):945-51. doi: 10.1097/00007632-199604150-00008. PMID: 8726198.

219. Linton SJ, Halldén K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. Clinical Journal of Pain. 1998;14(3):209-15. doi: 10.1097/00002508-199809000-00007.

220. Cats-Baril WL, Frymoyer JW. Identifying patients at risk of becoming disabled because of low-back pain. The Vermont Rehabilitation Engineering Center predictive model. Spine (Phila Pa 1976). 1991 Jun;16(6):605-7. doi: 10.1097/00007632-199106000-00001. PMID: 1830688.

221. Deyo RA, Diehl AK. Psychosocial predictors of disability in patients with low back pain. The Journal of rheumatology. 1988 Oct;15(10):1557-64. PMID: 2974489.

222. Dionne CE, Bourbonnais R, Frémont P, et al. A clinical return-to-work rule for patients with back pain. Cmaj. 2005 Jun 7;172(12):1559-67. doi: 10.1503/cmaj.1041159. PMID: 15939915.

223. Hazard RG, Haugh LD, Reid S, et al. Early physician notification of patient disability risk and clinical guidelines after low back injury: a randomized, controlled trial. Spine. 1997;Vol.22(24):2951-8p. doi: https://doi.org/10.1097/00007632-199712150-00019. PMID: 9431632.

224. McIntosh G, Frank J, Hogg-Johnson S, et al. Prognostic factors for time receiving workers' compensation benefits in a cohort of patients with low back pain. Spine (Phila Pa 1976). 2000 Jan 15;25(2):147-57. doi: 10.1097/00007632-200001150-00003. PMID: 10685477.

225. Grotle M, Vøllestad NK, Brox JI. Screening for yellow flags in first-time acute low back pain: reliability and validity of a Norwegian version of the Acute Low Back Pain Screening Questionnaire. The Clinical journal of pain. 2006 Jun;22(5):458-67. doi: 10.1097/01.ajp.0000208243.33498.cb. PMID: 16772801.

226. Fulton-Kehoe D, Stover BD, Turner JA, et al. Development of a brief questionnaire to predict long-term disability. J Occup Environ Med. 2008 Sep;50(9):1042-52. doi: 10.1097/JOM.0b013e31817d361e. PMID: 18784553.

227. Hancock MJ, Maher CG, Latimer J, et al. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. European spine journal. 2008;Vol.17(7):936-43p. doi: https://doi.org/10.1007/s00586-008-0679-9. PMID: 18427840.

228. Deyo RA, Diehl AK. Cancer as a cause of back pain: frequency, clinical presentation, and diagnostic strategies. J Gen Intern Med. 1988 May-Jun;3(3):230-8. doi: 10.1007/bf02596337. PMID: 2967893.

229. Keefe FJ, Brown GK, Wallston KA, et al. Coping with rheumatoid arthritis pain: catastrophizing as a maladaptive strategy. Pain. 1989 Apr;37(1):51-6. doi: 10.1016/0304-3959(89)90152-8. PMID: 2726278.

230. Waddell G, Newton M, Henderson I, et al. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993 Feb;52(2):157-68. doi: 10.1016/0304-3959(93)90127-b. PMID: 8455963.

231. Waddell G, McCulloch JA, Kummel E, et al. Nonorganic physical signs in low-back pain. Spine (Phila Pa 1976). 1980 Mar-Apr;5(2):117-25. doi: 10.1097/00007632-198003000-00005. PMID: 6446157.

232. Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. The Clinical journal of pain. 2003 Mar-Apr;19(2):80-6. doi: 10.1097/00002508-200303000-00002. PMID: 12616177.

233. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum. 2008 May 15;59(5):632-41. doi: 10.1002/art.23563. PMID: 18438893.

234. Khobrani M, Perona S, Patanwala AE. Effect of a legislative mandate on opioid prescribing for back pain in the emergency department. The American journal of emergency medicine. 2019 Nov;37(11):2035-8. doi: 10.1016/j.ajem.2019.02.031. PMID: 30824275.

235. Sun BC, Charlesworth CJ, Lupulescu-Mann N, et al. Effect of Automated Prescription Drug Monitoring Program Queries on Emergency Department Opioid Prescribing. Ann Emerg Med. 2018 Mar;71(3):337-47 e6. doi: 10.1016/j.annemergmed.2017.10.023. PMID: 29248333.

236. Alrwaily M, Timko M, Schneider M, et al. Treatment-Based Classification System for Low Back Pain: Revision and Update. Phys Ther. 2016 Jul;96(7):1057-66. doi: 10.2522/ptj.20150345. PMID: 26637653.

237. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Phys Ther. 1995 Jun;75(6):470-85; discussion 85-9. doi: 10.1093/ptj/75.6.470. PMID: 7770494.

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. Cherkin D, Balderson B, Wellman R, et al. Effect of Low Back Pain Risk-Stratification Strategy on Patient Outcomes and Care Processes: the MATCH Randomized Trial in Primary Care. J Gen Intern Med. 2018 Aug;33(8):1324-36. doi: 10.1007/s11606-018-4468-9. PMID: 29790073.

240. Hanley K, Zabar S, Altshuler L, et al. Opioid vs nonopioid prescribers: Variations in care for a standardized acute back pain case. Substance abuse. 2017 Jul-Sep;38(3):324-9. doi: 10.1080/08897077.2017.1319894. PMID: 28586281.

241. Hoppe JA, McStay C, Sun BC, et al. Emergency Department Attending Physician Variation in Opioid Prescribing in Low Acuity Back Pain. The western journal of emergency medicine. 2017 Oct;18(6):1135-42. doi: 10.5811/westjem.2017.7.33306. PMID: 29085548.

242. Webster BS, Cifuentes M, Verma S, et al. Geographic variation in opioid prescribing for acute, work-related, low back pain and associated factors: a multilevel analysis. American journal of industrial medicine. 2009 Feb;52(2):162-71. doi: 10.1002/ajim.20655. PMID: 19016267.

243. van der Windt D, Hay E, Jellema P, et al. Psychosocial interventions for low back pain in primary care: lessons learned from recent trials. Spine (Phila Pa 1976). 2008 Jan 1;33(1):81-9. doi: 10.1097/BRS.0b013e31815e39f9. PMID: 18165752.

Abbreviations

Acronym	Definition
4DSQ	four-dimensional symptom questionnaire
ACP	American College of Physicians
ADL	activity of daily living
ADS	Activities Discomfort Scale
AE	adverse event
AHRQ	Agency for Healthcare Research and Quality
ALBP	acute low back pain
ALBPSQ	Acute Low Back Pain Screening Questionnaire
APAP	acetaminophen
BMI	body mass index
BPFS	Back Pain Function Scale
CBD	cannabidiol
CDC	Centers for Disease Control and Prevention
CEA	cost effectiveness analysis
CES-D	Center for Epidemiologic Studies Depression Scale
CI	confidence interval
CNS	central nervous system
CPG	clinical practice guideline
CRP	C-reactive protein
CSQ	coping strategies questionnaire
CT	computed tomography
CUA	cost-utility analysis
DASS	Depression, Anxiety, and Stress Scale
DI	Disability Index
ED	emergency department
EPC	Evidence-based Practice Center
EQ-5D	EuroQol 5-dimension
EQ-5D-3L	EuroQol 5-dimension 3-level
ESI	epidural steroid injection
ESR	erythrocyte sedimentation rate
FABQ	fear-avoidance beliefs questionnaire
FDA	U.S. Food and Drug Administration
FDC	fixed dose combination
FREE	Fear Reduction Exercised Early
GI	gastrointestinal
GPE	Global Perceived Effect
HRQOL	health-related quality of life
HVLA	high velocity, low amplitude
ICER	incremental cost-effectiveness ratio
IM	intramuscular
IP	immediately post-treatment
IQR	interquartile range
IV	intravenous
KQ	Key Question
LBP	low back pain
LR	likelihood ratio
LR+	positive likelihood ratio
LR-	negative likelihood ratio
LRS	lumbosacral radicular syndrome
LS	lumbar spine
MCS	mental component scale
MCT	medium chain triglyceride
MD	mean difference
MED	morphine equivalent dose
MET	muscle energy technique

Acronym	Definition
MIS	minimal intervention strategy
MME	morphine milligram equivalents used
MPQ	McGill Pain Questionnaire
MR	muscle relaxant
MRI	magnetic resonance imaging
MT	manual therapy
NA	not applicable
NASEM	National Academies of Science, Engineering and Medicine
NC	not calculable
NHP	Nottingham Health Profile Mobility
NPRS	Numeric Pain Rating Scale
NR	not reported
NRS	Numeric Rating Scale
NRSI	nonrandomized study of interventions
NSAID	nonsteroid anti-inflammatory drug
ODI	Oswestry Disability Index
ÖMPSQ	Örebro Musculoskeletal Pain Screening Questionnaire
OMT	osteopathic manipulative treatment
PCS	physical component scale
PDMP	prescription drug monitoring programs
PICOTS	populations, interventions, comparators, outcomes, timing, and settings
PL	profile likelihood
PL-	negative profile likelihood
PL+	positive profile likelihood
PT	physical therapy
QALY	quality adjusted life years
QUIPS	Quality in Prognosis Studies
RCT	randomized controlled trial
RDQ	Roland-Morris Disability Questionnaire
ROB	risk of bias
RR	risk ratio
SAE	serious adverse event
SD	standard deviation
SF-12	12-item Short Form Questionnaire
SF-36	36-item Short Form Questionnaire
SMD	standardized mean difference
SOE	strength of evidence
SR	systematic review
TENS	transcutaneous electrical nerve stimulation
TPI	trigger point injection
UC	usual care
VAS	visual analogue scale
VDPQ	Vermont Disability Prediction Questionnaire