2018

Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: A network meta-analysis (Protocol)

Matthew K. Bagg

Andrew J. McLachlan

Christopher G. Maher

Steven J. Kamper

Christopher M. Williams

See next page for additional authors

This article was originally published as:
Bagg, M. K., McLachlan, A. J., Maher, C. G., Kamper, S. J., Williams, C. M., Henschke, N., Wand, B. M., Moseley, G. L., Hubscher, M., O'Connell, N. E., van Tulder, M. W., Nikolakopoulou, A., & McAuley, J. H. (2018). Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: A network meta-analysis (Protocol). *Cochrane Database of Systematic Reviews, 6*.

Original article available here:
https://doi.org/10.1002/14651858.CD013045

This article first published in the Cochrane Database of Systematic Reviews, June 2018. Available: https://doi.org/10.1002/14651858.CD013045

Bagg, M.K., McLachlan, A.J., Maher, C.G., Kamper, S.J., Williams, C.M., Henschke, N., Wand, B.M., Moseley, G.L., Hubscher, M., O'Connell, N.E., van Tulder, M.W., Nikolakopoulou, A., and McAuley, J.H. (2018) Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: A network meta-analysis (Protocol), *Cochrane Database of Systematic Reviews, 6.* doi: 10.1002/14651858.CD013045

Cochrane Library

Cochrane Database of Systematic Reviews

Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis (Protocol)

Bagg MK, McLachlan AJ, Maher CG, Kamper SJ, Williams CM, Henschke N, Wand BM, Moseley GL, Hübscher M, O'Connell NE, van Tulder MW, Nikolakopoulou A, McAuley JH

Bagg MK, McLachlan AJ, Maher CG, Kamper SJ, Williams CM, Henschke N, Wand BM, Moseley GL, Hübscher M, O'Connell NE, van Tulder MW, Nikolakopoulou A, McAuley JH.

Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis.

Cochrane Database of Systematic Reviews 2018, Issue 6. Art. No.: CD013045.

DOI: 10.1002/14651858.CD013045.

www.cochranelibrary.com

TABLE OF CONTENTS

[Intervention Protocol]

Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis

Matthew K Bagg[1,2,3], Andrew J McLachlan[4,5], Christopher G Maher[6], Steven J Kamper[6], Christopher M Williams[7,8], Nicholas Henschke[9], Benedict M Wand[10], G L Moseley[1,11], Markus Hübscher[1], Neil E O'Connell[12], Maurits W van Tulder[13], Adriani Nikolakopoulou[14], James H McAuley[1,15]

[1]Neuroscience Research Australia, Sydney, Australia. [2]Prince of Wales Clinical School, University of New South Wales, Sydney, Australia. [3]New College Village, University of New South Wales, Sydney, Australia. [4]Faculty of Pharmacy, University of Sydney, Sydney, Australia. [5]Centre for Education and Research on Ageing, Concord Repatriation General Hospital, Sydney, Australia. [6]Sydney School of Public Health, University of Sydney, Sydney, Australia. [7]School of Medicine and Public Health, University of Newcastle, Callaghan, Australia. [8]Hunter Medical Research Institute, New Lambton, Australia. [9]Sydney Medical School, University of Sydney, Sydney, Australia. [10]School of Physiotherapy, The University of Notre Dame Australia Fremantle, Perth, Australia. [11]Sansom Institute for Health Research, University of South Australia, Adelaide, Australia. [12]Health Economics Research Group, Institute of Environment, Health and Societies, Department of Clinical Sciences, Brunel University London, Uxbridge, UK. [13]Department of Health Sciences, Faculty of Science and Amsterdam Movement Science institute, Vrije Universiteit, Amsterdam, Netherlands. [14]Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland. [15]School of Medical Sciences, University of New South Wales, Sydney, Australia

Contact address: James H McAuley, Neuroscience Research Australia, Sydney, NSW, Australia. j.mcAuley@neura.edu.au.

Editorial group: Cochrane Back and Neck Group.
Publication status and date: New, published in Issue 6, 2018.

Citation: Bagg MK, McLachlan AJ, Maher CG, Kamper SJ, Williams CM, Henschke N, Wand BM, Moseley GL, Hübscher M, O'Connell NE, van Tulder MW, Nikolakopoulou A, McAuley JH. Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis. *Cochrane Database of Systematic Reviews* 2018, Issue 6. Art. No.: CD013045. DOI: 10.1002/14651858.CD013045.

ABSTRACT

This is a protocol for a Cochrane Review (Intervention). The objectives are as follows:

To answer the clinical question: 'what analgesic medicine shall I prescribe this patient with chronic low back pain to reduce their pain? '.

The objectives are 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.

BACKGROUND

Description of the condition

Low back pain (LBP) presents an enormous global health problem. It is the leading cause of disability (GBD Collaborators 2017a),

and fourth-leading cause of disease burden amongst non-communicable health conditions (GBD Collaborators 2017b). In the preceding 20 years, the amount of disability caused by LBP has remained stable; however the disease burden has increased. Recent-onset (acute) LBP has a favourable initial prognosis (Menezes Costa 2012). However, only one-third of people are recovered 12 weeks from pain onset and this changes little at 12 months from pain onset (Itz 2013). LBP persisting longer than 12 weeks is termed chronic low back pain (Furlan 2015; Treede 2015).

Technically, pain in the low back area is a symptom. Nonetheless, the usual use of the term LBP is to describe the condition of non-specific LBP (NSLBP). NSLBP is the most common form of LBP, commonly stated to be 90% of all cases (Koes 2006), and is so termed because there is no identifiable problem (in the back) that is causing the LBP symptoms (Maher 2017). Much less common conditions for which there is an identifiable problem include fracture, infection, malignancy (tumour), radiculopathy (sciatica involving problematic nerve function), spinal stenosis (narrowing of spinal openings) or inflammatory arthropathy (inflammatory joint problems) (Downie 2013; Enthoven 2016a; Henschke 2008; Koes 2006). The focus of this review is chronic non-specific LBP (CLBP).

Description of the intervention

Medicines prescribed to reduce pain intensity (analgesic medicines) are the most common form of treatment for CLBP (Gore 2012; Hart 2015; Ivanova 2011). Analgesic medicines work in various ways to reduce the intensity of pain but may also cause unwanted harmful effects. Clinical decisions to recommend particular pain-relieving medicines involve a balance between the anticipated reduction in pain and the possibility of side effects. This is particularly important when people have other relevant health conditions such as the inability to tolerate some medicines, or problems with kidney or liver function. Clinical use of these medicines varies across healthcare settings and geographical areas and is usually different to guideline recommendations (Amorin-Woods 2014; Bishop 2003; Cifuentes 2010; Gore 2012; Gouveia 2017; Hart 2015; Hoffmann 2013; Ivanova 2011; Mafi 2013; Ndlovu 2014; Webster 2005; Webster 2007). In this review, we will focus on the three most commonly prescribed classes of analgesic medicines: paracetamol (acetaminophen); non-steroidal anti-inflammatory drugs (NSAIDs); and opioid analgesics.

How the intervention might work

The human body creates its own natural opioid molecules (endogenous opioids), for a variety of functions. The nervous system contains receptors for these opioids on the exterior of cells (Kieffer 2009). When opioid molecules bind to receptors, this causes changes within the cells. Opioid analgesics, as they are termed in this review, are medicines made synthetically that match or are chemically similar to endogenous opioids. When taken, opioid analgesics attach to the opioid receptors on nervous system cells, causing a reduction in perceived pain (Kieffer 2009; Rachinger-Adam 2011). Opioids have other effects as well, depending on the location and type of the receptor. One common side effect from taking opioids is constipation, because the medicine also binds to receptors in the digestive system and interferes with digestion (Rachinger-Adam 2011); other side effects are that people taking opioids may develop tolerance (they require larger doses for the same effect) or dependence on the medicine (they experience withdrawal symptoms if a dose is missed), or both (Deyo 2015).

NSAIDs and paracetamol interfere with the natural inflammatory processes in the body, in particular the production of chemicals called prostaglandins (Brune 2015). Prostaglandins are themselves involved in several processes. One of these is pain perception, which they contribute to by increasing inflammation. They also contribute to blood-clotting processes and support the lining of the stomach (Brune 2015). An enzyme called cyclooxygenase (COX) helps to create prostaglandins. NSAIDs lessen the effect of COX type-1 (COX-1) and COX type-2 (COX-2), which reduces prostaglandins and, in turn, reduces the perception of pain. The unavoidable reduction in prostaglandins elsewhere (e.g. the stomach) is associated with side effects (e.g. stomach ulcers) (Brune 2015). Some NSAIDs selectively inhibit COX-2, which is less present in the stomach: these 'selective' NSAIDs are associated with diminished digestive system side effects (Davies 2000; Sostres 2013; Whittle 2000). Both selective and non-selective NSAIDs have known cardiovascular risks (Baigent 2013; McGettigan 2011).

Paracetamol also inhibits the COX-1 and COX-2 enzymes (Graham 2013; Jozwiak-Bebenista 2014), particularly COX-2 (Graham 2013; Hinz 2008; Hinz 2012). Paracetamol is generally a safe medicine to take. However, overdosing (exceeding the recommended daily dose) is associated with an increased risk of liver problems and death (Daly 2008; Graham 2013; Sheen 2002).

Why it is important to do this review

There is uncertainty about the comparative effectiveness of most analgesic medicines for CLBP. This is problematic given their wide use (Gore 2012; Hart 2015; Ivanova 2011), side effects (Baigent 2013; Deyo 2015; Humphreys 2017; Martell 2007; McGettigan 2011; Shaheed 2016), and the variety of available formulations. Evidence of comparative effectiveness - the analgesic effect and safety of a medicine in comparison to other medicines that may be prescribed for the same condition - is the information required most by clinicians. This information is presently insufficiently described in the literature. Systematic reviews of these medicines typically include limited data on comparative effectiveness (Chaparro 2014; Chou 2017; Enthoven 2016b; Saragiotto 2016b), and

focus on effects compared to placebo (Chaparro 2014; Chou 2007; Chung 2013; Enthoven 2016b; Furlan 2006; Furlan 2011; Kuijpers 2011; Machado 2017; Shaheed 2016; White 2011a), or aspects of clinical trial design (Furlan 2006; Furlan 2011). While many factors influence prescribing decisions (Bhamb 2006; Cabana 1999; Darlow 2012; Perrot 2008; Schers 2001), the lack of evidence of comparative effectiveness for many analgesics means that prescribing is less evidence-informed and more reliant on other factors such as clinical experience and patient preferences.

Network meta-analysis (NMA) can provide the information on comparative effectiveness that is needed to make informed clinical decisions. NMA is a simultaneous comparison of all competing treatments (Salanti 2012). The results of NMA are expressed as the effect of a treatment, on a particular outcome, compared to *every other* competing treatment. This review will provide a comprehensive evaluation of the comparative effectiveness of analgesic medicines for CLBP and will express these data in formats that are accessible to clinicians.

Description of network meta-analysis

Network meta-analysis (NMA) is an extension of pairwise meta-analysis that compares more than two treatments (Lu 2004). NMA uses networks to represent the evidence base. Each competing treatment is represented as a node (vertex) and each comparison between treatments is represented by lines (edges) connecting the treatments (Jansen 2011). As with pairwise meta-analysis, it is important that an NMA is conducted within a rigorous systematic review process to ensure that all relevant information is included. Information that is missing from the network, for example clinical trials of a particular treatment comparison, may lead to bias in the results (Mills 2013). A network with no missing information will likely still contain a mix of solid and blank (or absent) edges. Solid edges are termed 'direct' comparisons because these have been compared in clinical trials. Absent edges are termed 'indirect' comparisons - there have been no clinical trials of these comparisons (Bucher 1997; Jansen 2011).

NMA uses the 'direct' clinical trial data to 'indirectly' estimate the effects for the missing comparisons (blank edges) (Caldwell 2005; Cipriani 2009; Jansen 2011; Jansen 2013; Mills 2013; Song 2003; Sutton 2008). In this way, NMA can be said to fill in the gaps in the evidence base. This process involves fitting a single statistical model to the network, which simultaneously combines all the information from the clinical trials (the direct comparisons) together with the information about the network structure (in particular, the missing comparisons) (Efthimiou 2016; Krahn 2013; Salanti 2012).

The transitivity assumption underpins the validity of NMA. The assumption holds when the characteristics of the participants and trials included in the network are sufficiently similar across all of the direct comparisons in the network (Jansen 2013). From a researcher's perspective, transitivity means that if all treatments in the network were included in a single mega-trial, there is an equal likelihood that any one participant may be randomised to any of those treatments (Salanti 2012). From a participant's perspective, transitivity means that if they were a participant in any one of the trials in the network, they would have an equal chance of being a participant in any of the other trials in the network.

The results from NMA may be portrayed in formats that are useful for clinical decision-making. The results for each treatment are displayed in a league table, which may display results for two outcomes (e.g. analgesic effect and safety). This is helpful for clinicians, because they can see the evidence for a treatment's benefits alongside its risks. The league table displays all of the results, which means that clinicians have all of the required information in one place. These results can also be used to produce rankings of the effect of treatments on a particular outcome (Dominici 1999; Hoaglin 2011; Jansen 2011; Sutton 2008; van der Valk 2009).

OBJECTIVES

To answer the clinical question: 'what analgesic medicine shall I prescribe this patient with chronic low back pain to reduce their pain?'.

The objectives are 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.

METHODS

Criteria for considering studies for this review

Types of studies

We will include conventional and enriched-design parallel group RCTs and the first phase of cross-over RCTs. We will exclude cluster RCTs. Studies must compare at least two of the interventions of interest.

Types of participants

We will include studies of people with chronic, non-specific LBP defined as a primary area of pain between the twelfth rib and gluteal fold, with or without associated leg pain, for which the same episode has persisted for longer than three months (Furlan 2015; Treede 2015). We will exclude studies where subjects have leg pain that meets the definition of sciatica used by Koes 2007 or where low back pain is caused by pathological entities such as infection, neoplasm, metastasis, osteoporosis, inflammatory disease

or fractures. We will include studies that include participants with conditions other than CLBP (e.g. sciatica or ankylosing spondylitis) if separate data is reported for the CLBP participants or may be obtained from study authors.

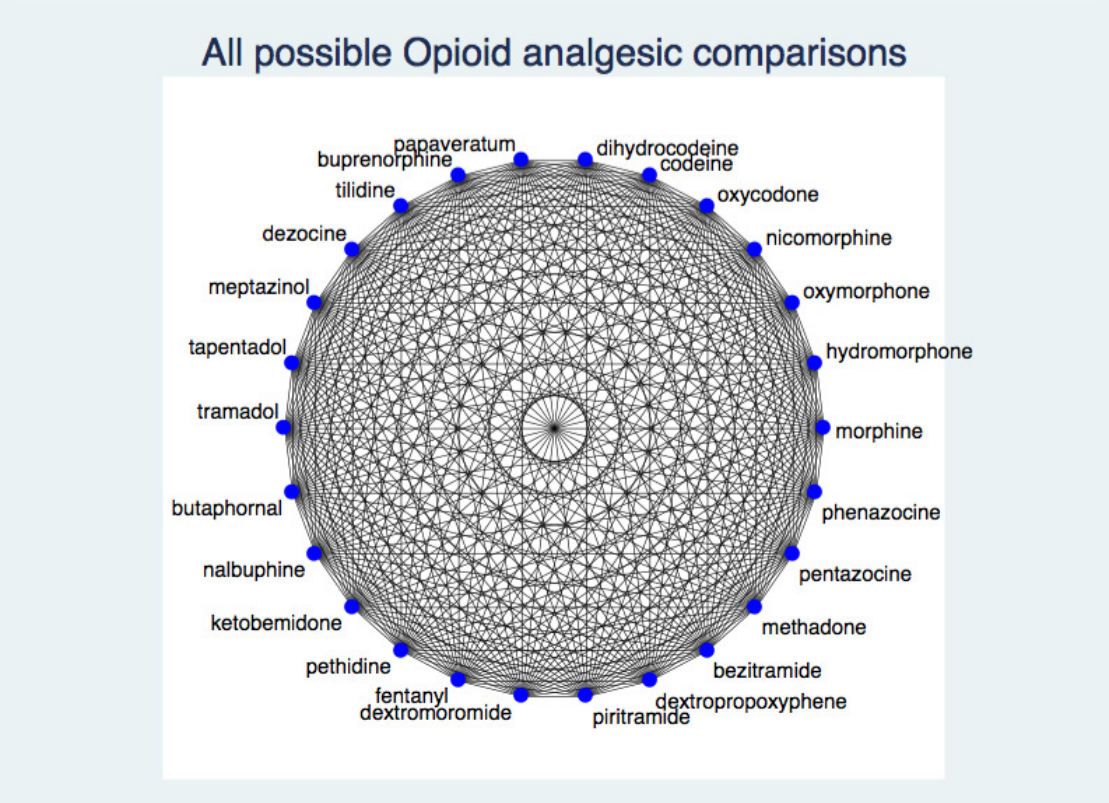
Types of interventions

We will include studies of analgesic medicines prescribed for the intended purpose of reducing pain intensity for someone with CLBP, defined as paracetamol (acetaminophen), NSAIDs and opioid analgesics. The interventions of interest are listed in Appendix 1 and include placebo and no treatment. Each medicine is a separate intervention (node) in the analysis. We will not combine different medicines into the same node. The NMA will compare every medicine to every other medicine. If we identify NSAIDs or opioid analgesics that are inadvertently not listed in Appendix 1, we will consider them as eligible and we will include them in the network after assessing their comparability with the pre-specified set of competing interventions. We will report the findings for these interventions in the results and the conclusions of the review.

We will include analgesic medicines delivered as mono or combination therapy (delivery of one or more agents at once for the same intended therapeutic effect, in single or multiple formulations) via all systemic routes of administration. We will exclude local therapeutic injections or topical applications and medicines or dosages that are not currently licensed for human use. We will include trials that contain non-pharmacological co-interventions (intended by the trial investigators) in one or more of the intervention arms. Such co-interventions will be categorised as i) physical therapies, or ii) psychological therapies, or iii) combinations of i) and ii), for the evaluation of transitivity.

We assume that all analgesic medicines, defined herein, are directly comparable treatments. In other words, we assume that the distribution of important characteristics (effect modifiers) is the same across all treatment comparisons (Salanti 2012). The network diagram in Figure 1 displays all possible pairwise comparisons between opioid analgesics. We have not included NSAIDs or paracetamol in this diagram for the sake of visual simplicity. All medicines are categorised according to the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) system. The ATC codes are listed in Appendix 1.

Figure 1. All theoretically possible comparisons between the opioid analgesic interventions of interest. NSAIDs and paracetamol are not included for the sake of simplicity.



Network nodes will be further defined using the relevant licensed dosing range (LDR) from the US Food and Drug Administration (FDA), the Australian Therapeutic Goods Administration (TGA) or the European Medicines Agency (EMA). For most medicines, this is equivalent to the standard dosing range (SDR). For older medicines, the SDR may be smaller than the LDR. The network will include these doses as separate nodes. Thus, for drug A there may be a single LDR node, whereas for drug B there may be a 'less than SDR' node, an SDR node and a 'greater than SDR' node. We will use the standard dosing range from the Australian Medicines Handbook (AMH) and the Prescriber's Digital Reference or MIMS when the medicine is not listed in the AMH (AMH 2017; MIMS 2017; PDR 2017). The placebo node is defined as any drug intervention that does not contain an active ingredient. The 'no treatment' node is defined as any trial arm that contains no investigator-intended treatment, and includes trial arms that comprise continuation of usual care or being placed on a wait-list.

Types of outcome measures

Primary outcomes

The co-primary outcomes are pain and safety.

- Pain is defined as pain intensity, measured at the time point closest to the end of treatment. Pain intensity may be measured with a continuous self-report scale (e.g. visual analogue scale (VAS) or numeric rating scale (NRS)), a rating scale within a composite measure of pain (e.g. McGill Pain Questionnaire), or an ordinal scale with greater than six levels (we will consider such ordinal scales to exhibit continuous properties). We will not exclude studies that use other measurement tools. We will estimate the relative ranking of the competing interventions according to their effect on pain intensity at the end of treatment.

- Safety is defined as the proportion of participants who experience a serious adverse effect during the treatment period. Adverse effects are described broadly as any of 'adverse event', 'adverse drug reaction', 'side effect', 'toxic effect' or 'complication' that are associated with the medicine under investigation. A serious adverse effect is that which causes a reduction in dose or cessation of treatment. No change or an increase in pain is not considered an adverse effect. We will estimate the relative ranking of the competing interventions according to their effect on safety.

We note that randomised controlled trials have limitations in the evaluation of medicine safety, particularly enriched designs (Furlan 2011). The study size, duration of treatment, length of follow-up and inclusion criteria are usually not appropriate to detect rare events or evaluate long-term adverse effects (Brewer 1999; Sills 1986; Sutton 2002). Data on adverse effects may be available from multiple sources (Sutton 2002), including observational cohort designs, which are not included in this review. The Cochrane Back and Neck Group is currently conducting a review of harms of opioids for chronic non-cancer pain (Furlan 2014), and other reviews of harms of NSAIDs have been published (for example, Baigent 2013).

Secondary outcomes

- Pain, defined as pain intensity, measured at 1-month post randomisation, provided the individual study treatment is complete. The outcome is otherwise defined as above.

- Functional ability, defined as low back specific function, measured at end of treatment. Functional ability may be measured with a continuous, self-report scale (e.g. Roland Morris Disability Questionnaire (RMDQ) or Oswestry Disability Index (ODI)), a rating scale within a composite measure (e.g. SF-36), or an ordinal scale with greater than six levels (we will consider such ordinal scales to exhibit continuous properties). We will not exclude studies that use other measurement tools. We will estimate the relative ranking of the competing interventions according to their effect on functional ability at the end of treatment.

Search methods for identification of studies

Electronic searches

We will search for all possible comparisons formed by the interventions of interest. We will search the following electronic databases from inception to current.

- Cochrane Back and Neck Group's Trials Register (through CENTRAL).
- Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library, current issue.
- MEDLINE (Ovid) (1946 to current).
- Embase (Ovid) (1980 to current).
- CINAHL (EBSCO) (1982 to current).
- ClinicalTrials.gov (ClinicalTrials.gov/ct2/home).
- WHO International Clinical Trial Registry Platform (WHO ICTRP) (apps.who.int/trialsearch/Default.aspx).

We have developed the search strategies specifically for this review, using the recommended back pain search terms (Furlan 2015), and specific terms for the interventions of interest and their combinations. The search strategy for MEDLINE is listed in Appendix 2.

Searching other resources

We will check the reference lists of eligible studies and previous systematic reviews to identify additional studies. We will contact the corresponding authors of the most recent systematic reviews on opioid analgesics, NSAIDs and paracetamol to enquire if they are aware of any additional studies in the area of their review. We will search for the protocols and registrations of all included studies.

Data collection and analysis

Selection of studies

Two reviewers from a panel of six (MKB, SK, BMW, AM, NH, CW) will independently screen identified studies for eligibility in two independent stages: i) title and abstract; ii) full text. Review authors will not be involved in the decision to include any study with which they have been involved. The reviewers will resolve disagreements through discussion, with arbitration by a third reviewer (JHM) if required. We will contact study authors up to three times to obtain additional information to determine study eligibility. We will place the study in 'Studies awaiting assessment' for this iteration of the review if no reply is received within six weeks. We will obtain translations of articles written in languages that cannot be read by the authors through colleagues or Cochrane resources such as Task Exchange. We will summarise the search process using a PRISMA flow diagram (Liberati 2009).

Data extraction and management

Two reviewers from a panel of six (MKB, BMW, MH, NOC, NH, CW) will independently extract data from included studies. Review authors will not extract data from any study in which they have had any involvement. We will use standardised, piloted, data extraction forms. Required data will be taken from previous Cochrane or non-Cochrane Reviews (conducted by the authors) when possible. We will scan these data for implausible values; and if found we will refer to the original article. The colleagues providing the translation of studies written in other languages will also extract data from these studies. The reviewers will resolve disagreements through discussion, with arbitration by a third reviewer (JHM) if required.

We will contact study authors up to three times to request additional data. We will consider the data unobtainable for this iteration of the review if no reply is received within six weeks. We will only extract data for the interventions of interest from multi-arm studies of many interventions. We will preference published over unpublished data in situations where data sources conflict, because published data have been through peer review.

Study methods

We will extract data on the trial design. This will include the number, country and setting of trial sites, sample size and total duration of the trial.

Participants

We will extract data on the individual study sample. This will include the diagnosis, duration of LBP episode, age, male/female ratio and co-morbidities, including alternate sites of pain. We will extract arm-level pain intensity at baseline (mean (SD)).

Interventions

We will extract data on the interventions. This will include the duration of the intervention, including duration of any washout, run-in or titration period; dosage regimen, including titration; and routes of administration and usage of rescue medication.

Outcomes

We will extract the type and dimensions of the scale or measure used to assess pain or functional ability and the time from randomisation at which the end of treatment data were obtained in the individual trials. We will extract the definition of 'adverse effect' and 'serious adverse effect' used in each study.

Results

We will extract the number of participants allocated to each intervention group and the proportion of adherence to the intervention, including the definition of adherence. We will extract the proportion of participants in each group who discontinued treatment due to an adverse effect. We will extract from each trial the event rates and descriptions of all reported adverse effects and serious adverse effects. We will extract the outcome score (preferred) for pain and functional ability or the change in outcome from baseline and the accompanying measure of variance (or available statistic to estimate these values) for each group at the time point closest to the end of treatment. We will also extract these data for pain at the time point closest to 1-month post-randomisation, provided study treatment is complete. We will extract the between-group differences in scores and the accompanying measure of variance if group-level data are not available. We will select and extract data from a single outcome measure, in studies with more than one relevant outcome measure for pain, in the following order: 100 mm VAS; 10 cm VAS; 11-point NRS; rating scale for pain intensity from a composite measure of pain; ordinal scale with more than six levels. We will select and extract data from a single outcome measure, in studies with more than one relevant outcome measure for functional ability, in the following order: ODI; RMDQ; rating scale for functional ability from a composite measure; ordinal scale with more than six levels.

Assessment of risk of bias in included studies

Two reviewers from a panel of six (MKB, CW, MH, NH, SK, NOC) will independently appraise outcome-level risk of bias for the domains of selection, performance, attrition, detection, reporting and other sources of bias, using the Cochrane 'Risk of bias' tool, version 5.1.0 (Higgins 2011), and recommendations of Furlan 2015 (Table 1; Table 2). Review authors will not appraise risk of bias for any study in which they have had any involvement. We will pilot-test the 'Risk of bias' assessment procedure on a small number of articles. 'Risk of bias' assessments will be taken from previous Cochrane Reviews or from previous reviews conducted by the authors when possible. We will add any missing outcome-level assessments.

We will determine single outcome-level risk of bias ratings for each study, using an adaptation of the criteria in Furukawa 2016 for the GRADE evaluation. Outcomes are at low overall risk when three or fewer domains are rated 'unclear' risk and no domains are rated 'high'; moderate overall risk if a single domain is rated as 'high' risk of bias, or no domain is rated as 'high' risk but four or more are rated as 'unclear'; and high overall risk in all other instances. The reviewers will resolve disagreements through discussion, with arbitration by a third reviewer (JHM) if required.

Measures of treatment effect

Relative treatment effects

We will convert all outcome data for pain and functional ability to common 0 to 100 point scales (mean (SD)). We will estimate the relative treatment effects of the competing interventions on pain and functional ability using weighted mean differences (WMD), with 95% confidence intervals (95% CIs); and on safety using risk ratios (RRs), with 95% CIs. We will present the results from NMA as WMDs (pain and functional ability) and RRs (safety), with 95% CIs, for each intervention compared to the reference (placebo).

We will interpret the meaningfulness of the effect of medicines on pain using the threshold of 10 points on a 0 to 100 point scale (Chou 2017). We will use the 95% CIs of the risk ratio for safety to judge meaningfulness on this outcome.

Relative treatment ranking

We will obtain hierarchies (rankings) of the effect of all interventions on pain and on safety using two measures: the probability to be at each possible rank; and the surface under the cumulative ranking curve (SUCRA) (Salanti 2011). We will present these hierarchies using rankograms, cumulative probability plots and clustered ranking graphs (Chaimani 2013; Salanti 2011).

Unit of analysis issues

Studies with multiple treatment groups (multi-arm trials)

We will fit the network meta-analysis models using multivariate meta-analysis in Stata (Higgins 2012; StataCorp 2015; White 2009; White 2011b; White 2012; White 2015), which accounts for the presence of multi-arm trials. We will combine intervention arms from the same trial that are within the same licensed dosing range, using the formulae in Table 7.7a of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Dealing with missing data

We will contact study authors to provide missing data. We will impute missing data that are required for meta-analysis (e.g. standard deviations), if they are not obtainable for this iteration of the review, using previously published methods (Wan 2014), and those described in Sections 7.7 and 16.1 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Assessment of heterogeneity

Assessment of heterogeneity within treatment comparisons

We will use clinical and methodological judgement to assess the presence of clinical and methodological heterogeneity within each pairwise comparison by comparing the trial and study population characteristics across all eligible trials. A judgement of excessive clinical or methodological heterogeneity within a particular pairwise comparison will prevent meta-analysis.

Assessment of transitivity across the network

Transitivity is the fundamental assumption underpinning the validity of network meta-analysis. It is a statistically untestable assumption, relying on a subjective assessment of clinical and methodological heterogeneity across the network. Transitivity assumes that there is no imbalance in the distribution of treatment effect modifiers across each of the comparisons in the network (Jansen 2013). Any imbalance threatens the transitivity assumption. We will use several methods to evaluate the transitivity assumption. We consider the following clinical and methodological factors to be potential treatment effect modifiers, in the absence of robust evidence to date for effect modification in LBP trials (Saragiotto 2016a).

- Baseline pain intensity
- Intended co-interventions, categorised as i) physical therapies ii) psychological therapies, iii) combinations of i) and ii)
- Sample size (Dechartres 2016)
- Enriched design

We will visually assess the distributions of these effect modifiers across all treatment comparisons in the network, using weighted network plots in Stata (Chaimani 2013; Salanti 2012). We anticipate that insufficient reporting of effect modifiers in individual studies and pairwise comparisons containing few studies will make assessment of the distribution of effect modifiers difficult (Cipriani 2013). We will interpret a clearly dissimilar distribution of an effect modifier across trials as a threat to transitivity. We will proceed with NMA in the case of minor dissimilarities; and explore the influence of the effect modifier on inconsistency/heterogeneity using network meta-regression or subgroup analyses (or both). We will consider excluding network nodes in the case of considerable dissimilarity and will consider not proceeding with NMA if intransitivity persists.

An equivalent expression of transitivity is that treatments missing from each trial in the network are missing at random (Lu 2006), which implies that there is no preference for a particular (set of) comparison treatment/s (Salanti 2012). We will calculate network geometry metrics for diversity and co-occurrence, using EcoSim Professional v1.2d (Entsminger 2014), to assess the presence of comparator preference bias. We will interpret a probability of interspecific encounter (PIE) index less than 0.75 and a statistically significant C-score ($P < 0.10$) as indicative of limited diversity and likely co-occurrence, respectively (Hurlbert 1971; Salanti 2008). We will construct network plots and weight the edges by the number of checkerboard units for each comparison if limited diversity or likely co-occurrence are identified. We will manually scrutinise comparisons with a large number of checkerboard units to determine whether this co-occurrence is justifiable (Salanti 2008). We will exclude network nodes (treatments) forming comparisons for which there is non-justifiable co-occurrence.

Assessment of reporting biases

We will assess small study effects in pairwise comparisons, using contour enhanced funnel plots (Peters 2008), when there are at least 10 studies available. Various factors contribute to the association between study effect size and funnel plot asymmetry. Contour-enhanced funnel plots assist interpretation of asymmetry that is due to publication bias and not other factors, such as lesser methodological quality. The contour lines imposed on the plot indicate levels of statistical significance. We will interpret an absence of studies in areas of non-significance as suggestive of publication bias for that pairwise comparison.

Data synthesis

Methods for direct treatment comparisons

We will perform pairwise random-effects meta-analyses in Stata or Review Manager 5 for each direct comparison for which there are at least two studies available (StataCorp 2015; Review Manager 2014).

Methods for indirect and mixed comparisons

We aim to perform network meta-analysis within a frequentist framework in Stata (Higgins 2012; White 2009; White 2011b; White 2012; White 2015). The multivariate meta-analysis models incorporate random effects for heterogeneity and for inconsistency. We will not stratify study inclusion to the analysis by risk of bias.

Assessment of statistical heterogeneity

Assumptions when estimating the heterogeneity

We will assume that the heterogeneity variance is different for each direct comparison in standard pairwise meta-analyses. We will assume that the heterogeneity variance is the same across the different comparisons in network meta-analysis.

Measures and tests for heterogeneity

We will test for the presence of statistical heterogeneity (variance in true effects) within each pairwise comparison using the Q statistic, with alpha less than 0.10 as we anticipate a small number of trials per comparison. We will calculate 95% prediction intervals for the pooled effects and interpret prediction intervals spanning greater than 15 points (on a 0 to 100 scale) on either side of the pooled effect as indicative of important heterogeneity. We will visually inspect the distribution of effect sizes in the forest plots and calculate the I^2 value to indicate the proportion of observed variance that is due to heterogeneity (Borenstein 2009). We will interpret I^2 greater than 50% as indicative of important heterogeneity. We will use the heterogeneity variances from the NMA models as measures of total network heterogeneity.

Assessment of statistical inconsistency

Assessment of inconsistency is only possible if there are closed loops of evidence in the network (all comparisons have direct evidence). We will rely on the assessment of transitivity to infer the presence of inconsistency if there are no closed loops. Inconsistency may not be detected for two reasons. Firstly, tests to detect inconsistency have low power (Higgins 2012; Krahn 2013; Song 2012; Veroniki 2014). Secondly, heterogeneity and inconsistency are interwoven. In situations where there is large heterogeneity, this may mask the presence of inconsistency (Song 2012; Veroniki 2013; Veroniki 2014).

Global approaches for evaluating consistency

We will evaluate consistency across the entire network using the 'design-by-treatment' interaction model (Higgins 2012; White 2012); and infer the presence of inconsistency based on P < 0.10.

Local approaches for evaluating consistency

We will evaluate consistency in closed loops with the loop-specific approach (Bucher 1997) and node-splitting (Dias 2010), using a threshold of P < 0.10 for either approach.

Strategy for investigating the sources of inconsistency

We will employ a staged approach to investigate any significant inconsistency that we encounter (Cipriani 2013; Salanti 2012). In the first instance, we will check for data extraction errors in the comparisons identified as inconsistent and those with important heterogeneity. Secondly, we will test whether the inconsistency may be explained using pre-specified covariates in network meta-regression and subgroup analyses, provided sufficient studies are available. Lastly, if there remains significant unexplained inconsistency, we will consider not proceeding with NMA. This judgement will involve the clinical and methodological evaluation of transitivity, the approaches to identify inconsistency and the knowledge that small amounts of inconsistency may be due to chance (Veroniki 2013).

Confidence in cumulative evidence

We will use the approach described by Salanti 2014 to construct judgements of confidence in each of the pairwise effects derived from NMA for pain and safety at end of treatment and confidence in the ranking of treatments with respect to their effect on pain and safety at end of treatment. We will likely use the Confidence in Network Meta-Analysis (CINeMA) web application for this process (ISPM 2017). We will consider the five GRADE domains: risk of bias; indirectness; inconsistency; imprecision; and publication bias. Initial judgements of confidence will be 'high', because the data come from randomised controlled trials, which are the ideal study design for the research question. The procedures for evaluation and downgrading of judgements differ slightly between those for confidence in the pairwise effects and confidence in the treatment rankings. These are described in Appendix 3.

Summary of findings

We will present 'Summary of findings' tables for pain and safety at end of treatment. The tables will be adapted from the current template used in Cochrane Reviews and will contain, at a minimum, the treatment effect, assumed risk (for safety only), quality of evidence (GRADE) and number of participants for each comparison in the network. An example 'Summary of findings' table is shown for a single comparison in Table 3. Development of methods for presentation of results of NMA is ongoing (e.g. Tan 2014; Veroniki 2016), and we anticipate the summary of findings in the finished review to incorporate these methodological advances where appropriate.

Subgroup analysis and investigation of heterogeneity

We will perform network meta-regression or subgroup analyses if sufficient studies are available, by investigating baseline pain intensity, sample size and enriched design as possible sources of inconsistency or heterogeneity. We may also investigate clinical or methodological factors identified during the review process that may threaten transitivity (Jansen 2012), as sources of inconsistency or heterogeneity, or both. We will assume for each network meta-regression model that the effects of the covariates are common (the same) across all comparisons in the network (Chaimani 2012; Dias 2013). This strong assumption is likely to make best use of the available data (Dias 2013; Efthimiou 2016). We specify the following assumptions about the direction of effect for each covariate.

- Baseline pain intensity (continuous): increasing magnitudes of the covariate increases the effect size between the intervention and comparator (compared to trials in which the covariate is less).
- Sample size (continuous): increasing magnitudes of the covariate reduces the effect size between the intervention and comparator.
- Enriched design (binary): the presence of the covariate increases the effect size for pain and decreases the effect size for safety, between the intervention and comparator (compared to trials in which the covariate is absent).

We plan to conduct a subgroup analysis if intended co-interventions are identified as threats to transitivity. We will group trials that test drug interventions alone separately from trials that include intended co-interventions, provided that sufficient data are available. We will assess small-study effects in pairwise comparisons using conventional funnel plots and across the entire network using comparison-adjusted funnel plots (Chaimani 2013).

Sensitivity analysis

We will conduct sensitivity analyses of pain and safety in which studies at high risk of bias are excluded, provided that the original network structure remains the same. We plan to investigate baseline pain intensity, sample size and enriched design in sensitivity analyses if sufficient data are not available for network meta-regression or subgroup analysis. We will dichotomise baseline pain intensity and sample size and exclude trials with baseline pain intensity higher than 70/100 (VAS) and sample size less than 50, respectively. We will also conduct sensitivity analyses of the analyses for pain at end of treatment and safety if we impute missing data for either of these outcomes.

ACKNOWLEDGEMENTS

Prof Georgia Salanti and A/Prof Andrea Cipriani for assistance with methodological development.

Dr Anna Chaimani for providing Stata code.

Dr Michel Rossignol, Dr Sarah Nevitt, Prof Roger Chou and Marilyn Walsh for their constructive feedback on the manuscript.

REFERENCES

Additional references

AMH 2017

Australian Medicines Handbook Pty Ltd. Australian Medicines Handbook 2017 (online). www.amhonline.amh.net.au/ accessed prior to 11 December 2017.

Amorin-Woods 2014

Amorin-Woods LG, Beck RW, Parkin-Smith GF, Lougheed J, Bremner AP. Adherence to clinical practice guidelines among three primary contact professions: a best evidence synthesis of the literature for the management of acute and subacute low back pain. *Journal of the Canadian Chiropractic Association* 2014;**58**(3):220–37.

Baigent 2013

Baigent C, Bhala N, Emberson J, Merhi A, Abramson S, Arber N, et al. Vascular and upper gastrointestinal effects of non-steroidal anti-inflammatory drugs: meta-analyses of individual participant data from randomised trials. *Lancet* 2013;**382**(9894):769–79.

Bhamb 2006

Bhamb B, Brown D, Hariharan J, Anderson J, Balousek S, Fleming MF. Survey of select practice behaviors by primary care physicians on the use of opioids for chronic pain. *Current Medical Research and Opinion* 2006;**22**(9):1859–65.

Bishop 2003

Bishop PB, Wing PC. Compliance with clinical practice guidelines in family physicians managing worker's compensation board patients with acute lower back pain. *Spine Journal* 2003;**3**(6):442–50.

Borenstein 2009

Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to meta-analysis*. Chichester, United Kingdom: John Wiley & Sons, Ltd., 2009.

Brewer 1999

Brewer T, Colditz GA. Postmarketing surveillance and adverse drug reactions. *JAMA* 1999;**281**(9):824.

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.

Bucher 1997

Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997;**50**(6):683–91.

Cabana 1999

Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud P-AC, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;**282**(15):1458–65.

Caldwell 2005

Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;**331**(7521):897–900.

Chaimani 2012

Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Research Synthesis Methods* 2012;**3**(2): 161–76.

Chaimani 2013

Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. *PLoS ONE* 2013;**8**(10):e76654.

Chaparro 2014

Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC. Opioids compared with placebo or other treatments for chronic low back pain: an update of the Cochrane Review. *Spine* 2014;**39**(7):556–63.

Chou 2007

Chou R, Hoffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians Clinical Practice Guideline. *Annals of Internal Medicine* 2007;**147**(7):492–504.

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.

Chung 2013

Chung JW, Zeng Y, Wong TK. Drug therapy for the treatment of chronic nonspecific low back pain: systematic review and meta-analysis. *Pain Physician* 2013;**16**(6): E685–704.

Cifuentes 2010
Cifuentes M, Webster B, Genevay S, Pransky G. The course of opioid prescribing for a new episode of disabling low back pain: Opioid features and dose escalation. *Pain* 2010; **151**(1):22–9.

Cipriani 2009
Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;**373**(9665):746–58.

Cipriani 2013
Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and Technical Challenges in Network Meta-analysis. *Annals of Internal Medicine* 2013;**159**(2):130–7.

Daly 2008
Daly FF, Fountain JS, Murray L, Graudins A, Buckley NA. Guidelines for the management of paracetamol poisoning in Australia and New Zealand - explanation and elaboration. *Medical Journal of Australia* 2008;**188**(5):296–301.

Darlow 2012
Darlow B, Fullen BM, Dean S, Hurley DA, Baxter GD, Dowell A. The association between health care professional attitudes and beliefs and the attitudes and beliefs, clinical management, and outcomes of patients with low back pain: a systematic review. *European Journal of Pain* 2012;**16**(1): 3–17.

Davies 2000
Davies NM, McLachlan AJ, Day RO, Williams KM. Clinical pharmacokinetics and pharmacodynamics of celecoxib: a selective cyclo-oxygenase-2 inhibitor. *Clinical Pharmacokinetics* 2000;**38**(3):225–42.

Dechartres 2016
Dechartres A, Trinquart L, Faber T, Ravaud P. Empirical evaluation of which trial characteristics are associated with treatment effect estimates. *Journal of Clinical Epidemiology* 2016;**77**:24–37.

Deyo 2015
Deyo RA, Von Korff M, Duhrkoop D. Opioids for low back pain. *BMJ* 2015;**350**:g96380.

Dias 2010
Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine* 2010;**29**(7-8):932–44.

Dias 2013
Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity-subgroups, meta-regression, bias, and bias-adjustment. *Medical Decision Making* 2013;**33**(5):618–40.

Dominici 1999
Dominici F, Parmigiani G, Wolpert RL, Hasselblad V. Meta-analysis of migraine headache treatments: combining information from heterogeneous designs. *Journal of the American Statistical Association* 1999;**94**(445):16–28.

Downie 2013
Downie A, Williams CM, Henschke N, Hancock MJ, Ostelo RW, de Vet HC, et al. Red flags to screen for malignancy and fracture in patients with low back pain: systematic review. *BMJ* 2013;**347**:f7095.

Efthimiou 2016
Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, et al. GetReal in network meta-analysis: a review of the methodology. *Research Synthesis Methods* 2016;**7**(3):236–63.

Enthoven 2016a
Enthoven WT, Geuze J, Scheele J, Bierma-Zeinstra SM, Bueving HJ, Bohnen AM, et al. Prevalence and "red flags" regarding specified causes of back pain in older adults presenting in general practice. *Physical Therapy* 2016;**96**(3): 305–12.

Enthoven 2016b
Enthoven WT, Roelofs PD, 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. DOI: 10.1002/14651858.CD012087

Entsminger 2014 [Computer program]
Entsminger GL. EcoSim Professional: Null modelling software for ecologists. Montrose: Acquired Intelligence Inc., Kesey–Bear, & Pinyon Publishing, 2014.

Furlan 2006
Furlan AD, Sandoval JA, Mailis-Gagnon A, Tunks E. Opioids for chronic noncancer pain-a meta-analysis of effectiveness and side effects. *CMAJ* 2006;**174**(11): 1589–94.

Furlan 2011
Furlan AD, Chaparro LE, Irvin E, Mailis-Gagnon A. A comparison between enriched and nonenriched enrollment randomized withdrawal trials of opioids for chronic noncancer pain. *Pain Research and Management* 2011;**16** (5):337–51.

Furlan 2014
Furlan AD, Irvin E, Kim J, Van Eerd D, Carnide N, Munhall C, et al. 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. DOI: 10.1002/14651858.CD011062

Furlan 2015
Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 Updated method guideline for systematic reviews in the Cochrane Back and Neck Group. *Spine* 2015;**40**(21):1660–73.

Furukawa 2016
Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* 2016;**6**(7):e010919. DOI: 10.1136/bmjopen-2015-010919

GBD Collaborators 2017a
GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. 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 2016. *Lancet* 2017;**390**(10100):1211–59.

GBD Collaborators 2017b

GBD 2016 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017;**390**(10100):1260–344.

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.

Graham 2013

Graham GG, Davies MJ, Day RO, Mohamudally A, Scott KF. The modern pharmacology of paracetamol: therapeutic actions, mechanism of action, metabolism, toxicity and recent pharmacological findings. *Inflammopharmacology* 2013;**21**(3):201–32.

Hart 2015

Hart OR, Uden RM, McMullan 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.

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, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2012

Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods* 2012;**3**(2):98–110.

Hinz 2008

Hinz B, Cheremina O, Brune K. Acetaminophen (paracetamol) is a selective cyclooxygenase-2 inhibitor in man. *The FASEB Journal* 2008;**22**(2):383–90.

Hinz 2012

Hinz B, Brune K. Paracetamol and cyclooxygenase inhibition: is there a cause for concern?. *Annals of the Rheumatic Diseases* 2012;**71**(1):20–5.

Hoaglin 2011

Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the ISPOR task force on indirect treatment comparisons good research practices: Part 2. *Value in Health* 2011;**14**(4): 429–37.

Hoffmann 2013

Hoffmann TC, Del Mar CB, Strong J, Mai J. Patients' expectations of acute low back pain management: implications for evidence uptake. *BMC Family Practice* 2013;**14**(1):7.

Humphreys 2017

Humphreys K. Avoiding globalisation of the prescription opioid epidemic. *Lancet* 2017;**390**(10093):437–39.

Hurlbert 1971

Hurlbert SH. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 1971;**52**(4):577–86.

ISPM 2017

Institute of Social and Preventive Medicine, University of Bern. CINeMA: Confidence in Network Meta-Analysis [Software]. www.cinema.ispm.ch accessed prior to 11 December 2017.

Itz 2013

Itz CJ, Geurts JW, Van Kleef M, Nelemans P. Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. *European Journal of Pain* 2013;**17**(1):5–15.

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 Journal* 2011; **11**(7):622–32.

Jansen 2011

Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: Report of the ISPOR task force on indirect treatment comparisons good research practices: Part 1. *Value in Health* 2011;**14**(4):417–28.

Jansen 2012

Jansen JP, Schmid CH, Salanti G. Directed acyclic graphs can help understand bias in indirect and mixed treatment comparisons. *Journal of Clinical Epidemiology* 2012;**65**(7): 798–807.

Jansen 2013

Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Medicine* 2013;**11**: 159.

Jozwiak-Bebenista 2014

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

Kieffer 2009
Kieffer BL, Evans CJ. Opioid receptors: from binding sites to visible molecules *in vivo*. *Neuropharmacology* 2009;**56 Suppl 1**:205–12.

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**(16):1313–7.

Krahn 2013
Krahn U, Binder H, König J. A graphical tool for locating inconsistency in network meta-analyses. *BMC Medical Research Methodology* 2013;**13**:35.

Kuijpers 2011
Kuijpers T, van Middelkoop M, Rubinstein SM, Ostelo R, Verhagen A, Koes BW, et al. A systematic review on the effectiveness of pharmacological interventions for chronic non-specific low-back pain. *European Spine Journal* 2011; **20**(1):40–50.

Liberati 2009
Liberati A, Altman DG, Tetzlaff J, Mulrow C, Ioannidis JP, Clarke M, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine* 2009;**151**(4):W65–W94.

Lu 2004
Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004;**23**(20):3105–24.

Lu 2006
Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 2006;**101**(474):447–59.

Machado 2017
Machado GC, Maher CG, Ferreira PH, Day RO, Pinheiro MB, Ferreira ML. Non-steroidal anti-inflammatory drugs for spinal pain: a systematic review and meta-analysis. *Annals of the Rheumatic Diseases* 2017;**76**(7):1269–78.

Mafi 2013
Mafi JM, McCarthy EP, Davis RB, Landon BE. Worsening trends in the management and treatment of back pain. *JAMA Internal Medicine* 2013;**173**(17):1573–81.

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

Martell 2007
Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, et al. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. *Annals of Internal Medicine* 2017;**146**(2): 116–27.

McGettigan 2011
McGettigan P, Henry D. Cardiovascular risk with non-steroidal anti-inflammatory drugs: Systematic review of population-based controlled observational studies. *PLoS Medicine* 2011;**8**(9):e1001098.

Menezes Costa 2012
Menezes Costa L da C, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ* 2012;**184** (11):1229–30.

Mills 2013
Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ* 2013;**2914** (May):10–5.

MIMS 2017
Haymarket Media Group Ltd. MIMS Online. www.mims.co.uk/ accessed prior to 11 December 2017.

Ndlovu 2014
Ndlovu M, Bedson J, Jones PW, Jordan KP. Pain medication management of musculoskeletal conditions at first presentation in primary care: analysis of routinely collected medical record data. *BMC Musculoskeletal Disorders* 2014; **15**:418.

PDR 2017
PDR, LLC. Prescriber's Digital Reference. www.pdr.net accessed prior to 11 December 2017.

Perrot 2008
Perrot S, Concas V, Allaert F, Laroche F. Deciding on analgesic prescription dosing for acute back pain: once daily or more? [Évaluation De La Prise De Décision De Prescription D'Un Antalgique En Prise Quotidienne Unique Ou Multiple Dans Les Rachialgies Aiguës]. *Presse Medicale* 2008;**37**(1):14–20.

Peters 2008
Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* 2008;**61**(10):991–6.

Rachinger-Adam 2011
Rachinger-Adam B, Conzen P, Azad SC. Pharmacology of peripheral opioid receptors. *Current Opinion in Anesthesiology* 2011;**24**(4):408–13.

Review Manager 2014 [Computer program]
Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3.5. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Salanti 2008
Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. *Annals of Internal Medicine* 2008;**148**(7):544–53.

Salanti 2011
Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: An overview and tutorial. *Journal of Clinical Epidemiology* 2011;**64**(2):163–71.

Salanti 2012
Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* 2012;**3**(2):80–97.

Salanti 2014
Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE* 2014;**9**(7):e99682.

Saragiotto 2016a
Saragiotto BT, Maher CG, Moseley AM, Yamato TP, Koes BW, Sun X, et al. A systematic review reveals that the credibility of subgroup claims in low back pain trials was low. *Journal of Clinical Epidemiology* 2016;**79**:3–9.

Saragiotto 2016b
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. DOI: 10.1002/14651858.CD012230

Schers 2001
Schers H, Wensing M, Huijsmans Z, van Tulder MW, Grol R. Implementation barriers for general practice guidelines on low back pain a qualitative study. *Spine* 2001;**26**(15): E348–E53.

Shaheed 2016
Shaheed CA, Maher CG, Williams KA, Day R, McLachlan AJ. Efficacy, tolerability, and dose-dependent effects of opioid analgesics for low back pain a systematic review and meta-analysis. *JAMA Internal Medicine* 2016;**176**(7): 958–68.

Sheen 2002
Sheen CL, Dillon JF, Bateman DN, Simpson KJ, Macdonald TM, Poisons S, et al. Paracetamol toxicity: epidemiology, prevention and costs to the health-care system. *QJM : Monthly Journal of the Association of Physicians* 2002;**95**(9): 609–19.

Sills 1986
Sills JM, Tanner LA, Milstien JB. Food and Drug Administration monitoring of adverse drug reactions. *American Journal of Hospital Pharmacy* 1986;**43**(11): 2764–70.

Song 2003
Song F, Altman DG, Glenny A-M, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;**326**(7387):472.

Song 2012
Song F, Clark A, Bachmann MO, Maas J. Simulation evaluation of statistical properties of methods for indirect and mixed treatment comparisons. *BMC Medical Research Methodology* 2012;**12**:138.

Sostres 2013
Sostres C, Gargallo CJ, Lanas A. Nonsteroidal anti-inflammatory drugs and upper and lower gastrointestinal mucosal damage. *Arthritis Research and Therapy* 2013;**15** (Suppl 3):1–8.

StataCorp 2015 [Computer program]
StataCorp. Stata. Version 14. College Station, TX, USA: StataCorp, 2015.

Sutton 2002
Sutton AJ, Cooper NJ, Lambert PC, Jones DR, Abrams KR, Sweeting MJ. Meta-analysis of rare and adverse event data. *Expert Review of Pharmacoeconomics & Outcomes Research* 2002;**2**(4):367–79.

Sutton 2008
Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *PharmacoEconomics* 2008;**26**(9):753–67.

Tan 2014
Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. *Journal of Clinical Epidemiology* 2014;**67**(6):672–80.

Treede 2015
Treede R-D, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. *PAIN* 2015;**156**(6):1003–7.

van der Valk 2009
van der Valk R, Webers CA, Lumley T, Hendrikse F, Prins MH, Schouten JS. A network meta-analysis combined direct and indirect comparisons between glaucoma drugs to rank effectiveness in lowering intraocular pressure. *Journal of Clinical Epidemiology* 2009;**62**(12):1279–83.

Veroniki 2013
Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *International Journal of Epidemiology* 2013;**42**(1):332–45.

Veroniki 2014
Veroniki AA, Mavridis D, Higgins JP, Salanti G. Characteristics of a loop of evidence that affect detection and estimation of inconsistency: a simulation study. *BMC Medical Research Methodology* 2014;**14**:106.

Veroniki 2016
Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *Journal of Clinical Epidemiology* 2016;**76**:193–9.

Wan 2014
Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology* 2014;**14**:135.

Webster 2005
Webster BS, Courtney TK, Huang YH, Matz S, Christiani DC. Brief report: physicians' initial management of acute low back pain versus evidence-based guidelines. Influence of sciatica. *Journal of General Internal Medicine* 2005;**20** (12):1132–5.

Webster 2007
Webster BS, Verma SK, Gatchel RJ. Relationship between early opioid prescribing for acute occupational low back pain and disability duration, medical costs, subsequent surgery and late opioid use. *Spine* 2007;**32**(19):2127–32.

White 2009
White IR. Multivariate random-effects meta-analysis. *Stata Journal* 2009;**9**(1):40–56.

White 2011a
White AP, Arnold PM, Norvell DC, Ecker E, Fehlings MG. Pharmacologic management of chronic low back pain. *Spine* 2011;**36**(21):S131–S43.

White 2011b
White IR. Multivariate random-effects meta-regression: Updates to mvmeta. *Stata Journal* 2011;**11**(2):255–70.

White 2012
White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods* 2012;**3**(2):111–25.

White 2015
White IR. Network meta-analysis. *Stata Journal* 2015;**15**(4):951–85.

Whittle 2000
Whittle BJ. COX-1 and COX-2 products in the gut: therapeutic impact of COX-2 inhibitors. *Gut* 2000;**47**(3):320–5.

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. Sources of risk of bias

Bias Domain	Source of Bias	PossibleAnswers
Selection	(1) Was the method of randomization adequate?	Yes/No/Unsure
Selection	(2) Was the treatment allocation concealed?	Yes/No/Unsure
Performance	(3) Was the patient blinded to the intervention?	Yes/No/Unsure
Performance	(4) Was the care provider blinded to the intervention?	Yes/No/Unsure
Detection	(5) Was the outcome assessor blinded to the intervention?	Yes/No/Unsure
Attrition	(6) Was the drop-out rate described and acceptable?	Yes/No/Unsure
Attrition	(7) Were all randomized participants analyzed in the group to which they were allocated?	Yes/No/Unsure
Reporting	(8) Are reports of the study free of suggestion of selective outcome reporting?	Yes/No/Unsure
Selection	(9) Were the groups similar at baseline regarding the most important prognostic indicators?	Yes/No/Unsure
Performance	(10) Were cointerventions avoided or similar?	Yes/No/Unsure
Performance	(11) Was the compliance acceptable in all groups?	Yes/No/Unsure
Detection	(12) Was the timing of the outcome assessment similar in all groups?	Yes/No/Unsure

Table 1. Sources of risk of bias (*Continued*)

Other	(13) Are other sources of potential bias unlikely?	Yes/No/Unsure

Furlan 2015

Table 2. Criteria for a judgment of "yes" for the sources of risk of bias

1	A random (unpredictable) assignment sequence. Examples of adequate methods are coin toss (for studies with 2 groups), rolling a dice (for studies with 2 or more groups) , drawing of balls of different colours, drawing of ballots with the study group labels from a dark bag, computer-generated random sequence, preordered sealed envelopes, sequentially-ordered vials, telephone call to a central office, and preordered list of treatment assignments Examples of inadequate methods are: alternation, birth date, social insurance/security number, date in which they are invited to participate in the study, and hospital registration number
2	Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient
3	Index and control groups are indistinguishable for the patients or if the success of blinding was tested among the patients and it was successful
4	Index and control groups are indistinguishable for the care providers or if the success of blinding was tested among the care providers and it was successful
5	Adequacy of blinding should be assessed for each primary outcome separately. This item should be scored "yes" if the success of blinding was tested among the outcome assessors and it was successful or: - for patient-reported outcomes in which the patient is the outcome assessor (e.g. pain, disability): the blinding procedure is adequate for outcome assessors if participant blinding is scored "yes" - for outcome criteria assessed during scheduled visit and that supposes a contact between participants and outcome assessors (e.g., clinical examination): the blinding procedure is adequate if patients are blinded, and the treatment or adverse effects of the treatment cannot be noticed during clinical examination - for outcome criteria that do not suppose a contact with participants (e.g., radiography, magnetic resonance imaging): the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed when assessing the main outcome - for outcome criteria that are clinical or therapeutic events that will be determined by the interaction between patients and care providers (e.g., cointerventions, hospitalization length, treatment failure), in which the care provider is the outcome assessor: the blinding procedure is adequate for outcome assessors if item "4" (caregivers) is scored "yes" - for outcome criteria that are assessed from data of the medical forms: the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed on the extracted data
6	The number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given. If the percentage of withdrawals and drop-outs does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias a "yes" is scored. (N.B. these percentages are arbitrary, not supported by literature)
7	All randomized patients are reported/analyzed in the group they were allocated to by randomization for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and cointerventions

Table 2. Criteria for a judgment "yes" for the sources of risk of bias (*Continued*)

8	All the results from all prespecified outcomes have been adequately reported in the published report of the trial. This information is either obtained by comparing the protocol and the report, or in the absence of the protocol, assessing that the published report includes enough information to make this judgment
9	Groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with neurological symptoms, and value of main outcome measure(s)
10	If there were no cointerventions or they were similar between the index and control groups
11	The reviewer determines if the compliance with the interventions is acceptable, based on the reported intensity, duration, number and frequency of sessions for both the index intervention and control intervention(s). For example, physiotherapy treatment is usually administered for several sessions; therefore it is necessary to assess how many sessions each patient attended. For single-session interventions (e.g. surgery), this item is irrelevant
12	Timing of outcome assessment should be identical for all intervention groups and for all primary outcome measures
13	Other types of biases. - Industry-sponsored trials. The conflict of interest (COI) statement should explicitly state that the researchers have had full possession of the trial process from planning to reporting without funders with potential COI having any possibility to interfere in the process. If, for example, the statistical analyses have been done by a funder with a potential COI, usually "unsure" is scored

Furlan 2015

Table 3. Summary of findings

Paracetamol, NSAIDs and opioid analgesics for chronic low back pain: network meta-analysis

Patient or population: non-specific chronic low back pain
Settings: primary and secondary care, (additional detail unknown at protocol stage)
Intervention: paracetamol, NSAIDs or opioid analgesics
Comparison: intervention A vs intervention B, of X possible comparisons (unknown at protocol stage)

Outcomes	Outcome type	Outcome measure	Comments
Pain	Continuous	Pain intensity on 0 to 100 mm VAS	Measured at end of treatment
Safety	Dichotomous	Proportion of participants that experience a serious adverse event during treatment	

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Table 3. Summary of findings (*Continued*)

Very low quality: we are very uncertain about the estimate.

APPENDICES

Appendix 1. Interventions of interest

Drug name	ATC code	Notes
morphine	N02AA01	
morphine combinations	N02AA51	
hydromorphone	N02AA03	
oxymorphone	N02AA	
nicomorphine	N02AA04	
oxycodone	N02AA05	
oxycodone and naloxone	N02AA55	
oxycodone and paracetamol	N02AJ17	
oxycodone and acetylsalicylic acid	N02AJ18	
oxycodone and ibuprofen	N02AJ19	
codeine	R05DA04	
codeine, combinations excl. psycholeptics	N02AA59	where these are NSAIDs, as defined herein, or Paracetamol
codeine and paracetamol	N02AJ06	
codeine and acetylsalicylic acid	N02AJ07	
codeine and ibuprofen	N02AJ08	
codeine and other non-opioid analgesics	N02AJ09	where these are NSAIDs, as defined herein, or Paracetamol

(*Continued*)

dihydrocodeine	N02AA08	
dihydrocodeine combinations	N02AA58	where these are combinations with NSAIDs, as defined herein, or Paracetamol
dihydrocodeine and paracetamol	N02AJ01	
dihydrocodeine and acetylsalicylic acid	N02AJ02	
dihydrocodeine and other non-opioid analgesics	N02AJ03	where these are NSAIDs, as defined herein, or Paracetamol
papaveretum	N02AA10	
buprenorphine	N02AE01	
tilidine	N02AX01	
dezocine	N02AX03	
meptazinol	N02AX05	
tapentadol	N02AX06	
tramadol	N02AX02	
tramadol and paracetamol	N02AJ13	
tramadol and dexketoprofen	N02AJ14	
tramadol and other non-opioid analgesics	N02AJ15	where these are NSAIDs, as defined herein, or Paracetamol
butaphornal	N02AF01	
nalbuphine	N02AF02	
ketobemidone	N02AB01	
pethidine	N02AB02	
pethidine, combinations excl. psycholeptics	N02AB52	where these are NSAIDs, as defined herein, or Paracetamol
fentanyl	N02AB03	
dextromoromide	N02AC01	
piritramide	N02AC03	

(*Continued*)

dextropropoxyphene	N02AC04	
bezitramide	N02AC05	
methadone	N07BC02	
methadone, combinations excl. psycholeptics	N02AC52	where these are NSAIDs, as defined herein, or Paracetamol
dextropropoxyphene, combinations excl. psycholeptics	N02AC54	where these are NSAIDs, as defined herein, or Paracetamol
pentazocine	N02AD01	
phenazocine	N02AD02	
phenylbutazone	M01AA01	
mofebutazone	M01AA02	
oxyphenbutazone	M01AA03	
piroxicam	M01AC01	
lornoxicam	M01AC02	
meloxicam	M01AC06	
ibuprofen	M01AE01	
naproxen	M01AE02	
ketoprofen	M01AE03	
fenoprofen	M01AE04	
flurbiprofen	M01AE09	
tiaprofenic acid	M01AE11	
oxaprozin	M01AE12	
dexibruprofen	M01AE14	
dexketoprofen	M01AE17	
mefenamic acid	M01AG01	

(*Continued*)

tolfenamic acid	M01AG02	
meclofenamic acid	M01AG04	
indomethacin	M01AB01	
sulindac	M01AB02	
tolmetin	M01AB03	
zomepirac	M01AB04	
diclofenac	M01AB05	
alclofenac	M01AB06	
etodolac	M01AB08	
ketorolac	M01AB15	
aceclofenac	M01AB16	
bufexamac	M01AB17	
celecoxib	M01AH01	
valdecoxib	M01AH03	
parecoxib	M01AH04	
etoricoxib	M01AH05	
nabumetone	M01AX01	
glucosamine	M01AX05	
GAG polysulfate	M01AX12	
nimesulide	M01AX17	
chondroitin sulphate	M01AX25	
paracetamol (acetaminophen)	N02BE01	
acetylsalicylic acid (aspirin)	N02BA01	
acetylsalicylic acid, combinations excl. psycholeptics	M01BA03	where these are NSAIDs, as defined herein, or Paracetamol

Appendix 2. Search strategies

Search Strategy for MEDLINE (Ovid):

Part A: Generic search for randomized controlled trials

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. comparative study.pt.
4. clinical trial.pt.
5. random*.ab,ti.
6. placebo.ab,ti.
7. drug therapy.fs.
8. trial.ab,ti.
9. groups.ab,ti.
10. or/1-9
11. (animals not (humans and animals)).sh.
12. 10 not 11

Part B: Specific search for low back, sacrum and coccyx problems

13. dorsalgia.ti,ab.
14. exp Back Pain/
15. backache.ti,ab.
16. ((lumb* or back) adj pain).ti,ab.
17. coccydynia.ti,ab.
18. sciatica.ti,ab.
19. spondylosis.ti,ab.
20. lumbago.ti,ab.
21. or/13-20

Part C: Specific search for other spinal disorders

22. Coccyx.sh
23. Lumbar Vertebrae.sh
24. Intervertebral disc.sh
25. Sacrum.sh
26. Intervertebral disc degeneration.sh
27. (disc adj degenerat*).ti,ab.
28. (disc adj prolapse*).ti,ab.
29. (disc adj herniat*).ti,ab.
30. spinal fusion.sh.
31. (facet adj joint*).ti,ab.
32. Intervertebral Disc Displacement.sh.
33. postlaminectomy.ti,ab.
34. or/22-33

Part D: Specific search for interventions of interest

35. morphine.sh or morphine.tw,kf.
36. hydromorphone.sh or hydromorphone.tw,kf.
37. oxycodone.sh or oxycodone.tw,kf.
38. oxymorphone.sh or oxymorphone.tw,kf.
39. nicomorphine.tw,kf.
40. codeine.sh or codeine.tw,kf.
41. exp naloxone/ or naloxone.tw,kf.
42. dihydrocodeine.tw,kf.
43. papaveratum.tw,kf.
44. buprenorphine.sh or buprenorphine.tw,kf.
45. tilidine.sh or tilidine.tw,kf.
46. dezocine.tw,kf.

47. meptazinol.sh or meptazinol.tw,kf.
48. tapentadol.tw,kf.
49. tramadol.sh or tramadol.tw,kf.
50. butaphornal.tw,kf.
51. nalbuphine.sh or nalbuphine.tw,kf.
52. ketobemidone.tw,kf.
53. meperidine.sh or meperidine.tw,kf. or pethidine.tw,kf.
54. fentanyl.sh or fentanyl.tw,kf.
55. dextromoramide.sh or dextromoramide.tw,kf.
56. piritramide.sh or piritramide.tw,kf.
57. dextropropoxyphene.sh or dextropropoxyphene.tw,kf.
58. bezitramide.tw,kf.
59. methadone.sh or methadone.tw,kf.
60. pentazocine.sh or pentazocine.tw,kf.
61. phenazocine.sh or phenazocine.tw,kf.
62. phenylbutazone.sh or phenylbutazone.tw,kf.
63. mofebutazone.tw,kf.
64. oxyphenbutazone.sh or oxyphenbutazone.tw,kf.
65. meloxicam.tw,kf.
66. piroxicam.sh or piroxicam.tw,kf.
67. lornoxicam.tw,kf.
68. ibuprofen.sh or ibuprofen.tw,kf.
69. naproxen.sh or naproxen.tw,kf.
70. ketoprofen.sh or ketoprofen.tw,kf.
71. fenoprofen.sh or fenoprofen.tw,kf.
72. flurbiprofen.sh or flurbiprofen.tw,kf.
73. tiaprofenic acid.tw,kf.
74. oxaprozin.tw,kf.
75. dexibuprofen.tw,kf.
76. dexketoprofen.tw,kf.
77. mefenamic acid.sh or mefenamic acid.tw,kf.
78. tolfenamic acid.tw,kf.
79. meclofenamic acid.sh or meclofenamic acid.tw,kf.
80. exp indomethacin/ or indometacin.tw,kf.
81. sulindac.sh or sulindac.tw,kf.
82. tolmetin.sh or tolmetin.tw,kf.
83. zomepirac.tw,kf.
84. diclofenac.sh or diclofenac.tw,kf.
85. alclofenac.tw,kf.
86. etodolac.sh or etodolac.tw,kf.
87. aceclofenac.tw,kf.
88. bufexamac.sh or bufexamac.tw,kf.
89. celecoxib.sh or celecoxib.tw,kf.
90. valdecoxib.tw,kf.
91. etoricoxib.tw,kf.
92. nabumetone.tw,kf.
93. exp glucosamine/ or glucosamine.tw,kf.
94. glucosaminoglycan polysulfate.tw,kf.
95. nimesulide.tw,kf.
96. chondroitin sulfate.tw,kf.
97. acetaminophen.sh or acetaminophen.tw,kf. or paracetamol.tw,kf.
98. aspirin.sh or aspirin.tw,kf. or acetylsalicylic acid.tw,kf.
99. or/35-98 (all interventions of interest)

Results
100. 21 or 34 (all back pain)
101. 99 and 100 (all back pain and all interventions of interest)
102. 12 and 101 (all RCTs of interventions of interest in back pain)

Appendix 3. The GRADE approach to evidence synthesis

The quality of evidence will be categorized as follows:

- High (⊙ ⊙ ⊙ ⊙): further research is very unlikely to change the confidence in the estimate of effect.
- Moderate (⊙ ⊙ ⊙ ◯): further research is likely to have an important impact in the confidence in the estimate of effect.
- Low (⊙ ⊙ ◯◯): further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- Very Low (⊙ ◯◯◯): any estimate of effect is very uncertain.

Initial judgements of confidence will be 'high', because the data come from randomised controlled trials. We will consider the five GRADE domains: risk of bias, indirectness, inconsistency, imprecision and publication bias. The evidence available under each of these domains will be graded according to slightly different procedures for confidence in the pairwise effects and confidence in the treatment rankings.

Judgement of confidence in the pairwise effects

1) Risk of bias. We will rate each direct comparison as at low, moderate or high risk of bias, assigning scores of 0, −1 and −2, respectively. We will consider direct comparisons at high risk of bias when > 25% of participants in the comparison are from studies at high overall risk of bias, low risk when > 50% of participants are from studies at low overall risk and moderate risk in all other instances. We will use the percentage contributions of all direct comparisons to each pairwise effect (from the contribution matrix) to construct weighted averages of these scores. Then we will base judgements of the confidence to be placed in each pairwise effect on a weighted average of the risk of bias of direct comparisons feeding into it. We will downgrade one level for $-1.5 <$ score < -0.5 and two levels for scores < -1.51, although this may change in practice if there is clear imbalance in the contribution of evidence that would render such action inappropriate.
2) Indirectness. We will evaluate the populations, interventions and outcomes in the studies contributing to each direct comparison for congruence with those specified for this review. We will consider downgrading a pairwise effect one level for indirectness if direct comparisons with important imbalances in populations, treatments or outcomes contribute the majority of information to it. In addition, we will downgrade all comparisons one additional level if there were concerns over the transitivity assumption identified prior to the analysis.
3) Inconsistency. We will evaluate each direct comparison for consistency in the direction and magnitude of the effect sizes from individual trials, considering the width of the prediction interval and magnitude of the heterogeneity parameter. We will downgrade direct comparisons one level if important heterogeneity is identified. Additionally, we will downgrade one level any direct comparisons that are implicated in loops with important incoherence or where there is a discrepancy between direct and indirect evidence.
4) Imprecision. We will downgrade due to imprecision based on consideration of the width of the confidence intervals. We will downgrade one level when these span either the null or the threshold for a clinically meaningful effect on pain intensity (10 points on a 0 to 100 scale (Chou 2017)) and two levels when the interval spans both. We will not consider sample sizes as there are no established criteria for this evaluation in the NMA context (Salanti 2014).
5) Publication bias. We will consider the likelihood of publication bias for each direct comparison, using both epidemiological judgement and the contour-enhanced funnel plots. We will downgrade one level if we consider it likely that studies have been conducted and not published.

Judgement of confidence in the treatment rankings

The difference between the judgement of confidence in the treatment rankings and that for the pairwise effect sizes is that this is a single judgement for the ranking evidence as a whole.

1) Risk of bias. We will use the scores of 0, −1, or −2 as previously for the judgements of overall risk of bias for each direct comparison and construct a weighted average for the entire network using the percentage contributions of each direct comparison to the entire network. We will downgrade one level for −1.5 < score < −0.5 and two levels for scores < −1.51, although this may change in practice if there is clear imbalance in the contribution of evidence that would render such action inappropriate.
2) Indirectness. We will use the above evaluation of congruence between study populations, interventions and outcomes. We will evaluate the contribution to the entire network by comparisons that are judged to exhibit indirectness. We will consider downgrading one level if there is an important contribution from one/more of these comparisons. In addition, we will downgrade all comparisons one additional level if there were concerns over the transitivity assumption identified prior to the analysis.
3) Inconsistency. We will consider the magnitude of the heterogeneity parameter from the NMA model and the result of the Chi^2 test for global inconsistency, acknowledging that we may fail to detect important global inconsistency due to the low power of such a test and the presence of large heterogeneity (Higgins 2012; Krahn 2013; Song 2012; Veroniki 2013; Veroniki 2014). We will downgrade one level if either heterogeneity or inconsistency are present and two levels if both are present.
4) Imprecision. Imprecision in ranking may be conceptualised as uncertainty in the probability that a treatment achieves a certain rank (Salanti 2014). We will evaluate the precision of the treatment rankings by examining the SUCRA values that are used to calculate these probabilities. We will deem rankings imprecise when there are similar probabilities for two/more treatment to be ranked at the same level. We will downgrade one level in this situation.
5) Publication bias. We will evaluate the likelihood of publication bias across the network, using epidemiological judgement, evidence of publication bias in direct comparisons and the comparison-adjusted funnel plot, acknowledging that plot asymmetry may due to real factors other than publication bias (Salanti 2014). We will downgrade one level if there is evidence of publication bias or small study effects.

CONTRIBUTIONS OF AUTHORS

MKB provided NMA methodology expertise and wrote the protocol.

AJM provided pharmacological and systematic review methodology expertise.

AN provided NMA methodology and statistical analysis expertise.

CGM, SK, CMW, NH, GLM, BMW, MH, NOC and MvT provided systematic review methodology and clinical area expertise.

JHM is the guarantor and provided clinical area, systematic review and NMA methodology expertise.

All authors read, contributed to and approved the final version of the manuscript.

DECLARATIONS OF INTEREST

CGM and AJM are conducting an investigator-initiated, NHMRC-funded trial of opioids for acute LBP (OPAL). The OPAL trial has no funding from pharmaceutical companies. They have both received supplementary research funding from Pfizer and GlaxoSmithKline for two investigator-initiated, NHMRC-funded trials of paracetamol and pregabalin for low back pain (PACE & PRECISE).

GLM has received lecturing fees from Kaiser Permanente for their Physical Therapy Residency programme, and consultancy fees from workers' compensation boards in Australia, North America and Europe for advice on policy matters concerning acute back pain and persistent musculoskeletal pain states. GLM has received consultancy fees from Port Adelaide Football Club for professional services regarding pain and performance. GLM has received travel and consultancy fees from Arsenal Football Club for professional services regarding pain and performance. GLM has received travel support and expenses from the International Olympic Committee for committee roles focusing on pain management in elite sport. GLM receives guest lecturing fees from Noigroup Australasia and has received guest lecturing fees from Agile Physiotherapy (USA) and Results Physiotherapy (USA), for professional development courses. GLM receives speaker's fees for lectures on pain and rehabilitation and royalties for books on pain and rehabilitation. GLM has received honoraria and travel support from Pfizer for two lectures and for participation in the ASIA Masterclass on Neuropathic Pain. These contributions did not mention drugs. GLM has also received travel support from Grünenthal to attend a European CRPS minimum standards of care summit. This contribution did not involve any drugs.

MKB has received travel support from the Chiropractor's Association of Australia for speaking engagements unrelated to this topic.

MvT and CGM are on the Editorial Board of Cochrane Back and Neck. MvT was Co-ordinating Editor until September 2017. Editors are required to conduct at least one Cochrane Review. This requirement ensures that they are aware of the processes and commitment needed to conduct reviews. None of the Editors are first authors. Any Editor who is a review author is excluded from editorial decisions on the review in which they are contributors.

BMW, MH, SK, NH, NOC, CMW, AN and JHM have no 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

- No sources of support supplied

External sources

- National Health and Medical Research Council, Australia.

Research Fellowship (CMW, SK), Principal Research Fellowship (RH, CGM, GLM)

- Australian Government Research Training Program Scholarship, Australia.

(MKB)

- UNSW Research Excellence Award, Australia.

(MKB)

- Neuroscience Research Australia PhD Candidature Supplementary Scholarship, Australia.

(MKB)