Cochrane Library

Cochrane Database of Systematic Reviews

Non-steroidal anti-inflammatory drugs for chronic low back pain (Review)

Enthoven WTM, Roelofs PDDM, Deyo RA, van Tulder MW, Koes BW

Enthoven WTM, Roelofs PDDM, Deyo RA, van Tulder MW, Koes BW.
Non-steroidal anti-inflammatory drugs for chronic low back pain.
Cochrane Database of Systematic Reviews 2016, Issue 2. Art. No.: CD012087.
DOI: 10.1002/14651858.CD012087.

www.cochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

Non-steroidal anti-inflammatory drugs for chronic low back pain

Wendy TM Enthoven[1], Pepijn DDM Roelofs[2], Richard A Deyo[3], Maurits W van Tulder[4], Bart W Koes[1]

[1]Department of General Practice, Erasmus Medical Center, Rotterdam, Netherlands. [2]Research Centre Innovations in Care, Rotterdam University of Applied Sciences, Rotterdam, Netherlands. [3]Department of Family Medicine, Dept. of Medicine, Dept. of Public Health & Preventive Medicine, Oregon Health and Science University, Portland, OR, USA. [4]Department of Health Sciences, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, Netherlands

Contact address: Wendy TM Enthoven, Department of General Practice, Erasmus Medical Center, PO Box 2040, Rotterdam, 3000CA, Netherlands. w.enthoven@erasmusmc.nl.

Editorial group: Cochrane Back and Neck Group.
Publication status and date: Edited (no change to conclusions), published in Issue 8, 2016.
Review content assessed as up-to-date: 23 June 2015.

Citation: Enthoven WTM, Roelofs PDDM, Deyo RA, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for chronic low back pain. *Cochrane Database of Systematic Reviews* 2016, Issue 2. Art. No.: CD012087. DOI: 10.1002/14651858.CD012087.

ABSTRACT

Background

Chronic back pain is an important health problem. Non-steroidal anti-inflammatory drugs (NSAIDs) are widely used to treat people with low back pain, especially people with acute back pain. Short term NSAID use is also recommended for pain relief in people with chronic back pain. Two types of NSAIDs are available and used to treat back pain: non-selective NSAIDs and selective COX-2 NSAIDs. In 2008, a Cochrane review identified a small but significant effect from NSAIDs compared to placebo in people with chronic back pain. This is an update of the Cochrane review published in 2008 and focuses on people with chronic low back pain.

Objectives

To determine if NSAIDs are more efficacious than various comparison treatments for non-specific chronic low back pain and if so, which type of NSAID is most efficacious.

Search methods

We searched CENTRAL, MEDLINE, EMBASE, PubMed and two clinical trials registry databases up to 24 June 2015 for randomized controlled trials (RCTs) published in English, German or Dutch. We also screened references cited in relevant reviews.

Selection criteria

We included RCTs (double-blind and single-blind) of NSAIDs used to treat people with chronic low back pain.

Data collection and analysis

Two review authors independently screened trials for inclusion in this Cochrane review according to the inclusion criteria. One review author extracted the data, and a second review author checked the data. Two review authors independently evaluated the risk of bias of all included trials. If data were clinically homogeneous, we performed a meta-analysis and assessed the quality of evidence using the GRADE approach.

Main results

We included 13 trials in this Cochrane review. Ten studies were at 'low' risk of bias. Six studies compared NSAIDs with placebo, and included 1354 participants in total. There is low quality evidence that NSAIDs are more effective than placebo, with a mean difference in pain intensity score from baseline of -6.97 (95% CI −10.74 to −3.19) on a 0 to 100 visual analogue scale (VAS) with a median follow-up of 56 days (interquartile range (IQR) 13 to 91 days). Four studies measured disability using the Roland Morris Disability Questionnaire. There is low quality evidence that NSAIDs are more effective than placebo on disability, with a mean difference from baseline of −0.85 (95% CI −1.30 to −0.40) on a scale from 0 to 24 with a median follow-up of 84 days (IQR 42 to 105 days). All six placebo controlled studies also reported adverse events, and suggested that adverse events are not statistically significant more frequent in participants using NSAIDs compared to placebo (RR 1.04, 95% CI 0.92 to 1.17). Due to the relatively small sample size and relatively short follow-up in most included trials, it is likely that the proportion of patients experiencing an adverse event is underestimated.

Two studies compared different types of non-selective NSAIDs, namely ibuprofen versus diclofenac and piroxicam versus indomethacin. The trials did not find any differences between these NSAID types, but both trials had small sample sizes. One trial reported no differences in pain intensity between treatment groups that used selective or non-selective NSAIDs. One other trial compared diflunisal with paracetamol and showed no difference in improvement from baseline on pain intensity score. One trial showed a better global improvement in favour of celecoxib versus tramadol.

One included trial compared NSAIDs with 'home-based exercise'. Disability improved more in participants who did exercises versus participants receiving NSAIDs, but pain scores were similar.

Authors' conclusions

Six of the 13 included RCTs showed that NSAIDs are more effective than placebo regarding pain intensity. NSAIDs are slightly more effective than placebo regarding disability. However, the magnitude of the effects is small, and the level of evidence was low. When we only included RCTs at low risk of bias, differences in effect between NSAIDs and placebo were reduced. We identified no difference in efficacy between different NSAID types, including selective versus non-selective NSAIDs. Due to inclusion of RCTs only, the relatively small sample sizes and relatively short follow-up in most included trials, we cannot make firm statements about the occurrence of adverse events or whether NSAIDs are safe for long-term use.

PLAIN LANGUAGE SUMMARY

Non-steroidal anti-inflammatory drugs for chronic low back pain

Review question

We assessed the evidence regarding the effect of non-steroidal anti-inflammatory drugs (NSAIDs) among people with chronic low back pain. NSAIDs were compared to placebo, other NSAIDs, other drugs or other kinds of treatment.

Background

Chronic low back pain is common and causes pain and disability. NSAIDs are often used to treat people with chronic low back pain and are available both over-the-counter and on prescription in different types and chemical entities.

Study characteristics

We collected all published randomized controlled trials evaluating the efficacy of NSAIDs until 24 June 2015. We included 13 trials which compared NSAIDs with placebo, other NSAIDs, other drugs or other treatment in people with chronic low back pain. Six trials compared NSAIDs with placebo, and included 1354 participants in total. Follow-up was between nine days and 16 weeks.

Key results

NSAIDs reduced pain and disability in people with chronic low back pain compared to placebo. However, the differences were small: 7 points on a 100-point scale for pain intensity. Regarding disability, people receiving NSAIDs scored 0.9 points better on a 0 to 24 disability scale. The number of adverse events was not significantly different between the people receiving NSAIDs and people receiving placebo, but larger studies of longer duration would be needed to identify rare or delayed adverse events, important drug interactions and adverse events occurring with prolonged use.

Different types of NSAIDs did not show significantly different effects. Three of the 13 included studies compared two different types of NSAIDs and none found any differences.

NSAIDs were also compared to other drug types: paracetamol, tramadol and pregabalin. There were no differences found between NSAIDs and paracetamol and pregabalin in either effect or adverse events. A single study comparing celecoxib with tramadol showed a better global improvement in peoples using celecoxib.

One trial compared NSAIDs with 'home-based exercise'. Regarding disability, people who did exercise improved more than people receiving NSAIDs, but pain scores were not statistically different.

Quality of the evidence

There was low quality evidence that NSAIDs are slightly more effective than placebo in chronic low back pain. The magnitude of the difference was small, and when we only accounted for trials of higher quality, these differences reduced.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [Explanation]

NSAIDs for people with chronic low back pain compared to placebo

Participant or population: people with chronic low back pain
Settings: General practice and outpatient clinic
Intervention: NSAIDs

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	Number of participants (trials)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	NSAIDs			
Change in pain intensity from baseline 100 mm VAS Follow-up: 9 to 112 days	Not estimable	The mean change in pain intensity from baseline in the intervention groups was **6.97 lower** (10.74 to 3.19 lower)	-	1354 (6 trials)	⊕⊕◯◯ **low**[1,2,3]
Change in disability from baseline RDQ 0 to 24 Follow-up: 4 to 16 weeks	Not estimable	The mean change in disability from baseline in the intervention groups was **0.85 lower** (1.30 to 0.40 lower)	-	1161 (4 trials)	⊕⊕◯◯ **low**[3,4,5]
Proportion of participants experiencing adverse events Follow-up: 9 to 112 days	**Study population**		**RR 1.04** (0.92 to 1.17)	1354 (6 trials)	⊕⊕◯◯ **low**[1,2,3]
	410 per 1000	**427 per 1000** (378 to 480)			
	Moderate				
	477 per 1000	**496 per 1000** (439 to 558)			

Sensitivity analysis: change in pain intensity from baseline 100 mm VAS Follow-up: 2 to 16 weeks	Not estimable	The mean sensitivity analysis change in pain intensity from baseline. in the intervention groups was **5.03 lower** (10.37 lower to 0.32 higher)	-	728 (3 trials)	⊕⊕⊕○ **moderate**[6]
Sensitivity analysis: change in disability from baseline RDQ 0 to 24 Follow-up: 6 to 16 weeks	Not estimable	The mean sensitivity analysis change in disability from baseline in the intervention groups was **0.41 lower** (1.04 lower to 0.23 higher)	-	654 (2 trials)	⊕⊕⊕○ **moderate**[7]
Sensitivity analysis: proportion of participants experiencing adverse events. Follow-up ≤ 16 weeks Follow-up: 2 to 16 weeks	**Study population**		**RR 0.93** (0.81 to 1.07)	728 (3 trials)	⊕⊕⊕○ **moderate**[6]
	536 per 1000	**498 per 1000** (434 to 573)			
	Moderate				
	522 per 1000	**485 per 1000** (423 to 559)			

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **RR:** risk ratio; **RDQ:** Roland Morris Disability Questionnaire. **VAS:** Visual Analogue Scale

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

[1]Allocation concealment was uncertain in most included trials, and randomization was uncertain in half of the included trials, therefore selection bias is likely. Five out of six trials had high drop-out rates, so attrition bias is likely, one level downgrade.
[2]Two out of six trials allowed co-interventions. Two trials included a 'flare design', one level downgrade.

[3]See funnel plot: we could not detect publication bias, no downgrade.
[4]Allocation concealment was uncertain in most included trials. All four trials had high drop-out rates, so attrition bias is highly likely, one level downgrade.
[5]One included trial allowed co-interventions. One trial included a 'flare design', one level downgrade.
[6]Allocation concealment and randomization were uncertain in all included trials, therefore selection bias is likely. Two out of three included trials had high drop-out rates, so attrition bias is likely, one level downgrade.
[7]Allocation concealment and randomization was uncertain in both trials, therefore selection bias is likely. Both trials had high drop-out rates, so attrition bias is likely, one level downgrade.

BACKGROUND

Description of the condition

Low back pain is a major health problem and has a reported lifetime prevalence of up to 84% (Cassidy 1998; Walker 2000). More than one quarter of North Americans have reported to have experienced low back pain within the previous three months (Deyo 2006) and low back pain is a leading cause of years lived with disability (Vos 2012). In the first three months, a large proportion of patients will recover, but most people still experience pain after one year (Itz 2013). Chronic low back pain is associated with more disability and these people make a great demand on the healthcare system (Webb 2003). Also, low back pain is the most common type of pain in people experiencing any chronic pain (Müller-Schwefe 2011a) and people with chronic low back pain use healthcare more compared to people with acute low back pain (Müller-Schwefe 2011b). For treatment, guidelines recommend staying active and exercising, if necessary with the use of analgesics. Non-steroidal anti-inflammatory drugs (NSAIDs) are one of the most frequently used analgesics in low back pain management (Gore 2012; Piccoliori 2013). People with acute low back pain can receive NSAIDs for their pain, and short term NSAID use is recommended for pain relief in people with chronic back pain (Airaksinen 2006).

Description of the intervention

Most guidelines on treatment of low back pain recommend using paracetamol as first choice, followed by NSAIDs if paracetamol is insufficient (Koes 2010). NSAIDs are widely available in several types and brands and both over-the-counter and on prescription. NSAID treatment is based on the analgesic and anti-inflammatory mechanisms of the drug, but is also associated with adverse events, such as gastro-intestinal (Sostres 2013; Wehling 2014) and cardiovascular events (Kearney 2006).

How the intervention might work

Cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2) are key enzymes in prostaglandin synthesis, which contribute to inflammation, pain and fever. NSAIDs inhibit the COX enzymes and can therefore inhibit the production of prostaglandins. Consequently this can reduce inflammation, pain and fever. COX-1 produces prostaglandins that also support platelets and protect the stomach lining. It also helps to maintain kidney function. COX-1 inhibition can raise the risk of renal insufficiency and gastro-intestinal adverse events, such as gastritis or stomach bleeding.
There are two types of NSAIDs: non-selective NSAIDs, which inhibit both COX-1 and COX-2 enzymes, and selective NSAIDs, which inhibit only the COX-2 enzyme. Both selective and non-selective NSAIDs are available for pain treatment, and the choice of NSAID is mostly based on the different possible known adverse events, convenience of use, and cost.
Non-selective or traditional NSAIDs have a higher risk compared to selective NSAIDs regarding gastro-intestinal adverse events (Sostres 2013) due to the inhibition of both COX enzymes. However, aside from these gastro-intestinal benefits of selective NSAIDs, there is a known cardiovascular risk from use of these NSAID types. Cardiovascular risks are also present in non-selective NSAIDs and should be taken into account when prescribing any NSAIDs (CNT Collaboration 2013; Trelle 2011).

Why it is important to do this review

This Cochrane review is one of a series of Cochrane reviews of NSAIDs for people with low back pain and is an update of a Cochrane review first published in 2008 (Roelofs 2008). The original review consisted of 65 randomized controlled trials (RCTs); for this update we decided to create a series of Cochrane reviews regarding NSAID use for acute back pain, chronic back pain and sciatica. Also, efficacy of treatment with NSAIDs can differ among these different types of back pain. This Cochrane review focuses on NSAIDs for treating people with chronic low back pain.

OBJECTIVES

To determine if NSAIDs are more efficacious than various comparison treatments for non-specific chronic low back pain and if so, which type of NSAID is most efficacious.

METHODS

Criteria for considering studies for this review

Types of studies

We included double-blinded and single-blinded randomized controlled trials (RCTs). We only included English, German or Dutch trials, as we had stated in the original Cochrane protocol.

Types of participants

We included participants aged 18 years or older, who were treated for non-specific chronic low back pain. We defined chronic low back pain as pain for at least 12 weeks. If the trial did not describe the duration of back pain, but labeled back pain as chronic, we included the trial. If a trial included mixed populations of acute, sub-acute or chronic low back pain, we only included these trials if

they presented chronic low back pain data separately. We excluded participants with sciatica or with specific low back pain caused by pathological entities, such as infection, neoplasm, metastasis, osteoporosis, rheumatoid arthritis or fractures.

Types of interventions

We included RCTs that assessed one or more types of NSAIDs. We permitted additional interventions if there was a contrast for NSAIDs in the trial. For example, we included trials that compared NSAIDs plus muscle relaxants versus muscle relaxants alone, but excluded trials that compared NSAIDs plus muscle relaxants versus paracetamol.

We excluded trials that used NSAIDs which are no longer available on the market, such as rofecoxib.

Types of outcome measures

Primary outcomes

Primary outcome measures were:

- pain intensity (e.g. visual analogue scale (VAS) or Numerical Rating Scale (NRS))
- global measure (e.g. overall improvement, proportion of participants that recover)
- back pain-specific functional status (e.g. Roland Disability Questionnaire, Oswestry Scale)
- return to work (e.g. return to work status, number of days off work)
- adverse events (proportion of participants experiencing adverse events)

Secondary outcomes

Secondary outcome measures were physiological outcomes (e.g. range of motion, spinal flexibility, degrees of straight leg raising or muscle strength) and generic functional status (e.g. Short Form 36 (SF-36), Nottingham Health Profile, Sickness Impact Profile). We also considered other symptoms, such as health care consumption.

Search methods for identification of studies

Electronic searches

We identified RCTs for inclusion by searching the following databases up to 24 June 2015:

- Cochrane Central Register of Controlled Trials (CENTRAL, the Cochrane Library, Issue 5 of 12, May 2015)
- MEDLINE (OvidSP, 1946 to June Week 2 2015)
- MEDLINE In-Process & Other Non-Indexed Citations (OvidSP, June 23, 2015)
- EMBASE (OvidSP, 1980 to 2015 Week 25)
- ClinicalTrials.gov
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP)
- PubMed

For this update, we conducted the literature searches annually between May 2012 and 24 June 2015. We added the trial registers (clinicaltrials.gov and WHO ICTRP) in 2013, MEDLINE In-Process & Other Non-Indexed Citations in 2014 and PubMed in 2015 to identify studies not in MEDLINE using the strategy recommended by Duffy 2014. We have presented the search strategies in Appendix 1, Appendix 2, Appendix 3, Appendix 4 and Appendix 5.

A research librarian from the Cochrane Back and Neck Review Group devised and performed these searches according to the guidelines of the Cochrane Back and Neck Review Group (Furlan 2009).

Searching other resources

After the electronic search, we screened systematic reviews regarding NSAIDs for chronic low back pain. We included articles that we had included in the previous version of this Cochrane review (Roelofs 2008).

Data collection and analysis

Selection of studies

Two review authors (BK and PR, or PR and WE) independently screened all search results. We excluded clearly ineligible studies based on title and abstract. We retrieved full-text articles of all remaining studies and two review authors screened these articles independently for inclusion. We resolved any disagreements regarding inclusion by consensus between the review authors.

Data extraction and management

One review author, WE, extracted the data, and a second review author, PR, checked the extracted data. The review authors extracted data on type and dose of NSAIDs, type of reference treatment, follow-up time, duration of current symptoms and the outcomes described above. If data were unavailable for data extraction due to use of a different format, we contacted the trial authors for further information. We resolved any disagreements through consensus between all review authors.

Assessment of risk of bias in included studies

Two review authors (WE and PR) independently evaluated the risk of bias of all included trials, using the criteria list recommended by

the Cochrane Back Review Group (Furlan 2009) and described in Appendix 6. We scored each of the criteria as either 'low', 'high' or 'unclear' risk. If we scored the criteria as unclear, we did not contact the trial authors for further information. We resolved any disagreements by consensus and consulted a third review author if disagreements persisted.

Measures of treatment effect

The primary outcome, pain intensity, is measured with the VAS or NRS on a scale from 0 to 100 and 0 to 10 respectively. Global improvement is measured by the proportion of participants that recovered. Disability is measured on different disability scales, (e.g. Roland Morris Disability Questionnaire (RDQ) on a 0 to 24 scale). Adverse events are measured by the proportion of participants experiencing any adverse event.

Dealing with missing data

We did not include data in this review that were not reported in the article and that we considered missing. If trials showed data in graphs instead of describing data in the text but were shown in graphs, we collected data from the graphs.

Assessment of heterogeneity

We assessed clinical heterogeneity for all included RCTs that reported similar outcomes. We judged the included trials based on setting, participants and intervention. If trials were clinically heterogeneous, we did not pool them. We assessed statistical heterogeneity using the Chi² test and I² statistic. If I² statistic values were greater than 50%, substantial heterogeneity could be present (Higgins 2011) and we pooled data using a random-effects model. When we suspected no, low or moderate heterogeneity, we used a fixed-effect model.

Assessment of reporting biases

We used funnel plots to investigate reporting bias when we included at least four trials in a particular comparison.

Data synthesis

We analysed dichotomous outcomes by calculating the relative risk (RR). We analysed continuous outcomes by calculating the mean difference (MD) when the same instrument was used to measure outcomes, or the standardized mean difference (SMD) when different instruments were used to measure the outcomes. We expressed uncertainty with 95% confidence intervals (95% CIs). We performed a meta-analysis if studies were clinically homogeneous (comparable population, intervention and outcomes among trials) using a fixed-effect model unless there was significant statistical heterogeneity, in which case we used a random-effects model. We used the I² and chi² test to assess statistical heterogeneity as suggested in the Cochrane handbook (Higgins 2011). If meta-analysis was not possible, we described the results from clinically comparable trials in the review text.

We assessed the overall quality of the evidence for each outcome using the GRADE approach, as recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011) and adapted in the updated CBRG method guidelines (Furlan 2009). Five factors that may have decreased the quality of the evidence were: study design and risk of bias, inconsistency of results, indirectness (not generalizable), imprecision (sparse data) and other factors (e.g. reporting bias). We downgraded the quality of the evidence for a specific outcome by one level according to the performance of the studies against each of these five factors. We assessed the overall quality of the evidence for each outcome as:

- **High quality evidence:** there are consistent findings among at least 75% of RCTs with low risk of bias, consistent, direct and precise data and no known or suspected publication biases. Further research is unlikely to change either the estimate or our confidence in the results.
- **Moderate quality evidence:** one of the domains is not met. Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- **Low quality evidence:** two of the domains are not met. 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 quality evidence:** three of the domains are not met. We are very uncertain about the results.
- **No evidence:** we did not identify any RCTs that addressed this outcome.

Subgroup analysis and investigation of heterogeneity

We performed subgroup analyses if both non-selective and selective NSAIDs were present. We split these results into non-selective and selective NSAIDs.

Sensitivity analysis

We performed a sensitivity analysis on the comparison between NSAIDs and placebo. We excluded trials at high risk of bias (less than six positive items on the 'Risk of bias' table) or trials with a 'flare design' from this analysis. A trial with a 'flare design' only includes participants who previously used NSAIDs and reported aggravated back complaints during a washout period.

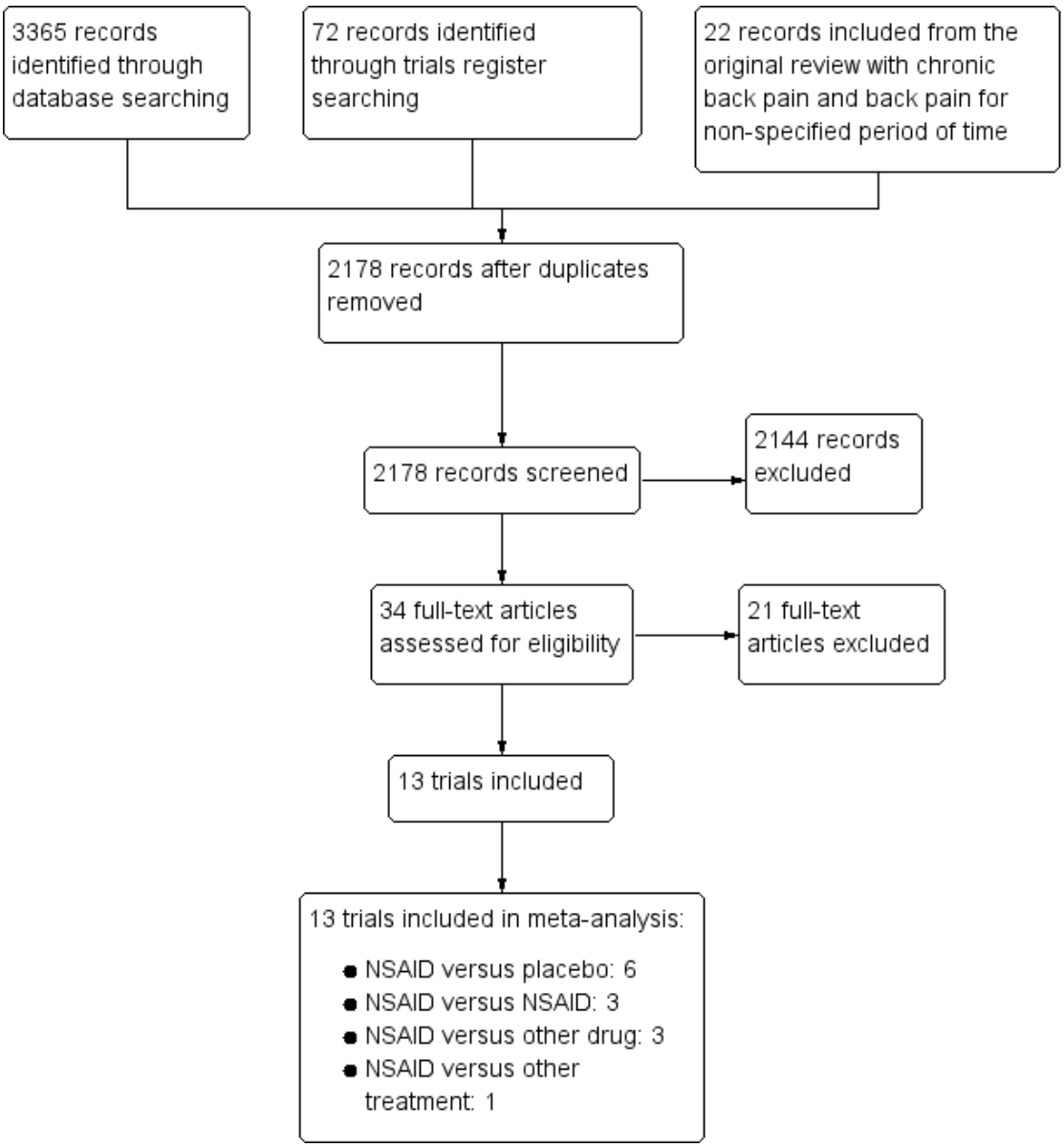
RESULTS

Description of studies

Results of the search

We identified a total of 3437 potential articles in the updated electronic search (Figure 1). After screening the titles and abstracts, we assessed full-text articles and included 13 trials. Amongst these were seven of the nine articles on chronic low back pain from Roelofs 2008. Two trials reported on rofecoxib, which was withdrawn from the market, and we excluded these trials from this review (Chrubasik 2003; Katz 2003).

Figure 1. study flow diagram.



Included studies

The sample size of the 13 included trials ranged from 28 to 1593 participants, with a total of 4807 included participants. Six trials compared NSAIDs versus placebo (Allegrini 2009; Berry 1982; Birbara 2003; Coats 2004; Katz 2011; Kivitz 2013). Three trials compared two different types of NSAIDs (Driessens 1994; Videman 1984; Zerbini 2005). One trial compared NSAIDs versus paracetamol (Hickey 1982), one trial compared NSAIDs versus tramadol (O'Donnell 2009) and one trial compared NSAIDs versus pregabalin (Romanò 2009). One trial compared exercise therapy versus NSAIDs (Shirado 2010).

Excluded studies

We have described the reasons for exclusion of studies in the 'Characteristics of excluded studies' table. We excluded most studies because it was unclear whether participants had chronic low back pain.

Risk of bias in included studies

We have presented the 'Risk of bias' assessment in Figure 2 and Figure 3. Ten of the 13 studies were considered having a low risk of bias. (Berry 1982; Birbara 2003; Coats 2004; Driessens 1994; Hickey 1982; Katz 2011; Kivitz 2013; O'Donnell 2009; Shirado 2010; Zerbini 2005).

Figure 2. 'Risk of bias' summary: review authors' judgements about each 'Risk of bias' item for each included trial.

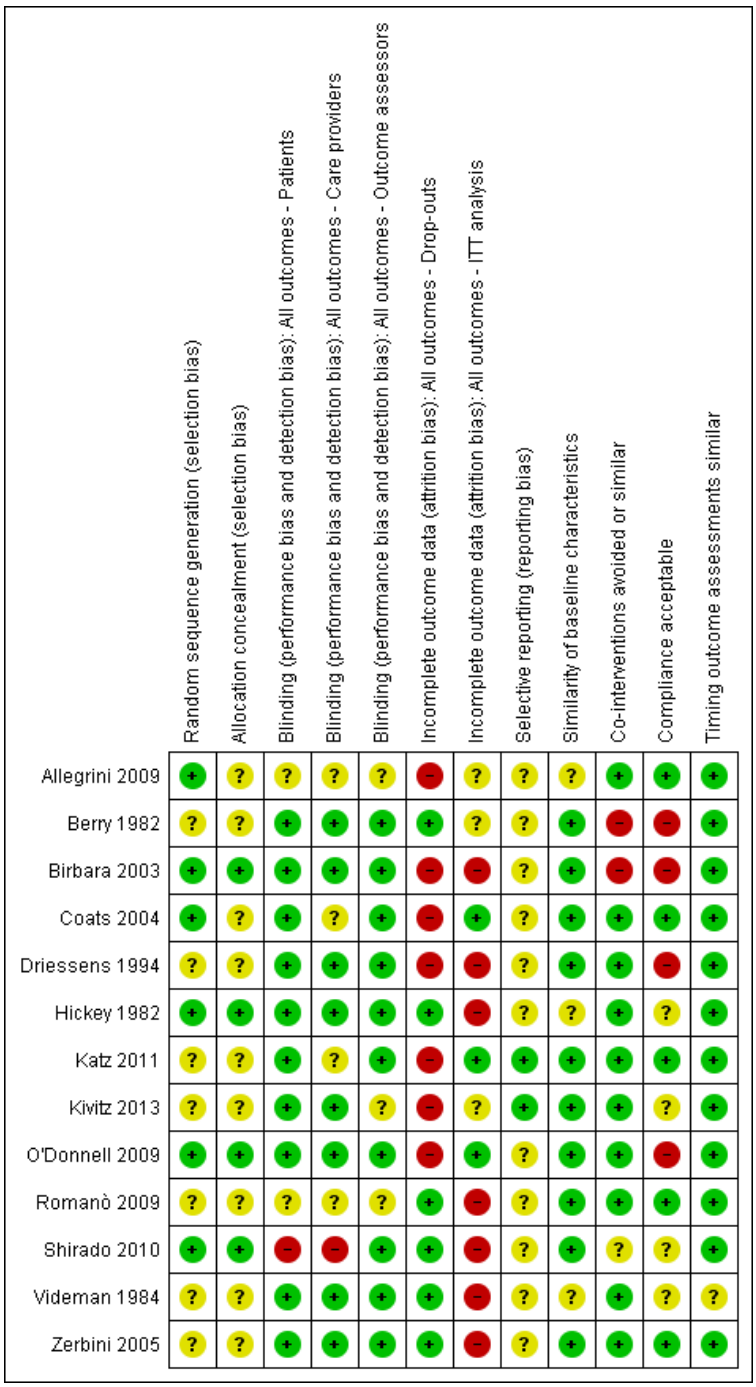
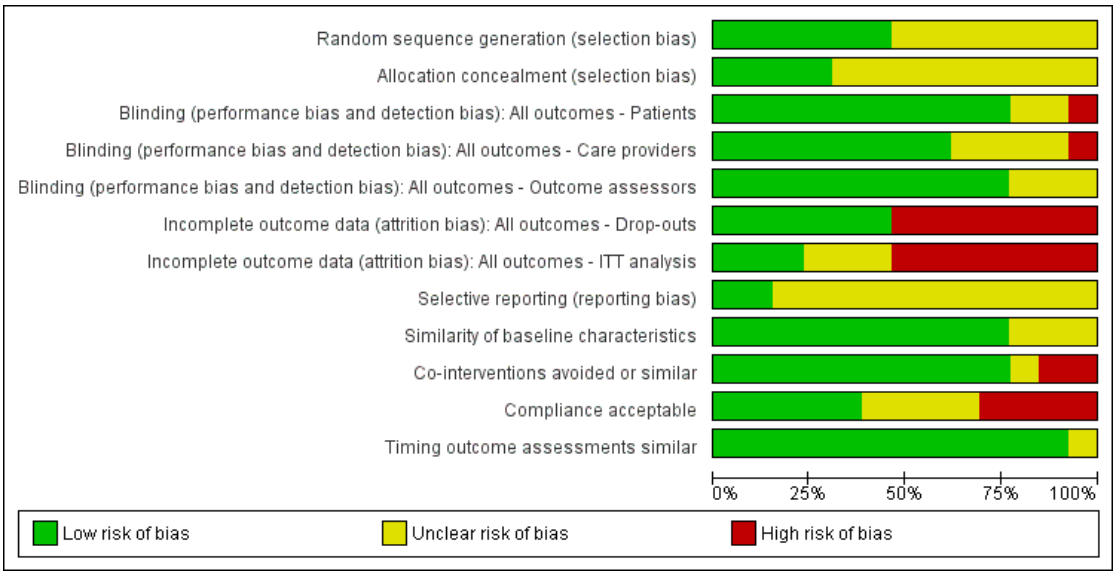


Figure 3. 'Risk of bias' graph: review authors' judgements about each 'Risk of bias' item presented as percentages across all included trials.



Allocation

Of the 13 included studies, six reported a randomization procedure (Allegrini 2009; Birbara 2003; Coats 2004; Hickey 1982; O'Donnell 2009; Shirado 2010). Of these six studies, only four also adequately described concealment of treatment allocation (Birbara 2003; Hickey 1982; O'Donnell 2009; Shirado 2010). Most studies did not report the method of randomization or allocation concealment and were scored as 'unclear' on these items.

Blinding

Seven included trials reported blinding of patients, care providers and outcome assessors (Berry 1982; Birbara 2003; Driessens 1994; Hickey 1982; O'Donnell 2009; Videman 1984; Zerbini 2005). The other six trials did not blind patients, care providers, or outcome assessors or they did not report on blinding.

Incomplete outcome data

Six trials reported low drop out rates (Berry 1982; Hickey 1982; Romanò 2009; Shirado 2010; Videman 1984; Zerbini 2005). The seven other studies reported drop-out rates higher than 20% (Allegrini 2009; Birbara 2003; Coats 2004; Driessens 1994; Katz 2011; Kivitz 2013; O'Donnell 2009).

Only three trials performed an intention to treat (ITT) analysis (Coats 2004; Katz 2011; O'Donnell 2009).

Selective reporting

Only two RCTs were registered in an accessible clinical trial registry (Katz 2011; Kivitz 2013) and had low risk of reporting bias.

Other potential sources of bias

Most studies showed similarity of baseline characteristics; only three RCTs did not report this (Allegrini 2009; Hickey 1982; Videman 1984).

Regarding co-interventions, only paracetamol as rescue medication was allowed; other types of medication were not. All but two trials avoided co-interventions (Berry 1982; Birbara 2003) and one trial did not state anything about co-interventions (Shirado 2010).

Nine trials reported compliance, and five trials had acceptable compliance (Allegrini 2009; Coats 2004; Katz 2011; Romanò 2009; Zerbini 2005). Four other trials had unacceptable compliance (Berry 1982; Birbara 2003; Driessens 1994; O'Donnell 2009).

Timing of outcome assessment was similar between the groups in almost all included trials.

We created funnel plots to assess risk of publication bias and for the analysis of NSAIDs versus placebo (Figure 4; Figure 5; Figure 6). We could not identify publication bias. We did not create any funnel plots for other comparisons, since less than four RCTs were available for this analysis.

Figure 4. Funnel plot of comparison: 1 NSAIDs versus placebo, outcome: 1.1 Change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 12 weeks.

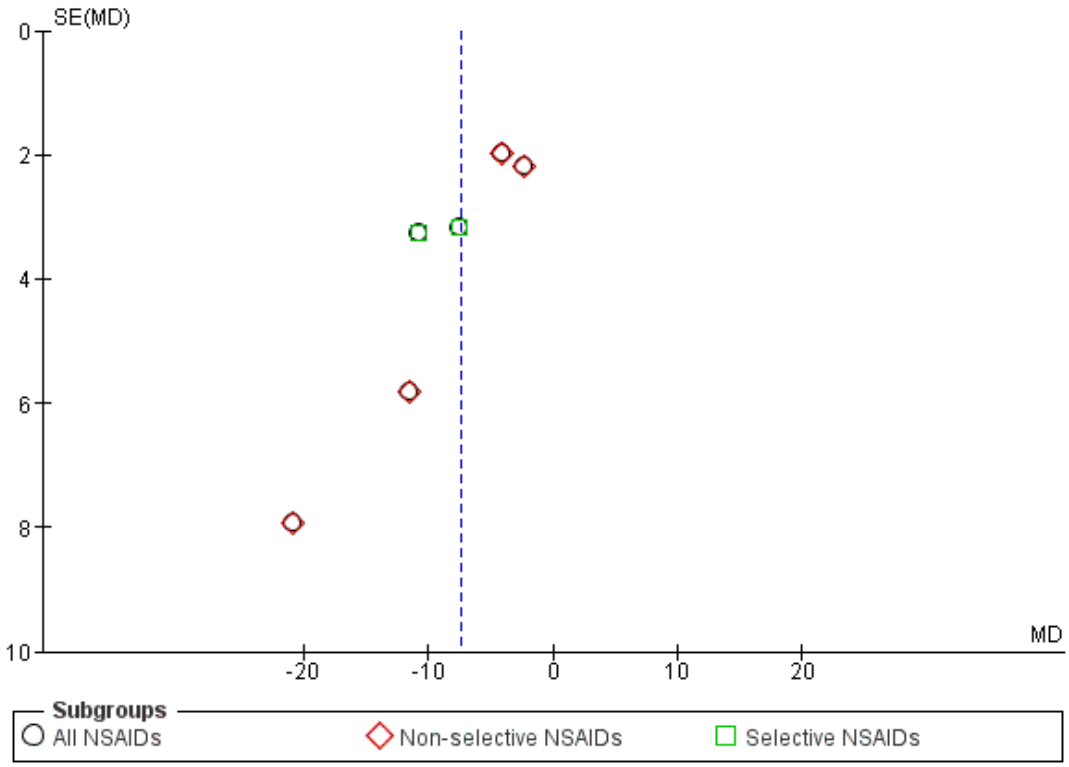


Figure 5. Funnel plot of comparison: 1 NSAIDs versus placebo, outcome: 1.2 Change in disability from baseline.

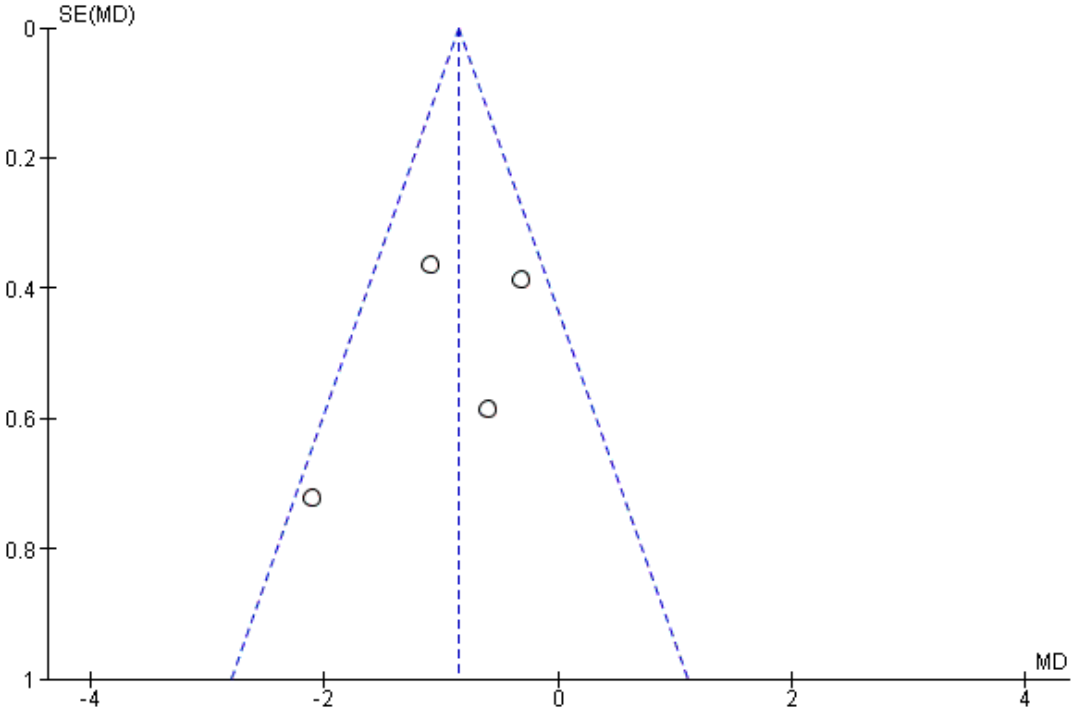
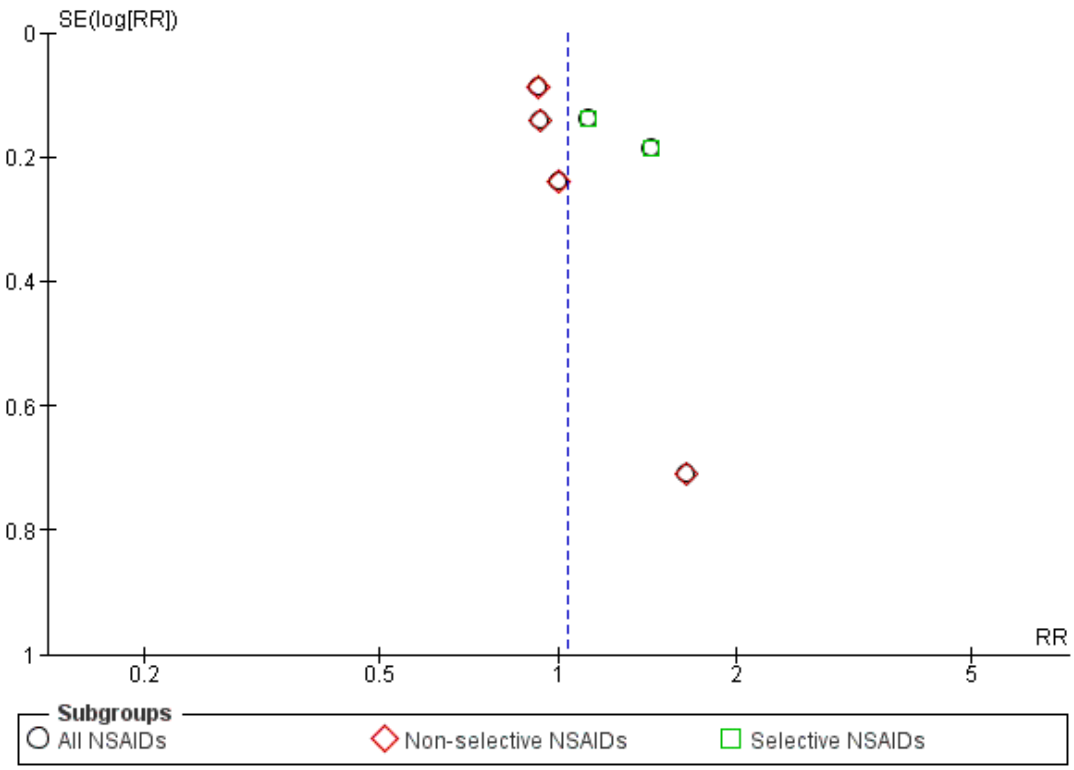


Figure 6. Funnel plot of comparison: 1 NSAIDs versus placebo, outcome: 1.3 Proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks.



Half of the included trials reported a potential conflict of interest. Three studies reported support from a pharmaceutical company (Birbara 2003; Hickey 1982; Zerbini 2005) and the authors of four RCTs had affiliations with a pharmaceutical company (Coats 2004; Katz 2011; Kivitz 2013; O'Donnell 2009). The remaining six RCTs did not report any potential conflict of interest.

Effects of interventions

See: **Summary of findings for the main comparison** NSAIDs for people with chronic low back pain
See: 'Summary of findings' table 1.

Efficacy of NSAIDs compared to placebo

Six RCTs compared NSAIDs with placebo (Allegrini 2009; Berry 1982; Birbara 2003; Coats 2004; Katz 2011; Kivitz 2013). Median follow-up was 56 days (IQR 13 to 91 days). Three of these trials reported short-term outcomes of four weeks or less (Allegrini 2009; Berry 1982; Coats 2004). The other three trials had a duration of follow-up of 12 or 16 weeks (Birbara 2003; Katz 2011; Kivitz 2013). Naproxen was the most common type of NSAID (Berry 1982; Katz 2011; Kivitz 2013), but piroxicam patch, etoricoxib and valdecoxib were also compared to placebo.

All RCTs reported pain intensity on a 100 mm VAS or 11-point numerical rating scale (NRS). The Chi² value for homogeneity of the mean difference (MD) was 10.41 (P 0.06) and I² statistic 52%, which suggests substantial statistical heterogeneity. This might be due to different types of NSAIDs used in the trials and we used a random-effects model to pool these data. The pooled mean difference in pain intensity score from baseline was −6.97 (95% CI −10.74 to −3.19; Analysis 1.1), indicating a statistically significant effect in favour of participants receiving NSAIDs compared to participants receiving placebo. The quality of this evidence was low (Summary of findings for the main comparison). When we split results into selective and non-selective NSAIDs versus placebo, there was still a substantial statistical heterogeneity among the trials considering non-selective NSAIDs, although three out of four RCTs used naproxen as trial medication. There was statistical homogeneity among the trials on selective NSAIDs. The effect of selective NSAIDs was somewhat larger and the effect of non-selective NSAIDs was smaller.

Four RCTs compared NSAIDs with placebo, with disability as outcome measure, measured with the Roland Morris Disability Questionnaire (RDQ) (Birbara 2003; Coats 2004; Katz 2011; Kivitz 2013) on a 0 to 24 scale. Median follow-up was 84 days (IQR 42 to 105 days). The Chi² value for homogeneity of the mean difference (MD) was 5.53 (P = 0.14) and the I² statistic was 46%, indicating moderate statistical heterogeneity among these trials. The pooled mean difference in disability from baseline was −0.85 (95% CI −1.30 to −0.40; Analysis 1.2). The quality of this evidence was low ('Summary of findings' table 1).

All trials also reported adverse events. The Chi² value for homogeneity of the RR for adverse events in all RCTs was 6.22 (P = 0.28) and the I² statistic value was 20%, indicating no statistical heterogeneity among the RCTs. The pooled RR for adverse events was 1.04 (95% CI 0.92 to 1.17; Analysis 1.3), indicating that adverse events were not statistically significant more present in participants using NSAIDs compared to placebo. Using the GRADE approach, we assessed the quality of evidence of these trials as low ('Summary of findings' table 1). Results did not change when we specified NSAIDs into selective and non-selective NSAIDs, although adverse events in selective NSAIDs showed a trend in favour of placebo. However, RCTs have low power in detecting uncommon and delayed adverse events. The sample sizes of most included trials were relatively small and duration of follow-up was relatively short. It is possible that not all adverse events had emerged, especially since most important adverse events are rare and can take weeks or months to present. Therefore, we cannot make firm statements about the difference in occurrence of adverse events between different NSAID types.

Of the trials that compared NSAIDs with placebo, we considered three trials at high risk of bias (Allegrini 2009; Birbara 2003; Coats 2004). The latter two trials used a 'flare design'. We performed a sensitivity analysis using the three RCTs which were at low risk of bias (Berry 1982; Katz 2011; Kivitz 2013). The difference between NSAIDs and placebo on pain intensity score (on 0 to 100 mm VAS) and the disability (measured with RDQ 0 to 24) became smaller and was no longer statistically significant; the difference in pain intensity score between NSAIDs and placebo was −5.03 (95% CI −10.37 to 0.32; Analysis 1.4) and for disability was −0.41 (95% CI −1.04 to 0.23; Analysis 1.5). We assessed the quality of evidence as moderate ('Summary of findings' table 1).

Efficacy of selective versus non-selective NSAIDs and non-selective versus non-selective NSAIDs

Two small RCTs compared two types of non-selective NSAIDs (Driessens 1994; Videman 1984). Driessens 1994 compared ibuprofen (1600 mg/day) and diclofenac (100 mg/day) for two weeks, Videman 1984 compared piroxicam (20 mg/day) and indomethacin (75 mg/day) for six weeks. Both trials found no significant difference between the two types of non-selective NSAIDs. The number of adverse events in Driessens 1994 was statistically significant higher in the diclofenac group. In Videman 1984 there was no statistically significant difference in experienced adverse events between the two trial groups. One other RCT, Zerbini 2005, compared a non-selective NSAID with a COX-2 inhibitor (diclofenac 150 mg/day versus etoricoxib 60 mg/day for four weeks). This trial included 440 participants in the analysis and found no significant difference in change in pain intensity from baseline between the non-selective NSAIDs and COX-2 inhibitor. The trial also did not find any differences in adverse events in general and specific gastrointestinal adverse events between the two trial groups.

Efficacy of NSAIDs versus other drugs

NSAIDs compared to other drug types are shown in Analysis 2.1 and Analysis 2.2. We did not pool these RCTs because the trials used different types of medication as comparison. Hickey 1982, which had with 30 participants, compared NSAIDs (diflunisal 1000 mg/day) with paracetamol (4000 mg/day). In this trial, NSAIDs were not significantly better than paracetamol and adverse events were not significantly more present in patients using NSAIDs compared to the other studied drugs.

O'Donnell 2009 included 1593 participants and compared NSAIDs (celecoxib 400 mg/day) with tramadol (200 mg/day) for six weeks. Results of global improvement (RR 1.26, 95% CI 1.16 to 1.38) and adverse events (RR 0.83, 95% CI 0.75 to 0.91) after six weeks both favoured celecoxib.

Romanò 2009 compared celecoxib with pregabalin and scored change in pain intensity from baseline to four weeks on a VAS score. There was no significant difference between the two trial groups and adverse events were similar in number in both celecoxib and pregabalin trial groups.

Efficacy of NSAIDs versus non-drug treatment

One RCT, Shirado 2010, compared NSAIDs with 'home-based exercise'. Improvement in functional status between baseline and eight weeks was significantly better in exercise participants then participants receiving NSAIDs, but there was no difference in pain intensity.

DISCUSSION

Summary of main results

In this Cochrane review we included 13 RCTs that assessed NSAID efficacy for the management of chronic low back pain. Six trials comparing NSAIDs with placebo showed low quality evidence that NSAIDs are more effective than placebo, with a mean difference in pain intensity score from baseline of -6.97 (95% CI

−10.74 to −3.19) on a 0 to 100 visual analogue scale (VAS) with a median follow-up of 56 days (IQR 13 to 91 days). There is also low quality evidence that NSAIDs are more effective than placebo on disability, with a mean difference from baseline of −0.85 (95% CI −1.30 to −0.40) on a scale from 0 to 24 with a median follow-up of 84 days (IQR 42 to 105 days). When only trials with low risk of bias were included in the analysis, the difference between NSAIDs and placebo was no longer significant. Adverse events were not significantly more present in the NSAIDs or placebo trial group, but this could be because we only included RCTs in this review, or the short duration of use and the short follow-up period in most included trials.

Studies comparing non-selective versus selective NSAIDs or comparing different types of non-selective NSAIDs were also limited available. All three included RCTs showed no significant effect between the different NSAID types.

Whether NSAIDs are more effective than other drugs or non-drug therapies for people with chronic low back pain remains unclear. A limited number of trials compared NSAIDs versus other drug treatments and all trials included different kind of drugs as comparator. One large RCT compared celecoxib to tramadol. Results of global improvement and adverse events were both in favour of celecoxib after six weeks.

Overall completeness and applicability of evidence

In this Cochrane review we used strict inclusion criteria regarding the duration of back pain, meaning that we only included trials that reported results on people with chronic low back pain. This means that fewer trials met the inclusion criteria of this Cochrane review, but it makes the review results more distinct for people with non-specific chronic low back pain.

Two included trials used a 'flare design'. These trials included participants who responded well to NSAIDs when they showed a worsening in back pain during a wash-out period. As these participants already responded well to NSAIDs, these trials are likely to have overestimated the effect of NSAIDs. It may also reduce the external validity since this is a select group of participants. When we excluded these RCTs from the analysis together with one other trial with a high risk of bias, the results changed. The magnitude of effect of NSAIDs became smaller and the difference was not statistically significant anymore.

Some included trials operationalized outcomes differently and not all trials included disability as outcome. None of the included trials mentioned return to work or other work outcomes, although this might be an important outcome in patients with chronic low back pain.

Almost all included RCTs mentioned adverse events. Most trials reported the overall number of adverse events, and some trials also mentioned specific gastrointestinal adverse events. Cardiovascular adverse events are rarely mentioned. However, these trials were powered to investigate treatment effects of the primary outcomes. As most important adverse events are rare and can take weeks or months to evolve, it is likely that sample sizes of these trials were too small and follow-up periods too short to draw clear conclusions from these trials regarding the risks for gastrointestinal and other adverse events of NSAIDs.

Quality of the evidence

Three included RCTs were considered high risk of bias. Even in the 10 other RCTs with low risk of bias other methodological shortcomings were present, such as no clear description of the randomization procedure, high drop-out rates and low or unclear compliance in the trial groups. Uncertain or low compliance makes it difficult to interpret the measured effect in the study and can both under- and overestimate the results found. The level of evidence, which we assessed using the GRADE approach, was low due to similar issues. The most common reasons for downgrading evidence were 'risk of bias' and 'imprecision' for the included trials.

Most trials had a follow-up period of at least four weeks, and only three trials had follow-up periods of less than four weeks (ranging from nine days to two weeks). NSAIDs are usually used for a short period of time. This short follow up period might not have consequences on our results, since effects are expected shortly after the start of the NSAIDs. Although it is difficult due to this short follow-up period to assess adverse events.

Included RCTs had different trial population sizes; four trials included less than 50 participants and may lack statistical power to detect differences in effects. Pooling may overcome this problem. However, the most important question is whether the effect is clinically relevant. The main finding that NSAIDs are more effective than placebo on pain intensity was based on a meta-analysis that showed a mean difference of 3.30 on a 0 to 100 scale. Although statistically significant, one could argue that this effect is too small to be clinically relevant.

A sensitivity analysis with a moderate quality of evidence showed that the positive effect of NSAIDs compared to placebo was reduced and no longer statistically significant when we only included RCTs in the analysis that were of low risk of bias.

Potential biases in the review process

We only included trials published in English, German or Dutch, which could have led to the exclusion of trials published in other languages from this Cochrane review. Reports on language bias show conflicting results (Higgins 2011; Jüni 2002; Moher 2003). It is not to be expected that inclusion of articles written in other languages will change the results in this review, especially since there seems to be a shift in publishing more articles in English and less frequent in other languages (Galandi 2006; Higgins 2011).

Only one review author extracted data and the second review author checked the extracted data. This could have led to a higher risk of error in data extraction.

Different types and chemical entities of NSAIDs are available, which makes it difficult to compare different NSAIDs. Regarding the comparison of NSAIDs versus placebo, we included both selective and non-selective NSAIDs. An analysis of two separate comparisons showed no differences in directions of the findings when we compared selective and non-selective NSAIDs separately with placebo.

Publication bias may have occurred, but this was difficult to assess due to the limited number of included trials. In particular the comparisons of different NSAID types or NSAIDs compared to other types of drugs we could not examine publication bias using a funnel plot. Half of the included trials were supported by or included authors from pharmaceutical companies. Clinical trials sponsored by pharmaceutical companies are less likely to be published and are more likely to have outcomes in favour of the sponsor (Lexchin 2003), which could have caused publication bias. Even when publication bias would have occurred, this will not change the found results. The found effect is already very small and not clinically relevant.

Agreements and disagreements with other studies or reviews

In the previous version of this Cochrane review, Roelofs 2008, we studied NSAIDs for people with sciatica, acute and chronic low back pain based on literature published from September 1998 to June 2007. These trials found a change in pain intensity in favour of NSAIDs compared to placebo. In this review update we found similar results, but the magnitude of the results in our review was smaller than found in Roelofs 2008. Adverse events were statistically more present in the NSAID group in Roelofs 2008, but we did not find a statistically significant difference in our review. Most trials included in this Cochrane review had a small sample size or short-time follow-up, or both, and were not suited to evaluate adverse events. A large meta-analysis on adverse events in RCTs (CNT Collaboration 2013) and observational data (Castellsague 2012) showed that adverse events are more present in participants using NSAIDs compared to placebo.

After 2008, several systematic reviews were published regarding NSAIDs as a therapeutic option in treating people with chronic low back pain. Pain scores between NSAIDs and placebo were often reported. In 2013, a review on NSAIDs showed that COX-2 selective NSAIDs were significantly more effective in reducing VAS score and disability measured with RDQ (Chung 2013). Four studies were included in Chung's analysis, of which we did not include two in this Cochrane review. We excluded one trial, Pallay 2004, from the previous version of this review because it is additional information to an earlier reported study that was already included in the review (Birbara 2003). Including both would lead to double counting. The other study, Katz 2003, reported on rofecoxib and was excluded from this review because it was withdrawn from the market. Kuijpers 2011 found similar results to Chung 2013 and concluded that there is low quality evidence that NSAIDs are more effective than placebo. This is comparable to findings in this Cochrane review. Chung 2013 also assessed disability and results were comparable to our findings..

Chung 2013 also evaluated selective and non-selective NSAIDs and found no differences in efficacy between these two groups. Two studies were analysed in the review; one of those was also examined in this review and found the same results. We excluded the other study used in Chung 2013 from this Cochrane review because it included rofecoxib.

AUTHORS' CONCLUSIONS

Implications for practice

For people with chronic low back pain there is low quality evidence that NSAIDs are slightly better in reducing pain and disability than placebo, but the effect is very small and possibly not clinically relevant. The low risk of bias studies showed no significant difference between NSAIDs and placebo. It is unclear whether NSAIDs are more effective than other drugs and there is no evidence to show that one NSAID type is more effective than other types.

Implications for research

The quality of evidence for NSAIDs compared to placebo in people with chronic low back pain is, at best, moderate. When studies are of higher quality, effects of NSAIDs become smaller or disappear. It is questionable whether or not additional research will change these findings and the estimate of effect. Especially since the observed differences in this study between NSAIDs and placebo are small and possibly not clinically relevant. In studies with flare designs, some participants respond to NSAID treatment. Therefore, it might be worthwhile to look into subgroups finding participants who are likely to respond well to NSAIDs.

ACKNOWLEDGEMENTS

We thank Shireen Harbin and Rachel Couban, Trials Search Coordinators of the Cochrane Back Review Group, who updated the literature searches. We also thank Rob Scholten for his contributions to the original Cochrane review (Roelofs 2008).

REFERENCES

References to studies included in this review

Allegrini 2009 *{published data only}*
Allegrini A, Nuzzo L, Pavone D, Tavella-Scaringi A, Giangreco D, Bucci M, et al. Efficacy and safety of piroxicam patch versus piroxicam cream in patients with lumbar osteoarthritis: A randomized, placebo-controlled study. *Arzneimittel-Forschung/Drug Research* 2009;**59**(8): 403–9.

Berry 1982 *{published data only}*
Berry H, Bloom B, Hamilton EBD, Swinson DR. Naproxen sodium, diflunisal, and placebo in the treatment of chronic back pain. *Annals of the Rheumatic Diseases* 1982;**41**(2): 129–32.

Birbara 2003 *{published data only}*
Birbara CA, Puopolo AD, Munoz DR, Sheldon EA, Mangione A, Bohidar NR, et al. Treatment of chronic low back pain with etoricoxib, a new cyclo-oxygenase-2 selective inhibitor: improvement in pain and disability--a randomized, placebo-controlled, 3-month trial. *The Journal of Pain* 2003;**4**(6):307–15.

Coats 2004 *{published data only}*
Coats TL, Borenstein DG, Nangia NK, Brown MT. Effects of valdecoxib in the treatment of chronic low back pain: results of a randomized, placebo-controlled trial. *Clinical Therapeutics* 2004;**26**(8):1249–60.

Driessens 1994 *{published data only}*
Driessens M, Famaey JP, Orloff S, Chochrad I, Cleppe D, de Brabanter G, et al. Efficacy and tolerability of sustained-release ibuprofen in the treatment of patients with chronic back pain. *Current Therapeutic Research* 1994;**55**(11): 1283–92.

Hickey 1982 *{published data only}*
Hickey RF. Chronic low back pain: a comparison of diflunisal with paracetamol. *The New Zealand Medical Journal* 1982;**95**(707):312–4.

Katz 2011 *{published data only}*
Katz N, Borenstein DG, Birbara C, Bramson C, Nemeth MA, Smith MD, et al. Efficacy and safety of tanezumab in the treatment of chronic low back pain. *Pain* 2011;**152** (10):2248–58.

Kivitz 2013 *{published data only}*
Kivitz AJ, Gimbel JS, Bramson C, Nemeth MA, Keller DS, Brown MT, et al. Efficacy and safety of tanezumab versus naproxen in the treatment of chronic low back pain. *Pain* 2013;**154**(7):1009–21.

O'Donnell 2009 *{published data only}*
O'Donnell JB, Ekman EF, Spalding WM, Bhadra P, McCabe D, Berger MF. The effectiveness of a weak opioid medication versus a cyclo-oxygenase-2 (COX-2) selective non-steroidal anti-inflammatory drug in treating flare-up of chronic low-back pain: results from two randomized, double-blind, 6-week studies. *Journal of International Medical Research* 2009;**37**(6):1789–802.

Romanò 2009 *{published data only}*
Romanò CL, Romanò D, Bonora C, Mineo G. Pregabalin, celecoxib, and their combination for treatment of chronic low-back pain. *Journal of Orthopaedics and Traumatology* 2009;**10**(4):185–91.

Shirado 2010 *{published data only}*
Shirado O, Doi T, Akai M, Hoshino Y, Fujino K, Hayashi K, et al. Multicenter randomized controlled trial to evaluate the effect of home-based exercise on patients with chronic low back pain: the Japan low back pain exercise therapy study. *Spine* 2010;**35**(17):E811–9.

Videman 1984 *{published data only}*
Videman T, Osterman K. Double-blind parallel study of piroxicam versus indomethacin in the treatment of low back pain. *Annals of Clinical Research* 1984;**16**(3):156–60.

Zerbini 2005 *{published data only}*
Zerbini C, Ozturk ZE, Grifka J, Maini M, Nilganuwong S, Morales R, et al. Efficacy of etoricoxib 60 mg/day and diclofenac 150 mg/day in reduction of pain and disability in patients with chronic low back pain: results of a 4-week, multinational, randomized, double-blind study. *Current Medical Research and Opinion* 2005;**21**(12):2037–49.

References to studies excluded from this review

Aoki 1983 *{published data only}*
Aoki T, Kuroki Y, Kageyama T, Irimajiri S, Mizushima Y, Yamamoto K. Multicentre double-blind comparison of piroxicam and indomethacin in the treatment of lumbar diseases. *European Journal of Rheumatology and Inflammation* 1983;**6**(3):247–52.

Babey-Dölle 1994 *{published data only}*
Babej-Dölle R, Freytag S, Eckmeyer J, Zerle G, Schinzel S, Schmeider G, et al. Parenteral dipyrone versus diclofenac and placebo in patients with acute lumbago or sciatic pain: randomized observer-blind multicenter study. *International Journal of Clinical Pharmacology and Therapeutics* 1994;**32** (4):204–9.

Borghi 2013 *{published data only}*
Borghi B, Aurini L, White PF, Mordenti A, Lolli F, Borghi R, et al. Long-lasting beneficial effects of periradicular injection of meloxicam for treating chronic low back pain and sciatica. *Minerva Anestesiologica* 2013;**79**(4):370–8.

Chang 2013 *{published data only}*
Chang WK, Wu HL, Yang CS, Chang KY, Liu CL, Chan KH, et al. Effect on pain relief and inflammatory response following addition of tenoxicam to intravenous patient-controlled morphine analgesia: a double-blind, randomized, controlled study in patients undergoing spine fusion surgery. *Pain Medicine* 2013;**14**(5):736–48.

Chrubasik 2003 *{published data only}*
Chrubasik S, Model A, Black A, Pollak S. A randomized double-blind pilot study comparing Doloteffin and Vioxx

in the treatment of low back pain. *Rheumatology* 2003;**42** (1):141–8.

Davoli 1989 ***{published data only}***
Davoli L, Ciotti G, Biondi M, Passeri M. Piroxicam-beta-cyclodextrin in the treatment of low-back pain. Controlled study vs etodolac. *Current Therapeutic Research, Clinical and Experimental* 1989;**46**:940–7.

Famaey 1998 ***{published data only}***
Famaey JP, Bruhwyler J, Géczy J, Vandekerckhove K, Appelboom T. Open controlled randomized multicenter comparison of nimesulide and diclofenacin the treatment of subacute and chronic low back pain. *Journal of Drug Assessment* 1998;**1**:349–68.

Ingpen 1969 ***{published data only}***
Ingpen ML. A controlled clinical trial of sustained-action dextropropoxyphene hydrochloride. *British Journal of Clinical Practice* 1969;**23**(3):113–5.

Jacobs 1968 ***{published data only}***
Jacobs JH, Grayson MF. Trial of anti-inflammatory agent (indomethacin) in low back pain with and without radicular involvement. *British Medical Journal* 1968;**3**(5611): 158–60.

Jaffé 1974 ***{published data only}***
Jaffé G. A double-blind, between-patient comparison of alclofenac ('Prinalgin') and indomethacin in the treatment of low back pain and sciatica. *Current Medical Research and Opinion* 1974;**2**(7):424–9.

Katz 2003 ***{published data only}***
Katz N, Ju WD, Krupa DA, Sperling RS, Bozalis Rodgers D, Gertz BJ, et al. Efficacy and safety of rofecoxib in patients with chronic low back pain: results from two 4-week, randomized, placebo-controlled, parallel-group, double-blind trials. *Spine* 2003;**28**(9):851–8.

Listrat 1990 ***{published data only}***
Listrat V, Dougados M, Chevalier X, Kramer F, Amor B. Comparison of the analgesic effect of tenoxicam after oral or intramuscular administration. *Drug Investigation* 1990;**2**: 51–2.

Matsumo 1991 ***{published data only}***
Matsumo S, Kaneda K, Nohara Y. Clinical evaluation of ketoprofen (Orudis) in lumbago: a double blind comparison with diclofenac sodium. *British Journal of Clinical Practice* 1991;**35**(7-8):266.

Merkulova 2013 ***{published data only}***
Merkulova DM, Onsin AA, Merkulov YA. Piascledin in the treatment of chronic dorsalgia. *Zhurnal Nevrologii i Psikhiatrii imeni S.S. Korsakova* 2013;**113**(9):18–22.

Peng 2014 ***{published data only}***
Peng YL, Wu QM, Li YF, Mu CY, Hu AH, Zhang Y. Temperature-controlled self-heated pain relief plaster for chronic nonspecific lower back pain: A prospective randomized controlled trial. *Chinese Journal of Evidence-Based Medicine* 2014;**2**:147–51.

Postacchini 1988 ***{published data only}***
Postacchini F, Facchini M, Palieri P. Efficacy of various forms of conservative treatment in low back pain: a comparative study. *Neuro-Orthopaedics* 1988;**6**:28–35.

Siegmeth 1978 ***{published data only}***
Siegmeth W, Sieberer W. A comparison of the short-term effects of ibuprofen and diclofenac in spondylosis. *Journal of International Medical Research* 1978;**6**(5):369–74.

Tavafian 2014 ***{published data only}***
Tavafian SS, Jamshidi AR, Mohammad K. Treatment of low back pain: randomized clinical trial comparing a multidisciplinary group-based rehabilitation program with oral drug treatment up to 12 months. *International Journal of Rheumatic Diseases* 2014;**17**(2):159–64.

Waikakul 1995 ***{published data only}***
Waikakul S, Soparat K. Effectiveness and safety of loxoprofen compared with naproxen in nonsurgical low back pain: a parallel study. *Clinical Drug Investigation* 1995; **10**(1):59–63.

Waikakul 1996 ***{published data only}***
Waikakul S, Danputipong P, Soparat K. Topical analgesics, indomethacin plaster and diclofenac emulgel for low back pain: a parallel study. *Journal of the Medical Association of Thailand* 1996;**79**(8):486–90.

Wetzel 2014 ***{published data only}***
Wetzel L, Zadrazil M, Paternostro-Sluga T, Authried G, Kozek-Langenecker S, Scharbert G. Intravenous nonopioid analgesic drugs in chronic low back pain patients on chronic opioid treatment: a crossover, randomised, double-blinded, placebo-controlled study. *European Journal of Anaesthesiology* 2014;**31**(1):35–40.

Additional references

Airaksinen 2006
Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *European Spine Journal* 2006;**15**(Suppl 2):S192–300. [PUBMED: 16550448]

Boutron 2005
Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable.. *Journal of Clinical Epidemiology* 2005;**58**: 1220–6.

Cassidy 1998
Cassidy JD, Carroll LJ, Côté P. The Saskatchewan health and back pain survey. The prevalence of low back pain and related disability in Saskatchewan adults. *Spine (Phila Pa 1976)* 1998 Sep 1;**23**(17):1860–6. [PUBMED: 9762743]

Castellsague 2012
Castellsague J, Riera-Guardia N, Calingaert B, Varas-Lorenzo C, Fourrier-Reglat A, Nicotra F, et al. Individual NSAIDs and upper gastrointestinal complications: a systematic review and meta-analysis of observational studies (the SOS project). *Drug Safety* 2012;**35**(12):1127–46. [PUBMED: 23137151]

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. [PUBMED: 24284847]

CNT Collaboration 2013
Coxib and traditional NSAID Trialists' (CNT) Collaboration, 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. [PUBMED: 23726390]

Deyo 2006
Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine (Phila Pa 1976)* 2006;**31**(23):2724–7. [PUBMED: 17077742]

Duffy 2014
Duffy S, Misso K, Noake C, Ross J, Stirk L. Supplementary searches of PubMed to improve currency of MEDLINE and MEDLINE In-Process searches via OvidSP. Kleijnen Systematic Reviews Ltd, York. Poster presented at the UK InterTASC Information Specialists' Sub-Group (ISSG) Workshop; 9 July 2014; Exeter, UK. https://medicine.exeter.ac.uk/media/universityofexeter/medicalschool/research/pentag/documents/Steven˙Duffy˙ISSG˙Exeter˙2014˙poster˙1.pdf (accessed 06/08/2014).

Furlan 2009
Furlan AD, Pennick V, Bombardier C, van Tulder M, Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;**34**(18): 1929–41.

Galandi 2006
Galandi D, Schwarzer G, Antes G. The demise of the randomised controlled trial: bibliometric study of the German-language health care literature, 1948 to 2004. *BMC Med Res Methodol* 2006;**6**:6–30. [MEDLINE: 16824217]

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. [PUBMED: 22304678]

Higgins 2011
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

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. [PUBMED: 22641374]

Jüni 2002
Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study.. *Int J Epidemiol.* 2002;**31** (1):115–23. [MEDLINE: 11914306]

Kearney 2006
Kearney P, Baigent C, Godwin J, Halls H, Emberson J, Patrono C. Do selective cyclo-oxygenase-2 inhibitors and traditional non-steroidal anti-inflammatory drugs increase the risk of atherothrombosis? Meta-analysis of randomised trials. *BMJ* 2006;**332**(7553):1302–8. [PUBMED: 16740558]

Koes 2010
Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *European Spine Journal* 2010;**19**(12):2075–94.

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. [PUBMED: 20680369]

Lexchin 2003
Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;**326**(7400):1167–70. [PUBMED: 12775614]

Moher 2003
Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess* 2003;**7**(41):1–90. [MEDLINE: 14670218]

Müller-Schwefe 2011a
Müller-Schwefe GH. European survey of chronic pain patients: results for Germany. *Current Medical Research and Opinion* 2011;**27**(11):2099–106. [PUBMED: 21933101]

Müller-Schwefe 2011b
Müller-Schwefe G, Freytag A, Höer A, Schiffhorst G, Becker A, Casser HR, et al. Healthcare utilization of back pain patients: results of a claims data analysis. *Journal of Medical Economics* 2011;**14**(6):816–23. [PUBMED: 21992218]

Pallay 2004
Pallay RM, Seger W, Adler JL, Ettlinger RE, Quaidoo EA, Lipetz R, et al. Etoricoxib reduced pain and disability and improved quality of life in patients with chronic low back pain: a 3 month, randomized, controlled trial. *Scandinavian Journal of Rheumatology* 2004;**33**(4):257–66. [PUBMED: 15370723]

Piccoliori 2013
Piccoliori G, Engl A, Gatterer D, Sessa E, in der Schmitten J, Abholz HH. Management of low back pain in general practice - is it of acceptable quality: an observational study

among 25 general practices in South Tyrol (Italy). *BMC Family Practice* 2013;**14**:148. [PUBMED: 24090155]

Sostres 2013
Sostres C, Gargallo CJ, Lanas A. Nonsteroidal anti-inflammatory drugs and upper and lower gastrointestinal mucosal damage. *Arthritis Research & Therapy* 2013;**15** (Suppl 3):S3. [PUBMED: 24267289]

Trelle 2011
Trelle S, Reichenbach S, Wandel S, Hildebrand P, Tschannen B, Villiger PM, et al. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ* 2011;**342**:c7086. [PUBMED: 21224324]

van Tulder 2003
van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003;**28** (12):1290–9.

Vos 2012
Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**(9859):2163–96. [PUBMED: 23245607]

Walker 2000
Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders* 2000;**13**(3):205–17.

Webb 2003
Webb R, Brammah T, Lunt M, Urwin M, Allison T, Symmons D. Prevalence and predictors of intense, chronic, and disabling neck and back pain in the UK general population. *Spine (Phila Pa 1976)* 2003;**28**(11):1195–202. [PUBMED: 12782992]

Wehling 2014
Wehling M. Non-steroidal anti-inflammatory drug use in chronic pain conditions with special emphasis on the elderly and patients with relevant comorbidities: management and mitigation of risks and adverse effects. *European Journal of Clinical Pharmacology* 2014;**70**(10):1159–72. [PUBMED: 25163793]

References to other published versions of this review

Roelofs 2008
Roelofs PDDM, Deyo RA, Koes BW, Scholten RJPM, van Tulder MW. Non-steroidal anti-inflammatory drugs for low back pain. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD000396.pub3]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Allegrini 2009

Methods	RCT
Participants	180 participants, 102 women and 78 men; mean age 51 years (range 19 to 78 years) Inclusion: symptomatic lumbar osteoarthritis with daily pain during daily activities defined as a score as 40 mm on a 100 mm VAS Exclusion: participants with known hypersensitivity or allergy to piroxicam or to other NSAIDs; participants using topical medications to the painful region and the use of steroids by any route within 7 days before inclusion
Interventions	NSAID (i): piroxicam patch 14mg/day, 8 consecutive days (N = 60) NSAID (ii): piroxicam 1% cream, 1.4g/day, 8 consecutive days (N = 60) Reference treatment (iii): placebo patch, 8 consecutive days (N = 60)
Outcomes	Responder (reduction of pain score of at least 30%) rate to the administered treatment after 9 days: (i) 60%, (ii) 62% and (iii) 34% Adverse events: (i) 5 participants; (ii) 3 participants; (iii) 3 participants
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer randomized
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Unclear risk	Not mentioned
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	Each group had a drop out rate of > 20%

Allegrini 2009 (*Continued*)

Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Unclear risk	Not mentioned
Selective reporting (reporting bias)	Unclear risk	No protocol
Similarity of baseline characteristics	Unclear risk	No table with baseline characteristics
Co-interventions avoided or similar	Low risk	Rescue medication: paracetamol, up to 1.5 g per day allowed
Compliance acceptable	Low risk	All included participants were compliant
Timing outcome assessments similar	Low risk	Timing was similar

Berry 1982

Methods	RCT, double blind, double-dummy, cross-over
Participants	37 participants, 24 women and 13 men; mean age 55 years (range 32 to 79); median disease duration of 3 years Inclusion: adult participants with chronic back pain (≥ 3 months) due to spondylosis, degenerative spinal disease, sciatica or pain of nonspecific cause Exclusion: pain due to malignant disorders, infective diseases, spondylolisthesis, an alkaline phosphatase level outside normal limits or an ESR > 25 mm/hour
Interventions	NSAID (i): naproxen sodium 1100 mg/day, 14 days (N = 37 in cross-over design) NSAID (ii): diflunisal 1000 mg/day, 14 days (N = 37 in cross-over design) Reference treatment (iii): Placebo of dummy naproxen sodium capsules and diflunisal tablets (N = 37 in cross-over design)
Outcomes	Global pain, night pain, pain on movement and pain on standing assessed on vertical 10 cm VAS Reduction of pain on (i), an increase of pain (iii), and no significant change on (ii) Adverse events: (i) 18 participants; (ii) 16 participants; (iii) 18 participants
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomization procedure not mentioned
Allocation concealment (selection bias)	Unclear risk	Allocation procedure not mentioned

Berry 1982 (*Continued*)

Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	There was < 20% drop out
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Unclear risk	Unclear whether or not all participants were analysed
Selective reporting (reporting bias)	Unclear risk	No protocol
Similarity of baseline characteristics	Low risk	Cross-over design
Co-interventions avoided or similar	High risk	Corsets, braces, physiotherapy and paracetamol were permitted as long as they were started before entry to the study and continued unchanged for the trial duration
Compliance acceptable	High risk	14 drug discontinuations in 37 people
Timing outcome assessments similar	Low risk	Timing was similar

Birbara 2003

Methods	RCT, double blind
Participants	319 participants, 190 women and 124 men; mean age 52 years Inclusion: participants 18 to 75 years, low back pain ≥ 3 months, at least the past 30 days user of NSAID or acetaminophen. Pain without radiation to an extremity and without neurological signs or with radiation but not below the knee; After wash out period: ≥ 40 mm on low back intensity scale, increase of 10 mm and worsening of patient global assessment of disease status by ≥ 1 point compared to first screening visit Exclusion: low back pain due to malignancy, inflammatory disease, osteoporosis, fibromyalgia, ochronosis, vertebral fracture, infection, juvenile scoliosis or congenital malformation. Surgery in the past 6 months, symptomatic depression, drugs or alcohol abuse within the past 5 years, opioid use more than 4 days in the previous month, corticosteroid injections in the previous 3 months

Interventions	NSAID (i): etoricoxib 60 mg/day, 12 weeks (N = 103) NSAID (ii): etoricoxib 90 mg/day, 12 weeks (N = 107) Reference treatment (iii): placebo (N = 109)
Outcomes	Mean difference (95% CI) pain intensity scale (100 mm VAS) at 12 weeks: (i versus iii) −10.45 (−16.77 to −4.14); (ii versus iii) −7.5 (−13.71 to −1.28) Mean difference (95% CI) LBP bothersomeness (4-point Likert scale) at 12 weeks: (i versus iii) −0.38 (−0.62 to −0.14); (ii versus iii) −0.33 (−0.57 to −0.09) Mean difference (95% CI) RDQ (0 to 24 point scale) over 12 weeks; (i versus iii) −2.42 (−3.87 to −0.98); (ii versus iii) −2.06 (−3.46 to −0.65) Adverse events: (i) 60 participants (14 withdrew), (ii) 56 participants (17 withdrew) (iii) 51 participants (10 withdrew)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer randomized
Allocation concealment (selection bias)	Low risk	Computer random allocation schedule
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	High drop-out rates, 33%, 28%, 41%
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	No ITT analysis
Selective reporting (reporting bias)	Unclear risk	No protocol
Similarity of baseline characteristics	Low risk	Basline characteristics similar

Co-interventions avoided or similar	High risk	Muscle relaxants, physical therapy, and chiropractic or alternative therapy (such as acupuncture) were permitted, if their use was stable for the month preceding the screening visit and was expected to remain stable for the trial duration
Compliance acceptable	High risk	Discontinuation in 6%, 11% and 26%
Timing outcome assessments similar	Low risk	Timing was similar

Coats 2004

Methods	RCT, double-blind; 'flare' design
Participants	293 participants, 166 women, 127 men; mean age 48.7 years Inclusion: participants ≥ 18 years with low back pain ≥ 3 months requiring regular use of analgesic medication. Flare criteria after washout period Exclusion: low back pain of neurologic aetiology or as the result of major trauma; surgical interventions for low back pain < 4 weeks prior to study entry; participants who had received corticosteroids or opioids < 90 days prior to the first dose of study medication; secondary cause of low back pain; pending workers' compensation claims; pregnancy or breastfeeding
Interventions	NSAID (i): valdecoxib 40 mg/day, 4 weeks (N = 148) Reference treatment (ii): placebo, 4 weeks (N = 143)
Outcomes	Mean change score on pain intensity scale (100 mm VAS) at 1 and 4 weeks: (i) 29.2 mm and 41.9 mm; (ii) 17.7 mm and 31.1 mm; (i versus ii) all P < 0.001 Adverse events: (i) 52 participants (1 withdrew); (ii) 35 participants (3 withdrew)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomly assigned using a computer generated list of random numbers
Allocation concealment (selection bias)	Unclear risk	Procedure is not described
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Unclear risk	Care providers were not mentioned in blinding procedure

Coats 2004 (*Continued*)

Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	In placebo group drop-out rate was 21%
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Low risk	Intention-to-treat (ITT) analysis was used
Selective reporting (reporting bias)	Unclear risk	No protocol
Similarity of baseline characteristics	Low risk	Baseline characteristics similar
Co-interventions avoided or similar	Low risk	Rescue medication: acetaminophen $\leq$ 2000 mg/d for $\leq$ 3 consecutive days only in the first week, thereafter participants requiring any additional rescue medication were to be withdrawn from the study
Compliance acceptable	Low risk	3 participants (2%) versus 1 participant (< 1%) withdrew
Timing outcome assessments similar	Low risk	Timing similar

Driessens 1994

Methods	RCT, double-blind, double-dummy
Participants	62 participants, 33 women, 29 men; mean age (SD) 52.6 (14.3) Inclusion: hospital outpatients, chronic back pain for at least 4 weeks and required NSAID treatment Exclusion: acute or chronic infections, neoplasm or metastases, other severe intercurrent systemic disease, sciatica, referred pain from other organs or believed to be of psychogenic origin, treatment with local corticosteroid injection within 4 weeks of study commencement, pregnancy, lactation, contraindications for NSAID therapy
Interventions	NSAID (i): ibuprofen sustained-release 1600 mg, plus placebo, 14 days (N = 30) NSAID (ii): diclofenac sustained-release 100 mg, plus placebo, 14 days (N = 32)
Outcomes	Mean (SD) overall change in clinical condition compared to baseline on a 9-point scale: (i) 6.0 (1.4) (ii) 5.3 (1.5) Adverse events: (i) 4 participants, (ii) 16 participants (P = 0.002)
Notes	
Risk of bias	

Driessens 1994 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomization procedure not described
Allocation concealment (selection bias)	Unclear risk	Allocation procedure not described
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	25% of the participants in the diclofenac dropped out
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	Withdrawn participants were not analysed
Selective reporting (reporting bias)	Unclear risk	No study protocol
Similarity of baseline characteristics	Low risk	Baseline characteristics similar
Co-interventions avoided or similar	Low risk	Rescue analgesia: 500 mg paracetamol with a maximum dose of 4000 mg/day
Compliance acceptable	High risk	12 participants withdrew during treatment period
Timing outcome assessments similar	Low risk	Timing similar

Hickey 1982

Methods	RCT, double-blind
Participants	30 participants, 26 women, 4 men Inclusion: incapacity due to low back pain, duration $\geq$ 6 months, age 21 to 75 years Exclusion: pain from intervertebral disc prolapse, suspected neoplastic disease, neurological disease, pregnancy, peptic ulceration or gastrointestinal haemorrhage, current treatment with systemic corticosteroids or anticoagulants, liver or kidney disease, haemopoi-

Hickey 1982 (*Continued*)

	etic disorders, history of sensitivity to salicylates or paracetamol, psychiatric problems
Interventions	NSAID (i): Diflunisal 1000 mg/day, 4 weeks (N = 16) Reference treatment (ii): paracetamol 4000 mg/day, 4 weeks (N = 14)
Outcomes	Number of participants with none or mild low back pain after 2 and 4 weeks: (i) 11, 13 (ii) 9, 7. Significantly more participants in (i) (10 out of 16) considered the therapy as good or excellent than in (ii) (4 out of 12) Adverse events: (i) 2 participants (ii) 1 participants
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomization prior to the trial
Allocation concealment (selection bias)	Low risk	Code-labelled drugs, code was not broken
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	Sixteen out of 16 participants and 13 out of 14 participants completed the trial
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	Two participants in the paracetamol group were not analysed in their allocation group
Selective reporting (reporting bias)	Unclear risk	No protocol
Similarity of baseline characteristics	Unclear risk	No baseline characteristics were shown
Co-interventions avoided or similar	Low risk	Only anti-hypertensive drug therapy was allowed, other drugs were forbidden
Compliance acceptable	Unclear risk	Compliance was not mentioned

Hickey 1982 (*Continued*)

Timing outcome assessments similar	Low risk	Timing similar

Katz 2011

Methods	RCT, double-blind
Participants	217 participants, 118 women, 99 men Inclusion: participants aged ≥ 18 years, body mass index ≤ 39 kg/m², nonradiculopathic low back pain for at least 3 months, required regular analgesic medication, analgesic medication > 4 days/week over the previous month, average pain intensity score ≥ 4 over previous 24 hours on 11-point numerical rating scale, minimum compliance of 4 entries in electronic daily pain diary over the 5 previous days Exclusion: radiculopathy in previous 2 years, secondary causes of back pain, surgical intervention for treatment of back pain, pregnancy, lactation, rheumatoid arthritis, seronegative spondyloarthropathy, Paget disease of spine, pelvis or femur, fibromyalgia, tumours or infections of spinal cord, cancer in previous 2 years other than cutaneous basal cell or squamous cell carcinoma, allergic reaction to monoclonal antibody or IgG-fusion protein, acetaminophen or NSAIDs, contraindications to NSAID therapy
Interventions	NSAID (i): naproxen 1000 mg daily and placebo single intravenous infusion, 12 weeks (N = 88) Reference treatment (ii): tanezumab single intravenous infusion 200 μg/kg and oral placebo daily, 12 weeks (N = 88) Reference treatment (iii): placebo single intravenous infusion and oral placebo daily, 12 weeks (N = 41)
Outcomes	Mean change in average low back pain intensity over previous 24 hours on 11-point numerical rating scale, at 6 weeks compared to baseline: (i versus iii) -2.5 versus -2.0 (P = 0.068) Adverse events: (i) 54 participants (3 withdrew); (ii) 50 participants (4 withdrew); (iii) 27 participants (2 withdrew)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not mentioned
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded, placebo tablets/injections

Katz 2011 (*Continued*)

Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Unclear risk	Unclear if care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	Drop out 32%
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Low risk	ITT was performed
Selective reporting (reporting bias)	Low risk	Trial was registered
Similarity of baseline characteristics	Low risk	Baseline characteristics were similar
Co-interventions avoided or similar	Low risk	Rescue medication acetaminophen with a maximum of 2000 mg per day and maximum 3 days per week
Compliance acceptable	Low risk	Nine people discontinued the trial
Timing outcome assessments similar	Low risk	Timing similar

Kivitz 2013

Methods	Randomized, double-blind, placebo and active-controlled trial
Participants	1359 participants, 714 women, 645 men Inclusion: duration of back pain of ≥ 3 months requiring regular use of analgesic medication (> 4 days per week for the past month), including immediate-release opioids (in which the average daily opioid dose (for a 7-day period) did not exceed a morphine equivalent dose of 30 mg/d) but excluding acetaminophen, gabapentin or pregabalin as the sole analgesics used for chronic low back pain; primary location of low back pain between the 12th thoracic vertebra and the lower gluteal folds, with or without radiation into the posterior thigh (Quebec Task Force on Spinal Disorders category 1 or 2); average low back pain intensity (LBPI) score of ≥ 4 (on an 11-point NRS) while receiving current treatment; and Patient's Global Assessment (PGA) of low back pain of fair, poor or very poor Exclusion: history of lumbosacral radiculopathy within the past 2 years, vertebral fracture, major trauma or back surgery in the past 6 months; significant cardiac, neurological, or other pain, or psychological conditions; known history of rheumatoid arthritis, seronegative spondyloarthropathy, Paget's disease of the spine, pelvis or femur, fibromyalgia, tumours or infections of the spinal cord; and any condition that might preclude NSAID

	use. Patients also were excluded if extended-release (ER) opioids or long-acting opioids such as oxycodone controlled release, oxymorphone ER, hydromorphone, transdermal fentanyl or methadone had been used within 3 months of screening
Interventions	NSAID (i): naproxen 1000 mg daily and placebo infusion at baseline, 8 weeks and 16 weeks (N = 295) Reference treatment (ii): placebo tablets daily and placebo infusion at baseline and 8 weeks, 16 weeks (N = 230) Reference treatment (iii): tanezumab 5 mg iv infusion over 5 minutes at baseline and 8 weeks, 16 weeks (N = 232) Reference treatment (iv): tanezumab 10 mg iv infusion over 5 minutes at baseline and 8 weeks, 16 weeks (N = 295) Reference treatment (v): tanezumab 20 mg iv infusion over 5 minutes at baseline and 8 weeks, 16 weeks (N= 295)
Outcomes	Least squares mean difference from baseline on a 11-point scale: (i versus iii) 0.08 (P = 0.688) Least squares mean difference from baseline on a 11-point scale: (i versus iv) −0.39 (P = 0.035) Least squares mean difference from baseline on a 11-point scale: (i versus v) −0.51 (P = 0.006) Adverse events: (i) 142 participants (10 withdrew), (ii) 120 participants (14 withdrew) , (iii) 141 participants (11 withdrew), (iv) 171 participants (19 withdrew), (v) 190 participants (28 withdrew)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not mentioned
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Unclear risk	Not mentioned for all examinations

Kivitz 2013 (*Continued*)

Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	All trial groups had high drop out rates
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Unclear risk	ITT and per protocol analysis used, unclear which analysis was used in what comparison
Selective reporting (reporting bias)	Low risk	Protocol present
Similarity of baseline characteristics	Low risk	Baseline characteristics were comparable
Co-interventions avoided or similar	Low risk	Only paracetamol up to 300 mg/day and max 3 days per week was allowed
Compliance acceptable	Unclear risk	Not mentioned
Timing outcome assessments similar	Low risk	Timing was similar

O'Donnell 2009

Methods	RCT, double-blind, double-dummy
Participants	2 studies; study 1: 791 participants, 462 women, 329 men; study 2: 802 participants, 450 women, 342 men Inclusion: participants aged ≥ 18 years, duration of back pain ≥ 12 weeks, requiring analgesics ≥ 4 days/week, back pain score of ≥ 4 on 11-point NRS at baseline Exclusion: back pain with neurologic aetiology, recent major trauma, due to visceral disorder, history of rheumatoid arthritis, spondyloarthropathy, spinal stenosis, malignancy, fibromyalgia, tumours or infections of the brain, spinal cord or peripheral nerves, herniated disc with neurological impairment in previous 2 years, psoriasis, seizure disorder, alcohol/analgesic/narcotic or other substance abuse in previous 2 years, asthma, allergic reactions on aspirin or NSAID, contraindications for NSAID use, surgical intervention for back pain in previous 6 months
Interventions	NSAID (i): celecoxib 400 mg/day, 6 weeks (study 1: N = 402; study 2: N = 396) Reference treatment (ii): tramadol 200 mg/day, 6 weeks (study 1: N = 389; study 2: N = 396)
Outcomes	At 6 weeks ≥ 30% improvement in pain from baseline, measured with 11-point numerical rating scale; study 1 (i versus ii) 63.2% versus 49.9% (P < 0.001); study 2 (i versus ii) 64.1% versus 55.1% (P = 0.008) Adverse events: study 1: (i) 191 participants (18 withdrew), (ii) 230 participants (72 withdrew); study 2: (i) 190 participants (21 withdrew), (ii) 224 participants (60 withdrew)
Notes	

O'Donnell 2009 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomization was computer generated
Allocation concealment (selection bias)	Low risk	Computerized schedule
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Double dummy, double blind
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Double dummy, double blind
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Double dummy, double blind
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	High risk	The tramadol group had a drop out rate > 20%
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	Low risk	ITT analysis
Selective reporting (reporting bias)	Unclear risk	No study protocol
Similarity of baseline characteristics	Low risk	Baseline characteristics were similar
Co-interventions avoided or similar	Low risk	No rescue medication allowed
Compliance acceptable	High risk	Non-compliance in 9.6% of celecoxib group and 15% in tramadol group
Timing outcome assessments similar	Low risk	Timing similar

Romanò 2009

Methods	RCT, cross-over design
Participants	42 participants, 20 women, 16 men Inclusion: low back pain ≥ 6 months due to disc prolapse, lumbar spondylosis or spinal stenosis or both, minimum VAS > 40 mm, age 18 to 75 Exclusion: Previous back surgery, diabetes, neurological disease, cardio-renal disease,

	history of gastric ulcers or intestinal bleeding, known allergy to drugs under study, alcohol/drugs abuse
Interventions	Each treatment lasted 4 weeks with 1 week discontinuation between treatments NSAID: (i) celecoxib approximately 3 to 6 mg/kg/day and placebo Reference treatment: (ii) pregabalin approximately 1 mg/kg/day and placebo Reference treatment: (iii) celecoxib and pregabalin
Outcomes	Mean (SD) pain reduction after 4 weeks on 100 mm VAS: (i) 5.6; (ii) 5; (iii) 17.7; Adverse events: (i) 4 participants (1 withdrew), (ii) 5 participants (1 withdrew), (iii) 7 participants (2 withdrew)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomization unclear
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Unclear risk	Not mentioned
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	Low dropout rate
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	Drop outs were excluded from data analysis
Selective reporting (reporting bias)	Unclear risk	No study protocol
Similarity of baseline characteristics	Low risk	Cross over design

Co-interventions avoided or similar	Low risk	Use of antidepressants or anticonvulsants or both, opioids, nonsteroidal anti-inflammatory drugs or muscle relaxants was not permitted
Compliance acceptable	Low risk	Individual drug consumption was measured and acceptable
Timing outcome assessments similar	Low risk	Timing similar

Shirado 2010

Methods	RCT
Participants	201 participants, 112 women, 89 men; mean age 42.2 years Inclusion: age 20 to 64 years, nonspecific chronic low back pain ≥ 3 months without radicular pain, ≥ 70° at straight leg raising test, negative femoral nerve stretching test, no superficial sensory deficits, muscle strength ≥ 4/5 Exclusion: low back pain due to tumours, infections, fractures, previous back surgery, severe osteoporosis, psychiatric disorders, liver and renal dysfunction, pregnancy, medication for cardiac failure, history of cerebrovascular accident or myocardial infarction, or both, in previous 6 months
Interventions	NSAID (i): 1 of the following 3 NSAIDs were prescribed: loxoprofen sodium 180 mg/day; diclofenac sodium 75 mg/day; zaltoprofen 240 mg/day, 12 weeks Reference treatment (ii): exercise programme with trunk muscle strengthening and stretching, 12 weeks
Outcomes	Mean change from baseline to 8 weeks on 100 mm VAS was not different between (i) and (ii), P = 0.33 Mean change from baseline to 8 weeks on RDQ in favour of (ii), P = 0.02
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated 4 block randomization
Allocation concealment (selection bias)	Low risk	Office manager concealed allocation
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	High risk	NSAIDs versus exercise
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	High risk	NSAIDs versus exercise

Shirado 2010 (*Continued*)

Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessor was blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	Two in exercise, 6 in NSAIDs.
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	No ITT analysis performed.
Selective reporting (reporting bias)	Unclear risk	study protocol not attainable
Similarity of baseline characteristics	Low risk	Baseline characteristics were similar
Co-interventions avoided or similar	Unclear risk	Rescue medication not mentioned
Compliance acceptable	Unclear risk	Compliance not mentioned
Timing outcome assessments similar	Low risk	Timing similar

Videman 1984

Methods	RCT, double-blind
Participants	28 outpatients, 11 women, 17 men; mean age 45 years Inclusion: chronic severe low back pain, age 25 to 76 years Exclusion: pregnant or nursing women, compensation claims, haematological, renal or hepatic disease, pre-existing radiological evidence of peptic ulcer, intolerance to indomethacin
Interventions	NSAID (i): piroxicam 20 mg/day, 6 weeks (N = 14) NSAID (ii): indometacin 75 mg/day, 6 weeks (N = 14)
Outcomes	Change of pain from baseline until 6 weeks: (i) 8.1 (ii) 9.4; no significant difference between groups Adverse events: (i) 8 participants (ii) 10 participants
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not mentioned

Videman 1984 (*Continued*)

Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Patients were blinded
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers were blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors were blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	Two out of 14 participants in one group were lost to follow-up
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	Complete case analysis
Selective reporting (reporting bias)	Unclear risk	No study protocol
Similarity of baseline characteristics	Unclear risk	No baseline characteristics shown
Co-interventions avoided or similar	Low risk	Only paracetamol as co-intervention up to 3000 mg
Compliance acceptable	Unclear risk	Compliance not mentioned
Timing outcome assessments similar	Unclear risk	Timing was unclear

Zerbini 2005

Methods	RCT, double-blind, double-dummy, 'flare design'
Participants	446 participants, 320 women, 126 men; mean age (SD) 51.9 (13.8) Inclusion: age 19 to 85 years, with chronic low back pain, regular users of analgesic medication, pain without radiation to an extremity and without neurological signs or pain with radiation to an extremity, but not below the knee and without neurological signs, after 1 week washout period LBP intensity $\leq$ 80 mm on 100 mm VAS scale
Interventions	NSAID (i) etoricoxib 60 mg/day, 4 weeks (N = 224) NSAID (ii) diclofenac 150 mg/day, 4 weeks (N = 222)
Outcomes	Mean difference (95% CI) pain intensity scale (100 mm VAS) at 4 weeks: (i, N = 222 versus ii, N = 218) 2.51 (−1.50 to 6.51) Mean difference (95% CI) RDQ (0 to 24) over 4 weeks: (i versus ii) −0.23 (−1.14 to

	0.67) Adverse events: (i) 79 participants (15 withdrew); (ii) 87 participants (12 withdrew)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomization procedure not described
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding (performance bias and detection bias): All outcomes - Patients All outcomes	Low risk	Double dummy
Blinding (performance bias and detection bias): All outcomes - Care providers All outcomes	Low risk	Care providers blinded
Blinding (performance bias and detection bias): All outcomes - Outcome assessors All outcomes	Low risk	Outcome assessors blinded
Incomplete outcome data (attrition bias): All outcomes - Drop-outs All outcomes	Low risk	9% and 11% drop out in both groups
Incomplete outcome data (attrition bias): All outcomes - ITT analysis All outcomes	High risk	Per protocol analysis
Selective reporting (reporting bias)	Unclear risk	No study protocol
Similarity of baseline characteristics	Low risk	Similar baseline characteristics
Co-interventions avoided or similar	Low risk	Paracetamol as rescue therapy
Compliance acceptable	Low risk	More than 95% compliance in both study groups
Timing outcome assessments similar	Low risk	Timing similar

Abbreviations: ESR: erythorcyte sedimentation rate, LBP: low back pain, NSAID: non-steroidal anti-inflammatory drug, RCT: randomized controlled trial, RDQ: Roland Morris Disability Questionnaire, SD: standard deviation, VAS: visual analogue scale.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Aoki 1983	Duration of back pain is unclear.
Babey-Dölle 1994	Only acute back pain included.
Borghi 2013	No comparison made, one study group.
Chang 2013	Patients were given intravenous infusion after spine surgery
Chrubasik 2003	Rofecoxib as study medication.
Davoli 1989	The trial only included participants with acute back pain.
Famaey 1998	The trial does not distinguish between participants with subacute and chronic back pain
Ingpen 1969	Duration of back pain is unclear.
Jacobs 1968	The study only included participants with acute back pain.
Jaffé 1974	Duration of back pain is unclear.
Katz 2003	Rofecoxib as study medication.
Listrat 1990	Duration of back pain is unclear.
Matsumo 1991	Inclusion > 1 month of back pain.
Merkulova 2013	Article in Russian.
Peng 2014	Article in Chinese.
Postacchini 1988	Inclusion > 2 months of back pain.
Siegmeth 1978	Participants selected based on radiological osteoarthritis, not on back pain
Tavafian 2014	NSAIDs were used in both groups as needed.
Waikakul 1995	Duration of back pain is unclear.
Waikakul 1996	Duration of back pain is unclear.
Wetzel 2014	Intravenous infusion in patients on chronic opioid treatment

DATA AND ANALYSES

Comparison 1. NSAIDs versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks.	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 All NSAIDs	6	1354	Mean Difference (IV, Random, 95% CI)	-6.97 [-10.74, -3.19]
1.2 Non-selective NSAIDs	4	847	Mean Difference (IV, Random, 95% CI)	-5.96 [-10.96, -0.96]
1.3 Selective NSAIDs	2	507	Mean Difference (IV, Random, 95% CI)	-9.11 [-13.56, -4.66]
2 Change in disability from baseline	4	1161	Mean Difference (IV, Fixed, 95% CI)	-0.85 [-1.30, -0.40]
3 Proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks.	6		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.1 All NSAIDs	6	1354	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.92, 1.17]
3.2 Non-selective NSAIDs	4	847	Risk Ratio (M-H, Fixed, 95% CI)	0.94 [0.82, 1.08]
3.3 Selective NSAIDs	2	507	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [1.00, 1.56]
4 Sensitivity analysis: change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks.	3	728	Mean Difference (IV, Random, 95% CI)	-5.03 [-10.37, 0.32]
5 Sensitivity analysis: change in disability from baseline	2	654	Mean Difference (IV, Fixed, 95% CI)	-0.41 [-1.04, 0.23]
6 Sensitivity analysis: proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks.	3	728	Risk Ratio (M-H, Fixed, 95% CI)	0.93 [0.81, 1.07]

Comparison 2. NSAIDs versus other drug treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Proportion of patients experiencing global improvement. Follow-up ≤ 6 weeks.	2		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
2 Proportion of patients experiencing adverse events. Follow-up ≤ 6 weeks.	3		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

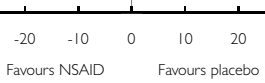
Analysis 1.1. Comparison 1 NSAIDs versus placebo, Outcome 1 Change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

Outcome: 1 Change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks.

Study or subgroup	NSAID N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 All NSAIDs							
Allegrini 2009	60	-28 (31.7)	59	-16.5 (31.7)		8.4 %	-11.50 [-22.89, -0.11]
Berry 1982	37	-11.5 (34)	37	9.4 (34)		5.1 %	-20.90 [-36.39, -5.41]
Birbara 2003	107	-7.5 (23.3)	109	0 (23.3)		18.2 %	-7.50 [-13.71, -1.29]
Coats 2004	148	-41.9 (27.7)	143	-31.1 (27.7)		17.7 %	-10.80 [-17.17, -4.43]
Katz 2011	88	-2.4 (11.6)	41	0 (11.6)		24.5 %	-2.40 [-6.70, 1.90]
Kivitz 2013	295	-4.1 (22.3)	230	0 (22.3)		26.1 %	-4.10 [-7.94, -0.26]
Subtotal (95% CI)	**735**		**619**			**100.0 %**	**-6.97 [-10.74, -3.19]**
Heterogeneity: Tau² = 10.35; Chi² = 10.41, df = 5 (P = 0.06); I² =52%							
Test for overall effect: Z = 3.62 (P = 0.00030)							
2 Non-selective NSAIDs							
Allegrini 2009	60	-28 (31.7)	59	-16.5 (31.7)		14.1 %	-11.50 [-22.89, -0.11]
Berry 1982	37	-11.5 (34)	37	9.4 (34)		8.7 %	-20.90 [-36.39, -5.41]
Katz 2011	88	-2.4 (11.6)	41	0 (11.6)		37.5 %	-2.40 [-6.70, 1.90]
Kivitz 2013	295	-4.1 (22.3)	230	0 (22.3)		39.7 %	-4.10 [-7.94, -0.26]
Subtotal (95% CI)	**480**		**367**			**100.0 %**	**-5.96 [-10.96, -0.96]**
Heterogeneity: Tau² = 12.54; Chi² = 6.70, df = 3 (P = 0.08); I² =55%							
Test for overall effect: Z = 2.34 (P = 0.019)							
3 Selective NSAIDs							
Birbara 2003	107	-7.5 (23.3)	109	0 (23.3)		51.2 %	-7.50 [-13.71, -1.29]
Coats 2004	148	-41.9 (27.7)	143	-31.1 (27.7)		48.8 %	-10.80 [-17.17, -4.43]
Subtotal (95% CI)	**255**		**252**			**100.0 %**	**-9.11 [-13.56, -4.66]**
Heterogeneity: Tau² = 0.0; Chi² = 0.53, df = 1 (P = 0.47); I² =0.0%							
Test for overall effect: Z = 4.02 (P = 0.000059)							
Test for subgroup differences: Chi² = 0.94, df = 2 (P = 0.63), I² =0.0%							



Analysis 1.2. Comparison 1 NSAIDs versus placebo, Outcome 2 Change in disability from baseline.

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

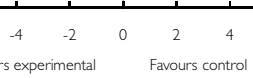
Outcome: 2 Change in disability from baseline

Study or subgroup	NSAID N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
Birbara 2003	107	-2.1 (5.3)	109	0 (5.3)		10.1 %	-2.10 [-3.51, -0.69]
Coats 2004	148	-1.1 (3.1)	143	0 (3.1)		39.7 %	-1.10 [-1.81, -0.39]
Katz 2011	88	-0.6 (3.1)	41	0 (3.1)		15.3 %	-0.60 [-1.75, 0.55]
Kivitz 2013	295	-0.32 (4.4)	230	0 (4.4)		35.0 %	-0.32 [-1.08, 0.44]
Total (95% CI)	**638**		**523**			**100.0 %**	**-0.85 [-1.30, -0.40]**

Heterogeneity: $Chi^2 = 5.53$, df = 3 (P = 0.14); $I^2 = 46\%$

Test for overall effect: Z = 3.72 (P = 0.00020)

Test for subgroup differences: Not applicable

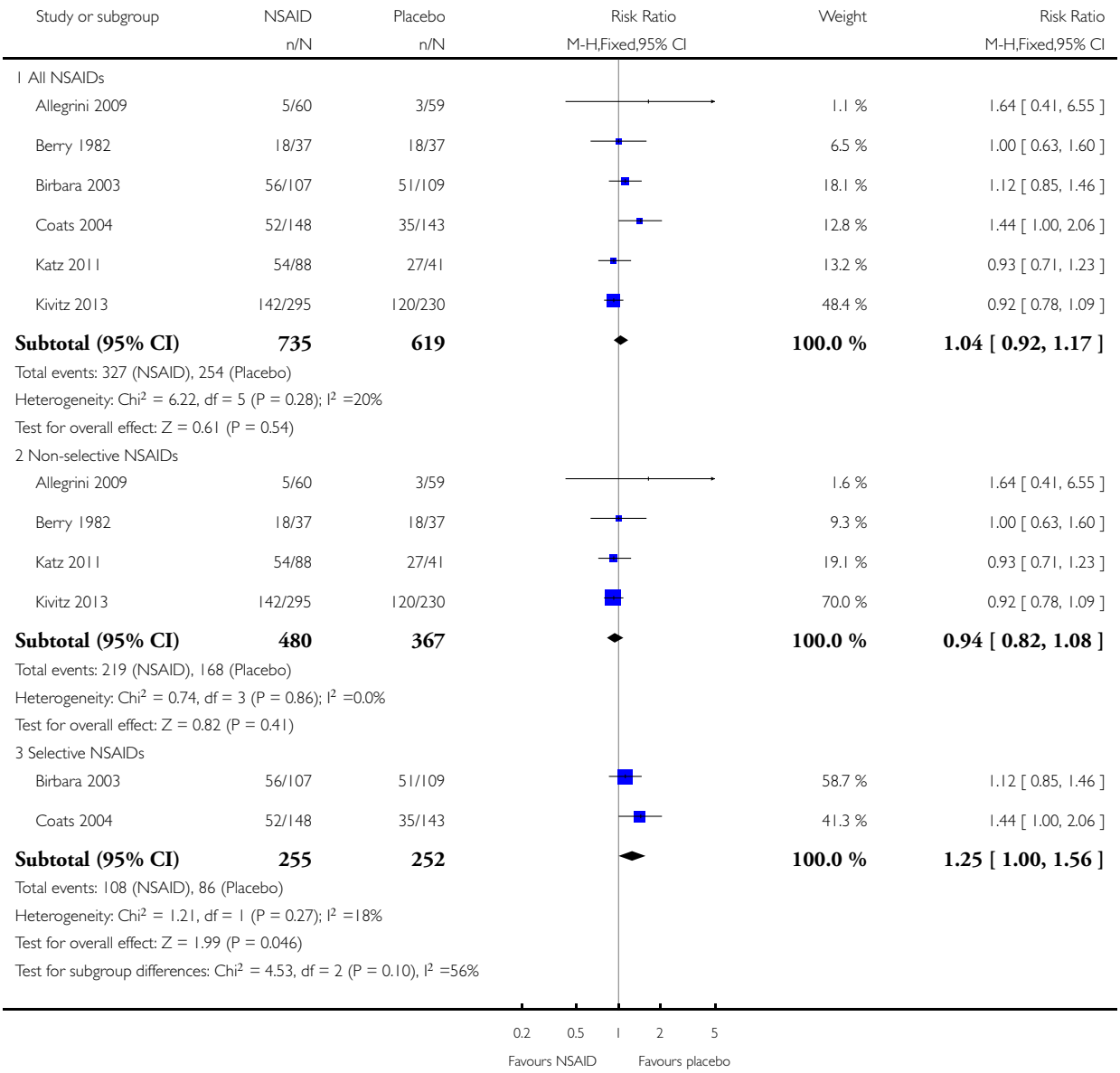


Analysis 1.3. Comparison 1 NSAIDs versus placebo, Outcome 3 Proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

Outcome: 3 Proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks.



Analysis 1.4. Comparison 1 NSAIDs versus placebo, Outcome 4 Sensitivity analysis: change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

Outcome: 4 Sensitivity analysis: change in pain intensity from baseline on 100 mm VAS. Follow-up ≤ 16 weeks.

Study or subgroup	NSAID N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Berry 1982	37	-11.5 (34)	37	9.4 (34)		10.0 %	-20.90 [-36.39, -5.41]
Katz 2011	88	-2.4 (11.6)	41	0 (11.6)		43.7 %	-2.40 [-6.70, 1.90]
Kivitz 2013	295	-4.1 (22.3)	230	0 (22.3)		46.3 %	-4.10 [-7.94, -0.26]
Total (95% CI)	**420**		**308**			**100.0 %**	**-5.03 [-10.37, 0.32]**

Heterogeneity: $Tau^2 = 12.20$; $Chi^2 = 5.10$, df = 2 (P = 0.08); $I^2 = 61\%$

Test for overall effect: Z = 1.84 (P = 0.065)

Test for subgroup differences: Not applicable

Analysis 1.5. Comparison 1 NSAIDs versus placebo, Outcome 5 Sensitivity analysis: change in disability from baseline.

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

Outcome: 5 Sensitivity analysis: change in disability from baseline

Study or subgroup	NSAID N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
Katz 2011	88	-0.6 (3.1)	41	0 (3.1)		30.4 %	-0.60 [-1.75, 0.55]
Kivitz 2013	295	-0.32 (4.4)	230	0 (4.4)		69.6 %	-0.32 [-1.08, 0.44]
Total (95% CI)	**383**		**271**			**100.0 %**	**-0.41 [-1.04, 0.23]**

Heterogeneity: $Chi^2 = 0.16$, df = 1 (P = 0.69); $I^2 = 0.0\%$

Test for overall effect: Z = 1.25 (P = 0.21)

Test for subgroup differences: Not applicable

Analysis 1.6. Comparison 1 NSAIDs versus placebo, Outcome 6 Sensitivity analysis: proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 1 NSAIDs versus placebo

Outcome: 6 Sensitivity analysis: proportion of patients experiencing adverse events. Follow-up ≤ 16 weeks.

Study or subgroup	NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Berry 1982	18/37	18/37		9.5 %	1.00 [0.63, 1.60]
Katz 2011	54/88	27/41		19.4 %	0.93 [0.71, 1.23]
Kivitz 2013	142/295	120/230		71.1 %	0.92 [0.78, 1.09]
Total (95% CI)	**420**	**308**		**100.0 %**	**0.93 [0.81, 1.07]**

Total events: 214 (NSAID), 165 (Placebo)

Heterogeneity: $Chi^2 = 0.10$, df = 2 (P = 0.95); $I^2 = 0.0\%$

Test for overall effect: Z = 0.99 (P = 0.32)

Test for subgroup differences: Not applicable

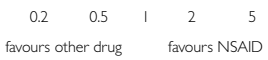
Analysis 2.1. Comparison 2 NSAIDs versus other drug treatment, Outcome 1 Proportion of patients experiencing global improvement. Follow-up ≤ 6 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 2 NSAIDs versus other drug treatment

Outcome: 1 Proportion of patients experiencing global improvement. Follow-up ≤ 6 weeks.

Study or subgroup	NSAID n/N	Other drug n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
Hickey 1982	13/16	7/12		1.39 [0.82, 2.37]
O'Donnell 2009	504/798	392/785		1.26 [1.16, 1.38]



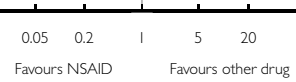
Analysis 2.2. Comparison 2 NSAIDs versus other drug treatment, Outcome 2 Proportion of patients experiencing adverse events. Follow-up ≤ 6 weeks..

Review: Non-steroidal anti-inflammatory drugs for chronic low back pain

Comparison: 2 NSAIDs versus other drug treatment

Outcome: 2 Proportion of patients experiencing adverse events. Follow-up ≤ 6 weeks.

Study or subgroup	NSAID n/N	Other drug n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
Hickey 1982	2/16	1/12		1.50 [0.15, 14.68]
O'Donnell 2009	381/798	454/785		0.83 [0.75, 0.91]
Roman 2009	4/36	5/36		0.80 [0.23, 2.74]



APPENDICES

Appendix 1. CENTRAL search strategy

Last searched 24 June 2015. Line 34 is added and line 42 is revised.
#1 MeSH descriptor: [Back Pain] explode all trees
#2 dorsalgia
#3 backache
#4 lumbar next pain or coccyx or coccydynia or spondylosis
#5 MeSH descriptor: [Spine] explode all trees
#6 MeSH descriptor: [Spinal Diseases] explode all trees
#7 lumbago and discitis and disc near herniation
#8 spinal fusion
#9 spinal neoplasms
#10 facet near joints
#11 MeSH descriptor: [Intervertebral Disk] explode all trees
#12 postlaminectomy
#13 arachnoiditis
#14 failed near back
#15 MeSH descriptor: [Cauda Equina] explode all trees
#16 lumbar near vertebra*
#17 spinal near stenosis
#18 slipped near (disc* or disk*)
#19 degenerat* near (disc* or disk*)
#20 stenosis near (spine or root or spinal)
#21 displace* near (disc* or disk*)
#22 prolap* near (disc* or disk*)
#23 MeSH descriptor: [Sciatic Neuropathy] explode all trees
#24 sciatic*
#25 back disorder*
#26 back near pain
#27 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or # 20 or #21 or #22 or #23 or #24 or #25 or #26
#28 nsaid*
#29 MeSH descriptor: [Anti-Inflammatory Agents, Non-Steroidal] explode all trees
#30 MeSH descriptor: [Cyclooxygenase Inhibitors] explode all trees
#31 MeSH descriptor: [Cyclooxygenase 2 Inhibitors] explode all trees
#32 non-steroidal anti inflammat*
#33 non-steroidal anti-inflammat*
#34 (cyclooxygenase or cyclo-oxygenase) next/3 inhibitor*
#35 aspirin
#36 acetylsalicyl*
#37 carbasalate calcium
#38 diflunisal
#39 aceclofenac
#40 alclofenac
#41 diclofenac
#42 indometacin or indomethacin
#43 sulindac
#44 meloxicam
#45 piroxicam
#46 dexibuprofen
#47 dexketoprofen

#48 fenoprofen
#49 flurbiprofen
#50 ibuprofen
#51 ketoprofen
#52 naproxen
#53 tiapro*
#54 metamizol
#55 phenylbutazone
#56 phenazone
#57 propyphenazone
#58 celecoxib
#59 etoricoxib
#60 nabumeton
#61 parecoxib
#62 rofecoxib
#63 celecoxib
#64 valdecoxib
#65 lumiracoxib
#66 parecoxib
#67 vioxx
#68 celebrex
#69 bextra
#70 prexige
#71 arcoxia
#72 etodolac
#73 floctafenine
#74 meclofenam*
#75 meloxicam
#76 oxaprozin
#77 piroxicam
#78 tenoxicam
#79 tolmetin
#80 #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69 or #70 or #71 or #72 or #73 or #74 or #75 or #76 or #77 or #78 or #79
#81 #27 and #80
#82 #81 in Trials

May 2012 strategy. In 2015, Line 77 and 66 are removed (duplicate with line 52 and 59), disc degeneration and prolapse are removed from line 8 (captured in line 20 and 23), and sciatica is removed from line 5 (captured in line 25).

#1 MeSH descriptor Back Pain explode all trees
#2 dorsalgia
#3 backache
#4 MeSH descriptor Low Back Pain explode all trees
#5 (lumbar next pain) or (coccyx) or (coccydynia) or (sciatica) or (spondylosis)
#6 MeSH descriptor Spine explode all trees
#7 MeSH descriptor Spinal Diseases explode all trees
#8 (lumbago) or (discitis) or (disc near degeneration) or (disc near prolapse) or (disc near herniation)
#9 spinal fusion
#10 spinal neoplasms
#11 facet near joints
#12 MeSH descriptor Intervertebral Disk explode all trees
#13 postlaminectomy
#14 arachnoiditis 36

#15 failed near back
#16 MeSH descriptor Cauda Equina explode all trees
#17 lumbar near vertebra*
#18 spinal near stenosis
#19 slipped near (disc* or disk*)
#20 degenerat* near (disc* or disk*)
#21 stenosis near (spine or root or spinal)
#22 displace* near (disc* or disk*)
#23 prolap* near (disc* or disk*)
#24 MeSH descriptor Sciatic Neuropathy explode all trees
#25 sciatic*
#26 back disorder*
#27 back near pain
#28 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27)
#29 nsaid*
#30 MeSH descriptor Anti-Inflammatory Agents, Non-Steroidal explode all trees
#31 MeSH descriptor Cyclooxygenase Inhibitors explode all trees
#32 MeSH descriptor Cyclooxygenase 2 Inhibitors explode all trees
#33 non-steroidal anti inflammat*
#34 non-steroidal anti-inflammat*
#35 aspirin
#36 acetylsalicyl*
#37 carbasalate calcium
#38 diflunisal
#39 aceclofenac
#40 alclofenac
#41 diclofenac
#42 indometacin
#43 sulindac
#44 meloxicam
#45 piroxicam
#46 dexibuprofen
#47 dexketoprofen
#48 fenoprofen
#49 flurbiprofen
#50 ibuprofen
#51 ketoprofen
#52 naproxen
#53 tiapro*
#54 metamizol
#55 phenylbutazone
#56 phenazone
#57 propyphenazone
#58 celecoxib
#59 etoricoxib
#60 nabumeton
#61 parecoxib
#62 rofecoxib
#63 celecoxib
#64 valdecoxib
#65 lumiracoxib
#66 etoricoxib

#67 parecoxib
#68 vioxx
#69 celebrex
#70 bextra
#71 prexige
#72 arcoxia
#73 etodolac
#74 floctafenine
#75 meclofenam*
#76 meloxicam
#77 naproxen
#78 oxaprozin
#79 piroxicam
#80 tenoxicam
#81 tolmetin
#82 (#29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54 OR #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61 OR #62 OR #63 OR #64 OR #65 OR #66 OR #67 OR #68 OR #69 OR #70 OR #71 OR #72 OR #73 OR #74 OR #75 OR #76 OR #77 OR #78 OR #79 OR #80 OR #81)
#83 (#28 AND #82), from 2007 to 2012

Appendix 2. MEDLINE search strategy

Last searched 24 June 2015. Line 3 and 61 are added and line 6, 22, 29, and 39 are revised.

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. pragmatic clinical trial.pt.
4. comparative study.pt.
5. clinical trial.pt.
6. randomi#ed.ab.
7. placebo.ab,ti.
8. drug therapy.fs.
9. randomly.ab,ti.
10. trial.ab,ti.
11. groups.ab,ti.
12. or/1-11
13. (animals not (humans and animals)).sh.
14. 12 not 13
15. dorsalgia.ti,ab.
16. exp Back Pain/
17. backache.ti,ab.
18. (lumbar adj pain).ti,ab.
19. coccyx.ti,ab.
20. coccydynia.ti,ab.
21. sciatica.ti,ab.
22. exp sciatic neuropathy/
23. spondylosis.ti,ab.
24. lumbago.ti,ab.
25. back disorder$.ti,ab.
26. or/15-25
27. exp Anti-Inflammatory Agents, Non-Steroidal/
28. nsaids.mp.

29. non-steroidal antiinflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
30. non-steroidal anti-inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
31. aspirin.mp. or exp Aspirin/
32. acetylsalicyl$.mp.
33. exp Salicylic Acid/
34. carbasalate calcium.mp.
35. diflunisal.mp. or exp Diflunisal/
36. aceclofenac.mp.
37. alclofenac.mp.
38. diclofenac.mp. or exp Diclofenac/
39. (indometacin or indomethacin).mp. or exp Indomethacin/
40. sulindac.mp. or exp Sulindac/
41. meloxicam.mp.
42. piroxicam.mp. or exp Piroxicam/
43. dexibuprofen.mp.
44. dexketoprofen.mp.
45. fenoprofen.mp. or exp Fenoprofen/
46. flurbiprofen.mp. or exp Flurbiprofen/
47. ibuprofen.mp. or exp Ibuprofen/
48. ketoprofen.mp. or exp Ketoprofen/
49. naproxen.mp. or exp Naproxen/
50. tiapro$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
51. metamizol.mp. or exp Dipyrone/
52. phenylbutazone.mp. or exp Phenylbutazone/
53. phenazone.mp. or exp Antipyrine/
54. propyphenazone.mp.
55. celecoxib.mp.
56. etoricoxib.mp.
57. nabumeton.mp.
58. parecoxib.mp.
59. or/27-58
60. exp cyclooxygenase inhibitors/ or exp cyclooxygenase 2 inhibitors/
61. ((cyclooxygenase or cyclo-oxygenase) adj3 inhibitor*).mp.
62. rofecoxib.mp.
63. celecoxib.mp.
64. valdecoxib.mp.
65. lumiracoxib.mp.
66. etoricoxib.mp.
67. parecoxib.mp.
68. vioxx.mp.
69. celebrex.mp.
70. bextra.mp.
71. prexige.mp.
72. arcoxia.mp.
73. etodolac.mp. or exp Etodolac/
74. floctafenine.mp.
75. exp Meclofenamic Acid/
76. meclofenamate.mp.
77. meloxicam.mp.
78. oxaprozin.mp.

79. piroxicam.mp. or exp Piroxicam/
80. tenoxicam.mp.
81. tolmetin.mp. or exp Tolmetin/
82. or/60-81
83. 59 or 82
84. 14 and 26 and 83
85. limit 84 to yr=2014-2015
86. limit 84 to ed=20140410-20150624
87. 85 or 86

May 2012 strategy. Line 77 is removed in 2015 (duplicate with line 49).

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. comparative study.pt.
4. clinical trial.pt.
5. randomized.ab.
6. placebo.ab,ti.
7. drug therapy.fs.
8. randomly.ab,ti.
9. trial.ab,ti.
10. groups.ab,ti.
11. or/1-10
12. (animals not (humans and animals)).sh.
13. 11 not 12
14. dorsalgia.ti,ab.
15. exp Back Pain/
16. backache.ti,ab.
17. exp Low Back Pain/
18. (lumbar adj pain).ti,ab.
19. coccyx.ti,ab.
20. coccydynia.ti,ab.
21. sciatica.ti,ab.
22. sciatic neuropathy/
23. spondylosis.ti,ab.
24. lumbago.ti,ab.
25. back disorder$.ti,ab.
26. or/14-25 33294
27. exp Anti-Inflammatory Agents, Non-Steroidal/
28. nsaids.mp.
29. non-steroidal anti inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
30. non-steroidal anti-inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
31. aspirin.mp. or exp Aspirin/
32. acetylsalicyl$.mp.
33. exp Salicylic Acid/
34. carbasalate calcium.mp.
35. diflunisal.mp. or exp Diflunisal/
36. aceclofenac.mp.
37. alclofenac.mp.
38. diclofenac.mp. or exp Diclofenac/
39. indometacin.mp. or exp Indomethacin/
40. sulindac.mp. or exp Sulindac/
41. meloxicam.mp.

42. piroxicam.mp. or exp Piroxicam/
43. dexibuprofen.mp.
44. dexketoprofen.mp.
45. fenoprofen.mp. or exp Fenoprofen/
46. flurbiprofen.mp. or exp Flurbiprofen/
47. ibuprofen.mp. or exp Ibuprofen/
48. ketoprofen.mp. or exp Ketoprofen/
49. naproxen.mp. or exp Naproxen/
50. tiapro$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
51. metamizol.mp. or exp Dipyrone/
52. phenylbutazone.mp. or exp Phenylbutazone/
53. phenazone.mp. or exp Antipyrine/
54. propyphenazone.mp.
55. celecoxib.mp.
56. etoricoxib.mp.
57. nabumeton.mp.
58. parecoxib.mp.
59. or/27-58
60. exp cyclooxygenase inhibitors/ or exp cyclooxygenase 2 inhibitors/
61. rofecoxib.mp.
62. celecoxib.mp.
63. valdecoxib.mp.
64. lumiracoxib.mp.
65. etoricoxib.mp.
66. parecoxib.mp.
67. vioxx.mp.
68. celebrex.mp.
69. bextra.mp.
70. prexige.mp.
71. arcoxia.mp.
72. etodolac.mp. or exp Etodolac/
73. floctafenine.mp.
74. exp Meclofenamic Acid/
75. meclofenamate.mp.
76. meloxicam.mp.
77. naproxen.mp. or exp Naproxen/
78. oxaprozin.mp.
79. piroxicam.mp. or exp Piroxicam/
80. tenoxicam.mp.
81. tolmetin.mp. or exp Tolmetin/
82. or/60-81
83. 59 or 82
84. 13 and 26 and 83
85. limit 84 to yr="2007 - 2012"
86. limit 84 to ed=20070601-20120524
87. 85 or 86

Appendix 3. MEDLINE In-Process & Other Non-Indexed Citations

Last searched 24 June 2015. Line 3 is added, line 6, 27, 37, and 58 are revised.

1. randomized controlled trial.ti,ab.
2. controlled clinical trial.ti,ab.
3. pragmatic.ti,ab.
4. comparative study.ti,ab.
5. clinical trial.ti,ab.
6. randomi#ed.ab.
7. placebo.ab,ti.
8. drug therapy.fs.
9. randomly.ab,ti.
10. trial.ab,ti.
11. groups.ab,ti.
12. or/1-11
13. dorsalgia.ti,ab.
14. Back Pain.ti,ab.
15. backache.ti,ab.
16. (lumbar adj pain).ti,ab.
17. coccyx.ti,ab.
18. coccydynia.ti,ab.
19. sciatica.ti,ab.
20. sciatic neuropathy.ti,ab.
21. spondylosis.ti,ab.
22. lumbago.ti,ab.
23. back disorder$.ti,ab.
24. or/13-23
25. Anti-Inflammatory Agents, Non-Steroidal.mp.
26. nsaids.mp.
27. non-steroidal antiinflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
28. non-steroidal anti-inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
29. aspirin.mp.
30. acetylsalicyl$.mp.
31. Salicylic Acid.mp.
32. carbasalate calcium.mp.
33. diflunisal.mp.
34. aceclofenac.mp.
35. alclofenac.mp.
36. diclofenac.mp.
37. (indomethacin or indometacin).mp.
38. sulindac.mp.
39. meloxicam.mp.
40. piroxicam.mp.
41. dexibuprofen.mp.
42. dexketoprofen.mp.
43. fenoprofen.mp.
44. flurbiprofen.mp.
45. ibuprofen.mp.
46. ketoprofen.mp.
47. naproxen.mp.

48. tiapro$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
49. metamizol.mp.
50. phenylbutazone.mp.
51. phenazone.mp.
52. propyphenazone.mp.
53. celecoxib.mp.
54. etoricoxib.mp.
55. nabumeton.mp.
56. parecoxib.mp.
57. or/25-56
58. ((cyclooxygenase or cyclo-oxygenase) adj3 inhibitor*).mp.
59. rofecoxib.mp.
60. celecoxib.mp.
61. valdecoxib.mp.
62. lumiracoxib.mp.
63. etoricoxib.mp.
64. parecoxib.mp.
65. vioxx.mp.
66. celebrex.mp.
67. bextra.mp.
68. prexige.mp.
69. arcoxia.mp.
70. etodolac.mp.
71. floctafenine.mp.
72. Meclofenamic Acid.mp.
73. meclofenamate.mp.
74. meloxicam.mp.
75. oxaprozin.mp.
76. piroxicam.mp.
77. tenoxicam.mp.
78. tolmetin.mp.
79. or/58-78
80. 57 or 79
81. 12 and 24 and 80
82. limit 81 to yr=2014-2015
83. limit 81 to ed=20140410-20150624
84. 82 or 83

April 2014 search strategy

1. randomized controlled trial.ti,ab.
2. controlled clinical trial.ti,ab.
3. comparative study.ti,ab.
4. clinical trial.ti,ab.
5. randomized.ab.
6. placebo.ab,ti.
7. drug therapy.fs.
8. randomly.ab,ti.
9. trial.ab,ti.
10. groups.ab,ti.
11. or/1-10
12. dorsalgia.ti,ab.
13. Back Pain.ti,ab.
14. backache.ti,ab.

15. (lumbar adj pain).ti,ab.
16. coccyx.ti,ab.
17. coccydynia.ti,ab.
18. sciatica.ti,ab.
19. sciatic neuropathy.ti,ab.
20. spondylosis.ti,ab.
21. lumbago.ti,ab.
22. back disorder$.ti,ab.
23. or/12-22
24. Anti-Inflammatory Agents, Non-Steroidal.mp.
25. nsaids.mp.
26. non-steroidal anti inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
27. non-steroidal anti-inflammat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
28. aspirin.mp.
29. acetylsalicyl$.mp.
30. Salicylic Acid.mp.
31. carbasalate calcium.mp.
32. diflunisal.mp.
33. aceclofenac.mp.
34. alclofenac.mp.
35. diclofenac.mp.
36. indomethacin.mp.
37. sulindac.mp.
38. meloxicam.mp.
39. piroxicam.mp.
40. dexibuprofen.mp.
41. dexketoprofen.mp.
42. fenoprofen.mp.
43. flurbiprofen.mp.
44. ibuprofen.mp.
45. ketoprofen.mp.
46. naproxen.mp.
47. tiapro$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
48. metamizol.mp.
49. phenylbutazone.mp.
50. phenazone.mp.
51. propyphenazone.mp.
52. celecoxib.mp.
53. etoricoxib.mp.
54. nabumeton.mp.
55. parecoxib.mp.
56. or/24-55
57. (cyclooxygenase inhibitors or cyclooxygenase 2 inhibitors).mp.
58. rofecoxib.mp.
59. celecoxib.mp.
60. valdecoxib.mp.
61. lumiracoxib.mp.
62. etoricoxib.mp.
63. parecoxib.mp.
64. vioxx.mp.

65. celebrex.mp.
66. bextra.mp.
67. prexige.mp.
68. arcoxia.mp.
69. etodolac.mp.
70. floctafenine.mp.
71. Meclofenamic Acid.mp.
72. meclofenamate.mp.
73. meloxicam.mp.
74. oxaprozin.mp.
75. piroxicam.mp.
76. tenoxicam.mp.
77. tolmetin.mp.
78. or/57-77
79. 56 or 78
80. 11 and 23 and 79

Appendix 4. EMBASE search strategy

Last searched 24 June 2015. The study design filter, line 38, and line 46 are revised and line 68 is added.

1. Randomized Controlled Trial/ (374656)
2. exp Controlled Clinical Trial/ (511712)
3. Controlled Study/ (4627382)
4. Double Blind Procedure/ (121249)
5. Single Blind Procedure/ (20436)
6. crossover procedure/ (43275)
7. placebo/ (258120)
8. allocat$.mp. (105697)
9. assign$.mp. (262956)
10. blind$.mp. (343130)
11. ((control$ or compar$ or prospectiv$ or clinical) adj25 (trial or study)).mp. (7800092)
12. (crossover or cross-over).mp. (81850)
13. factorial$.mp. (50965)
14. (followup or follow-up).mp. (1253262)
15. placebo$.mp. (339829)
16. random$.mp. (1133643)
17. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp. (222737)
18. volunteer$.mp. (196350)
19. or/1-18 (8994276)
20. exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/ (21299526)
21. human/ or normal human/ or human cell/ (15984909)
22. 20 and 21 (15952556)
23. 20 not 22 (5346970)
24. 19 not 23 (6914940)
25. dorsalgia.mp. (102)
26. back pain.mp. (59723)
27. exp BACKACHE/ (73517)
28. (lumbar adj pain).mp. (1626)
29. coccyx.mp. (800)
30. coccydynia.mp. (120)
31. sciatica.mp. (4597)

32. exp ISCHIALGIA/ (5449)
33. spondylosis.mp. (7198)
34. lumbago.mp. (1454)
35. or/25-34 (93822)
36. exp Nonsteroid Antiinflammatory Agent/ (444580)
37. nsaids.mp. (24138)
38. non-steroidal anti-inflammator$.mp. (16629)
39. exp Acetylsalicylic Acid/ (161086)
40. acetylsalicyl$.mp. (163662)
41. carbasalate calcium.mp. or exp Carbasalate Calcium/ (242)
42. diflunisal.mp. or exp DIFLUNISAL/ (2399)
43. aceclofenac.mp. or exp ACECLOFENAC/ (1287)
44. alclofenac.mp. or exp ALCLOFENAC/ (355)
45. diclofenac.mp. or exp DICLOFENAC/ (32204)
46. exp INDOMETACIN/ or (indometacin or indomethacin).mp. (70465)
47. sulindac.mp. or exp SULINDAC/ (6849)
48. meloxicam.mp. or exp MELOXICAM/ (4723)
49. exp PIROXICAM/ or piroxicam.mp. (10561)
50. dexibuprofen.mp. or exp DEXIBUPROFEN/ (212)
51. dexketoprofen.mp. or exp DEXKETOPROFEN/ (463)
52. exp FENOPROFEN/ or fenoprofen.mp. (2484)
53. flurbiprofen.mp. or exp FLURBIPROFEN/ (6927)
54. ibuprofen.mp. or exp IBUPROFEN/ (39286)
55. ketoprofen.mp. or exp KETOPROFEN/ (10969)
56. naproxen.mp. or exp NAPROXEN/ (22293)
57. tiapro$.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (1330)
58. metamizol.mp. or exp Dipyrone/ (6416)
59. phenylbutazone.mp. or exp PHENYLBUTAZONE/ (11876)
60. phenazone.mp. or exp PHENAZONE/ (5587)
61. exp PROPYPHENAZONE/ or propyphenazone.mp. (829)
62. celecoxib.mp. or exp CELECOXIB/ (17414)
63. etoricoxib.mp. or exp ETORICOXIB/ (2236)
64. exp Nabumetone/ or nabumeton.mp. (1837)
65. parecoxib.mp. or exp PARECOXIB/ (1501)
66. or/36-65 (464519)
67. exp Cyclooxygenase 2 Inhibitor/ (41240)
68. ((cyclooxygenase or cyclo-oxygenase) adj3 inhibitor*).mp. (27816)
69. rofecoxib.mp. or exp ROFECOXIB/ (9957)
70. valdecoxib.mp. or exp VALDECOXIB/ (2464)
71. lumiracoxib.mp. or exp LUMIRACOXIB/ (1046)
72. etoricoxib.mp. or exp ETORICOXIB/ (2236)
73. parecoxib.mp. or exp PARECOXIB/ (1501)
74. vioxx.mp. (2888)
75. celebrex.mp. (2353)
76. bextra.mp. (569)
77. prexige.mp. (174)
78. arcoxia.mp. (276)
79. etodolac.mp. or exp ETODOLAC/ (2403)
80. floctafenine.mp. or exp FLOCTAFENINE/ (216)
81. exp Meclofenamic Acid/ (2319)
82. meclofenam$.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (2769)

83. oxaprozin.mp. or exp OXAPROZIN/ (658)
84. exp PIROXICAM/ or piroxicam.mp. (10561)
85. tenoxicam.mp. or exp TENOXICAM/ (1889)
86. tolmetin.mp. or exp TOLMETIN/ (2406)
87. or/67-86 (62118)
88. 66 or 87 (469269)
89. 24 and 35 and 88 (3792)
90. limit 89 to yr="2014 - 2015" (394)
91. limit 89 to em=201414-201525 (396)
92. 90 or 91 (453)

Study design and animal filter used in the April 2014 search. The animal filter is revised in 2013 and line 31 is revised in 2014.

1 Clinical Article/
2 exp Clinical Study/
3 Clinical Trial/
4 Controlled Study/
5 Randomized Controlled Trial/
6 Major Clinical Study/
7 Double Blind Procedure/
8 Multicenter Study/
9 Single Blind Procedure/
10 Phase 3 Clinical Trial/
11 Phase 4 Clinical Trial/
12 crossover procedure/
13 placebo/
14 or/1-13
15 allocat$.mp.
16 assign$.mp.
17 blind$.mp.
18 (clinic$ adj25 (study or trial)).mp.
19 compar$.mp.
20 control$.mp.
21 cross?over.mp.
22 factorial$.mp.
23 follow?up.mp.
24 placebo$.mp.
25 prospectiv$.mp.
26 random$.mp.
27 ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp.
28 trial.mp.
29 (versus or vs).mp.
30 or/15-29
31 14 or 30
32 exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/
33 human/ or normal human/ or human cell/
34 32 and 33
35 32 not 34
36 31 not 35

May 2012 search strategy

1. Clinical Article/
2. exp Clinical Study/
3. Clinical Trial/
4. Controlled Study/
5. Randomized Controlled Trial/

6. Major Clinical Study/
7. Double Blind Procedure/
8. Multicenter Study/
9. Single Blind Procedure/
10. Phase 3 Clinical Trial/
11. Phase 4 Clinical Trial/
12. crossover procedure/
13. placebo/
14. or/1-13
15. allocat$.mp.
16. assign$.mp.
17. blind$.mp.
18. (clinic$ adj25 (study or trial)).mp.
19. compar$.mp.
20. control$.mp.
21. cross?over.mp.
22. factorial$.mp.
23. follow?up.mp.
24. placebo$.mp.
25. prospectiv$.mp.
26. random$.mp.
27. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp.
28. trial.mp.
29. (versus or vs).mp.
30. or/15-29
31. 14 and 30
32. human/
33. Nonhuman/
34. exp ANIMAL/
35. Animal Experiment/
36. 33 or 34 or 35
37. 32 not 36
38. 31 not 36
39. 37 and 38
40. 38 or 39
41. dorsalgia.mp.
42. back pain.mp.
43. exp BACKACHE/
44. (lumbar adj pain).mp.
45. coccyx.mp.
46. coccydynia.mp.
47. sciatica.mp.
48. exp ISCHIALGIA/
49. spondylosis.mp.
50. lumbago.mp.
51. exp Low Back Pain/
52. or/41-51
53. exp Nonsteroid Antiinflammatory Agent/
54. nsaids.mp.
55. non-steroidal anti-inflammatory.mp.
56. exp Acetylsalicylic Acid/
57. acetylsalicyl$.mp.
58. carbasalate calcium.mp. or exp Carbasalate Calcium/

59. diflunisal.mp. or exp DIFLUNISAL/
60. aceclofenac.mp. or exp ACECLOFENAC/
61. alclofenac.mp. or exp ALCLOFENAC/
62. diclofenac.mp. or exp DICLOFENAC/
63. exp INDOMETACIN/ or indometacin.mp.
64. sulindac.mp. or exp SULINDAC/
65. meloxicam.mp. or exp MELOXICAM/
66. exp PIROXICAM/ or piroxicam.mp.
67. dexibuprofen.mp. or exp DEXIBUPROFEN/
68. dexketoprofen.mp. or exp DEXKETOPROFEN/
69. exp FENOPROFEN/ or fenoprofen.mp.
70. flurbiprofen.mp. or exp FLURBIPROFEN/
71. ibuprofen.mp. or exp IBUPROFEN/
72. ketoprofen.mp. or exp KETOPROFEN/
73. naproxen.mp. or exp NAPROXEN/
74. tiapro$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]
75. metamizol.mp. or exp Dipyrone/
76. phenylbutazone.mp. or exp PHENYLBUTAZONE/
77. phenazone.mp. or exp PHENAZONE/
78. exp PROPYPHENAZONE/ or propyphenazone.mp.
79. celecoxib.mp. or exp CELECOXIB/
80. etoricoxib.mp. or exp ETORICOXIB/
81. exp Nabumetone/ or nabumeton.mp.
82. parecoxib.mp. or exp PARECOXIB/
83. or/53-82
84. exp Cyclooxygenase 2 Inhibitor/
85. rofecoxib.mp. or exp ROFECOXIB/
86. valdecoxib.mp. or exp VALDECOXIB/
87. lumiracoxib.mp. or exp LUMIRACOXIB/
88. etoricoxib.mp. or exp ETORICOXIB/
89. parecoxib.mp. or exp PARECOXIB/
90. vioxx.mp.
91. celebrex.mp.
92. bextra.mp.
93. prexige.mp.
94. arcoxia.mp.
95. etodolac.mp. or exp ETODOLAC/
96. floctafenine.mp. or exp FLOCTAFENINE/
97. exp Meclofenamic Acid/
98. meclofenam$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]
99. oxaprozin.mp. or exp OXAPROZIN/
100. exp PIROXICAM/ or piroxicam.mp.
101. tenoxicam.mp. or exp TENOXICAM/
102. tolmetin.mp. or exp TOLMETIN/
103. or/84-102
104. 83 or 103
105. 40 and 52 and 104
106. limit 105 to yr="2007 - 2012"
107. limit 105 to em=200712-201220 1071
108. 106 or 107

Appendix 5. Search strategies for clinical trials registries and PubMed

ClinicalTrials.gov

Last searched 24 June 2015.
Basic search: "back pain" and NSAIDS, received from 10 April 2014 to 24 June 2015.
May 2012 search strategy.
Condition: back pain AND Intervention: NSAID

WHO ICTRP

Last searched 24 June 2015.
Basic search: back pain and NSAIDS; we reviewed results from 2014 to 2015.
May 2012 search strategy.
Condition: back pain AND Intervention: NSAID

PubMed

Searched 24 June 2015.
((nsaids OR non-steroidal anti-inflammator* OR non-steroidal antiinflammator* OR aspirin OR acetylsalicyl* OR salicylic acid OR carbasalate calcium OR diflunisal OR aceclofenac OR alclofenac OR diclofenac OR indomethacin OR indometacin OR sulindac OR meloxicam OR piroxicam OR dexibuprofen OR dexketoprofen OR fenoprofen OR flurbiprofen OR ibuprofen OR ketoprofen OR naproxen OR tiapro* OR metamizol OR phenylbutazone OR phenazone OR propyphenazone OR celecoxib OR etoricoxib OR nabumeton OR parecoxib OR cyclooxygenase inhibitor* OR cyclo-oxygenase inhibitor* OR rofecoxib OR celecoxib OR valdecoxib OR lumiracoxib OR etoricoxib OR parecoxib OR vioxx OR celebrex OR bextra OR prexige OR arcoxia OR etodolac OR floctafenine OR Meclofenamic Acid OR meclofenamate OR meloxicam OR oxaprozin OR piroxicam OR tenoxicam OR tolmetin) AND (back pain OR sciatica OR lumbar pain OR lumbago OR dorsalgia OR backache OR back disorder*) AND (pubstatusaheadofprint OR publisher[sb] or pubmednotmedline[sb]))

Appendix 6. Criteria for assessing risk of bias for internal validity

Random sequence generation (selection bias)

Selection bias (biased allocation to interventions) due to inadequate generation of a randomized sequence

There is a low risk of selection bias if the investigators describe a random component in the sequence generation process such as: referring to a random number table, using a computer random number generator, coin tossing, shuffling cards or envelopes, throwing dice, drawing of lots, minimization (minimization may be implemented without a random element, and this is considered to be equivalent to being random).
There is a high risk of selection bia*s* if the investigators describe a non-random component in the sequence generation process, such as: sequence generated by odd or even date of birth, date (or day) of admission, hospital or clinic record number; or allocation by judgement of the clinician, preference of the participant, results of a laboratory test or a series of tests, or availability of the intervention.

Allocation concealment (selection bias)

Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment

There is a low risk of selection bias if the participants and investigators enrolling participants could not foresee assignment because investigators used one of the following, or an equivalent method, to conceal allocation: central allocation (including telephone,

internet-based and pharmacy-controlled randomization); sequentially numbered drug containers of identical appearance; or sequentially numbered, opaque, sealed envelopes.

There is a high risk of bias if participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: using an open random allocation schedule (e.g. a list of random numbers); assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or non-opaque or not sequentially numbered); alternation or rotation; date of birth; case record number; or other explicitly unconcealed procedures.

Blinding of participants

Performance bias due to knowledge of the allocated interventions by participants during the study

There is a low risk of performance bias if trial investigators ensured blinding of participants and it was unlikely that the blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is unlikely to be influenced by lack of blinding.

Blinding of personnel/care providers (performance bias)

Performance bias due to knowledge of the allocated interventions by personnel/care providers during the trial

There is a low risk of performance bias if trial investigators ensured blinding of personnel and it was unlikely that blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding.

Blinding of outcome assessor (detection bias)

Detection bias due to knowledge of the allocated interventions by outcome assessors

There is low risk of detection bias if trial investigators ensured the blinding of the outcome assessment and it was unlikely that blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding, or:

- for patient-reported outcomes in which the patient was the outcome assessor (e.g. pain, disability): there is a low risk of bias for outcome assessors if there is a low risk of bias for participant blinding (Boutron 2005)
- for outcome criteria that are clinical or therapeutic events that will be determined by the interaction between patients and care providers (e.g. co-interventions, length of hospitalisation, treatment failure), in which the care provider is the outcome assessor: there is a low risk of bias for outcome assessors if there is a low risk of bias for care providers (Boutron 2005)
- for outcome criteria that are assessed from data from medical forms: there is a low risk of bias if the treatment or adverse effects of the treatment could not be noticed in the extracted data (Boutron 2005).

Incomplete outcome data (attrition bias)

Attrition bias due to amount, nature or handling of incomplete outcome data

There is a low risk of attrition bias if there were no missing outcome data; reasons for missing outcome data were unlikely to be related to the true outcome (for survival data, censoring unlikely to be introducing bias); missing outcome data were balanced in numbers, with similar reasons for missing data across groups; for dichotomous outcome data, the proportion of missing outcomes compared with the observed event risk was not enough to have a clinically relevant impact on the intervention effect estimate; for continuous outcome data, the plausible effect size (difference in means or standardised difference in means) among missing outcomes was not enough to have a clinically relevant impact on observed effect size, or missing data were imputed using appropriate methods (if drop-outs are very large, imputation using even "acceptable" methods may still suggest a high risk of bias) (van Tulder 2003). The percentage

of withdrawals and drop-outs should not exceed 20% for short-term follow-up and 30% for long-term follow-up and should not lead to substantial bias (these percentages are commonly used but arbitrary, not supported by literature) (van Tulder 2003).

Selective reporting (reporting bias)

Reporting bias due to selective outcome reporting

There is low risk of reporting bias if the study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way, or if the study protocol is unavailable but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon).

There is a high risk of reporting bias if not all of the study's pre-specified primary outcomes have been reported; one or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; one or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); one or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; the study report fails to include results for a key outcome that would be expected to have been reported for such a study.

Group similarity at baseline (selection bias)

Bias due to dissimilarity at baseline for the most important prognostic indicators.

There is low risk of bias if groups are similar at baseline for demographic factors, value of main outcome measure(s), and important prognostic factors (examples in the field of back and neck pain are duration and severity of complaints, vocational status, percentage of patients with neurological symptoms) (van Tulder 2003).

Co-interventions (performance bias)

Bias because co-interventions were different across groups

There is low risk of bias if there were no co-interventions or they were similar between the index and control groups (van Tulder 2003).

Compliance (performance bias)

Bias due to inappropriate compliance with interventions across groups

There is low risk of bias if compliance with the interventions was acceptable, based on the reported intensity/dosage, duration, number and frequency for both the index and control intervention(s). For single-session interventions (e.g. surgery), this item is irrelevant (van Tulder 2003).

Intention-to-treat-analysis

There is low risk of bias if all randomized patients were reported/analysed in the group to which they were allocated by randomization.

Timing of outcome assessments (detection bias)

Bias because important outcomes were not measured at the same time across groups

There is low risk of bias if all important outcome assessments for all intervention groups were measured at the same time (van Tulder 2003).

Other bias

Bias due to problems not covered elsewhere in the table

There is a low risk of bias if the study appears to be free of other sources of bias not addressed elsewhere (e.g. study funding).

WHAT'S NEW

Last assessed as up-to-date: 23 June 2015.

Date	Event	Description
12 July 2016	Amended	Data extraction of the Kivitz article was not correct and has been adjusted. Conclusions have not change

HISTORY

Review first published: Issue 2, 2016

Date	Event	Description
24 June 2015	New search has been performed	We added the following databases to the search strategy: ClinicalTrials.gov, the World Health Organization International Clinical Trials Registry Platform (2013) , MEDLINE In-Process & Other Non-Indexed Citations (2014) and PubMed (2015)

CONTRIBUTIONS OF AUTHORS

BW Koes, PDDM Roelofs and WTM Enthoven screened titles and abstracts. WTM Enthoven and PDDM Roelofs performed methodological quality assessments, data extraction and data analyses. WTM Enthoven wrote the initial draft of the manuscript and all review authors critically reviewed the manuscript.

DECLARATIONS OF INTEREST

Wendy TM Enthoven has no known conflicts of interest.

Pepijn DDM Roelofs has no known conflicts of interest.

Richard A Deyo has no known conflicts of interest.

Maurits W van Tulder has no known conflicts of interest.

Bart W Koes has no known conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- Dutch Arthritis Foundation, Netherlands.

This Cochrane review is partly funded by a Dutch Arthritis Foundation programme grant.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We excluded NSAIDs which are no longer available on the market, such as rofecoxib, from this Cochrane review. We had not previously stated this in the Cochrane protocol.

INDEX TERMS

Medical Subject Headings (MeSH)

Anti-Inflammatory Agents, Non-Steroidal [adverse effects; *therapeutic use]; Chronic Pain [*drug therapy]; Diclofenac [therapeutic use]; Disability Evaluation; Ibuprofen [therapeutic use]; Indomethacin [therapeutic use]; Low Back Pain [*drug therapy]; Pain Measurement; Piroxicam [therapeutic use]; Randomized Controlled Trials as Topic

MeSH check words

Humans