RevMan: review – overview; 638917102400482353 (version 1.1)

Status: UNPUBLISHED DRAFT

Pharmacological treatments for low back pain in adults: an overview of Cochrane Reviews

Editors: Cochrane Back and Neck Group

Editors: Cochrane Musculoskeletal Group

Contact Person: Aidan G Cashin (a.cashin@neura.edu.au)

Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia

Aidan G Cashin[1][2]Benedict M Wand[3]Neil E O'Connell[4]Hopin Lee[5][6]

Rodrigo RN Rizzo[1][2]Matthew K Bagg[1][7][8]Edel O'Hagan[1][7]

Christopher G Maher[9][10]Andrea D Furlan[11]Maurits W van Tulder[12]

James H McAuley[1][2]

[1] Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia

[2] School of Health Sciences, Faculty of Medicine & Health, University of New South Wales, Sydney, Australia

[3] School of Physiotherapy, The University of Notre Dame Australia, Fremantle, Australia

[4] Department of Health Sciences, Centre for Health and Wellbeing Across the Lifecourse, Brunel University London, Uxbridge, UK

[5] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Oxford, UK

[6] School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

[7] Prince of Wales Clinical School, Faculty of Medicine, The University of New South Wales, Sydney, Australia

[8] New College Village, University of New South Wales, Sydney, Australia

[9] Sydney School of Public Health, The University of Sydney, Sydney, Australia

[10] Institute for Musculoskeletal Health, The University of Sydney and Sydney Local Health District, Sydney, Australia

[11] Institute for Work & Health, Toronto, Canada

[12] Department of Health Sciences, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, Netherlands

Citation

Cashin AG, Wand BM, O'Connell NE, Lee H, Rizzo RRN, Bagg MK, O'Hagan E, Maher CG, Furlan AD, van Tulder MW, McAuley JH. Pharmacological treatments for low back pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews TBD, Issue TBD. Art. No.: CD013815. DOI: 10.1002/14651858.CD013815.

Dates

Revision published: Issue TBD, TBD (TBD)

Version published (citation changed): Issue TBD, TBD (TBD)

Review first published: Issue TBD, TBD

Protocol first published: N/A

Abstract

Background

Low back pain (LBP) is a global health problem, which is increasing with the ageing and growing population. Pharmacological interventions are the treatment option most used by people with LBP to manage their pain. Cochrane Reviews have investigated the effects of pharmacological interventions for treating LBP and are available to decision-makers through the Cochrane Library. However, there are multiple reviews, of varying currency, scope and methodology which may inhibit decision-makers' access and use of this evidence.

Objectives

To summarise the evidence from Cochrane reviews of the efficacy, effectiveness, and safety of systemic pharmacological interventions for adults with non-specific LBP.

Methods

We searched the Cochrane Database of Systematic Reviews (from inception to 3 June 2021) to identify reviews investigating pharmacological interventions for people with LBP. We included reviews of randomised controlled trials that included adults (≥18 years) with non-specific LBP. Two Overview authors independently assessed eligibility, extracted data and assessed the quality of the reviews and certainty of the evidence using AMSTAR 2 and GRADE tools respectively. We were primarily interested in placebo comparisons reporting data on our main outcomes pain intensity, function and safety. We presented the short-term efficacy on pain intensity in a 'Summary of findings' table and presented the results of the remaining comparisons and outcomes at each time point in an 'Overview of reviews' table. We also presented a 'Summary of results' table which assigned each intervention to a category highlighting the effect size and certainty of the evidence.

Main results

We included seven Cochrane reviews that included 103 unique trials that randomised 22,238 participants. Based on the AMSTAR 2 assessment of methodological quality we have high confidence in the findings of seven reviews, moderate confidence in the findings of one review, and low confidence in the findings of one review. The reviews reported data on six distinct medicines or medicine classes: paracetamol, non-steroidal anti-inflammatory drugs (NSAIDs), muscle relaxants, benzodiazepines, opioids, and antidepressants. Three reviews included participants with acute or sub-acute LBP and five reviews included participants with chronic low back pain.

We have summarised the published evidence as outlined in the Cochrane reviews and present key findings for our primary comparison and outcome pain intensity for acute and chronic low back pain.

Acute LBP

Paracetamol

There was high certainty evidence for no evidence of difference between paracetamol and placebo for reducing pain intensity (MD of 0.49/100 (95% Confidence Interval (CI) -1.99 to 2.97)) at short-term follow up (≤ 3 months postintervention).

NSAIDs

There was moderate certainty evidence for a small between group difference favouring NSAIDs compared to placebo for reducing pain intensity (MD of -7.29/100 (95% CI -10.98 to -3.61)) at short-term follow up.

Muscle Relaxants and Benzodiazepines

There was moderate certainty evidence for a small between group difference favouring antispasmodic muscle relaxants compared to placebo for a reduced risk of not getting pain relief (RR of 0.58 (95% CI 0.45 to 0.76)) at short-term follow up.

Opioids

No reviews aimed to identify evidence for people with acute LBP.

Antidepressants

No evidence was identified for participants with acute LBP.

Chronic LBP

Paracetamol

No evidence was identified for participants with chronic LBP.

NSAIDs

There was low certainty evidence for a small between group difference favouring NSAIDs compared to placebo for reducing pain intensity (MD of -6.97/100 (95% CI -10.74 to -3.19)) at intermediate term follow up (> 3 months and ≤ 12 months postintervention).

Muscle Relaxants and Benzodiazepines

There was low certainty evidence for a small between group difference favouring benzodiazepines compared to placebo for reducing the risk of not getting pain relief (RR of 0.71 (95% CI 0.54 to 0.93)) at short-term follow up.

Opioids

There was high certainty evidence for a small between group difference favouring tapentadol compared to placebo for reducing pain intensity (MD of -8.00/100 (95% CI -1.22 to -0.38)), moderate certainty evidence for a small between group difference favouring strong opioids compared to placebo for reducing pain intensity (SMD -0.43 (95% CI -0.52 to -0.33)), low certainty evidence for a medium between group difference favouring tramadol compared to placebo for reducing pain intensity (SMD of -0.55 (95% CI -0.66 to -0.44)), and very low certainty evidence for a small between group difference favouring buprenorphine compared to placebo for reducing pain intensity (SMD -0.41 (95% CI -0.57 to -0.26)), all at short-term follow up.

Antidepressants

There was low certainty evidence for no evidence of difference for antidepressants (all types) compared to placebo for reducing pain intensity (SMD of -0.04 (95% CI -0.25 to 0.17)) at short-term follow up.

The certainty in the evidence for all other comparisons ranged from low to very low and provides insufficient evidence to either support or refute the use of these interventions. Clinical heterogeneity, and variation between trials in the outcome measures used and measurement timing prevented quantitative synthesis for many comparisons in the original reviews.

Authors' conclusions

We found no moderate or high certainty evidence that any investigated pharmacological intervention provided a large or medium effect on pain intensity for acute or chronic LBP compared to placebo. For acute LBP, we found evidence that NSAIDs and muscle relaxants may provide a small effect on pain, and no evidence of difference for paracetamol. For chronic LBP, we found evidence that NSAIDs and opioids may provide a small effect on pain. Substantial caution is required when considering comparisons with low and very low certainty evidence for clinical or policy decisions. There is a clear need for high-quality randomised controlled trials to resolve uncertainties about the efficacy, effectiveness, and safety of pharmacological interventions.

Plain language summary

Pharmacological treatments for low back pain in adults: an Overview of Cochrane Reviews

Review question

What is the evidence from Cochrane reviews on the most effective and safe medicines for adults with non-specific low back pain?

Why is this important?

Low back pain is a common and debilitating health condition. In most cases, the cause or causes of low back pain cannot be reliably identified and is described as ‘non-specific’ low back pain. Physicians commonly prescribe medicines to treat low back pain. There are multiple types of medicines and medicine classes available, for example, opioid analgesics, nonsteroidal anti-inflammatory drugs (NSAIDs), and paracetamol. There is a need to provide accessible, high-quality information on the size of the effect and certainty of the evidence for medicines to treat low back pain.

What did we do?

Cochrane systematic reviews of interventions aim to assess the benefits and harms of interventions used in healthcare and health policy decisions based on up-to-date research studies. We searched for all Cochrane systematic reviews that assessed the benefits and harms of medicines for adults with non-specific low back pain to produce an overview of Cochrane evidence.

What evidence did we find?

As of June 2021, we found seven reviews (that included 103 unique trials with 22,238 participants). Most reviews (5/7, 71%) were assessed as having high methodological quality. The reviews reported data on six distinct medicines or medicine classes: paracetamol, NSAIDs (for example, ibruprofen), muscle relaxants (for example cyclobenzaprine), benzodiazepines (for example diazepam), opioids (for example

tapentadol), and antidepressants (for example paroxetine). Most (5/7) reviews included participants reporting low back pain lasting greater than 6 weeks. The certainty of evidence ranged from very low to high.

We found moderate certainty evidence that NSAIDs may provide a small short-term (≤ 3 months postintervention) effect on pain for acute low back pain, moderate certainty evidence that muscle relaxants may provide a small short-term effect on pain for acute low back pain, and high certainty evidence for no evidence of difference to placebo for paracetamol in the short-term for acute low back pain. We found moderate certainty evidence that NSAIDS may provide a small intermediate term (> 3 months and ≤ 12 months postintervention) effect on pain for chronic low back pain and moderate certainty evidence that opioids may provide a small short-term effect on pain for chronic low back pain. The certainty in the effect estimates for the remaining comparisons and outcomes is limited due to low and very low certainty evidence.

What does this mean?

There is no moderate or high certainty evidence that medicines provide large or medium size effects on the outcomes pain for people with non-specific low back pain. NSAIDs, muscle relaxants and opioids may provide a small and possibly not important effect on pain for people with low back pain. Paracetamol does not provide an effect on pain for people with acute low back pain and antidepressants may not provide an effect on pain for people with chronic low back pain. Physicians should discuss the possibility for a small effect on pain with increased risk for harm when considering different medicines for treating low back pain. Funders and researchers should prioritise identifying medicines that provide clinically meaningful benefits to people with low back pain.

Background

Description of the condition

Low back pain (LBP) is a common health condition that has major impacts on function and quality of life (Koes 2006). It is estimated that 7.2% (95% confidence interval (CI) 6.4% to 8.0%) of people across the globe have LBP at any time (Abajobir 2017), and an estimated 38.0% might experience significant LBP during their lifetime (Hoy 2012). LBP is comparatively more common in people aged 40 to 69 years (Hoy 2012), and in those experiencing socioeconomic disadvantage (Schofield 2012). Amongst all diseases and injuries included in the Global Burden of Disease Study, LBP has been the leading cause of reduced function since 1990. For example, 57.65 million years lived with disability (95% CI 40.82 to 75.88 million) were attributed to LBP in 2016 (Abajobir 2017). LBP is also an increasing cause of overall disease burden. It was the eleventh largest cause of disease burden for women and the seventeenth for men in 1990, but the seventh largest cause for women and tenth for men by 2017 (Kyu 2018).

The prognosis of acute LBP is typically favourable (Menezes Costa 2012). However, 30 to 40% of people report symptoms beyond three months (Henschke 2008), at which time they are considered to have chronic LBP. The cause or causes underlying the development and persistence of LBP are unknown (Koes 2007; Maher 2017), or cannot be reliably identified (Hancock 2007), in approximately 85% of cases in primary care. The label ‘non-specific’ LBP means it is not currently possible to attribute the clinical presentation to any specific disease process (e.g. infection, inflammatory condition, cancer) or structural pathology (e.g. fracture, nerve root compression) (Koes 2007; Maher 2017). Research is ongoing to reliably determine the cause(s) of LBP in those cases currently labelled 'non-specific', though the impact that specific labels would have on management and outcomes is unclear.

There are significant economic consequences associated with non-specific LBP. For example, LBP incurred the third highest costs of any health condition in the USA in 2013 (USD 87.6 billion 95% CI 67.5 billion to 94.1 billion) (Dieleman 2016). This increased to the highest costs in 2016 (USD 134.5 billion 95% CI, 122.4 billion to 146.9 billion)

(Dieleman 2020). In the UK, the total direct healthcare costs for an individual with chronic LBP are double that for someone without chronic LBP, matched by age, sex and geographic region (Hong 2013). Data from Australia suggest that non-specific chronic LBP is the most common health-related reason for early retirement (Schofield 2008), resulting in income poverty for this group (Schofield 2012).

Description of the interventions

Pharmacological interventions are the treatment option most used for LBP (Carey 2009; Gore 2012; Hart 2015; Ivanova 2011). There are multiple classes of these interventions, including opioid analgesics, nonsteroidal anti-inflammatory drugs (NSAIDS), muscle relaxants, benzodiazepines, antidepressant medicines, anticonvulsant medicines, and systemic corticosteroids. Opioid analgesics and NSAIDs are the most used classes amongst the countries for which data on usage are available (Australia, Italy, Portugal, UK, USA). The relative usage of the less common medicine classes varies across these countries (Gore 2012; Gouveia 2017; Michaleff 2012; Piccoliori 2013).

Pharmacological interventions are used to improve pain and physical function and may achieve this through numerous pathways such as reducing muscular spasm, modulating sensory nerve function or altering the availability of signalling chemicals in the brain. This Overview of Cochrane reviews focuses on systemic pharmacological interventions used to improve pain and physical function in people with LBP.

How the intervention might work

Pharmacological interventions for LBP are designed to act on various neurobiological targets within the body. The mechanisms by which different pharmacological interventions might improve pain and function are not fully understood and differ across medicine classes. We present below commonly proposed mechanism(s) for each class of medicine.

NSAID and Paracetamol (Acetaminophen)

NSAIDs and paracetamol act on cyclo-oxygenase (COX) enzymes to interfere with natural inflammatory processes (Brune 2015). Specifically, they reduce the production of prostaglandins — signal chemicals that modulate inflammation, nociception and other autonomic processes (Jóźwiak-Bebenista 2014). Longer-term use of these medicines may be associated with increased risk of cardiovascular events (e.g. stroke) (Roberts 2016), and when taken together, gastrointestinal bleeding (McCrae 2018; Anderson 2022). Certain NSAIDs, depending on degree of COX-2 selectivity, are associated with increased risk of gastrointestinal side effects (e.g. stomach ulcers) (van der Linden 2009).

Muscle Relaxants

Muscle relaxants are a broad class of chemically varied medicines grouped together by their shared function (Trevor 2018, Cashin 2021). The two main categories discern between antispasmodic medicines, commonly prescribed for the treatment of muscle spasm associated with muscle injury, and antispastic medicines, commonly prescribed to reduce heightened muscle tone (spasticity) (Cashin 2021). Muscle relaxants are thought to act on the central nervous system, or in some cases, the skeletal muscle cell (Trevor 2018, Witenko 2014). Each muscle relaxant medicine has different clinical uses, mechanism(s) of action and associated side effects, although feelings of dizziness, drowsiness and nausea are common to all muscle relaxants (See 2008). Certain muscle relaxants (e.g., carisoprodol) are associated with an increased risk of misuse and dependency (Cashin 2021).

Benzodiazepines

Benzodiazepines act on the central nervous system, increasing the effects of the neurotransmitter gamma-aminobutyric acid (GABA). Although benzodiazepines share

functional muscle relaxing properties similar to muscle relaxants (Trevor 2018), they are not classified by the US Food and Drug Administration as muscle relaxants and are considered separately by some clinical guidelines (Qaseem 2017). Benzodiazepines produce strong sedative effects and are associated with problems with addiction, overdose and withdrawal (Bachhuber 2016; Hood 2014).

Opioid analgesics

Opioid analgesic medicines act on the naturally occurring (endogenous) opioid receptors in the nervous system, to reduce the contribution of nociceptive (danger-signalling) information to the pain experience (Rivat 2016). Opioid medicines are often classified as either weak (e.g., codein, tramadol) or strong (e.g., oxycodone, tapentadol) relating to their relative potency. Opioid medicines may cause adverse effects; commonly constipation, nausea and sedation, depending on the location and type of receptor (Kalso 2004). Longer-term use can contribute to opioid tolerance (requiring progressively higher doses), possible dependence, and death (Deyo 2015).

Antidepressants

Antidepressant medicines are another class of medicines of varied chemical structure, subclassified by their function. These medicines act on neurotransmitters in the brain. This is thought to produce analgesic effects independent of their effects on depression (Cohen 2001; Micó 2006), although the precise mechanisms are unclear (Harmer 2017). Categories of antidepressants prescribed to treat pain in order of perceived effectiveness include serotonin-norepinephrine re-uptake inhibitors (SNRIs, e.g. duloxetine), tricyclic antidepressants (TCAs, e.g. amitriptyline), and selective serotonin reuptake inhibitors (SSRIs, e.g. sertraline) (Ferraro 2021). People with LBP may also be prescribed these medicines to improve sleep and reduce depression or anxiety. Side effects differ between the categories, although drowsiness, dry mouth and dizziness are common (Chou 2010).

Anticonvulsants

Anticonvulsant medicines act across several sites in the central nervous system. The analgesic action of anticonvulsant medicines is thought to occur through limiting neuronal excitation and enhancing inhibition, although the precise mechanisms are unclear (Maizels 2005). Anticonvulsant medicines have a long history of off-label use in pain conditions. Common dose-related side effects include drowsiness and dizziness (Derry 2019).

Systemic Corticosteroids

Corticosteroids are a class of medicines that are structurally similar to the naturally occurring human adrenal hormone cortisol, considered an important regulator of homeostasis (Chou 2016; van der Laan 2008). These medicines mimic the physiological actions of cortisol to produce a wide range of effects, including both anti-inflammatory and immunosuppressive effects. Corticosteroid medicines differ by their relative potency, duration and mechanism(s) of action. Short-term use of corticosteroids is associated with increased rates of sepsis, venous thromboembolism, hyperglycaemia and fracture (Waljee 2017).

Why it is important to do this overview

Pharmacological interventions are the interventions most used by people with LBP to manage their pain. People with LBP, clinicians, researchers and health policymakers need accessible, high-quality information on the effect size and certainty of the evidence for efficacy, effectiveness, and safety of pharmacological interventions (Chou 2018a, Chou 2018b, Lim 2019). Cochrane Reviews have investigated the effects of pharmacological interventions and are available to decision-makers through the

Cochrane Library. There are multiple reviews, of varying currency, scope and methodology. This may inhibit decision-makers' access to this evidence.

There is a need to systematically synthesise this evidence into a single accessible Overview. Overviews, or systematic reviews of systematic reviews, allow multiple systematic reviews on similar or related topics, to be systematically brought together for appraisal and synthesis of results (Hunt 2018). An Overview should improve access to high-quality information and describe the currency and scope of the information. This may support people with LBP, clinicians and policymakers to use this evidence in their health decision-making (Hunt 2018). Information on currency, scope and methodology across the reviews may also support researchers, funders and policy decision makers to identify important evidence gaps for conducting updates of reviews or planning prospective reviews.

Objectives

To summarise the evidence from Cochrane reviews of the efficacy, effectiveness, and safety of systemic pharmacological interventions for adults with non-specific LBP.

Methods

Criteria for considering reviews for inclusion

Types of reviews

We included all Cochrane reviews of randomised controlled trials (RCTs) on pharmacological interventions for people with non-specific LBP published in the Cochrane Library. We excluded reviews that include randomised and non-randomised designs, unless the data for the randomised designs was available separately and excluded Cochrane reviews withdrawn or superseded. We identified Cochrane review protocols and listed them as ongoing reviews for future updates.

Types of participants

Participants were adults, 18 years or older, with non-specific LBP (e.g. non-radicular LBP, with or without non-specific degenerative changes), of any duration. LBP is defined as a primary area of pain between the twelfth rib and gluteal fold, with or without associated leg pain (Koes 2006). We excluded systematic reviews that included participants with spinal stenosis (back and leg pain associated with narrowing of the spinal canal), LBP caused by known structural or pathological processes (e.g. nerve root compression, osteoporosis, fractures, infection, neoplasm, metastasis) or specific medical conditions (e.g. pregnancy, inflammatory disease) (Koes 2007; Maher 2017), unless the review reported results for non-specific LBP separately. We excluded reviews that included participants younger than 18 years unless they reported separate results for the participants 18 years or older.

Types of interventions and comparisons

We included systemic pharmacological interventions, used with the intent to improve pain and function, for people with LBP. We considered systemic pharmacological interventions broadly as any medicine that affects the body as a whole, rather than individual parts or organs, that may be used with the intent to improve pain and function. We made no restriction on route of administration or dose. We also included combinations of pharmacological interventions.

Comparisons of interest were:

- pharmacological intervention versus placebo/sham intervention (efficacy comparisons)

- different forms of the same pharmacological intervention (e.g., selective NSAID versus a non-selective NSAID) (effectiveness comparisons)
- pharmacological intervention versus a different type of pharmacological intervention (effectiveness comparisons)
- pharmacological intervention versus a non-pharmacological intervention (effectiveness comparisons)

Types of outcome measures

The outcomes reflect the core outcome set for non-specific LBP (Chiarotto 2015), and recommendations of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) (Dworkin 2005). We made no restriction on time of measurement. We grouped outcomes into a short-term period (≤ 3 months postintervention), an intermediate-term period (> 3 months and ≤ 12 months postintervention), and a long-term period (> 12 months postintervention). We included the outcome measure closest to the midpoint of the period in cases where a review reported outcome data for multiple time points within a period, or measured the outcome at different time periods.

Primary outcomes

1. Pain, defined as pain intensity, assessed on a continuous self-report scale (e.g. a visual analogue scale (VAS), numerical rating scale (NRS), the brief pain inventory (BPI) (Cleeland 1989) or other validated measure), or in dichotomous format (e.g. as the proportion of participants in each group who attained a predetermined threshold of improvement).

2. Physical function, defined as back-pain related function, assessed through continuous self-report scales (e.g. Roland-Morris Disability Questionnaire (RMDQ) (Roland 1983), Oswestry Disability Index (ODI) (Fairbank 1980)), functional testing protocols or other validated quantitative measures.

3. Safety, defined as adverse events including amongst others: incidence and severity of adverse events, trial withdrawal due to adverse events and incidence of serious adverse events, as described by the systematic review.

Secondary outcomes

1. Participant ratings of improvement, defined as global perceived effect, assessed with a validated tool (e.g., Patient Global Impression of Change Scale (Guy 1976)).

2. Health-related quality of life, assessed with a validated tool (e.g. the 36-Item Short Form Health Survey (SF-36) (Ware 2000)).

3. Workplace participation, defined as days to return-to-work, days of absenteeism or days of reduced work activities.

Search methods for identification of reviews

Electronic searches

We conducted a sensitive search of the Cochrane Database of Systematic Reviews (The Cochrane Library, current issue) using a combination of Medical Subject Headings (MeSH) and keywords (Appendix 1), without restriction up to issue 5 of 12, 2021. The search strategy is presented in Appendix 1. We managed retrieved citations using EndNote 2017 and Covidence.

Data collection and analysis

Selection of reviews

We assessed in two stages the eligibility of identified Cochrane Reviews. Two authors (AGC and RRNR) independently screened the results of the electronic search by title and

abstract against the inclusion criteria. We obtained the full texts of reviews meeting these criteria and two authors (AGC and RRNR) independently screened them again to confirm inclusion. We planned to use a third Overview author to resolve discrepancies when the two first authors could not reach consensus; however, no discrepancies occurred. We provide a PRISMA flow diagram documenting the screening and review selection process; see Figure 1.

Data extraction and management

A pilot data extraction form was designed and piloted by four authors (AGC, BMW, NEOC, RRNR). Two authors (AGC and RRNR) independently extracted data using the finalised data extraction form. We planned to involve an independent third Overview author to resolve disagreements; however, this option was not required. The data extraction form included the following details:

Review Characteristics:

- Objectives of the review
- Dates of publication, most recent search and planned update
- Resources searched
- Number of included trials
- Characteristics of included participants (e.g. duration of pain, pain severity, sex, age, race, comorbidities, prior treatment history (to the extent possible))
- Description of interventions and comparisons
- Outcomes and time points assessed
- Details of meta-analyses, if applicable

Statistical summaries:

- Point estimates, 95% CIs and accompanying measures of heterogeneity for the pooled estimates of intervention effects; for all relevant comparisons at all available time points (e.g. risk ratios (RRs), risk difference (RD), odds ratios (ORs), number needed to treat for an additional beneficial effect (NNTB) or additional harmful effect (NNTH), mean differences (MDs), standardised mean difference (SMD))
- Results of responder analyses, including prespecified criteria for response and power calculation
- Results from exploration of heterogeneity, including subgroup analyses/meta-regression and whether these were prespecified
- Results from sensitivity analyses, including details of the approach taken and whether these were prespecified
- The judgements of risk of bias in the evidence, including details of the approach used (e.g. Cochrane ROB tool)
- The judgements of certainty in the evidence, including details of the approach used (e.g. GRADE)

We planned to contact the authors of included reviews if we could not extract the required information from the reports. We did not plan on contacting authors of individual studies included in the reviews.

Assessment of methodological quality of included reviews

Quality of included reviews

Two authors (AGC and RRNR) independently assessed the methodological quality of included systematic reviews using the AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews) (Shea 2017). The 16 items in the AMSTAR 2 instrument provides a broad assessment of systematic review quality that, taken together, inform a judgement of

confidence in the review findings (see Appendix 2). We resolved discrepancies through consensus or recourse to a third author (NEOC). The AMSTAR 2 assessments were also used to identify consistency of review methods and conduct as well as to identify areas for improvement.

We considered seven items recommended by Shea 2017 (item 2, protocol registered before commencement of the review; item 4, adequacy of the literature search; item 7, justification for excluding individual studies; item 9, risk of bias from individual studies being included in the review; item 11, appropriateness of meta-analytical methods; item 13, consideration of risk of bias when interpreting the results of the review; item 15, assessment of presence and likely impact of publication bias) as critical when forming an overall judgement on the quality of the included systematic review (Shea 2017). We defined a rating of:

- High overall confidence in the results of the review if there were *no or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
- Moderate if there were *more than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
- Low if there were *one critical flaw with or without non-critical weaknesses*: the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.
- Critically low if there were *more than one critical flaw with or without non-critical weaknesses*: the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies.

Risk of bias of primary studies included in reviews

We reported the 'risk of bias' assessments for the primary studies in each included systematic review. We did not repeat or update these assessments. We reported the 'risk of bias' tool used, including details regarding dimensions assessed (e.g. allocation concealment, participant blinding), and results of the assessments.

Certainty of evidence in included reviews

We reported, where available, the GRADE judgement of certainty for each core comparison for our primary outcomes (Balshem 2011). The GRADE approach uses five considerations (risk of bias, inconsistency, imprecision, indirectness and publication bias) to assess the certainty of the body of evidence for each outcome. First, two authors (AGC and RRNR) independently extracted the GRADE assessments for each systematic review for each independent outcome. Second, for reviews which did not report GRADE assessments, two Overview authors (AGC and RRNR) independently conducted GRADE assessments of certainty in the evidence using a checklist for the primary outcomes and placebo comparisons (Meader 2014). We resolved discrepancies through consensus. We planned to involve an independent third Overview author to resolve disagreements; however, this option was not required.

When required, we used the following to assign GRADE judgements:

- Serious study limitations: we downgraded once if less than 50% of studies were at low risk of bias across all risk of bias criteria.
- Inconsistency: we downgraded once if point estimates varied widely across studies, confidence intervals showed minimal or no overlap, statistical tests for heterogeneity were statistically significant, or the I^2 statistic was greater than 50%.
- Indirectness: we downgraded once if greater than 50% of participants were outside the target group.
- Imprecision: we downgraded once if there were fewer than 400 participants for continuous outcomes and fewer than 300 events for dichotomous data.

- Publication bias: we downgraded once where there was direct evidence of publication bias or if estimates of effect based on small scale, industry sponsored studies raised suspicion of publication bias.

GRADE judgements indicate the following degree of certainty in the conclusions of the systematic review.

- High: very certain that the true effect lies close to that of the estimate of the effect.
- Moderate: moderately certain in the effect estimate – the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- Low: certainty in the effect estimate is limited – the true effect may be substantially different from the estimate of the effect.
- Very Low: very little certainty in the effect estimate and the true effect is likely to be substantially different from the estimate of the effect.

Overlap between reviews

Following recommended guidance (Hennessy 2020), we examined the degree of overlap of primary studies in the included reviews. This involved creating a citation matrix of the primary studies (rows) included in each review (columns) to calculate the corrected covered area (CCA) (Pieper 2014). Pieper 2014 suggest interpreting CCA values lower than five to indicate slight overlap and CCA values greater than or equal to 15 to indicate high overlap.

Data synthesis

We presented data from each systematic review for each primary and secondary outcome in order of certainty (i.e. high-certainty evidence first, followed by moderate-certainty evidence, etc.); for efficacy (intervention vs placebo) comparisons, followed by effectiveness comparisons; at each level of follow-up (i.e. short-term, followed by intermediate-term, etc.). We presented narrative descriptions of results only when statistical outcome data was not available. We stratified the data by the duration of LBP observed in the included studies; acute (0 to 6 weeks); sub-acute (6 to 12 weeks); chronic (> 12 weeks); mixed (multiple symptom durations grouped together, e.g. acute and subacute or subacute and chronic); and unclear (symptom duration not reported).

We did not conduct any novel statistical synthesis of data or make any indirect comparisons. We planned to convert effect sizes, where possible, to common scales to facilitate interpretation.

We classified the size of the effect for the mean between group difference for the outcomes pain and function based on the definitions from the American College of Physicians and the American Pain Society (Chou 2017).

- Large effect: >20 points on a 0-100 scale or >0.8 SMD
- Medium effect: >10-20 points on a 0-100 scale or >0.5 to 0.8 SMD
- Small effect: 5-10 points on a 0-100 scale or 0.2 to 0.4 SMD
- No evidence of difference: boundaries of the 95% confidence interval span both sides of the line of no effect
- Harmful: boundaries of the 95% confidence interval fall completely within harm

We presented the short-term efficacy of the intervention compared to placebo on pain intensity in a 'Summary of findings' table, as described in Chapter V of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019). We presented the results of the remaining primary and secondary outcomes at each time point in an 'Overview of reviews' table. We also included two 'Summary of results' tables highlighting the size and certainty of the evidence, by considering both the effect size and GRADE rating for the outcomes of pain and function at the short-term follow up.

Results

The initial search of the *Cochrane Library* (June 3, issue 5, 2021) identified 1427 Cochrane review records. We excluded 1398 records after review of title and abstracts and excluded a further 18 records after full-text assessment (Figure 1). Seven reviews were deemed eligible for inclusion (Chaparro 2013, Enthoven 2016, Santos 2015, Saragiotto 2016, Urquhart 2008, van Tulder 2003, van der Gaag 2020). Reasons for exclusion included: wrong intervention (4), wrong patient population (3), wrong route of administration (5), wrong study design (3), Cochrane review withdrawn (3) and 1 review was excluded because it had been updated and replaced with two separate reviews, both of which were included in this Overview (Appendix 3). We identified three review protocols as potentially eligible for future updates once published, details of these protocols can be found in Table 1.

Description of included reviews

A detailed description of the characteristic of the included reviews is presented in Table 2.

The seven reviews included 22,238 participants across 103 unique RCTs. The number of included RCTs and participants in each review ranged from 2 (Saragiotto 2016) to 32 RCTs (van der Gaag 2020) and 722 (Urquhart 2008) to 5540 participants (Chaparro 2013). The median (IQR) year of review publication was 2015 (2010 to 2016) with most (86%, 6/7) reviews published before 2017. Six reviews searched trial registry records but none of the reviews included outcome data extracted directly from trial registry records. When reported by the systematic review, only a small proportion of the included RCTs in each review were prospectively registered (e.g., 2/13 [15%] Enthoven 2016 and 3/32 [9%] van der Gaag 2020), however many RCTs were published before trial registry platforms were established and registration was mandatory (Cashin 2021). None of the reviews reported any direct funding perceived to be a conflict of interest. Five of the seven reviews reported the funding of included RCTs. Of these five reviews, all reported that half or more of the included RCTs were either funded by a pharmaceutical company or declared relationships with a pharmaceutical company. This is reflective of previous systematic reviews which found that the majority of trials of pharmacological interventions are industry funded (Barden 2006; Bourgeois 2010). Although trials funded by a drug or device company have been shown to be more likely to have positive conclusions and statistically significant results (Lundh 2017), there can be substantial variation in the degree to which funding or the declared relationships can impact the validity and magnitude of the study findings (Chopra 2003).

One review included only acute to sub-acute LBP (<12 weeks) (van der Gaag 2020), three reviews included only chronic LBP (>12 weeks) (Chaparro 2013, Enthoven 2016, Santos 2015), and three reviews did not restrict the duration of LBP included (Saragiotto 2016, Urquhart 2008, van Tulder 2003), however Saragiotto 2016 only identified RCTS including people with acute LBP (≤ 6 weeks duration) and Urquhart 2008 only identified RCTS including people with chronic LBP (>12 weeks duration). One review restricted inclusion to RCTs with participants reporting moderate-severe LBP, defined as pain ≥ 4 on a 0 to 10 pain scale (Santos 2015). All nine reviews included pain as the primary outcome. Two reviews included patient-reported pain relief as a primary outcome measure with categorisation into “responder” groups reporting more than 30% and/or 50% pain relief (Chaparro 2013, Santos 2015). One review reported pain as a dichotomous effect measure – the risk of experiencing no pain relief using risk ratios – where risk ratios smaller than one indicate that the chance of “not getting pain relief” is less in the intervention group compared to the comparator (van Tulder 2003). Other commonly reported primary outcome measures included back pain-specific function, global measure of improvement, safety (adverse events), and return to work. No reviews provided clear definitions for how adverse events or serious adverse events were operationalised as outcomes in the review. Only one review reported that serious adverse events were considered as defined by each included RCT (Saragiotto 2016). We found that most reviews were not able to report data across each of the pre-planned outcomes due to a lack of adequate data.

No reviews discussed issues related to health equity or considered the social determinants of health when synthesising and interpreting the evidence. This, in part, could be because of incomplete reporting of sociodemographic characteristics from the included RCTs. Chaparro 2013 highlight that “many studies neglected to report other parameters affecting outcomes, such as duration of pain prior to enrolment, employment or compensation status or poor response to previous treatment”. Only two reviews considered the representativeness of the evidence reporting concerns generalising the evidence beyond the restricted and limited participant population (Chaparro 2013, Saragiotto 2016).

Interventions

The seven reviews reported on six pharmacological interventions or intervention classes, paracetamol (Saragiotto 2016), NSAIDs (Enthoven 2016, van der Gaag 2020), muscle relaxants (van Tulder 2003), benzodiazepines (van Tulder 2003), opioids (Chaparro 2013, Santos 2015), and antidepressants (Urquhart 2008). The most investigated intervention classes were NSAIDs (45 RCTS, 10163 participants), opioids (19 RCTs, 8653 participants), and muscle relaxants (26 RCTs, 2538 participants). Two reviews reported on pharmacological interventions administered orally (Santos 2015, Saragiotto 2016), and five through multiple routes of administration (Chaparro 2013, Enthoven 2016, Santos 2015, Urquhart 2008, van der Gaag 2020). Treatment duration ranged from a single injection to 24 weeks.

Comparisons

All reviews included placebo as a pre-specified comparator, and two reviews considered placebo as the only comparator (Saragiotto 2016, Urquhart 2008). The second most common comparator was other pharmacological interventions (5/7, 71% reviews) (Chaparro 2013, Enthoven 2016, Santos 2015, van der Gaag 2020, van Tulder 2003) followed by other non-pharmacological interventions (2/7, 29% reviews) (Enthoven 2016, van der Gaag 2020).

We found that most reviews were unable to report across all of their pre-planned comparisons and outcomes because of a lack of adequate data. In addition to a lack of data, heterogeneity in reported outcomes and comparisons limited the ability for reviews to conduct all pre-planned meta-analyses (7/7, 100% reviews) (Chaparro 2013, Enthoven 2016, Santos 2015, Saragiotto 2016, Urquhart 2008, van Tulder 2003, van der Gaag 2020), subgroup analyses (5/7, 71% reviews) (Chaparro 2013, Enthoven 2016, Santos 2015, Saragiotto 2016, van Tulder 2003), and inspection of small study bias using funnel plots (2/7, 29% reviews) (Chaparro 2013, Enthoven 2016).

Overlap between reviews

We identified three overlapping RCTs which were included in more than one review. The CCA was 0.5% suggesting very minimal overlap between reviews (Pieper 2014).

Quality of evidence

We found all seven reviews employed formal tools to assess risk of bias (Table 3): two used the Cochrane ‘Risk of bias’ tool (Higgins 2011) (Santos 2015, Saragiotto 2016); one used the 11 criteria for internal validity recommended by the Cochrane Back Review Group (van Tulder 1997) (van Tulder 2003); one review used the 11 criteria for methodological quality recommended by the Cochrane Back Review Group (van Tulder 2003a) (Urquhart 2008); two used the 12 criteria for risk of bias recommended by the Cochrane Back Review Group (Furlan 2009) (Chaparro 2013, Enthoven 2016); and one used the 12 criteria for risk of bias recommended by the Cochrane Back Review Group Cochrane Back and Neck Group (Furlan 2015) (van der Gaag 2020).

All reviews included at least one RCT assessed at unclear or high risk of bias across the investigated domains. Failure to report intention-to-treat analysis (attrition bias) (30/103, 29% RCTs) and inadequate allocation concealment (selection bias) (19/103, 18% RCTs) were the most common contributors to high risk of bias across the studies included in the

seven reviews. Most RCTs were rated as low risk of bias for blinding participants, personnel (performance bias) (67/103, 65% RCTs) and outcome assessors (detection bias) (63/103, 61% RCTs).

Four reviews used the GRADE approach to rate the overall certainty of the evidence (Chaparro 2013, Enthoven 2016, Saragiotto 2016, van der Gaag 2020). We conducted additional GRADE assessments for comparisons of 23 pharmacological interventions to placebo for primary outcomes pain, function and safety across five reviews (Chaparro 2013, Enthoven 2016, Santos 2015, Urquhart 2008, van Tulder 2003). The most common reasons for downgrading were study limitations and imprecision.

Methodological quality of included reviews

Results of the AMSTAR 2 assessment showed that we have high confidence in the findings of five reviews (Enthoven 2016, Santos 2015, Saragiotto 2016, van der Gaag 2020, van Tulder 2003), moderate confidence in the findings of one review (Chaparro 2013), and low confidence in the findings of one review (Urquhart 2008) (Table 4). One review did not assess the potential impact of risk of bias in individual studies on the results of the meta-analysis and did not provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review (Chaparro 2013). One review did not report on the sources of funding for the studies included in the review and did not carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review (Urquhart 2008).

Effect of interventions

A ‘Summary of findings’ table for the short-term efficacy of pharmacological interventions compared to placebo on our primary outcome of pain intensity is provided in Table 5.

An ‘Overview of reviews’ table for all other comparisons and outcomes for each pharmacological intervention or intervention class is provided in Table 6; Table 7; Table 8; Table 9; Table 10.

A ‘Summary of results’ table highlighting the effect size and certainty of the evidence for the primary outcomes (pain and function) and placebo comparisons for short-term follow-up is provided in Table 11; Table 12.

Data, where available, for each primary and secondary outcomes for the remaining effectiveness comparisons for all interventions is reported in Appendix 4; Appendix 5; Appendix 6.

Very few reviews reported data for intermediate term follow-up (> 3 months and ≤ 12 months postintervention) and no reviews reported data for long-term follow up (> 12 months postintervention). Outcome data are therefore presented below for short-term (≤ 3 months postintervention) follow up unless otherwise stated.

Pharmacological intervention versus placebo

Paracetamol (acetaminophen)

One Cochrane review, judged at high quality (Saragiotto 2016), included evidence on the effects of paracetamol compared to placebo. Saragiotto 2016 included two trials with a total sample size of 1785 participants with acute LBP. No trials were identified for participants with subacute or chronic low back pain.

Acute LBP

Pain: Saragiotto 2016 reported a pooled analysis of one three-arm study (two comparisons, n = 1516) investigating time-contingent prescription of paracetamol, as required prescription of paracetamol, and placebo. The review reported no evidence of difference between paracetamol and placebo (MD of 0.49 on a 0 to 100 pain intensity scale (95% CI -1.99 to 2.97, I^2 = 0)), which they rated as high certainty evidence.

Physical function: Saragiotto 2016 reported a pooled analysis of one three-arm study (two comparisons, n = 1516) investigating time-contingent prescription of paracetamol, as required prescription of paracetamol, and placebo. The review reported no evidence of difference between paracetamol and placebo (MD of 0.05 on a 0 to 24 Roland Morris Disability questionnaire (95% CI -0.50 to 0.60, I^2 = 0)), which they rated as high certainty evidence.

Safety: Saragiotto 2016 reported a pooled analysis of one three-arm study (two comparisons, n = 1516) investigating time-contingent prescription of paracetamol, as required prescription of paracetamol, and placebo. The review reported no evidence for an increased risk of experiencing an adverse event (RR of 1.07 (95% CI 0.86 to 1.33, I^2 = 0)) or a serious adverse event (RR of 0.90 (95% CI 0.30 to 2.67, I^2 = 0)) between paracetamol and placebo, which they rated as high certainty evidence.

Participant ratings of improvement: Saragiotto 2016 reported a pooled analysis of one three-arm study (two comparisons, n = 1511) investigating time-contingent prescription of paracetamol, as required prescription of paracetamol, and placebo. The review reported no evidence for a difference between paracetamol and placebo (MD of -0.10 on a -5 to 5 global perceived effect scale (95% CI -0.33 to 0.13, I^2 = 0)), which they rated as high certainty evidence.

Health-related quality of life: Saragiotto 2016 reported a pooled analysis of one three-arm study (two comparisons, n = 1145) investigating time-contingent prescription of paracetamol, as required prescription of paracetamol, and placebo. The review reported no evidence for a difference between paracetamol and placebo on the on the 12-item Short Health Survey physical component (MD of -0.79 (95% CI -1.94 to 0.36, I^2 = 0)) and on the 12-item Short Health Survey mental component (MD of -0.6 (95% CI -1.38 to 0.17, I^2 = 0)), which they rated as high certainty evidence.

Workplace participation: Saragiotto 2016 did not report data on workplace participation because it was not an outcome of interest in the review.

NSAIDs

Two Cochrane reviews, judged at high quality (Enthoven 2016, van der Gaag 2020), included evidence on the effects of NSAIDs compared to placebo. van der Gaag 2020 included 32 trials with a total sample size of 5356 participants with acute LBP. Enthoven 2016 included 13 trials with a total sample size of 1354 participants with chronic LBP. Enthoven 2016 only reported outcome data at ≤16 weeks follow up (median [IQR] follow-up was 84 days [42 to 105 days]) which we classified as intermediate follow-up (> 3 months and ≤ 12 months postintervention).

Acute LBP

Pain: van der Gaag 2020 reported a pooled analysis of four studies (five comparisons, n = 815) investigating NSAIDs compared to placebo. The review reported a small between group difference favouring NSAIDs (MD of -7.29 on a 0 to 100 pain intensity scale (95% CI -10.98 to -3.61, I^2 =35.18)), which they rated as moderate certainty evidence. van der Gaag 2020 narratively reported the results of one study (n =240) for the intermediate follow-up (> 3 months and ≤ 12 months postintervention), reporting no evidence for difference between NSAIDs and placebo on reducing pain intensity.

Physical function: van der Gaag 2020 reported a pooled analysis of two studies (three comparisons, n = 471) investigating NSAIDs compared to placebo. The review reported a small between group difference favouring NSAIDs (MD of -2.02 on a 0 to 24 Roland Morris Disability questionnaire (95% CI -2.89 to -1.15, I^2 = 0)), which they rated as high certainty evidence. van der Gaag 2020 narratively reported the results of one study (n = 240) for the intermediate follow-up (> 3 months and ≤ 12 months postintervention), reporting no evidence for difference between NSAIDs and placebo on physical function.

Safety: van der Gaag 2020 reported a pooled analysis of six studies (eight comparisons, n = 1394) investigating NSAIDs compared to placebo. The review reported no evidence of an increased risk of experiencing an adverse event (RR of 0.86 (95% CI 0. 63 to 1.18, I^2 = 0)), between NSAIDs and placebo which they rated as very low certainty evidence.

Participant rating of improvement: van der Gaag 2020 reported a pooled analysis of five studies (seven comparisons, n = 1201) investigating NSAIDs compared to placebo. The review reported an increased risk for experiencing global improvement for NSAIDs compared to placebo (RR of 1.40 (95% CI 1.12 to 1.75, I^2 = 51.64)), which they rated as low certainty evidence.

Health-related quality of life: van der Gaag 2020 did not report data on Health-related quality of life because it was not an outcome of interest in the review.

Workplace participation: van der Gaag 2020 reported data from one study (one comparison, n = 266) investigating NSAIDs compared to placebo. The review reported no evidence for an increased risk for workplace participation (RR of 1.48 (95% CI 0.98 to 2.23)) between NSAIDs and placebo, which they rated as very low certainty evidence.

Chronic LBP

Pain: Enthoven 2016 reported a pooled analysis of six studies (six comparisons, n = 1354) investigating NSAIDs compared to placebo at intermediate follow-up. The review reported a small between group difference favouring NSAIDs (MD of -6.97 on a 0 to 100 pain intensity scale (95% CI -10.74 to -3.19, I^2 = 51.96)), which they rated as low certainty evidence. The same review reported pooled analyses for non-selective NSAIDs compared to placebo (4 studies, 4 comparisons, n = 847) and selective NSAIDs (2 studies, 2 comparisons, n = 507) compared to placebo at intermediate follow-up. The review reported a small between group difference in favour of non-selective NSAIDs (MD of -5.96 on a 0 to 100 pain intensity scale (95% CI -10.96 to -0.96, I^2 = 55.25)) and selective NSAIDs (MD -9.11 on a 0 to 100 pain intensity scale (95% CI -13.56 to -4.66, I^2 = 0)), which we rated as low and moderate certainty evidence respectively.

Physical function: Enthoven 2016 reported a pooled analysis of four studies (four comparisons, n = 1161) investigating NSAIDs compared to placebo at intermediate follow-up. The review reported a small between group difference favouring NSAIDs (MD of -0.85 on a 0 to 24 Roland Morris Disability questionnaire (95% CI -1.30 to -0.40, I^2 = 45.78)), which they rated as low certainty evidence.

Safety: Enthoven 2016 reported a pooled analysis of six studies (six comparisons, n = 1354) investigating NSAIDs compared to placebo at intermediate follow-up. The review reported no evidence of experiencing an increased risk for an adverse event (RR of 1.04 (95% CI -0.92 to 1.17, I^2 = 19.68)), between NSAIDs and placebo, which they rated as low certainty evidence. The same review reported pooled analyses for non-selective NSAIDs (4 studies, 4 comparisons, n = 847) and selective NSAIDs (2 studies, 2 comparisons, n = 507) compared to placebo at intermediate follow-up. The review found no evidence of an increased risk for experiencing an adverse event with non-selective NSAIDs (RR of 0.94 (95% CI -0.82 to 1.08, I^2 = 0)) and an increased risk of experiencing an adverse event with selective NSAIDs (RR 1.25 (95% CI 1.00 to 1.56, I^2 = 17.54)), which we rated as low and moderate certainty evidence respectively.

Participant rating of improvement: Enthoven 2016 was unable to identify data on the outcome participant rating of improvement.

Health-related quality of life: Enthoven 2016 did not report data on Health-related quality of life because it was not an outcome of interest in the review.

Workplace participation: Enthoven 2016 was unable to identify data on the outcome workplace participation.

Muscle Relaxants and Benzodiazepines

One Cochrane review, judged at high quality (van Tulder 2003), included evidence on the effects of muscle relaxants (antispasmodics and antispastics) and benzodiazepines compared to placebo. van Tulder 2003 included 31 trials with a total sample size of 2884 participants with acute and chronic LBP. Most included trials (n=24, 80%) were on acute LBP.

Acute LBP

Pain: van Tulder 2003 reported a pooled analysis of three studies (three comparisons, n = 244) investigating antispasmodic muscle relaxants compared to placebo. The review reported a reduced risk of not getting pain relief for antispasmodic muscle relaxants compared to placebo (RR of 0.58 (95% CI 0.45 to 0.76, I^2 = 0)), which we rated as moderate certainty evidence. The same review narratively reported the results of two high quality studies (n = 220) of antispastic muscle relaxants compared to placebo and one low quality study (n = 50) of benzodiazepines compared to placebo. The review reported that both antispastic muscle relaxants and benzodiazepines were more effective than placebo for pain relief.

Physical function: van Tulder 2003 reported a pooled analysis of three studies (three comparisons, n = 251) investigating antispasmodic muscle relaxants compared to placebo. The review reported a reduced risk of not improving physical function for antispasmodic muscle relaxants compared to placebo (RR of 0.55 (95% CI 0.40 to 0.77, I^2 = 0)), which we rated as moderate certainty evidence.

Safety: van Tulder 2003 reported a pooled analysis of eight studies (eight comparisons, n = 724) investigating antispasmodic muscle relaxants compared to placebo. The review reported an increased risk of experiencing an adverse event (RR of 1.50 (95% CI 1. 14 to 1.98, I^2 = 0)) between antispasmodic muscle relaxants compared to placebo, which we rated as moderate certainty evidence.

Participant rating of improvement: van Tulder 2003 reported a pooled analysis of four studies (four comparisons, n = 323) investigating antispasmodic muscle relaxants compared to placebo. The review reported no evidence for a difference in the risk of not experiencing improvement (RR of 0.68 (95% CI 0.41 to 1.13, I^2 = 33.74)) between antispasmodic muscle relaxants compared to placebo. The same review narratively reported the results from one high quality trial (one comparisons, n = 200) investigating antispastic muscle relaxants compared to placebo. The review reported that antispastic muscle relaxants were more effective compared to placebo on increased participant ratings of improvement.

Health-related quality of life: van Tulder 2003 did not report data on Health-related quality of life because it was not an outcome of interest in the review.

Workplace participation: van Tulder 2003 was unable to identify data on the outcome workplace participation.

Chronic LBP

Pain: van Tulder 2003 narratively reported the results from two high quality trials (n = 219) investigating antispasmodic muscle relaxants compared to placebo. The review reported conflicting results on whether antispasmodic muscle relaxants are more effective than placebo for pain relief. The same review reported a pooled analysis of two studies (two comparisons, n = 146) investigating benzodiazepines compared to placebo. The review reported a reduced risk of not getting pain relief for benzodiazepines compared to placebo (RR of 0.71 (95% CI 0.54 to 0.93, I^2 = 0)), which we rated as low certainty evidence.

Physical function: van Tulder 2003 was unable to identify data on the outcome physical function.

Safety: van Tulder 2003 reported a pooled analysis of two studies (two comparisons, n = 246) investigating antispasmodic muscle relaxants compared to placebo. The review reported no evidence of difference in the risk of experiencing an adverse event (RR of

1.02 (95% CI 0.67 to 1.57)) between antispasmodic muscle relaxants and placebo, which we rated as low certainty evidence.

Participant rating of improvement: van Tulder 2003 narratively reported the results from two high quality studies (two comparisons, n = 219) investigating antispasmodic muscle relaxants compared to placebo. The review reported that antispasmodic muscle relaxants were more effective than placebo on participant ratings of improvement. The same review reported a pooled analysis of two studies (two comparisons, n = 151) investigating benzodiazepines compared to placebo. The review reported a reduced risk for not experiencing an improvement for antispasmodic muscle relaxants compared to (RR of 0.63 (95% CI 0.42 to 0.97, I^2 = 16.75)).

Health-related quality of life: van Tulder 2003 did not report data on Health-related quality of life because it was not an outcome of interest in the review.

Workplace participation: van Tulder 2003 was unable to identify data on the outcome workplace participation.

Opioids

Two Cochrane reviews, one judged at high quality (Santos 2015) and one judged at moderate quality (Chaparro 2013), included evidence on the effects of opioids compared to placebo. Chaparro 2013 included 15 trials with a total sample size of 5540 participants with chronic LBP. Santos 2015 included 4 trials with a total sample size of 4094 participants with chronic musculoskeletal pain (e.g., chronic LBP, osteoarthritis). Neither review aimed to identify studies including participants with acute LBP.

Chronic LBP

Pain: Santos 2015 reported the results of one study (one comparison, n = 637) investigating tapentadol compared to placebo. The review reported a small between group difference favouring tapentadol (MD of -0.80 on a 0 to 10 pain intensity scale (95% CI -1.22 to -0.38)), which we rated as high certainty evidence. Chaparro 2013 reported a pooled analyses for tramadol (five studies, five comparisons, n = 1378) and strong opioids (six studies, six comparisons, n = 1887) compared to placebo. The review reported a medium between group difference favouring tramadol (SMD of -0.55 (95% CI -0.66 to -0.44, I^2 = 85.88)) and small between group difference favouring strong opioids (SMD -0.43 (95% CI -0.52 to -0.33, I^2 = 0)), which they rated at low and moderate certainty evidence respectively. Chaparro 2013 also reported a pooled analysis for buprenorphine (two studies, two comparisons, n = 653) compared to placebo which we reanalysed following the detection of an error. The review found a small between group difference favouring buprenorphine (SMD -0.41 (-0.57 to -0.26, I^2 = 0)), which we rated as very low certainty evidence.

Both reviews Chaparro 2013 and Santos 2015 also reported responder analyses for pain intensity. Santos 2015 reported the results of one study (one comparison, n = 632) investigating tapentadol compared to placebo. The review reported an increased risk for a 50% reduction in pain intensity favouring tapentadol (RR of 1.43 (95% CI 1.07 to 1.91)), which we rated as high certainty evidence. Chaparro 2013 reported a pooled analysis of two studies investigated buprenorphine compared to placebo (two comparisons, n = 594). The review reported an increased likelihood for a 30% reduction in pain intensity favouring buprenorphine (OR of 1.49 (95% CI 1.08 to 2.06, I^2 = 69.25)), which we rated as low certainty evidence. The same review also reported the results for one study (one comparison, n = 498) comparing buprenorphine to placebo. The review reported an increased likelihood of experiencing a 50% reduction in pain favouring buprenorphine (OR of 1.39 (95% CI 0.97 to 1.99)), which we rated as low certainty evidence.
Finally, Chaparro 2013 reported a pooled analysis of three studies (three comparisons, n = 819) investigating strong opioids compared to placebo. The review reported an increased likelihood in experiencing a 30% reduction in pain intensity favouring strong opioids (OR of 1.91 (95% CI 1.41 to 2.58, I^2 = 38.47)), which they rated as moderate certainty evidence. The same review reported a pooled analysis of two studies (two

comparison, n = 750) comparing strong opioids to placebo. The review reported an increased likelihood in experiencing a 50% reduction in pain intensity favouring strong opioids (OR of 1.89 (95% CI 1.34 to 2.66)), which they rated as very low certainty evidence.

Physical function: Chaparro 2013 reported a pooled analyses for tramadol (five studies, five comparisons, n = 1348), buprenorphine (one studies, one comparison, n = 101), and strong opioids (four studies, five comparisons, n = 1375) compared to placebo. The review reported a small between group difference favouring tramadol (SMD of -0.18 (95% CI -0.29 to -0.07, I^2 = 0)), which they rated at moderate certainty evidence, a small between group difference favouring buprenorphine (SMD -0.14 (-0.53 to -0.25)), which they rated as very low certainty evidence, and a small between group difference favouring strong opioids (SMD -0.26 (95% CI -0.37 to -0.15, I^2 = 0)), which they rated as moderate certainty evidence.

Safety: Santos 2015 reported the results of one study (one comparisons, n = 637) investigating tapentadol compared to placebo. The review reported an increased risk in experiencing an adverse event (RR of 1.27 (95% CI 1.14 to 1.41)), an increased risk of experiencing a serious adverse event (RR of 2.34 (95% CI 0.61 to 8.97)) and an increased risk for withdrawal to treatment due to an adverse event (RR 3.41 (95% CI 1.96 to 5.94)) for tapentadol compared to placebo, which they rated as high, moderate and high certainty evidence respectively. Chaparro 2013 reported safety data for specific adverse events, most commonly nausea, headaches, constipation, dizziness and somnolence for opioids (all types) compared to placebo. The review reported 10 studies (10 comparisons, n =3747) investigating nausea, 10 studies (10 comparisons, n=3747) investigating headaches, nine studies (nine comparisons, n=3493) investigating constipation, nine studies (nine comparisons, n =3493) investigating dizziness, and eight studies (eight comparisons, n=3257) investigating somnolence. The review reported a small between group difference in risk for experiencing nausea (RD 0.10 95% CI 0.07 to 0.14, I^2 = 62.6)), headaches (RD 0.03 (95% CI 0.01 to 0.05, I^2 = 32.22)), constipation (RD 0.07 (95% CI 0.04 to 0.11, I^2=77.74)), dizziness (RD 0.08 (95% CI 0.05 to 0.11, I^2 = 67.81)), and somnolence (RD 0.06 (95% CI 0.03 to 0.09, I^2 = 65.78)) for opioids compared to placebo, all of which we rated at low certainty evidence.

Participant rating of improvement: Chaparro 2013 was unable to identify data on the outcome workplace participation. Santos 2015 did not report separate data on participants with chronic LBP for the outcome participant rating of improvement to be included in this review.

Health-related quality of life: Chaparro 2013 did not report data on Health-related quality of life because it was not an outcome of interest in the review. Santos 2015 did not report separate data for participants with chronic LBP on the outcome health-related quality of life to be included in this review.

Workplace participation: Chaparro 2013 was unable to identify data on the outcome workplace participation. Santos 2015 did not report data on workplace participation because it was not an outcome of interest in the review.

Antidepressants

One Cochrane review, judged at low quality (Urquhart 2008), included evidence on the effects of antidepressants compared to placebo. Urquhart 2008 included 10 trials with a total sample size of 722 participants with chronic LBP. No trials were identified for participants with acute or subacute LBP. We have low overall confidence in the results from this systematic review because of one critical and one non-critical flaw. Therefore, this review may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

Chronic LBP

Pain: Urquhart 2008 reported a pooled analysis of six studies (nine comparisons, n = 376) investigating antidepressants (all types) compared to placebo. The review reported

no evidence of difference between groups on pain intensity (SMD of -0.04 (95% CI -0.25 to 0.17, I^2 = 0)), which we rated as low certainty evidence. The same review reported pooled analyses for SSRI antidepressants (three studies, three comparisons, n = 199) and TCA (three studies, four comparisons, n = 148) compared to placebo. The review reported no evidence of a difference between SSRI antidepressants (SMD of 0.11 (95% CI -0.17 to 0.39, I^2 = 0)) and TCA (SMD -0.10 (95% CI -0.51 to 0.31, I^2 = 32.31)) on pain intensity, which we rated as moderate and very low certainty evidence respectively.

Physical function: Urquhart 2008 reported a pooled analysis of two studies (two comparisons, n = 132) investigating antidepressants (all types) compared to placebo. The review reported no evidence of a difference on physical function (SMD of -0.06 (95% CI -0.40 to 0.29, I^2 = 0)), which we rated as low certainty evidence.

Safety: Urquhart 2008 did not report data on safety because it was not an outcome of interest in the review.

Participant rating of improvement: Urquhart 2008 was unable to identify data on the outcome participant rating of improvement.

Health-related quality of life: Urquhart 2008 was unable to identify data on the outcome health-related quality of life.

Workplace participation: Urquhart 2008 was unable to identify data on the outcome workplace participation.

Discussion

Summary of main results

Our main objective was to summarise the evidence from Cochrane reviews of systemic pharmacological interventions for adults with non-specific LBP on pain, function and safety. We synthesised the results of published Cochrane reviews and identified significant gaps in the evidence for a number of our comparisons of interest as well as a degree of inconsistency in approaches taken to evaluate the evidence in the included Cochrane reviews.

We included seven reviews including a total of 22,238 participants across 103 unique RCTs on paracetamol (Saragiotto 2016), NSAIDs (Enthoven 2016, van der Gaag 2020), muscle relaxants (van Tulder 2003), benzodiazepines (van Tulder 2003), opioids (Chaparro 2013, Santos 2015), and antidepressants (Urquhart 2008), mostly (5/7, 71%) for people with sub-acute or chronic LBP. All seven reviews included pain as the primary outcome and included placebo as a primary prespecified comparator. Overall, the quality of the reviews was high, we have high confidence in the results of five of the seven reviews based on the AMSTAR 2 results (Shea 2017). We have moderate confidence in the results from one review and low confidence in the results of one review based on the AMSTAR 2 evaluation.

Despite the overall high methodological quality of included reviews, we found the evidence within the included reviews to be of varying certainty. Four reviews formally rated the certainty of the evidence using the GRADE approach (Chaparro 2013, Enthoven 2016, Saragiotto 2016, van der Gaag 2020). We conducted additional GRADE assessments for five reviews for missing assessments of placebo comparisons for primary outcomes pain, function and safety (Chaparro 2013, Enthoven 2016, Santos 2015, Urquhart 2008, van Tulder 2003). The majority of the evidence was rated as low or very low certainty. Evidence from the included reviews indicates that most trials of pharmacological interventions provide potentially biased estimates and suggest only small reductions on pain in the short-term if any effect at all. Data on function is reported less often than pain and effects are typically smaller and often not observed.

For the outcome of pain intensity in acute LBP, we found moderate certainty evidence that NSAIDs provide a small effect, moderate certainty evidence that muscle relaxants provide a small effect and high certainty evidence for no evidence of difference to placebo for paracetamol. For the outcome of function in acute LBP, we found high certainty

evidence that NSAIDs provide a small effect, moderate certainty evidence that muscle relaxants provide a small effect and high certainty for no evidence of difference for paracetamol. There is little evidence available for the effects of pharmacological interventions in acute LBP beyond short-term follow up.

For the outcome of pain intensity in chronic LBP, we found moderate certainty evidence that selective NSAIDs and strong opioids provide a small effect, high certainty that tapentadol (opioid) provides a small effect, and moderate certainty evidence for no evidence of difference to placebo for SSRIs (antidepressants). For the outcome of function in chronic LBP, we found moderate certainty evidence that both strong opioids and tramadol (opioid) provide a small effect. Again, there is little evidence available for the effects of pharmacological interventions in chronic LBP beyond short-term follow up.

We found that most reviews were not able to report data across each of the pre-planned outcomes due to a lack of adequate data. Further, many of the reviews were unable to conduct quantitative syntheses due to clinical heterogeneity in the participants and comparisons reported in the included trials as well as inconsistency in the type and timing of outcome measurement.

Without valid definitions and consensus on what constitutes a minimal clinically important effect, we chose to describe the magnitude of the effect and the certainty of the evidence when discussing the findings in this Overview. Clinicians should establish the clinical importance of the effects required by their patients when interpreting the effect size and certainty of the evidence of pharmacological interventions during treatment discussions. This should include appropriate consideration from the recipients of care for the proposed benefit, safety, costs, risks, and inconveniences of therapy, rather than benchmarking against an arbitrary value (Ferreira 2013).

Overall completeness and applicability of evidence

This Overview summarises published Cochrane reviews of all RCTs examining systemic pharmacological interventions for adults with non-specific LBP. However, 6 of the 7 reviews were published more than five years ago (Chaparro 2013, Enthoven 2016, Santos 2015, Saragiotto 2016, Urquhart 2008, van Tulder 2003), two of which were published more than 10 years ago (Urquhart 2008, van Tulder 2003). There are likely additional RCTs now published that might alter the results of the reviews, in particular those relating to muscle relaxants and antidepressants (Cashin 2021, Ferraro 2021). There is a need to update a number of the Cochrane reviews. There are also several pharmacological intervention classes where Cochrane reviews are not available (e.g., anticonvulsants, systemic corticosteroids), or with very few RCTs available (e.g., paracetamol).

Although this Overview aimed to consider all durations of low back pain, most reviews included participants with sub-acute or chronic low back pain. In addition to less reviews, there were also fewer medicine classes investigated for people with acute low back pain, only paracetamol, NSAIDs and muscle relaxants were investigated in this population.

Outcome measures were inconsistent, and different measures were used at different times between RCTs and between reviews. For example, only two reviews assessed quality of life, although very few data were available (Santos 2015, Saragiotto 2016). There is a need for trialists and review authors to consider the core outcome set for non-specific LBP (Chiarotto 2015), and recommendations of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) (Dworkin 2005). Very few RCTs provided data on longer-term follow-up, currently it is unclear whether the investigated interventions have sustained benefits or long-term harms.

The definition and reporting of adverse events within each review was limited, making it difficult to assess safety for each pharmacological intervention. We found that none of the reviews provided a definition for how adverse events were considered, and when reported, the description of adverse events was vague or incomplete. Although reporting of harms in primary studies is often inadequate (Ioannidis 2009), systematic reviews can compound this problem by failing to report harms or by doing so inadequately (Zorzela 2014). Further, commonly used methods to assess benefits in systematic reviews may

not be appropriate to be used to assess harms (Qureshi 2021). For example, systematic reviewers might reach incorrect conclusions if they focus on evidence of harms found in published reports of RCTs. This is partly because RCTs are often designed to minimize adverse events (e.g., by excluding patients with medical or psychological comorbidities) and are not commonly powered to detect differences in adverse events, particularly serious (rare) adverse events, which would require larger samples and longer-term follow up. Reviews of RCTs may therefore be misleading if they do not identify any differences in adverse events (suggesting safety where this might not be case). Valid and reliable syntheses of evidence of harms requires different types of data, and different methods for synthesis compared with evidence of benefit. Together, these limitations highlight clear gaps in the evidence base of safety for pharmacological interventions. Considering these gaps, evidence on adverse events for many common analgesic medicines could be leveraged from other populations (e.g., osteoarthritis) until more robust data for LBP becomes available.

None of the included reviews reported comprehensive data on the included participants (e.g., demographic and clinical characteristics including baseline pain intensity). Without an adequate description of the included participants, it is difficult to establish for whom the evidence is applicable (i.e., the target population). More comprehensive reporting of the participants' characteristics in RCTs and the reviews that summarise them will help assess the applicability and potential generalisability of the evidence. The PROGRESS-Plus acronym serves to help RCT and review authors identify and report participant characteristics that stratify health opportunities and outcomes (O'Neill 2014).

Given the number of different pharmacological interventions, heterogeneity and low certainty of the evidence, and gaps in the current literature, it is not surprising that pharmacological intervention prescription practice varies between clinicians. In the absence of a robust evidence base, guidelines and clinical treatment will continue to be based upon other considerations including clinician experience, cost, adverse effects, regulatory approvals and established local practices.

Quality of the evidence

We used AMSTAR 2 in our evaluation of quality in the included systematic reviews. Five of the seven reviews were judged to have overall high confidence, one as moderate confidence, and one as low confidence in the results of the review. Only one review (Urquhart 2008) did not satisfy a critical domain, the review authors did not carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review. Cochrane reviews are generally regarded as having high methodological rigor and more complete reporting than non-Cochrane reviews (Dosenovic 2018, Goldkuhle 2018, Page 2016). Similar to Pollock 2017, we found that not all Cochrane reviews are high quality, at least not to current standards. However, we acknowledge that standards for conducting and reporting reviews has evolved over time, and at the time of publication, each of these reviews had gone through the Cochrane editorial process and peer review. Further, several of the included Cochrane reviews were published before methodological and reporting standards had been developed (e.g., GRADE and PRISMA), which could partly explain this finding. Finally, we did not assess the recency of publication when evaluating quality. That many of the included reviews were published greater than five years ago may decrease our confidence in their findings.

Potential biases in the overview process

We conducted the Overview according to the published protocol (Cashin 2020). We used a broad and inclusive search strategy, which was designed under expert guidance by the Cochrane Back and Neck Review Group. This was an Overview of Cochrane reviews and the search was conducted across all years up to June 2021, within the *Cochrane Database of Systematic Reviews* to identify published reviews and planned or ongoing reviews (protocols). Given the sensitive search strategy, it is reasonable to suggest this Overview offers a current summation of Cochrane reviews investigating the effect of pharmacological interventions in adults with LBP.

Several of the included reviews (Chaparro 2013, Enthoven 2016, Saragiotto 2016, van Tulder 2003, Urquhart 2008, van der Gaag 2020) were authored by members of this Overview team (MWvT, ADF, CGM). As such, there may have been a risk of potential bias with review and appraisal of this work. We minimised this risk by allocating data extraction and quality assessment to members of the author team who were not authors on the original reviews (AGC, RNNR).

We included only Cochrane reviews; there are other more recent systematic reviews on pharmacological interventions for LBP published outside of the Cochrane Library, but we are unable to comment on what biases this might introduce. Results and outcomes reported in non-Cochrane reviews may have showed different results from those presented here, though it is worth noting that non-Cochrane reviews are generally of lower quality than Cochrane reviews (Goldkuhle 2018, Page 2016, Pollock 2017).

Finally, like all overviews, we were reliant on the reporting quality of the included reviews in addition to the RCTs that they synthesised. It is possible for instance, that problems with reporting quality in the original RCTs filtered through to the systematic review and finally to the overview level. For example, all reviews explicitly stated that they included participants with non-specific low back pain. However, inadequate, or opaque reporting of the original RCTs may have meant that some RCTs could have included more heterogenous populations including radicular low back pain. In addition, we were reliant on the reporting of GRADE judgements by the included reviews. Given that GRADE assessments include an element of subjectivity, it is possible that the reviews may have used slightly different thresholds for making GRADE judgements, and as a result, some reviews may have judged the same evidence as higher certainty than others.

Agreements and disagreements with other studies or reviews

We found no published overviews of pharmacological interventions for managing LBP in adults. One review was identified which investigated recent systematic reviews of RCTs covering pharmacological interventions for chronic LBP (Koes 2018). Despite this review including both Cochrane and non-Cochrane reviews, the conclusions were consistent with ours; “The overall impression of the efficacy of pharmacological treatments for patients with chronic low back pain is rather sobering. The effects on pain reduction and improvement of function are commonly small to moderate and short lasting when compared to placebo. At the same time, the various types of drugs are not without side-effects”. The authors also highlighted the low certainty of the evidence due to systemic methodological shortcomings of the included RCTs.

Other published overviews have focused on pain relief for a specific medicine class, i.e., paracetamol (Abdel Shaheed 2021), or have conducted systematic reviews of reviews and high-quality RCTs to provide evidence to inform clinical guidelines (e.g., Chou 2017) and a Lancet LBP series (Foster 2018). There was considerable overlap with the Cochrane reviews included in these reviews with our current Overview. Despite slight variations in interpretations of the clinically relevance and certainty in the data, the reviews report consistent conclusions with this Overview and highlight common issues related to the outcomes measured and inadequate methodological conduct of included RCTs.

Authors' conclusions

Implications for practice

This Overview summarises the evidence from Cochrane reviews of RCTs of systemic pharmacological interventions for adults with non-specific LBP, and can be used by researchers, clinicians, and policy makers to assist them in decision-making and knowledge translation. We found evidence that NSAIDs and muscle relaxants may provide a small effect on pain and function, and no evidence of difference for paracetamol for acute LBP. We found no evidence for the use of opioids or any other medicines for acute LBP. For chronic LBP, we found evidence that

NSAIDs and opioids may provide a small effect on pain. We acknowledge that some of the evidence from these reviews is more than 10 years old and implications for practice may change when newer RCTs are included.

While there are some discrepancies between the recommendations from current international clinical practice guidelines for the pharmacological treatment of LBP, a substantial proportion of recommendations were consistent with the evidence from our Overview (Oliveira 2018). Most, but not all guidelines recommend NSAIDs and weak opioids for acute LBP, and NSAIDs and antidepressants for chronic low back pain (Oliveira 2018). Data from this Overview cannot contribute to the recommendation of weak opioids for acute LBP because no reviews aimed to provide relevant data. Further the recommendation for antidepressants in some guidelines does not reflect the results found in this Overview.

Overall, the available evidence suggests that pharmacological interventions for adults with non-specific LBP appear to be only marginally effective or ineffective and carry an increased risk of adverse events. There is a clear need to prioritise new effective and cost-effective treatment strategies to better help people with LBP.

Implications for research

There is a need to update most of the published Cochrane reviews and complete the three published Cochrane review protocols on pharmacological interventions for LBP. We recommend that these review updates follow updated guidance from the Cochrane Handbook for conducting systematic reviews of interventions (Higgins 2019). Further, updated guidance from this review group could improve the consistency of methods applied by review authors.

New RCTs investigating pharmacological interventions should follow the core outcome set for non-specific LBP (Chiarotto 2015) and the recommendations by IMMPACT (Dworkin 2005) to improve the synthesis of results and compatibility between trials. Trialists should also adhere to methodological safeguards to reduce bias and transparently report their findings following the Consolidated Standards of Reporting Trials statement (Schulz 2010). It is important that new RCTs clearly and comprehensively describe the characteristics of the included participants such as demographics and clinical characteristics, to better understand the study population included in the RCT. Currently, it is unclear who the available evidence is applicable to.

There are substantially fewer comparative studies for pharmacological interventions, additional comparative studies would enable us to draw firmer conclusions about which treatments are most effective. The use of network meta-analysis could also offer information to help guide clinical decision making regarding which medicine is most effective for acute and chronic LBP (Wewege 2020). More research is also needed to better understand whether combining pharmacological interventions is associated with incremental benefits, and which combinations and sequences are the most effective (Chou 2017). Finally, further research is required to determine which patients are most likely to benefit from pharmacological interventions. Currently most RCTs are underpowered to explore subgroup effects. Research initiatives which focus on identifying which patients respond more favourably to specific classes of pharmacological interventions may help individualise care for people with LBP and optimise treatment effectiveness.

Acknowledgements

We would like to thank the following people for their useful comments during peer-review of the Overview protocol:

Anne Asher, Antti Malmivaara, Dawid Pieper, Tim Carey, Roger Chou & Nuala Livingstone.

Contributions of authors

AGC, BMW, NEO, JHM conceived the idea for an Overview of pharmacological treatments for low back pain. AGC and RNNR conducted the search, screening, data extraction, AMSTAR 2 and GRADE assessments. AGC wrote the first draft included drafting the tables and results of this Overview. All authors provided review and feedback to the draft versions and approved the final version published.

Declarations of interest

MKB has received support from his institution (UNSW) for conference travel that is unrelated to the present work, from the Chiropractor's Association of Australia to speak about pain rehabilitation and from the Memorial University of Newfoundland to speak about engagement with research evidence, including evidence about medicines. MKB's salary was provided by scholarships. MKB is first author on the Cochrane Review 'Paracetamol, NSAIDs or opioid analgesics for chronic low back pain: a network meta-analysis' and was not involved in any decisions about this review in the Overview.

CGM has received competitive grants from government agencies and industry to support his research. As an invited speaker at conferences, he has had his expenses covered and also received small gifts such as a box of chocolates or a bottle of wine. He has received honoraria for marking theses, reviewing grants and preparing talks. CGM has published multiple papers on low back pain, some of which may be referenced in the review, and is on the Editorial Board of Cochrane Back and Neck review group. Mitigation of conflict of interest: CGM was not involved in editorial decisions on this review.

ADF is co-ordinating Editor of the Cochrane Back and Neck review group. Mitigation of conflict of interest: ADF will not be involved in decisions to approve or reject this review. She has published multiple papers on low back pain, some of which may be referenced in the review. The institutions where ADF works have received various grants from external organisations, including government and public institutions in Ontario, Canada, and the UK.

MWvT is on the Editorial Board of Cochrane Back and Neck review group. MWvT was co-ordinating Editor until September 2017. Mitigation of conflict of interest: MWvT will not be involved in editorial decisions on this review. MWvT has published multiple papers on low back pain, some of which may be referenced in the review. MWvT has no additional competing interest; all research funding comes from non-profit, governmental funding agencies, and all funding (including travel and stay expenses) were paid to the VU University.

HL has consulted for Cancer Council Australia; and has received funding from the Australian Health and Medical Research Council (grant no. APP1126767), and Center for Effective Global Action (CEGA) & Berkeley Initiative for Transparency in the Social Sciences (BITSS).

BMW has received payment for lectures on the non-pharmacological management of chronic low back pain. He has received honoraria for marking theses related to low back pain.

AGC, BMW, NEO, RNRR, EO and JHM have no known declarations.

This research project did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Sources of support

Internal sources

- Brunel University London, UK
  (NEO)
- Institute for Work & Health, Canada
  (ADF)

External sources

- Australian Government Research Training Program, Australia
  Postgraduate Research Scholarship (MKB, EO)

- National Health and Medical Research Council, Australia
  Research Fellowship (CGM), Program Grant (HL, CGM), Centre for Research Excellence (CGM)
- Neuroscience Research Australia (NeuRA), Australia
  PhD Candidature Supplementary Scholarship (AGC, MKB, EO)
- University of New South Wales, Australia
  Research Excellence Award (MKB)
- University of New South Wales Prince of Wales Clinical School, Australia
  Postgraduate Research Scholarship (AGC)

Differences between protocol and review

One additional review author joined the review team: Rodrigo RN Rizzo.

We did not use the planned definition of a 10-point reduction in pain intensity and disability as the minimally important difference. We instead used the definitions from the American College of Physicians and the American Pain Society (Chou 2017) to improve interpretation of the size of the effect, without making judgements on what would be considered minimally important by an individual.

Appendices

Appendix 1. Search strategy: The Cochrane Library

#1 MeSH descriptor back pain explode all trees
#2 MeSH descriptor pain explode all trees
#3 (back or spine or spinal) adj2 pain
#4 lumbar* or lumbo*
#5 backache* or back ache*
#6 (#1 OR #2 OR #3 OR #4 OR #5)

(Limited to Cochrane Reviews and Cochrane protocols)

Appendix 2. AMSTAR-2 assessment criteria

AMSTAR-2 is a 16-item critical appraisal tool to assist in identifying high quality systematic reviews. There is no summary score but an overall rating based on weaknesses across 7 critical domains*.

1. Did the research questions and inclusion criteria for the review include the components of PICO?

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?*

3. Did the review authors explain their selection of the study designs for inclusion in the review?

4. Did the review authors use a comprehensive literature search strategy?*

5. Did the review authors perform study selection in duplicate?

6. Did the review authors perform data extraction in duplicate?

7. Did the review authors provide a list of excluded studies and justify the exclusions?*

8. Did the review authors describe the included studies in adequate detail?

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?*

10. Did the review authors report on the sources of funding for the studies included in the review?

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?*

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?*

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?*

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

Ratings in overall confidence in the results of the review are as follows:

High - Zero or one non-critical weakness: The systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest

Moderate - More than one non-critical weakness*: The systematic review has more than one weakness, but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.

Low - One critical flaw with or without non-critical weaknesses: The review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

Critically low - More than one critical flaw with or without non-critical weaknesses: The review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies.

Appendix 3. Reasons for excluded study

Name of review	Reason for exclusion
Dagenais 2007	Wrong intervention
Derry 2014	Wrong patient population
Derry 2016	Wrong route of administration
Derry 2014a	Wrong route of administration
Derry 2015	Wrong route of administration
Furlan 2014	Wrong study design
Haroutounian 2012	Wrong patient population
Noble 2010	Wrong study design
Oltean 2014	Wrong intervention
Quigley 2013	Cochrane review withdrawn
Roelofs 2008	Cochrane review updated
Samuel 2012	Wrong intervention
Seidel 2013	Wrong patient population
Soares 2014	Wrong study design
Staal 2008	Wrong route of administration
Waseem 2011	Wrong route of administration
Wiffen 2010	Cochrane review withdrawn

Wiffen 2011	Cochrane review withdrawn
Zaina 2016	Wrong intervention

[Enter text here]

Appendix 4. Effectiveness comparison (Different forms of the same pharmacological intervention (e.g., selective NSAID versus a non-selective NSAID))

Outcome data is for short-term (≤ 3 months postintervention) unless otherwise stated.

NSAIDs

Acute LBP

Pain: One review performed a pooled analysis of two studies (two comparisons, n = 437) investigating selective versus non-selective NSAIDs and reported a MD of -2.6 (95% CI -9.23 to 4.03, I^2 = 56.53, low certainty evidence) on a 0 to 100 pain intensity scale (van der Gaag 2020) favouring selective NSAIDs. The same review narratively reported the results of thirteen studies (thirteen comparisons, n = 1823) investigating different types of non-selective NSAIDs and found no clear or clinically meaningful differences on pain intensity.

Physical function: One review narratively reported the results of five studies (five comparisons, n = 1006) investigating different types of non-selective NSAIDs and found no clear or clinically meaningful differences on function (van der Gaag 2020). The same review narratively reported the results from two studies (two separate comparisons, n = 444) investigating selective versus non-selective NSAIDs and reported conflicting results for improvements in function.

Safety: One review performed a pooled analysis of two studies (two comparisons, n = 444) investigating selective versus non-selective NSAIDs and reported a RR of 0.97 (95% CI 0.63 to 1.50, I^2 = 22.13, very low certainty evidence) on risk of adverse events (van der Gaag 2020). The same review narratively reported the results of fourteen studies (fourteen comparisons, n = 2337) investigating different types of non-selective NSAIDs and found no clear difference between treatments in the proportion of participants experiencing adverse events.

Participant rating of improvement: One review narratively reported the results of seven studies (seven comparisons, n = 987) investigating different types of non-selective NSAIDs and one study (one comparison, n = 333) investigating selective versus non-selective NSAIDs and found no clear or clinically meaningful differences on participant ratings of improvement (van der Gaag 2020).

Health-related quality of life: We did not find any reviews providing useable data or evidence for the effects of pharmacological interventions on this outcome for this comparison.

Workplace participation: One review narratively reported the results of one study (one comparison, n = 30) investigating different types of non-selective NSAIDs and found no differences for return to work (van der Gaag 2020).

Chronic LBP

Pain: One review narratively reported the results of two studies (two separate comparisons, n = 90) investigating different types of non-selective NSAIDs and one study (one comparison, n = 440) investigating selective versus non-selective NSAIDs and found no evidence of differences on pain intensity (Enthoven 2016).

Physical function: We did not find any reviews providing useable data or evidence for the effects of pharmacological interventions on this outcome for this comparison.

Safety: One review narratively reported the results of two studies (two separate comparisons, n = 90) investigating different types of non-selective NSAIDs and one study (one comparison, n = 440) investigating selective versus non-selective NSAIDs and found no evidence of differences in experienced adverse events (Enthoven 2016).

Opioids

Chronic LBP

Pain: One review reported data on one study (one comparison, n = 641) investigating tapentadol versus oxycodone and reported a RR of 1.16 (95% CI 0.89 to 1.51) on 50% reduction in pain intensity and a MD of 0 (95% CI -0.40 to 0.40) on a 0 to 10 pain intensity scale (Santos 2015).

Safety: One review reported the results of one study (one comparison, n = 646) investigating tapentadol versus oxycodone and reported a RR of 0.89 (95% CI 0.82 to 0.96) for the risk of experiencing and adverse event, reported a RR of 0.66 (95% CI 0.26 to 1.67) for the risk of experiencing a serious adverse event, and reported a RR of 0.49 (95% CI 0.37 to 0.66) for risk of withdrawal due to an adverse event (Santos 2015).

Appendix 5. Effectiveness comparison (Pharmacological intervention versus a different type of pharmacological intervention (e.g. NSAID versus opioid))

Outcome data is for short-term (≤ 3 months postintervention) unless otherwise stated.

NSAIDS

Acute LBP

Pain: One review performed a pooled analysis of two studies (two comparisons, n = 289) investigating NSAIDs versus paracetamol and reported a SMD of -0.12 (95% CI -0.35 to 0.12, I^2 = 0, low certainty evidence) (van der Gaag 2020). The same review narratively reported the results of four studies (four comparisons, n = 391) investigating NSAIDs versus other drugs and reported no clinically meaningful differences between the groups.

Phycial function: One review narratively reported the results of one study (one comparison, n = 219) investigating NSAIDs versus paracetamol and reported no clear differences between the groups (van der Gaag 2020).

Safety: One review narratively reported the results of two studies (two comparisons, n = 289) and found low certainty evidence that NSAIDs led to a greater proportion of participants experiencing an adverse event compared to paracetamol (van der Gaag 2020). The same review narratively reported the results of four studies (four comparisons, n = 391) and found that those who took NSAIDs were more likely to report adverse events than those who took other drugs.

Participant rating of improvement: One review performed a pooled analysis of two studies (two comparisons, n = 162) investigating NSAIDs versus other drugs and reported a RR 1.01 (95% CI 0.81 to 1.25, I^2 = 0, moderate certainty evidence) (van der Gaag 2020).

Workplace participation: One review narratively reported the results of one study (three comparisons, n = 45) investigating NSAIDs versus paracetamol and reported no clear differences between groups (van der Gaag 2020).

Chronic LBP

Safety: One review reported data on one study (one comparison, n = 28) investigating NSAIDs versus paracetamol, one study (one comparison, n = 1583) investigating NSAIDs versus tramadol, and one study (one comparison, n = 72) investigating pregabalin and reported a RR of 1.50 (95% CI 0.15 to 14.68), a RR of 0.86 (95% CI 0.75 to 0.91), and a RR of 0.80 (95% CI 0.23 to 2.74) for the risk of reporting adverse events respectively (Enthoven 2016).

Participant rating of improvement: One review reported data on one study (one comparison, n = 28) investigating NSAIDs versus paracetamol and one study (one comparison, n = 1583) investigating NSAIDs versus tramadol and reported a RR of 1.39 (95% CI 0.82 to 2.37) and a RR of 1.26 (95% CI 1.16 to 1.38) respectively (Enthoven 2016).

Workplace participation: We did not find any reviews providing useable data or evidence for the effects of pharmacological interventions on this outcome for this comparison.

Opioids

Chronic LBP

Physical function: One review reported the results of one study (one comparison, n = 56) investigating opioids versus antidepressants and reported a SMD of -0.11 (95% CI -0.63 to 0.42, very low certainty evidence) (Chaparro 2015).

Antidepressants

Chronic LBP

Pain: One review performed a pooled analysis of two studies (two comparisons, n = 272) investigating opioids versus antidepressants and reported a SMD of 0.21 (95% CI -0.03 to 0.45, I^2 = 0, very low certainty evidence) (Urquhart 2008). The same review reported the results of one study (one comparison, n = 1583) investigating tramadol versus celecoxib (a NSAID) and reported a RR of 0.82 (95% CI 0.76 to 0.90) for reducing pain intensity and an OR of 0.63 (95% CI 0.52 to 0.77, very low certainty evidence) for a 30% reduction in pain intensity.

Appendix 6. Effectiveness comparison (Pharmacological intervention versus a non-pharmacological intervention (e.g., NSAID versus spinal manipulative therapy))

Outcome data is for short-term (≤ 3 months postintervention) unless otherwise stated.

NSAIDs

Acute LBP

Pain: One review narratively reported the results of four studies (six comparisons, n = 353) and found very low certainty evidence that most studies (3/4) showed no clinically meaningful difference between NSAIDs and spinal manipulation at short-term follow up, and no clear difference in the intermediate term follow up (van der Gaag 2020).

Physical function: One review narratively reported the results of two studies (two comparisons, n = 193) and found conflicting results; one study found a clinically meaningful difference between NSAIDs and spinal manipulation in favour of spinal

manipulation and the other did not (very low certainty evidence) (van der Gaag 2020). The same review reported no clear difference in the intermediate term follow up.

Safety: One review narratively reported the results of two studies (two comparisons, n = 189) and found very low certainty evidence of no clear difference between NSAIDs and spinal manipulation on risk of adverse events (van der Gaag 2020).

Participant rating of improvement: One review narratively reported the results of two studies (three comparisons, n = 180) and found conflicting results; one study found a clinically meaningful difference between NSAIDs and spinal manipulation in favour of spinal manipulation and the other did not, very low certainty evidence (van der Gaag 2020). The same review reported no clear difference in the intermediate term follow up.

Workplace participation: One review narratively reported the results of one study (one comparison, n = not reported) and found low certainty evidence of no clear difference between NSAIDs and spinal manipulation on workplace participation (van der Gaag 2020).

Chronic LBP

Pain: One review narratively reported the results from one study (one comparison, n = 201) and reported that there was no difference in pain reduction between NSAIDs and home-based exercise (Enthoven 2016).

Phycial function: One review narratively reported the results from one study (one comparison, n = 201) and reported that functional status was better with home-based exercise compared to NSAIDs (Enthoven 2016).

References

References to included reviews

Chaparro 2013

Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC. Opioids compared to placebo or other treatments for chronic low-back pain. Cochrane Database of Systematic Reviews 2013, Issue 8.

Enthoven 2016

Enthoven WTM, Roelofs PDDM, Deyo RA, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for chronic low back pain. Cochrane Database of Systematic Reviews 2016, Issue 2.

Santos 2015

Santos J, Alarcão J, Fareleira F, Vaz Carneiro A, Costa J. Tapentadol for chronic musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2015, Issue 5.

Saragiotto 2016

Saragiotto BT, Machado GC, Ferreira ML, Pinheiro MB, Abdel Shaheed C, Maher CG. Paracetamol for low back pain. Cochrane Database of Systematic Reviews 2016, Issue 6.

Urquhart 2008

Urquhart DM, Hoving JL, Assendelft WJJ, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. Cochrane Database of Systematic Reviews 2008, Issue 1.

van der Gaag 2020

van der Gaag WH, Roelofs PDDM, Enthoven WTM, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for acute low back pain. Cochrane Database of Systematic Reviews 2020, Issue 4.

van Tulder 2003

van Tulder MW, Touray T, Furlan AD, Solway S, Bouter LM. Muscle relaxants for non-specific low-back pain. Cochrane Database of Systematic Reviews 2003, Issue 2.

References to excluded reviews

Bagg 2018

Bagg M K, McLachlan A J, Maher C G, Kamper S J, Williams C M, Henschke N, Wand B M, Moseley G L, Hübscher M, O'Connell N E, al et. Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis. Cochrane Database of Systematic Reviews 2018, Issue 6.

Bezerra 2014

Bezerra D M, El Dib R, Vidal E I O, De Barros G A M, Chou R, Fukushima F. Anticonvulsants for chronic low-back pain. Cochrane Database of Systematic Reviews 2014, Issue 6 .

Chou 2016a

Chou R, Pinto R Z, Fu R, Lowe R A, Henschke N, Dana T. Systemic corticosteroids for radicular and non-radicular low back pain. Cochrane Database of Systematic Reviews 2016, Issue 12.

Dagenais 2007

Dagenais S, Yelland MJ, Del Mar C, Schoene ML. Prolotherapy injections for chronic low-back pain. Cochrane Database of Systematic Reviews 2007, Issue 2.

Derry 2014

Derry CJ, Derry S, Moore RA. Caffeine as an analgesic adjuvant for acute pain in adults. Cochrane Database of Systematic Reviews 2014, Issue 12.

Derry 2014a

Derry S, Matthews PRL, Wiffen PJ, Moore RA. Salicylate-containing rubefacients for acute and chronic musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2014, Issue 11.

Derry 2015

Derry S, Moore RA, Gaskell H, McIntyre M, Wiffen PJ. Topical NSAIDs for acute musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2015, Issue 6.

Derry 2016

Derry S, Conaghan P, Da Silva JAP, Wiffen PJ, Moore RA. Topical NSAIDs for chronic musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2016, Issue 4.

Furlan 2014

Furlan AD, Irvin E, Kim J, Van Eerd D, Carnide N, Munhall C, Fortune M, Mahood Q, van Tulder MW. Impact of long-term opioid use for chronic non-cancer pain on

misuse, abuse or addiction, overdose, falls and fractures. Cochrane Database of Systematic Reviews 2014, Issue 4.

Haroutounian 2012

Haroutounian S, McNicol ED, Lipman AG. Methadone for chronic non-cancer pain in adults. Cochrane Database of Systematic Reviews 2012, Issue 11.

Noble 2010

Noble M, Treadwell JR, Tregear SJ, Coates VH, Wiffen PJ, Akafomo C, Schoelles KM, Chou R. Long-term opioid management for chronic noncancer pain. Cochrane Database of Systematic Reviews 2010, Issue 1.

Oltean 2014

Oltean H, Robbins C, van Tulder MW, Berman BM, Bombardier C, Gagnier JJ. Herbal medicine for low-back pain. Cochrane Database of Systematic Reviews 2014, Issue 12.

Quigley 2013

Quigley C. Hydromorphone for acute and chronic pain. Cochrane Database of Systematic Reviews 2013, Issue 10.

Roelofs 2008

Roelofs PDDM, Deyo RA, Koes BW, Scholten RJPM, van Tulder MW. Non-steroidal anti-inflammatory drugs for low back pain. Cochrane Database of Systematic Reviews 2008, Issue 1.

Samuel 2012

Samuel S, David KS, Gray RJ, Tharyan P. Fusion versus conservative management for low-grade isthmic spondylolisthesis. Cochrane Database of Systematic Reviews 2012, Issue 10.

Seidel 2013

Seidel S, Aigner M, Ossege M, Pernicka E, Wildner B, Sycha T. Antipsychotics for acute and chronic pain in adults. Cochrane Database of Systematic Reviews 2013, Issue 8.

Soares 2014

Soares A, Andriolo RB, Atallah ÁN, da Silva EMK. Botulinum toxin for myofascial pain syndromes in adults. Cochrane Database of Systematic Reviews 2014, Issue 7.

Staal 2008

Staal JB, de Bie R, de Vet HCW, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain. Cochrane Database of Systematic Reviews 2008, Issue 3.

Waseem 2011

Waseem Z, Boulias C, Gordon A, Ismail F, Sheean G, Furlan AD. Botulinum toxin injections for low-back pain and sciatica. Cochrane Database of Systematic Reviews 2011, Issue 1.

Wiffen 2010

Wiffen PJ, Collins S, McQuay HJ, Carroll D, Jadad A, Moore RA. Anticonvulsant drugs for acute and chronic pain. Cochrane Database of Systematic Reviews 2010,

Issue 1.

Wiffen 2011

Wiffen PJ, McQuay HJ, Edwards J, Moore RA. Gabapentin for acute and chronic pain. Cochrane Database of Systematic Reviews 2011, Issue 3.

Zaina 2016

Zaina F, Tomkins-Lane C, Carragee E, Negrini S. Surgical versus non-surgical treatment for lumbar spinal stenosis. Cochrane Database of Systematic Reviews 2016, Issue 1.

Additional references

Abajobir 2017

Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study. Lancet 2017;390(10100):1211-59.

Abdel Shaheed 2021

Abdel Shaheed C, Ferreira GE, Dmitritchenko A, McLachlan AJ, Day RO, Saragiotto B, et al. The efficacy and safety of paracetamol for pain relief: an overview of systematic reviews. Medical Journal of Australia 2021;214(7):324-331.

Anderson 2022

DB Anderson, Abdel Shaheed C. Medications for Treating Low Back Pain in Adults. Evidence for the Use of Paracetamol, Opioids, Nonsteroidal Anti-inflammatories, Muscle Relaxants, Antibiotics, and Antidepressants: An Overview for Musculoskeletal Clinicians. JOSPT 2022;52(7):425-431.

Bachhuber 2016

Bachhuber M, Hennessy S, Cunningham C, Starrels J. Increasing benzodiazepine prescriptions and overdose mortality in the United States, 1996-2013. American Journal of Public Health 2016;106(4):686-8.

Balshem 2011

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of Clinical Epidemiology 2011;64:401-6.

Barden 2006

Barden J, Derry S, McQuay HJ, Moore AR. Bias from industry trial funding? A framework, a suggested approach, and a negative result. Pain 2006;121(3):207-218.

Bourgeois 2010

Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. Ann Intern Med 2010;153(3):158-166.

Brune 2015

Brune K, Patrignani P. New insights into the use of currently available non-steroidal anti-inflammatory drugs. Journal of Pain Research 2015;8:105-18.

Carey 2009

Carey TS, Freburger JK, Holmes GM, Castel L, Darter J, Agans R, et al. A long way to go: Practice patterns and evidence in chronic low back pain care. Spine 2009;7(34):718-24.

Cashin 2020

Cashin AG, Wand BM, O'Connell NE, Lee H, Bagg MK, O'Hagan E, et al. Pharmacological treatments for low back pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews 2020, Issue 11.

Cashin 2021

Cashin AG, Folly T, Bagg MK, Wewege MA, Jones MD, Ferraro MC et al . Efficacy, acceptability, and safety of muscle relaxants for adults with non-specific low back pain: systematic review and meta-analysis. BMJ 2021;374:n1446.

Chiarotto 2015

Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. European Spine Journal 2015;24(6):1127-42.

Chopra 2003

Chopra SS. Industry Funding of Clinical Trials: Benefit or Bias? JAMA 2003;290(1):113-114.

Chou 2010

Chou R. Pharmacological management of low back pain. Drugs 2010;70(4):387-402.

Chou 2016

Chou R, Pinto R, Fu R, Lowe R, Henschke N, Duna T. Systemic corticosteroids for radicular and non-radicular low back pain. Cochrane Database of Systematic Reviews 2016, Issue 12. Art. No: CD012450. [DOI: 10.1002/14651858.CD012450]

Chou 2017

Chou R, Deyo R, Friedly J, Skelly A, Weimer M, Fu R, et al. Systemic pharmacologic therapies for low back pain: a systematic review for an American College of Physicians clinical practice guideline. Annals of Internal Medicine 2017;166(7):480-92.

Chou 2018a

Chou L, Ranger TA, Peiris W, Cicuttini FM, Urquhart DM, Sullivan K, et al. Patients' perceived needs of health care providers for low back pain management: a systematic scoping review. Spine 2018;18(4):691-711.

Chou 2018b

Chou L, Ranger TA, Peiris W, Cicuttini FM, Urquhart DM, Sullivan K, et al. Patients' perceived needs for medical services for non-specific low back pain: A systematic scoping review. PLOS One 2018;13(11):1-29.

Cleeland 1989

Cleeland CS. Measurement of pain by subjective report. In: Chapman CR, Loeser JD, editors(s). Issues in pain measurement. Advances in pain research and management. Vol. 12. New York: Raven Press, 1989:391-403.

Cohen 2001

Cohen SP, Abdi S. New developments in the use of tricyclic antidepressants for the management of pain. Current Opinion in Anaesthesiology 2001;14(5):505-11.

Covidence [Computer program]

Covidence. Melbourne, Australia: Veritas Health Innovation, Accessed prior to 24 November 2020. Available at covidence.org.

Derry 2019

Derry S, Bell RF, Straube S, Wiffen PJ, Aldington D, Moore RA. Pregabalin for neuropathic pain in adults. Cochrane Database of Systematic Reviews 2019, Issue 1. Art. No: CD007076. [DOI: 10.1002/14651858.CD007076.pub3]

Deyo 2015

Deyo RA, Von Korff M, Duhrkoop D. Opioids for low back pain. BMJ 2015;350:g6380.

Dieleman 2016

Dieleman JL, Baral R, Birger M, Bui AL, Bulchis A, Chapin A, et al. US spending on personal health care and public health, 1996-2013. Journal of the American Medical Association 2016;316(24):2627-46.

Dieleman 2020

Dieleman JL, Cao J, Chapin A, Chen C, Li Z, Liu A, et al. US health care spending by payer and health condition, 1996-2016. Journal of the American Medical Association 2020;323(9):863-84.

Dosenovic 2018

Dosenovic S, Jelicic Kadic A, Vucic K, Markovina N, Pieper D, Puljak L. Comparison of methodological quality rating of systematic reviews on neuropathic pain using AMSTAR and R-AMSTAR. BMC Med Res Methodol 2018;18:37.

Dworkin 2005

Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113(1):9-19.

EndNote 2017 [Computer program]

EndNote. Clarivate Analytics, 2017.

Fairbank 1980

Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66(8):271-273.

Ferraro 2021

Ferraro MC, Bagg MK, Wewege MA, Cashin AG, Leake HB, Rizzo RRN, et al . Efficacy, acceptability, and safety of antidepressants for low back pain: a systematic review and meta-analysis. Syst Rev 2021;10:62.

Ferreira 2013

Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, et al. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and

physiotherapy for chronic low back pain: a benefit-harm trade-off study. J Clin Epidemiol 2013;66(12):1397-1404.

Foster 2018

Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 2018;391(10137):2368-2383.

Furlan 2009

Furlan AD, Pennick V, Bombardier C, van Tulder M, Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine 2009;34(18):1929-1941.

Furlan 2015

Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al, for the Editorial Board of the Cochrane Back, Neck Group. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine 2015;40(21):1660-73.

Goldkuhle 2018

Goldkuhle M, Narayan VM, Weigl A, Dahm P, Skoetz N. A systematic assessment of Cochrane reviews and systematic reviews published in high-impact medical journals related to cancer. BMJ Open 2018;8:e020869.

Gore 2012

Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Use and costs of prescription medications and alternative treatments in patients with osteoarthritis and chronic low back pain in community-based settings. Pain Practice 2012;12(7):550-60.

Gouveia 2017

Gouveia N, Rodrigues A, Ramiro S, Eusébio M, Machado PM, Canhão H, et al. The use of analgesic and other pain-relief drugs to manage chronic low back pain: results from a national survey. Pain Practice 2017;17(3):353-65.

Guy 1976

Guy W. ECDEU assessment manual for psychopharmacology . US Government Printing Office, 1976.

Hancock 2007

Hancock MJ, Maher CG, Latimer J, Spindler MF, McAuley JH, Laslett M, et al. Systematic review of tests to identify the disc, SIJ or facet joint as the source of low back pain. European Spine Journal 2007;16(10):1539-50.

Harmer 2017

Harmer CJ, Duman RS, Cowen PJ. How do antidepressants work? New perspectives for refining future treatment approaches. Lancet Psychiatry 2017;4(5):409-18.

Hart 2015

Hart OR, Uden RM, McMullen JE, Ritchie MS, Williams TD, Smith BH. A study of National Health Service management of chronic osteoarthritis and low back pain. Primary Health Care Research & Development 2015;16(2):157-66.

Hennessy 2020

Hennessy EA, Johnson BT. Examining overlap of included studies in meta-reviews: Guidance for using the corrected covered area index. Research synthesis methods 2020;11:134-145.

Henschke 2008

Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. BMJ 2008;337:a171.

Higgins 2011

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

Higgins 2019

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. 2nd edition. Chichester (UK): John Wiley & Sons Ltd., 2019.

Hong 2013

Hong J, Reed C, Novick D, Happich M. Costs associated with treatment of chronic low back pain. Spine 2013;38(1):75-82.

Hood 2014

Hood S, Norman A, Hince D, Melichar J, Hulse G. Benzodiazepine dependence and its treatment with low dose flumazenil. British Journal of Clinical Pharmacology 2014;77(2):285-94.

Hoy 2012

Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis & Rheumatism 2012;64(6):2028-37.

Hunt 2018

Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: planning a relevant research question and objective for an overview. Systematic Reviews 2018;7(1):39.

Ioannidis 2009

Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;141(10):781-788.

Ivanova 2011

Ivanova JI, Birnbaum HG, Schiller M, Kantor E, Johnstone BM, Swindle RW. Real-world practice patterns, health-care utilization, and costs in patients with low back pain: the long road to guideline-concordant care. Spine 2011;7(11):622-32.

Jóźwiak-Bebenista 2014

Jóźwiak-Bebenista M, Nowak JZ. Paracetamol: mechanism of action, applications and safety concern. Acta Poloniae Pharmaceutica 2014;71(1):11-23.

Kalso 2004

Kalso E, Edwards JE, Moore RA, McQuay HJ. Opioids in chronic non-cancer pain: systematic review of efficacy and safety. Pain 2004;112(3):372-80.

Koes 2006

Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006;332(7555):1430-4.

Koes 2007

Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. BMJ 2007;334(7607):1313-7.

Koes 2018

Koes BW, Backes D, Bindels PJ. Pharmacotherapy for chronic non-specific low back pain: current and future options. Expert Opinion on Pharmacotherapy 2018;19(6):537-545.

Kyu 2018

Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1859-1922.

Lim 2019

Lim YZ, Chou L, Au RT, Seneviwickrama KM, Cicuttini FM, Briggs AM, et al. People with low back pain want clear, consistent and personalised information on prognosis, treatment options and self-management strategies: a systematic review. Journal of Physiotherapy 2019;65(3):124-35.

Lundh 2017

Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. Cochrane Database of Systematic Reviews 2017, Issue 2.

Maher 2017

Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017;389(10070):736-47.

Maizels 2005

Maizels M, Mccarberg B. Antidepressants and antiepileptic drugs for chronic non-cancer pain. American Family Physician 2005;71(3):483-90.

McCrae 2018

McCrae JC, Morrison EE, MacIntyre IM, Dear JW, Webb DJ. Long-term adverse effects of paracetamol - a review. Br J Clin Pharmacol 2018;84(10):2218-2230.

Meader 2014

Meader N, King K, Llewellyn A, Norman G, Brown J, Rodgers M, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. BMC Systematic Reviews 2014;3:82.

Menezes Costa 2012

Menezes Costa LC, Maher C, Hancock M, McAuley J, Herbert R, Costa L. The prognosis of acute and persistent low back pain: a meta-analysis. Canadian

Medical Association Journal 2012;184(11):613-24.

Michaleff 2012

Michaleff ZA, Harrison C, Britt H, Lin CW, Maher CG. Ten-year survey reveals differences in GP management of neck and back pain. European Spine Journal 2012;21(7):1283-9.

Micó 2006

Micó JA, Ardid D, Berrocoso E, Eschalier A. Antidepressants and pain. Trends in Pharmacological Sciences 2006;27(7):348-54.

O'Neill 2014

O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. Journal of clinical epidemiology 2014;67:56-64.

Oliveira 2018

Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot JF, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. Eur Spine J 2018;27(11):2791-2803.

Page 2016

Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. PLoS medicine 2016;13(5):e1002028.

Piccoliori 2013

Piccoliori G, Engl A, Gatterer D, Sessa E, in der Schmitten J, Abholz HH. Management of low back pain in general practice - is it of acceptable quality: an observational study among 25 general practices in South Tyrol (Italy). BMC Family Practice 2013;14(1):148.

Pieper 2014

Pieper D, Antoine S, Mathes T, Neugebauer E, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. Journal of Clinical Epidemiology 2014;67(4):368-75.

Pollock 2017

Pollock M, Fernandes RM, Hartling L. Evaluation of AMSTAR to assess the methodological quality of systematic reviews in overviews of reviews of healthcare interventions. BMC Medical Research Methodology 2017;17:1-13.

Qaseem 2017

Qaseem A, Wilt T, McLean R, Forciea M. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Annals of Internal Medicine 2017;166(7):514-30.

Qureshi 2021

Qureshi R, Mayo-Wilson E, Li T. Summaries of harms in systematic reviews are unreliable Paper 1: An introduction to research on harms. Journal of Clinical Epidemiology 2021;143:186-196.

Rivat 2016

Rivat C, Ballantyne J. The dark side of opioids in pain management: basic science explains clinical observation. Pain Reports 2016;1(2):e570.

Roberts 2016

Roberts E, Delgado Nunes V, Buckner S, Latchem S, Constanti M, Miller P, et al. Paracetamol: not as safe as we thought? A systematic literature review of observational studies. Ann Rheum Dis 2016;75(3):552-559.

Roland 1983

Roland M, Morris R. A study of the natural history of back pain, part 1: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-144.

Schofield 2008

Schofield DJ, Shrestha RN, Passey ME, Earnest A, Fletcher SL. Chronic disease and labour force participation among older Australians. Medical Journal of Australia 2008;189(8):447-50.

Schofield 2012

Schofield DJ, Callander EJ, Shrestha RN, Percival R, Kelly SJ, Passey ME. Labor force participation and the influence of having back problems on income poverty in Australia. Spine 2012;37(13):1156-63.

Schulz 2010

Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.

See 2008

See S, Ginzburg R. Choosing a skeletal muscle relaxant. American Family Physician 2008;78(3):365-70.

Shea 2017

Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, 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;358:j4008.

Trevor 2018

Trevor A, Katzung B, Masters S. Katzung and Trevor's pharmacology: examination and board review. 12th edition. New York (NY): McGraw-Hill Education, 2018.

van der Laan 2008

van der Laan S, Meijer O. Pharmacology of glucocorticoids: beyond receptors. European Journal of Pharmacology 2008;585(2-3):483-91.

van der Linden 2009

van der Linden MW, van der Bij S, Welsing P, Kuipers EJ, Herings RM. The balance between severe cardiovascular and gastrointestinal events among users of selective and non-selective non-steroidal anti-inflammatory drugs. Annals of the Rheumatic Diseases 2009;68:668-673.

van Tulder 1997

van Tulder MW, Assendelft WJ, Koes BW, Bouter LM, Editorial Board of the Cochrane Collaboration Back Review Group. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for spinal disorders. Spine 1997;22(20):2323-2330.

van Tulder 2003a

van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine 2003;28(12):1290-1299.

Waljee 2017

Waljee A, Rogers M, Lin P, Singal A, Stein J, Marks R et al. Short term use of oral corticosteroids and related harms among adults in the United States: population based cohort study. BMJ 2017;357:j1415.

Ware 2000

Ware JE. SF-36 health survey update. Spine 2000;25(24):3130-3139.

Wewege 2020

Wewege MA, Bagg MK, Jones MD, McAuley JK, The ANiMALIA investigators. Analgesic medicines for adults with low back pain: protocol for a systematic review and network meta-analysis. Systematic Reviews 2020;9:255.

Witenko 2014

Witenko C, Moorman-Li R, Motycka C, Duane K, Hincapie-Castillo J, Leonard P, et al. Considerations for the appropriate use of skeletal muscle relaxants for the management of acute low back pain. Pharmacy & Therapeautics 2014;39(6):427-35.

Zorzela 2014

Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, Joffe A, et al. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ 2014;348:f7668.

Figures and tables

Additional tables

Table 1

Reference	Review Aim	Dates/notes
Bagg 2018	To determine the analgesic effects, safety, effect on function, and relative rank according to analgesic effect, safety and effect on function of a single course of opioid analgesics, NSAIDs or paracetamol or combinations of these medicines.	Published 09 June 2018
Bezerra 2014	To assess the effectiveness and safety of anticonvulsants for the management of chronic low-back pain, with or without radiculopathy.	Published 23 June 2014
Chou 2016a	To determine the benefits and harms of systemic corticosteroids compared with placebo or no systemic corticosteroid for patients with acute, subacute, or chronic radicular or non-radicular low back pain.	Published 05 December 2016

Details of ongoing reviews

Table 2

Review	Date of last search	Total number of studies	Population	Interventions	Comparisons	Outcomes planned

		(total number of participants)				
Chaparro 2013	October 2012	15 (5540)	Adults (≥18 years) with chronic (≥12 weeks), non-specific low back pain with or without leg pain	Opioids	Placebo, other drugs	Primary: pain, function, global improvement, proportion of patients reporting 30% and 50% pain relief Secondary: work-related disability, treatment related adverse events, healthcare usage, non-opioid medication consumption, addiction, overdose-related events
Enthoven 2016	June 2015	13 (4807)	Adults (≥18 years) with chronic (≥12 weeks), non-specific low back pain	Non-steroidal anti-inflammatory drugs	Placebo, NSAID, other drugs, other non-drug treatments	Primary: pain, global measure of improvement, back pain-specific functional status, return to work, adverse events Secondary: physiological outcomes, generic functional status, healthcare consumption
Santos 2015	March 2014	4 (4094)	Adults (≥18 years) with chronic (≥12 weeks), moderate-severe (≥4/10 NRS) musculoskeletal pain (including non-specific low back pain)	Opioids (tapentadol)	Placebo, other drugs (oxycodone)	Primary: pain, safety Secondary: patient global impression of change, quality of life scores, requirements for breakthrough analgesia, functional health status and wellbeing, sleep

						evaluation, withdrawal rate, adverse events
Saragiotto 2016	August 2015	2 (1785)	People with acute (<6 weeks), non-specific low back pain	Paracetamol	Placebo	Primary: pain, disability Secondary: quality of life, function, adverse events, global impression of recovery, sleep quality, patient adherence, use of rescue medication
Urquhart 2008	November 2008	10 (722)	Adults (≥18 years) with non-specific low back pain with or without leg pain	Antidepressants	Placebo	Primary: pain, overall improvement, proportion of patients recovered, back pain-specific functional status, return to work, depression Secondary: physiological outcomes, generic functional status
van der Gaag 2020	January 2020	32 (5356)	Adults (≥18 years) with acute (<12 weeks) non-specific low back pain with or without leg pain	Non-steroidal anti-inflammatory drugs	Placebo, NSAID, paracetamol, other drug, non-drug treatment	Primary: pain, back pain-specific functional status, global measure of improvement, adverse events, return to work Secondary: none
van Tulder 2003	October 2002	30 (2884)	People with non-specific low back pain with or without leg pain	Muscle relaxants (antispasmodics, antispastics), Benzodiazepines	Placebo, NSAIDs, other muscle relaxants, placebo + analgesics/NSAIDs	Primary: pain, global measure of improvement, back specific function, return to work, physiological outcomes, generic functional status

						Secondary: none

Characteristics of included reviews

Table 3

Review	Number of studies assessed	GRADE	Methodological quality assessment tool	Risk of bias assessment (from review authors)
Chaparro 2013	15	Yes	2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group (Furlan 2009)	Random sequence generation: low risk 10/15 Allocation concealment: low risk 6/15 Blinding (participants): low risk 14/15 Blinding (providers): low risk 8/15 Blinding (outcome assessors): low risk 2/15 Incomplete outcome data (drop-outs): low risk 0/15 Incomplete outcome data (ITT): low risk 12/15 Similarity of baseline characteristics: low risk 11/15 Co-interventions avoided or similar: low risk 14/15 Compliance acceptable: low risk 4/15 Timing of outcome assessment similar: low risk 14/15 Free from selective reporting: low risk 9/15
Enthoven 2016	13	Yes	2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group (Furlan 2009)	Random sequence generation: low risk 6/13 Allocation concealment: low risk 4/13 Blinding (participants): low risk 10/13 Blinding (providers): low risk 8/13 Blinding (outcome assessors): low risk 10/13 Incomplete outcome data (drop-outs): low risk 6/13 Incomplete outcome data (ITT): low risk

				3/13 Similarity of baseline characteristics: low risk 10/13 Co-interventions avoided or similar: low risk 10/13 Compliance acceptable: low risk 5/13 Timing of outcome assessment similar: low risk 12/13 Selective reporting: low risk 2/13
Santos 2015	4	No	Cochrane 'risk of bias' tool 1.0 (Higgins 2011)	Random sequence generation: low risk 4/4 Allocation concealment: low risk ¾ Blinding (participants, providers): low risk ¾ Blinding (outcome assessors): low risk ¾ Incomplete outcome data: low risk 0/4 Selective reporting: low risk 4/4 Duration: low risk 4/4 Outcomes: low risk 2/4 Size: low risk 4/4
Saragiotto 2016	2	Yes	Cochrane 'risk of bias' tool 1.0 (Higgins 2011)	Random sequence generation: low risk ½ Allocation concealment: low risk ½ Blinding (participants, providers): low risk1/2 Blinding (outcome assessors): low risk ½ Incomplete outcome data: low risk 2/2 Selective reporting: low risk ½
Urquhart 2008	10	No	Updated Method Guidelines for Systematic Reviews in the Cochrane Collaboration Back Review Group (van Tulder 2003a)	Random sequence generation: low risk 5/10 Allocation concealment: low risk 4/10 Blinding (participants): low risk 10/10

					Blinding (providers): low risk 9/10 Blinding (outcome assessors): low risk 9/10 Incomplete outcome data (drop-outs): low risk 3/10 Incomplete outcome data (ITT): low risk 8/10 Similarity of baseline characteristics: low risk 7/10 Co-interventions avoided or similar: low risk 3/10 Compliance acceptable: low risk 3/10 Timing of outcome assessment similar: low risk 9/10
	van der Gaag 2020	32	Yes	2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group (Furlan 2015)	Random sequence generation: low risk 12/32 Allocation concealment: low risk 10/32 Blinding (participants, providers): low risk 12/32 Blinding (outcome assessors): low risk 12/32 Incomplete outcome data: low risk 10/32 Selective reporting: low risk 3/32 Other bias: low risk 32/32
	van Tulder 2003	30	No	Method Guidelines for Systematic Reviews in the Cochrane Collaboration Back Review Group for Spinal Disorderd (van Tulder 1997)	Random sequence generation: low risk 6/30 Allocation concealment: low risk 2/30 Blinding (participants): low risk 28/30 Blinding (providers): low risk 28/30 Blinding (outcome assessors): low risk 28/30 Incomplete outcome data (drop-outs): low risk 20/30 Incomplete outcome data (ITT): low risk 12/30

				Similarity of baseline characteristics: low risk 17/30 Co-interventions avoided or similar: low risk 12/30 Compliance acceptable: low risk 5/30 Timing of outcome assessment similar: low risk 27/30

Risk of bias in the included reviews

Table 4

AMSTAR 2 item	Cochrane review						
	Chaparro 2013	Enthoven 2016	Santos 2015	Saragiotto 2016	Urquhart 2008	van der Gaag 2020	van Tulder 2003
1. *Did the research questions and inclusion criteria for the review include the components of PICO?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. *Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. *Did the review authors explain their selection of the study designs for inclusion in the review?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4. *Did the review authors use a comprehensive literature search strategy?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5. *Did the review authors perform study selection in duplicate?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6. *Did the review authors perform data extraction in duplicate?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. *Did the review authors provide a list of excluded studies and justify the exclusions?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8. *Did the review authors describe the included studies in adequate detail?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9. *Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. *Did the review authors report on the sources of funding for the studies included in the review?*	Yes	Yes	Yes	Yes	No	Yes	Yes
11. I*f meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. *If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?*	No	Yes	Yes	Yes	Yes	Yes	Yes
13. *Did the review authors account for RoB in primary studies when interpreting/discussing the results of the review?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14. *Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?*	No	Yes	Yes	Yes	Yes	Yes	Yes

15. *If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?*[*]	Yes	Yes	Yes	Yes	No	Yes	Yes
16. *Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rating of overall confidence	**Moderate**	**High**	**High**	**High**	**Low**	**High**	**High**

AMSTAR 2 quality assessment

Table 5

Pharmacological interventions for low back pain in adults

Comparison medicine vs placebo ≤3 months postintervention (short-term)

Outcome	Low back pain duration	Intervention and placebo	Relative effect (95% CI)	I^2	Number of trials (participants)	Certainty of the evidence (GRADE)	Comments
Pain	acute	Paracetamol	MD, 0.49 (-1.99 to 2.97)	0	1 (1516)	High	
Pain	acute	NSAID	MD, -7.29 (-10.98 to -3.61)	35.18	4 (815)	Moderate	Downgraded for study limitations
Pain	acute	muscle relaxants (non-benzodiazepine antispasmodic)	RR, 0.58 (0.45 to 0.76)	0	3 (244)	Moderate	Downgraded for study limitations
Pain	chronic	benzodiazepine	RR, 0.71 (0.54 to 0.93)	0	2 (146)	Low	Downgraded for study limitations, imprecision
Pain	chronic	antidepressant	SMD, -0.04 (-0.25 to 0.17)	0	6 (376)	Low	Downgraded for study limitations, imprecision
Pain	chronic	Antidepressant (SSRI)	SMD, 0.11 (-0.17 to 0.39)	0	3 (199)	Moderate	Downgraded for imprecision
Pain	chronic	Antidepressant (TCA)	SMD, -0.1 (-0.51 to 0.31)	32.31	3 (148)	Very low	Downgraded for study limitations, inconsistency, imprecision
Pain	chronic	Opioid (tramadol)	SMD, -0.55 (-0.66 to -0.44)	85.88	5 (1378)	Low	Downgraded for study limitations, inconsistency
Pain	chronic	Opioid (buprenorphine)	SMD, -0.41 (-0.57 to -0.26)	99.49	2 (653)	Low	Downgraded for study limitations
Pain (30% reduction)	chronic	Opioid (buprenorphine)	OR, 1.49 (1.08 to 2.06)	66.95	2 (594)	Low	Downgraded for study limitations, inconsistency
Pain (50% reduction)	chronic	Opioid (buprenorphine)	OR, 1.39 (0.97 to 1.99)	NA	1 (498)	Low	Downgraded for study limitations, imprecision
Pain	chronic	Opioid (strong)	SMD, -0.43 (-0.52 to -0.33)	0	6 (1887)	Moderate	Downgraded for study limitations
Pain (30% reduction)	chronic	Opioid (strong)	OR, 1.91 (1.41 to 2.58)	38.47	3 (819)	Moderate	Downgraded for study limitations
Pain (50%	chronic	Opioid (strong)	OR, 1.89 (1.34 to	81.45	2 (750)	Very low	Downgraded for study limitations,

reduction)			2.66)				inconsistency, imprecision
Pain	chronic	Opioid (enriched)	MD, -21.34 (-22.77 to 19.91)	90.41	3 (382)	Low	Downgraded for study limitations, inconsistency
Pain	chronic	Opioid (tapentadol)	MD* -8.00 (-12.2 to -3.8)	NA	1 (637)	High	
Pain (50% reduction)	chronic	Opioid (tapentadol)	RR, 1.43 (1.07 to 1.91)	NA	1 (632)	High	
Pain	chronic	NSAIDs	MD, -6.97 (-10.74 to -3.14)	51.96	6 (1354)	Low	≤16 weeks; Downgraded for study limitations, other factors
Pain	chronic	NSAIDs (non-selective)	MD, -5.96 (-10.96 to -0.96)	55.25	4 (847)	Low	≤16 weeks; Downgraded for study limitations, inconsistency
Pain	chronic	NSAIDs (selective)	MD, -9.11 (-13.54 to -4.66)	0	2 (507)	Moderate	≤16 weeks; Downgraded for study limitations

All MD expressed on 0-100 scale.

* Converted from 0-10 to 0-100 scale by multiplying by 10.

'Summary of findings' table

Table 6

Paracetamol for low back pain in adults ≤3 months postintervention (short-term)							
Outcome	**Low back pain duration**	**Intervention and comparison**	**Relative effect (95% CI)**	**I^2**	**Number of trials (participants)**	**Certainty of the evidence (GRADE)**	**Comments**
Function	acute	Paracetamol versus placebo	MD[1], 0.05 (-0.50 to 0.60)	0	1 (1506)	High	
Safety (adverse events)	acute	Paracetamol versus placebo	RR, 1.07 (0.86 to 1.33)	0	1 (1624)	High	
Safety (serious adverse events)	acute	Paracetamol versus placebo	RR, 0.9 (0.30 to 2.67)	0	1 (1643)	High	
Health-related quality of life (physical)	acute	Paracetamol versus placebo	MD, -0.79 (-1.94 to 0.36)	0	1 (1145)	High	
Health-related quality of life (mental)	acute	Paracetamol versus placebo	MD, -0.60 (-1.38, 0.17)	0	1 (1145)	High	
Participant rating of improvement	acute	Paracetamol versus placebo	MD[2], -0.1 (-0.33 to 0.13)	0	1 (1511)	High	

[1] 0-24 scale

[2] -5 to +5 scale

'Overview of reviews' table, paracetamol

Table 7

NSAIDs for low back pain in adults ≤3 months postintervention (short-term)							
Outcome	**Low back pain duration**	**Intervention and comparison**	**Relative effect (95% CI)**	**I^2**	**Number of trials (participants)**	**Certainty of the evidence (GRADE)**	**Comments**
Pain	acute	NSAID (selective)	MD, -2.6 (-9.23 to	56.53	2 (437)	Low	Downgraded for study limitations,

		versus NSAID (non-selective)	4.03)				inconsistency
Pain	acute	NSAID versus paracetamol	SMD, -0.12 (-0.35 to 0.12)	0	2 (289)	Low	Downgraded for study limitations, imprecision
Function	acute	NSAID versus placebo	MD[1], -2.02 (-2.89 to -1.15)	0	2 (471)	High	
Function	acute	NSAID (selective) versus NSAID (non-selective)	MD[2], -7 (-13.15 to -0.85)	NA	1 (104)	Moderate	Downgraded for study limitations
Function	chronic	NSAID versus placebo	MD[1], -0.85 (-1.30 to -0.40)	45.78	4 (1161)	Low	Downgraded for study limitations, other factors
Safety (adverse events)	acute	NSAID versus placebo	RR, 0.86 (0.63 to 1.18)	0	6 (1394)	Very low	Downgraded for study limitations, inconsistency, indirectness, imprecision
Safety (adverse events)	acute	NSAID (selective) versus NSAID (non-selective)	RR, 0.97 (0.63 to 1.50)	22.13	2 (444)	Very low	Downgraded for study limitations, imprecision
Safety (adverse events)	chronic	NSAID versus placebo	RR, 1.04 (0.92 to 1.17)	19.68	6 (1354)	Low	Downgraded for study limitations, other factors
Safety (adverse events)	chronic	NSAID (non-selective) versus placebo	RR, 0.94 (0.82 to 1.08)	0	4 (847)	Low	Downgraded for study limitations, other factors
Safety (adverse events)	chronic	NSAID (selective) versus placebo	RR, 1.25 (1.00 to 1.56)	17.54	2 (507)	Moderate	Downgraded for study limitations
Safety (adverse events)	chronic	NSAID versus paracetamol	RR, 1.5 (0.15 to 14.68)	NA	1 (28)	Not reported	
Safety (adverse events)	chronic	NSAID versus tramadol	RR, 0.83 (0.75 to 0.91)	NA	1 (1583)	Not reported	
Safety (adverse events)	chronic	NSAID versus pregabalin	RR, 0.8 (0.23 to 2.74)	NA	1 (72)	Not reported	
Participant rating of improvement	acute	NSAID versus placebo	RR, 1.4 (1.12 to 1.75)	51.64	5 (1201)	Low	Downgraded for inconsistency, indirectness
Participant rating of improvement	acute	NSAID versus other drug	RR, 1.01 (0.81 to 1.25)	0	2 (162)	Moderate	Downgraded for imprecision
Participant rating of improvement	chronic	NSAID versus paracetamol	RR, 1.39 (0.82 to 2.37)	NA	1 (28)	Not reported	
Participant rating of improvement	chronic	NSAID versus tramadol	RR, 1.26 (1.16 to 1.38)	NA	1 (1583)	Not reported	
Workplace participation	acute	NSAID versus placebo	RR, 1.48 (0.98 to 2.23)	NA	1 (266)	Very low	Downgraded for study limitations, imprecision

MD on 0-100 scale unless otherwise indicated

[1] 0-24 scale

[2] 0-50 scale

'Overview of reviews' table, NSAIDs

Table 8

Muscle relaxants and benzodiazepines for low back pain in adults ≤3 months postintervention (short-term)							
Outcome	**Low back pain duration**	**Intervention and comparison**	**Relative effect (95% CI)**	**I^2**	**Number of trials (participants)**	**Certainty of the evidence (GRADE)**	**Comments**
Pain	acute	Non-benzodiazepine antispasmodic + analgesic/NSAID versus placebo + analgesic/NSAID	RR, 0.64 (0.37 to 1.09)	84.08	2 (469)	Not reported	
Function	acute	Non-benzodiazepine antispasmodic versus placebo	RR, 0.55 (0.40 to 0.70)	0	3 (251)	Moderate	Downgraded for study limitations
Safety (adverse events)	acute	Non-benzodiazepine antispasmodic versus placebo	RR, 1.5 (1.14 to 1.98)	0	8 (724)	Moderate	Downgraded for study limitations
Safety (adverse events)	chronic	Non-benzodiazepine antispasmodic versus placebo	RR, 1.02 (0.67 to 1.57)	0	2 (246)	Low	Downgraded for study limitations, imprecision
Safety (adverse events)	acute	Non-benzodiazepine antispasmodic + analgesic/NSAID versus placebo + analgesic/NSAID	RR, 1.3 (0.62 to 2.75)	83.9	3 (506)	Not reported	
Participant rating of improvement	acute	Non-benzodiazepine antispasmodic versus placebo	RR, 0.68 (0.41 to 1.13)	33.74	4 (323)	Not reported	
Participant rating of improvement	acute	Non-benzodiazepine antispasmodic + analgesic/NSAID versus placebo + analgesic/NSAID	RR, 0.37 (0.08 to 1.77)	79.84	2 (148)	Not reported	
Participant rating of improvement	chronic	Benzodiazepine versus placebo	RR, 0.63 (0.42 to 0.97)	16.75	2 (151)	Not reported	

'Overview of reviews' table, muscle relaxants and benzodiazepines

Table 9

Opioids for low back pain in adults ≤3 months postintervention (short-term)							
Outcome	**Low back pain duration**	**Intervention and comparison**	**Relative effect (95% CI)**	**I^2**	**Number of trials (participants)**	**Certainty of the evidence (GRADE)**	**Comments**
Pain	chronic	Tramadol versus celecoxib	RR, 0.82 (0.76 to 0.90)	NA	1 (1583)	Not reported	
Pain (30% reduction)	chronic	Tramadol versus celecoxib	OR, 0.63 (0.52 to 0.77)	NA	1 (1583)	Very low	Downgraded for study limitations, indirectness, imprecision
Pain	chronic	Opioids versus antidepressants	SMD, 0.21 (-0.03 to 0.45)	0	2 (272)	Very low	Downgraded for study limitations, imprecision
Pain	chronic	Tapentadol versus oxycodone	MD[1], 0 (-0.4 to 0.4)	NA	1 (not reported)	Not reported	
Pain (50% reduction)	chronic	Tapentadol versus oxycodone	RR, 1.16 (0.89 to 1.51)	NA	1 (641)	Not reported	
Function	chronic	Tramadol versus placebo	SMD, -0.18	0	5 (1348)	Moderate	Downgraded for study limitations

			(-0.29 to -0.07)				
Function	chronic	Buprenorphine versus placebo	SMD, -0.14 (-0.53 to 0.25)	NA	1 (101)	Very low	Downgraded for study limitations, imprecision
Function	chronic	Opioids (strong) versus placebo	SMD, -0.26 (-0.37 to -0.15)	0	4 (1375)	Moderate	Downgraded for study limitations
Function	chronic	Opioids versus antidepressants	SMD, -0.11 (-0.63 to 0.42)	NA	1 (56)	Very low	Downgraded for study limitations, imprecision
Safety (adverse events)	chronic	Tapentadol versus placebo	RR, 1.27 (1.14 to 1.41)	NA	1 (637)	High	
Safety (serious adverse events)	chronic	Tapentadol versus placebo	RR, 2.34 (0.61 to 8.97)	NA	1 (637)	Moderate	Downgraded for imprecision
Safety (nausea)	chronic	Opioids (all types) versus placebo	RD, 0.10 (0.07 to 0.14)	62.6	10 (3747)	Low	Downgraded for study limitations and inconsistency
Safety (headaches)	chronic	Opioids (all types) versus placebo	RD, 0.03 (0.01 to 0.05)	32.22	10 (3747)	Low	Downgraded for study limitations and inconsistency
Safety (constipation)	chronic	Opioids (all types) versus placebo	RD, 0.07 (0.04 to 0.11)	77.74	9 (3493)	Low	Downgraded for study limitations and inconsistency
Safety (dizziness)	chronic	Opioids (all types) versus placebo	RD, 0.08 (0.05 to 0.11)	67.81	9 (3494)	Low	Downgraded for study limitations and inconsistency
Safety (somnolence)	chronic	Opioids (all types) versus placebo	RD, 0.06 (0.03 to 0.09)	65.78	8 (3257)	Low	Downgraded for study limitations and inconsistency
Safety (adverse events)	chronic	Tapentadol versus oxycodone	RR, 0.89 (0.82 to 0.96)	NA	1 (646)	Not reported	
Safety (serious adverse events)	chronic	Tapentadol versus oxycodone	RR, 0.66 (0.26 to 1.67)	NA	1 (646)	Not reported	
Safety (withdrawal due to adverse events)	chronic	Tapentadol versus placebo	RR 3.41 (1.96 to 5.94)	NA	1 (637)	High	
Safety (withdrawal due to adverse events)	chronic	Tapentadol versus oxycodone	RR 0.49 (0.37 to 0.66)	NA	1 (646)	Not reported	

[1] MD 0-10 scale

'Overview of reviews' table, opioids

Table 10

Antidepressants for low back pain in adults ≤3 months postintervention (short-term)							
Outcome	**Low back pain duration**	**Intervention and comparison**	**Relative effect (95% CI)**	**I^2**	**Number of trials (participants)**	**Certainty of the evidence (GRADE)**	**Comments**
Function	chronic	Antidepressant versus placebo	SMD, -0.06 (-0.40 to 0.29)	0	2 (132)	Low	Downgraded for study limitations, imprecision

'Overview of reviews' table, antidepressants

Table 11

			Small effect		**Harmful**

	Large effect	Medium effect		No evidence of difference	
High certainty evidence			*Opioid (tapentadol)*	**Paracetamol**	
Moderate certainty evidence			**NSAIDs, muscle relaxants** *Opioids (strong), NSAIDs (selective),*	*Antidepressants (SSRIs)*	
Low certainty evidence		*Opioid (tramadol)*	*NSAIDs, NSAIDs (non-selective), Opioid (buprenorphine)*	*Antidepressants,*	
Very low certainty evidence				*Antidepressants (TCAs),*	

The size of the effect for the mean between group difference are based on the definitions from the American College of Physicians and the American Pain Society (Chou 2017). Large effect, >20 points on a 0-100 scale or >0.8 SMD; Medium effect, >10-20 points on a 0-100 scale or >0.5 to 0.8 SMD; Small effect, 5-10 points on a 0-100 scale or 0.2 to 0.4 SMD; No evidence of difference, boundaries of the 95% confidence interval span both sides of the line of no effect; Harmful, boundaries of the 95% confidence interval fall completely within harm. **Acute LBP = bold,** *Chronic LBP = italics*

'Summary of results', pain outcome table

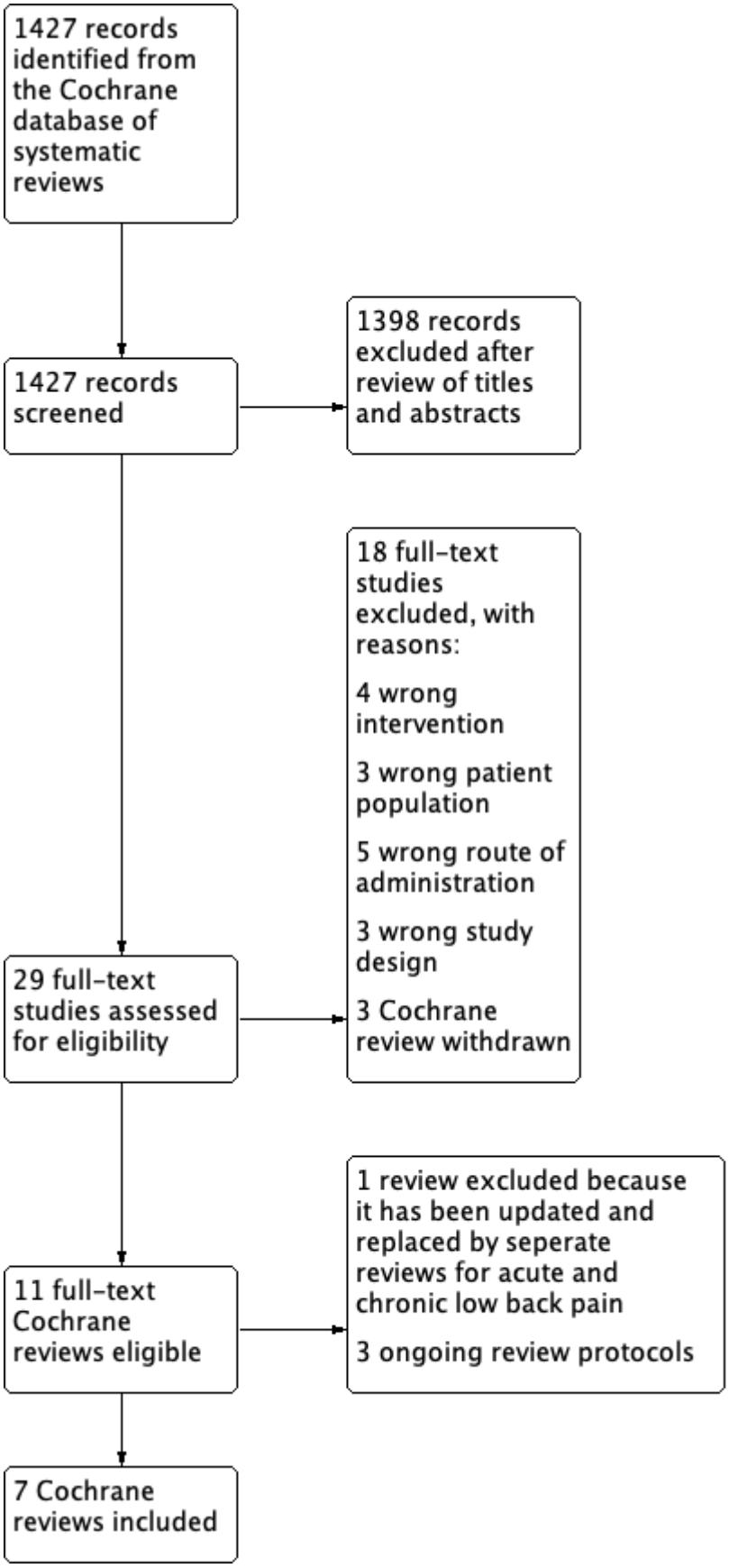
Table 12

	Large effect	Medium effect	Small effect	No evidence of difference	Harmful
High certainty evidence			**NSAIDs**	**Paracetamol**	
Moderate certainty evidence			**Muscle relaxants** *Opioid (tramadol), Opioid (strong)*		
Low certainty evidence			*NSAIDs*	*Antidepressant*	
Very low certainty evidence				*Opioid (buprenorphine)*	

The size of the effect for the mean between group difference are based on the definitions from the American College of Physicians and the American Pain Society (Chou 2017). Large effect, >20 points on a 0-100 scale or >0.8 SMD; Medium effect, >10-20 points on a 0-100 scale or >0.5 to 0.8 SMD; Small effect, 5-10 points on a 0-100 scale or 0.2 to 0.4 SMD; No evidence of difference, boundaries of the 95% confidence interval span both sides of the line of no effect; Harmful, boundaries of the 95% confidence interval fall completely within harm. **Acute LBP = bold,** *Chronic LBP = italics*

'Summary of results', function outcome table

Figure 1



Flow diagram