Cochrane Library

**Cochrane** Database of Systematic Reviews

# Opioids compared to placebo or other treatments for chronic low-back pain (Review)

Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC

Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC.
Opioids compared to placebo or other treatments for chronic low-back pain.
*Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No.: CD004959.
DOI: 10.1002/14651858.CD004959.pub4.

**www.cochranelibrary.com**




WILEY

# TABLE OF CONTENTS

[Intervention Review]

# Opioids compared to placebo or other treatments for chronic low-back pain

Luis Enrique Chaparro[1], Andrea D Furlan[2], Amol Deshpande[3], Angela Mailis-Gagnon[4], Steven Atlas[5], Dennis C Turk[6]

[1]Anesthesiology Department, Hospital Pablo Tobon Uribe, Medellin, Colombia. [2]Institute for Work & Health, Toronto, Canada. [3]TWH-Comprehensive Pain Unit, University Health Network, Toronto, Canada. [4]Department of Medicine, Toronto Western Hospital Comprehensive Pain Program, Toronto, Canada. [5]Medical Practices Evaluation Center, Massachusetts General Hospital, Boston, MA, USA. [6]Department of Anesthesiology and Pain Medicine, University of Washington, Seattle, Washington, USA

**Contact:** Luis Enrique Chaparro, Anesthesiology Department, Hospital Pablo Tobon Uribe, Medellin, Colombia. luisdr74@yahoo.com.

**Editorial group:** Cochrane Back and Neck Group.
**Publication status and date:** Edited (no change to conclusions), published in Issue 5, 2014.

**Citation:** Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC. Opioids compared to placebo or other treatments for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No.: CD004959. DOI: 10.1002/14651858.CD004959.pub4.



## ABSTRACT

### Background

The use of opioids in the long-term management of chronic low-back pain (CLBP) has increased dramatically. Despite this trend, the benefits and risks of these medications remain unclear. This review is an update of a Cochrane review first published in 2007.

### Objectives

To determine the efficacy of opioids in adults with CLBP.

### Search methods

We electronically searched the Cochrane Back Review Group's Specialized Register, CENTRAL, CINAHL and PsycINFO, MEDLINE, and EMBASE from January 2006 to October 2012. We checked the reference lists of these trials and other relevant systematic reviews for potential trials for inclusion.

### Selection criteria

We included randomized controlled trials (RCTs) that assessed the use of opioids (as monotherapy or in combination with other therapies) in adults with CLBP that were at least four weeks in duration. We included trials that compared non-injectable opioids to placebo or other treatments. We excluded trials that compared different opioids only.

### Data collection and analysis

Two authors independently assessed the risk of bias and extracted data onto a pre-designed form. We pooled results using Review Manager (RevMan) 5.2. We reported on pain and function outcomes using standardized mean difference (SMD) or risk ratios with 95% confidence intervals (95% CI). We used absolute risk difference (RD) with 95% CI to report adverse effects.

### Main results

We included 15 trials (5540 participants). Tramadol was examined in five trials (1378 participants); it was found to be better than placebo for pain (SMD -0.55, 95% CI -0.66 to -0.44; *low quality evidence*) and function (SMD -0.18, 95% CI -0.29 to -0.07; *moderate quality evidence*). Transdermal buprenorphine (two trials, 653 participants) may make little difference for pain (SMD -2.47, 95%CI -2.69 to -2.25; *very low quality evidence*), but no difference compared to placebo for function (SMD -0.14, 95%CI -0.53 to 0.25; *very low quality evidence*). Strong opioids (morphine, hydromorphone, oxycodone, oxymorphone, and tapentadol), examined in six trials (1887 participants), were better

than placebo for pain (SMD -0.43, 95%CI -0.52 to -0.33; *moderate quality evidence*) and function (SMD -0.26, 95% CI -0.37 to -0.15; *moderate quality evidence*). One trial (1583 participants) demonstrated that tramadol may make little difference compared to celecoxib (RR 0.82, 95% CI 0.76 to 0.90; *very low quality evidence*) for pain relief. Two trials (272 participants) found no difference between opioids and antidepressants for either pain (SMD 0.21, 95% CI -0.03 to 0.45; *very low quality evidence*), or function (SMD -0.11, 95% -0.63 to 0.42; *very low quality evidence*). The included trials in this review had high drop-out rates, were of short duration, and had limited interpretability of functional improvement. They did not report any serious adverse effects, risks (addiction or overdose), or complications (sleep apnea, opioid-induced hyperalgesia, hypogonadism). In general, the effect sizes were medium for pain and small for function.

**Authors' conclusions**

There is some evidence (*very low to moderate quality*) for short-term efficacy (for both pain and function) of opioids to treat CLBP compared to placebo. The very few trials that compared opioids to non-steroidal anti-inflammatory drugs (NSAIDs) or antidepressants did not show any differences regarding pain and function. The initiation of a trial of opioids for long-term management should be done with extreme caution, especially after a comprehensive assessment of potential risks. There are no placebo-RCTs supporting the effectiveness and safety of long-term opioid therapy for treatment of CLBP.

## PLAIN LANGUAGE SUMMARY

**Opioids for the treatment of chronic low-back pain**

**Review question**

We reviewed the evidence about the effect of opioids on pain and function among people with chronic low-back pain (CLBP).

**Background**

Opioids are pain relievers that act on the central nervous system. People with low-back pain (LBP) use these drugs to relieve pain. We examined whether the use of opioids for at least four weeks was better or worse than other treatments of CLBP.

**Study characteristics**

We searched for trials, both published and unpublished, up to October 2012. We included fifteen trials which included 5540 participants and compared opioids against a placebo (fake medication) or other drugs that have been used for LBP. Most people included in the trials were aged 40 to 50 years and all reported at least moderate pain across the low-back area. The trials included a slightly higher proportion of women. Most of the trials followed the patients during three months and were supported by the pharmaceutical industry.

**Key results**

In general, people that received opioids reported more pain relief and had less difficulty performing their daily activities in the short-term than those who received a placebo. However, there is little data about the benefits of opioids based on objective measures of physical functioning. We have no information from randomized controlled trials supporting the efficacy and safety of opioids used for more than four months. Furthermore, the current literature does not support that opioids are more effective than other groups of analgesics for LBP such as anti-inflammatories or antidepressants.

This review partially supports the effectiveness of several opioids for CLBP relief and function in the short-term. However, the effectiveness of prescribing of these medications for long-term use is unknown and should take into consideration the potential for serious adverse effects, complications, and increased risk of misuse, abuse, addiction, overdose, and deaths.

As expected, side effects are more common with opioids but non-life-threatening with short-term use. Insufficient data prevented making conclusions about the side-effect profile of opioids versus other type of analgesics (for example, antidepressants or anti-inflammatories).

**Quality of the evidence**

The quality of evidence in this review ranged between "very low" and "moderate". The review results should be interpreted with caution and may not be appropriate in all clinical settings. High quality randomized trials are needed to address the long-term (months to years) risks and benefits of opioid use in CLBP, their relative effectiveness compared with other treatments, and to better understand which people may benefit most from this type of intervention.

## Summary of findings for the main comparison. Strong opioids compared to placebo for chronic low-back pain

**Strong opioids compared to placebo for chronic low-back pain**

**Patient or population:** people with chronic low-back pain
**Settings: Outpatient pain management**
**Intervention:** strong opioids compared to placebo

| Outcomes | Illustrative comparative risks* (95% CI) | | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Assumed risk | Corresponding risk | | | | |
| | Control | Strong opioids compared to placebo | | | | |
| **Mean pain intensity** on a numerical scale. For example<br><br>0 (no pain) to 10 (maximum pain) | * The baseline for the most representative study (Buynak 2010) is 7.6 (SD 1.33) | The mean mean pain intensity in the intervention groups was<br><br>**0.43 standard deviations lower** (0.52 to 0.33 lower) | Not applicable | 1887 (6 studies) | ⊕⊕⊕⊖ **moderate** [1,2,3] | The magnitude of this difference is in the range of small to moderate. |
| **Disability** (higher ratings mean greater disability).<br><br>Various instruments were used, for example 0% (no disability) to 100% (bed-ridden) | * The baseline for the most representative study (Buynak 2010) is 6.7 (SD 1.61). | The mean disability (higher ratings mean greater disability) in the intervention groups was<br><br>**0.26 standard deviations lower** (0.37 to 0.15 lower) | Not applicable | 1375 (4 studies) | ⊕⊕⊕⊖ **moderate** [2,4,5] | The magnitude of this difference is small. |
| **At least 30% of pain relief or moderate relief** | **Study population** | | **OR 1.91** (1.41 to 2.58) | 819 (3 studies) | ⊕⊕⊕⊖ **moderate** [6,7] | The magnitude of this OR is large. |
| | **327 per 1000** | **481 per 1000** (406 to 556) | | | | |
| **At least 50% of pain relief** | **Study population** | | **OR 1.89** (1.34 to 2.66) | 750 (2 studies) | ⊕⊖⊖⊖ **very low** [8,9,10] | The magnitude of this OR is large. |
| | **236 per 1000** | **369 per 1000** (293 to 451) | | | | |
| **Side effects - Somnolence** | **Study population** | | | 2346 (5 studies) | ⊕⊖⊖⊖ **very low** [10,11] | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **25 per 1000** | **86 per 1000** (45 to 125) | **RD: 6%** (2% to 10%) | | | This difference is not clinically important (< 10%). |
| **Side effects - Nausea** | **Study population** | | **RD: 12%** (5% to 19%) | 2346 (5 studies) | ⊕⊕⊖⊖ **low** [11] | This difference is clinically important (> 10%). |
| | **102 per 1000** | **223 per 1000** (151 to 291) | | | | |
| **Side effects - Constipation** | **Study population** | | **RD: 11%** (4% to 19%) | 2346 (5 studies) | ⊕⊖⊖⊖ **very low** [10,11,12] | This difference is clinically important (> 10%). |
| | **36 per 1000** | **148 per 1000** (76 to 226) | | | | |

* Of the included trials for this outcome, we chose the study that is a combination of the most representative study population and has the largest weighting in the overall result in Revman (Buynak 2010). This figure represents the baseline mean in the control group of this particular study.
**CI:** Confidence interval; **RD:** Risk difference; **OR:** Odds ratio; **SD:** Standard deviation.

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] Four trials had low risk of selection bias and one trial (Chu 2012) was unclear. All five trials had low risk of performance bias, and low risk of reporting bias. However, all five trials suffered from high risk of attrition bias, and some trials also had high risk of detection bias because it was unclear if the outcome assessor were blinded.
[2] $I^2 = 0\%$
[3] See Figure 1. Funnel plot could not demonstrate bias.
[4] Selection bias: three trials low risk of bias, and one trial unclear (Chu 2012). All four trials had low risk of performance bias. Detection bias was unclear in 3 trials, except Khoromi 2007. Attrition bias was judged high in all four trials. Reporting bias was not a problem in any trial.
[5] See Figure 2. Funnel plot could not demonstrate bias.
[6] All trials had risk of attrition bias and performance bias. One trial was unclear about randomization method.
[7] Total number of events was 335.
[8] Both trials had high risk of attritiion bias. Both trials are unclear about performance bias. One trial was not clear about method of randomization.
[9] Heterogeneity: $Chi^2 = 5.39$, df = 1 (P = 0.02); $I^2 = 81\%$.
[10] Total number of events was < 300.
[11] All trials had high risk of attrition bias. Most trials had a problem with performance bias, and one trial was not clear about method of randomization.
[12] Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 33.50$, df = 4 (P < 0.00001); $I^2 = 88\%$.

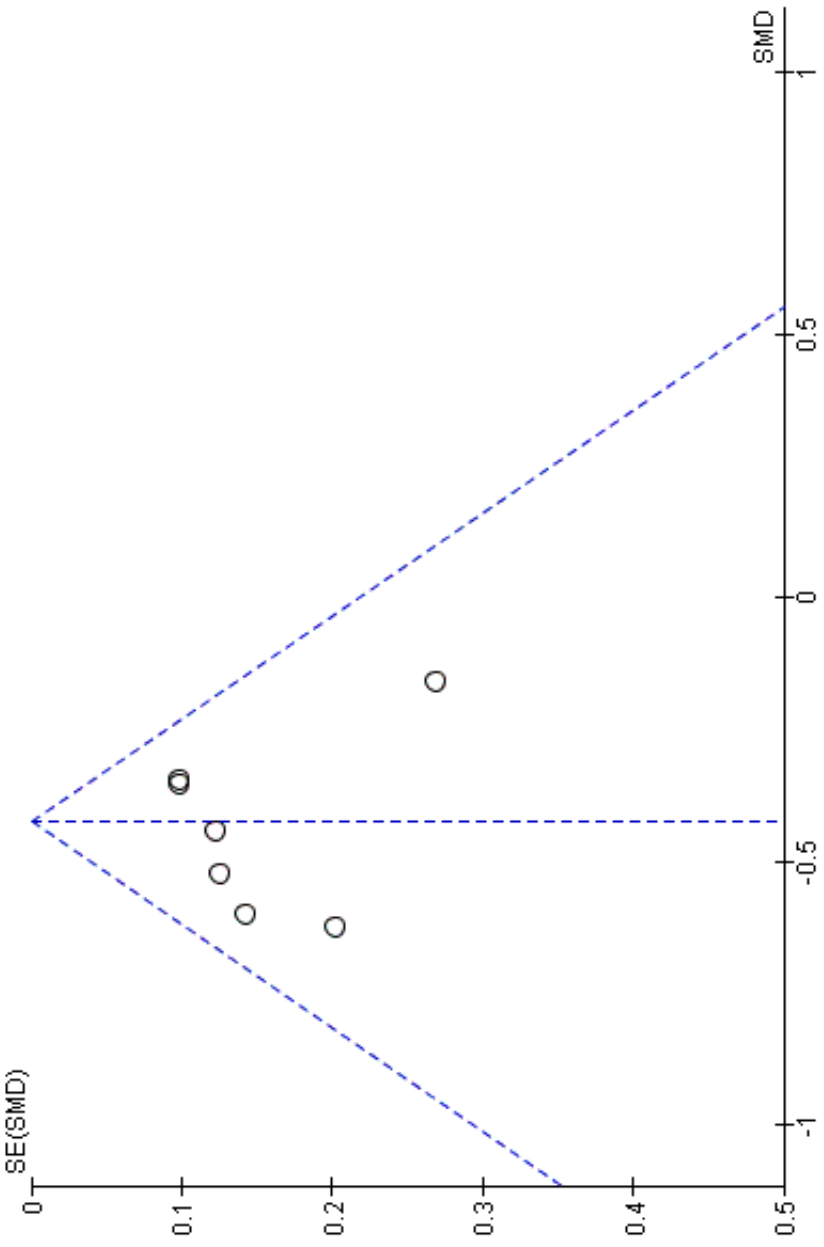



**Figure 1. Funnel plot of comparison: 3 Strong opioids compared to placebo, outcome: 1.3 Mean pain intensity.**

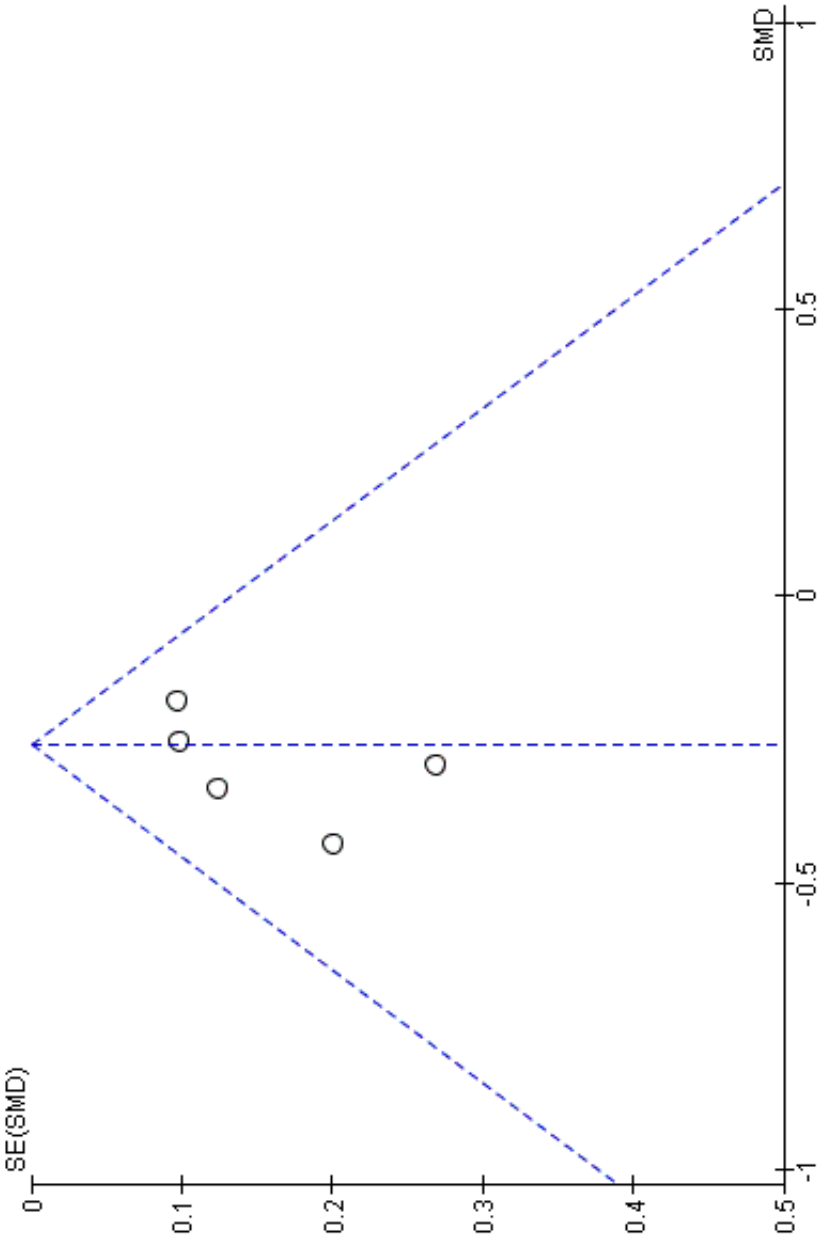



**Figure 2. Funnel plot of comparison: 3 Strong opioids compared to placebo, outcome: 3.4 Disability.**

# BACKGROUND

## Description of the condition

Low back pain (LBP) is the main cause of pain, disability, social and financial cost throughout the world (Volinn 1997; Vos 2012). Approximately 80% of people will experience at least one episode of acute back pain during their lifetime (Cassidy 1998). Almost one quarter of North Americans are estimated to have experienced an episode of LBP within the previous three months (Deyo 2006). Although an early review concluded that 80% to 90% of people with chronic low back pain (CLBP) improve by 12 weeks (Shekelle 1995), a proportion continue to report symptoms over several months and even years. In one study, one-third of people with CLBP continued to be symptomatic after 12 months (Thomas 1999). More recent reviews suggest that the prevalence of LBP is around 23% (Vos 2012). Moreover, a substantial proportion of people with back pain will have recurrences even after the resolution of initial symptoms (Von Korff 1996).

### CLBP and functionality

LBP is the main cause of disability-adjusted life years (DALYs) worldwide (Vos 2012) and the prevalence of CLBP-related disability is estimated at 11% (Vos 2012). Individuals with CLBP not only experience personal distress, but also present with significant sleep disorders and disability (Gore 2012). According to an early study (Spitzer 1987), fewer than 50% of individuals with CLBP who missed work for more than 12 weeks actually returned to work. An absence of two years from employment was associated with almost no chance of returning to work.

## Description of the intervention

### Therapeutic options

The vast majority of CLBP treatments are directed towards symptomatic and functional improvement rather than cure. Patients may be offered a variety of treatment regimens as either monotherapy or a combination of therapies. Treatments may include medications and physical modalities (for example, transcutaneous electrical nerve stimulation (TENS), massage therapy, work hardening), rehabilitation, or injection therapy (such as, epidurals, facet joint blocks, and trigger point injections) that are directed specifically at potential anatomic causes for CLBP. A proportion of individuals with CLBP will undergo surgery to alleviate their symptoms. Despite general acceptance of lumbar discectomy, with or without decompression, and lumbar fusion (with or without instrumentation), the actual success rates for symptomatic and functional improvement have been variable, with surgical 'failure' rates estimated between 10% and 40% (Fritsch 1996; Ostelo 2003). Furthermore, the results are similar with surgery or pharmacological therapy (Peul 2007). These individuals often return to the pool of patients with CLBP, and they often experience poor outcomes regardless of future treatment. Medications play an important role in the management of CLBP and generally fall into four broad categories: non-steroidal anti-inflammatory drugs (NSAIDs), antidepressants, muscle relaxants, and analgesics including opioids.

Opioids are generally classified as either weak or strong. These terms refer to relative efficacy rather than potency; weak opioids exhibit a ceiling to their analgesic effect, limited principally by increased adverse reactions. The use of opioids remains a controversial issue in the management of chronic non-cancer pain (CNCP) (Furlan 2010), and CLBP in particular (Turk 2011). The American College of Physicians & The American Pain Society consensus guidelines for the treatment of LBP recommend opioids for the short-term management of severe and disabling LBP that has had no response with anti-inflammatories or acetaminophen. Notably, this guideline was published in 2007 and includes only a few trials (Chou 2007). In contrast, the American Geriatrics Society Guidelines have suggested that given the problems of NSAIDs and cyclooxygenase-2 (COX-2) inhibitors, opioids should be considered first line treatment for moderate-to-severe pain in older adults (Ferrell 2009). However, recent evidence links the abuse of opioids to negative social consequences (Bohnert 2011).

### Controversies with use of opioids

Although many clinicians believe that opioids offer a valuable tool in the management of CNCP, there is still a large group of practitioners who remain hesitant, or even opposed to, the use of these medications. A survey of Canadian physicians exploring attitudes towards opioid use for chronic pain confirmed that 35% of general practitioners and 23% of palliative care physicians would never use opioids for the management of severe CNCP (Morley-Foster 2003). A recent study of opioid prescribing stratified across the United States by region and by medical specialty found that 41.5% of respondents prescribed long-term opioids in fewer than 20% of their CNCP patients (Wilson 2013). Clinicians reluctant to prescribe opioids to treat people with CNCP believe that side-effects (Wilson 2013), somnolence resulting in poor function, the risk of abuse (Von Korff 2004), and general ineffectiveness of opioids may outweigh any potential benefit. Several trials have demonstrated that rather than underlying pathology, characteristics such as age, depression, personality disorder, and substance abuse, distinguished patients with CLBP who were on opioids from those who were receiving non-opioid treatments (Turk 1997; Breckenridge 2003; Edlund 2007). These trials continue to contribute to the confusion and uncertainty regarding the indications and actual benefits of opioids in CLBP. A recent survey among Canadian primary care physicians revealed that the most common fears for opioid prescription were abuse, overdose, and early prescription renewals (Wenghofer 2011).

## How the intervention might work

Current evidence suggests that opioids are effective for the treatment of CNCP in the short-term (Furlan 2010), irrespective of somatic or neuropathic etiology. The diverse mechanisms of action of opioids across the central and peripheral nervous system can be the reason for unpredictable responses to these medications. More importantly, they can lead to the potential development of adverse effects, including development of addictive behaviour.

## Why it is important to do this review

This is an update of a Cochrane review that was published in 2007 (Deshpande 2007). The original review included only four RCTs. Three of the trials included tramadol and a fourth trial evaluated morphine and oxycodone in an open-label fashion.

# OBJECTIVES

Our primary objective was to determine whether opioids were effective in improving pain, or function, or both, in individuals with CLBP.

Our secondary objectives were to determine the effectiveness of opioids in:

1. Patients with CLBP with or without prior spinal surgery;
2. Patients with CLBP with or without radicular symptoms (patients with symptoms radiating into the buttock or leg irrespective of radiological or electrophysiological evidence);
3. Patients with CLBP managed with tramadol;
4. Patients with CLBP managed with transdermal buprenorphine;
5. Patients with CLBP managed with strong opioids.

# METHODS

## Criteria for considering studies for this review

### Types of studies

We included published RCTs with a blinded assessment of outcomes that compared any opioid to placebo or any other drug with analgesic properties. We had no restriction on the language of the publication.

### Types of participants

#### *Inclusion criteria*

We included male and female participants, aged 18 years or older, that had persistent pain in the low-back for at least 12 weeks, with or without radiating symptoms to the legs or prior low-back surgery (failed back surgery syndrome).

We defined LBP as pain occurring below the lower ribs and above the gluteal folds, including the buttocks. We defined failed back surgery syndrome as back pain, leg pain, or both, lasting longer than six months from the date of surgical intervention, or pain that began prior to one year from the date of intervention, after the individual had achieved symptomatic relief.

#### *Exclusion criteria*

We excluded patients with cancer, infections, inflammatory arthritic conditions (including osteoarthritis [OA]) or compression fractures. We also excluded trials where < 50% of participants had CLBP or study authors failed to report results separately for this specific cohort

### Types of interventions

We included trials that examined the use of any opioid prescribed in an outpatient setting, for a period of one month or longer. We considered trials with opioids given by oral, transdermal, mucosal (nasal or rectal), or intramuscular routes, administered either alone or in combination with other interventions, such as: pharmacological therapy (for example, anti-inflammatories, antidepressants, sedatives), physical modalities (for example, TENS, chiropractic), exercise, or alternative pain management techniques (for example, acupuncture).

We required opioids to be prescribed for a period of one month or longer to provide relevant feedback to the clinician and identify trials that may simulate actual clinical practice patterns. We excluded trials that examined opioids given by intravenous route, including implantable pumps, due to the invasive nature of the therapy and its limited clinical relevance in the outpatient setting. We did not assess the effectiveness of opioids used in neuraxial implantable pumps as this has been discussed elsewhere (Noble 2008).

We considered trials with the following comparisons:

1. Opioids compared to placebo;
2. Opioids compared to no treatment;
3. Opioids compared to non-pharmacological treatments;
4. Opioids compared to other pharmacological agents, alone or in combination (for example, NSAIDs, muscle relaxants, anti-depressants);
5. Opioids given in combination with other pharmacological agents (for example, NSAIDs, muscle relaxants, anti-depressants) or non-pharmacological treatments compared to other pharmacological or non-pharmacological treatments, either alone or in combination.

We excluded trials where comparisons were made between opioids.

### Types of outcome measures

#### *Primary outcomes*

Trials must have reported on at least one of four primary outcome measures for efficacy:

1. Pain ratings: verbal rating scale, visual analog scale or final visit pain score.
2. Function: Oswestry Disability Index (ODI), Roland-Morris Disability Questionnaire (RMDQ) or Quebec Back Pain Disability Scale (QBPDS).
3. Global improvement: patient satisfaction or quality of life improvements.
4. Proportion of patients reporting 30% or 50% pain relief.

#### *Secondary outcomes*

1. Work-related disability: time on compensation, return-to-work, or productivity.
2. Treatment-related adverse effects.
3. Others: healthcare usage, non-opioid medication consumption, addiction, or overdose-related events.

We grouped outcome measures according to the timing of post-randomization follow-up: very short-term (less than one month), short-term (between one and three months), intermediate (greater than three but less than six months) and long-term (longer than six months).

## Search methods for identification of studies

We searched the following databases for relevant trials: MEDLINE (OVID) 1966 to Oct 2012; EMBASE (OVID) 1980 to Oct 2012; Cochrane Library, Central Register of Controlled Trials (Wiley) 2012, Issue 10; PsycINFO (OVID) 1967 to Oct 2012; and CINAHL (Ebsco) 1982 to Oct 2012. We performed electronic searches with the assistance of an experienced librarian, using the sensitive searches recommended by the Cochrane Back Review Group (Furlan 2009). We have presented the search strategy for MEDLINE in Appendix 1. We adapted this search strategy as indicated to search the other databases (see Appendix 2). We examined references provided in the trials we identified from the database search and relevant

systematic reviews for further trials. We also tracked the citations of identified relevant trials.

## Data collection and analysis

We followed the methods recommended by the Cochrane Back Review Group (Furlan 2009).

### Selection of studies

Two teams of two authors each (AF and LEC; LEC and AD) independently screened titles, abstracts, and keywords of trials that we identified by the search strategies to determine if the references met the inclusion criteria. We obtained the full text of trials that either appeared to meet criteria or for which we considered their inclusion was uncertain. We screened these articles for inclusion and we resolved any disagreements through discussion.

### Data extraction and management

Three authors (LEC, AD, AF) independently extracted data, using the standardized forms developed by the Cochrane Back Review Group, on characteristics of participants, intervention group, clinical setting, method of recruitment, interventions, primary and secondary outcomes, opioid abuse or addiction, side effects, country of study, and sponsorship of study. If data were not available in a format that was appropriate for data extraction, we contacted the authors of the trial for further clarification. We resolved any disagreements through discussion.

### Data synthesis

Given the similarities in populations, methodology, interventions and outcomes, we pooled data from trials comparing opioids to placebo (using Review Manager (RevMan)). We performed meta-analyses (both fixed-effect and random-effects methods) on the outcomes of pain, function, and side effects. If we noted a significant statistical discrepancy between methods, we reported the more conservative result. We reported the results of pain and function from the pooled data as standardized mean difference (SMD) with a 95% confidence interval (CI). We reported side effects using absolute risk differences (RD) with a 95% CI.

We used the GRADE approach, as recommended by the Cochrane Collaboration (Higgins 2011) and by the updated Cochrane Back Review Group method guidelines (Furlan 2009). Following GRADE guidelines, we categorized the quality of evidence as follows:

· High: further research is very unlikely to change the confidence in the estimate of effect.

· Moderate: further research is likely to have an important impact in the confidence in the estimate of effect.

· Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

· Very low: any estimate of effect is very uncertain.

We graded the evidence available on specific domains as follows:

#### *1. Study design*

In this review we only included randomized, controlled, double-blinded trials.

#### *2. Risk of bias*

Three authors (LC, AD and AF) independently evaluated the risk of bias of the selected articles, based on criteria described in the Cochrane Back Review Group's updated methods guidelines (Furlan 2009). We scored each criterion as "low risk", "high risk", or "unclear". We have presented the description of this evaluation in Appendix 3. We examined all trials for five types of bias:

1. Selection (random sequence generation, allocation concealment, group similarities at baseline)
2. Performance (blinding of participants, blinding of healthcare providers, co-interventions, and compliance with intervention)
3. Attrition (dropouts and intention-to-treat (ITT) analysis)
4. Measurement (blinding of the outcome assessors and timing of outcome assessment)
5. Reporting bias (selective reporting)

We used the overall risk of bias for each trial in the GRADE synthesis. When all trials were judged as "low risk of bias" for all five categories, we did not downgrade the evidence. We downgraded the evidence by one point when less than three categories were judged "high or unclear". We downgraded the evidence by 2 points when four or more categories were judged "high or unclear".

#### *3. Inconsistency*

Inconsistency refers to an unexplained heterogeneity of results. Widely differing estimates of the treatment effect (such as, heterogeneity or variability in results) across trials suggest true differences in underlying treatment effect. Inconsistency may arise from differences in: populations (for example, drugs may have larger relative effects in sicker populations), interventions (for example, larger effects with higher drug doses), or outcomes (for example, diminishing treatment effect with time). This item does not apply when there is only one trial. We downgraded the quality of evidence by one point when the heterogeneity or variability in results was large (for example: $I^2 > 80\%$). We downgraded by two points when the heterogeneity or variability in results was large and there was inconsistency arising from populations, interventions, or outcomes.

#### *4. Indirectness*

We assessed whether the question being addressed in this systematic review was different from the available evidence regarding the population, intervention, comparator, or an outcome. We downgraded the quality of evidence by one point when there was indirectness in only one area; and by two levels when there was indirectness in two or more areas.

#### *5. Imprecision*

Results are imprecise when trials include relatively few patients and few events and thus have wide CIs around the estimate of the effect.

For dichotomous outcomes, we considered imprecision for either of the following two reasons:

(1) There was only one trial. When there was more than one trial, the total number of events was < 300 (a threshold rule-of-thumb value) (Mueller 2007).

(2) 95% CI around the pooled or best estimate of effect included both (1) no effect and (2) appreciable benefit or appreciable harm. The threshold for "appreciable benefit" or "appreciable harm" is a relative risk reduction (RRR) or relative risk increase (RRI) > 25%. We downgraded the quality of the evidence by one point when there was imprecision due to (1) or (2); or by two levels when there was imprecision due to (1) and (2).

For continuous outcomes, we considered imprecision for either of the following two reasons:

(1) There was only one trial. When there was more than one trial, the total population size was < 400 (a threshold rule-of-thumb value; using the usual α and β, and an effect size of 0.2 SD, representing a small effect).

(2) 95% CI included no effect and the upper or lower CI crosses an effect size (standardized mean difference) of 0.5 in either direction. We downgraded the quality of the evidence by one point when there was imprecision due to (1) or (2); or by two points when there was imprecision due to (1) and (2).

#### *6. Publication bias*

Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of trials. We downgraded the quality of the evidence by one point when the funnel plot suggested publication bias.

#### *7. Magnitude of the effect*

We did not assess this in the review.

#### *8. Dose response gradient*

We did not assess this in the review.

#### *9. Influence of all plausible residual confounding*

We did not assess this in the review.

We prepared the summary of findings tables following published guidelines from the Cochrane Collaboration (Higgins 2011).

We used GRADEprofiler 3.6 to prepare the GRADE tables and Summary of Findings Tables.

# RESULTS

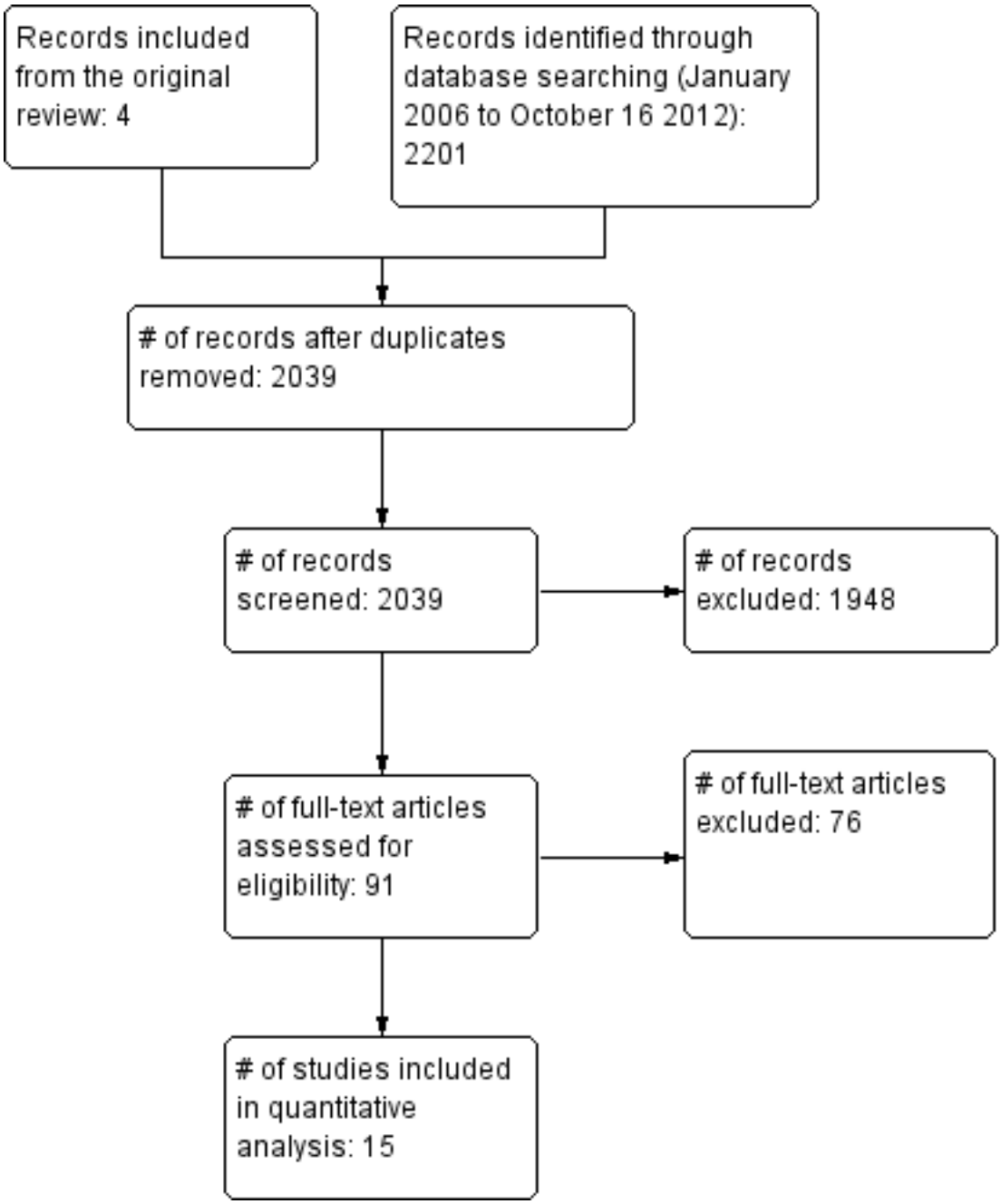
## Description of studies

We have listed this information in the Characteristics of included studies and Characteristics of excluded studies tables.

### Results of the search

We identified 2201 references through the literature search, which three authors (LC, AF, AD) screened by title and abstract. We obtained full-text articles for 91 studies and 12 RCTs met the inclusion criteria. We included three of the four trials that we included in the original review; we excluded one study (Jamison 1998) because it was not blinded. In total, we included 15 trials (5540 participants) in this review (Figure 3).

**Figure 3. Study flow diagram.**




We searched for trials registered on the clinicaltrials.gov website and we identified one trial evaluating the combination oxycodone + naltrexone (NCT01571362); one trial for hydromorphone (NCT01455519); one trial for the combination oxycodone naloxone (NCT01358526); and two trials using hydrocodone (NCT01789970 - NCT01081912) (see Characteristics of ongoing studies).

### Included studies

#### *Study identification*

We included 15 RCTs in this review update. Six trials evaluated either tramadol alone (Schnitzer 2000; Vorsanger 2008; O'Donnell 2009; Uberall 2012) or the combination tramadol/acetaminophen (Ruoff 2003; Peloso 2004). One trial evaluated a drug with a similar mechanism of action, tapentadol (Buynak 2010). Two trials focused on morphine (Khoromi 2007; Chu 2012); two on oxymorphone (Hale 2007; Katz 2007); two trials investigated the effect of transdermal buprenorphine (Gordon 2010; Steiner 2011); and two papers evaluated the effectiveness of oxycodone (Webster 2006) or hydromorphone (Hale 2010) for CLBP.

All included trials were performed in a placebo-controlled fashion, except for one publication that reported two trials with identical methodology and used celecoxib in the control arm (O'Donnell 2009). All trials were conducted in the United States (Buynak 2010; Chu 2012; Hale 2007; Hale 2010; Katz 2007; Khoromi

2007; O'Donnell 2009; Peloso 2004; Ruoff 2003; Schnitzer 2000; Steiner 2011; Vorsanger 2008; Webster 2006), except for one in Canada (Gordon 2010) and another in Germany (Uberall 2012). We summarized the study characteristics of included trials in the Characteristics of included studies section.

**Tramadol compared to placebo**

Five RCTs, including 1378 participants, examined the use of tramadol compared to placebo (Schnitzer 2000; Ruoff 2003; Peloso 2004; Vorsanger 2008; Uberall 2012). Uberall 2012 used tramadol as the active control arm and evaluated the efficacy of flupirtine (a centrally-acting, non-opioid agent) for treatment of people with CLBP.

The five trials were similar in their reported demographics including age (mean age ranged between 47.1 (Schnitzer 2000) and 58.5 (Uberall 2012)); sex (female = 64.1% (Peloso 2004); female = 63.2% (Ruoff 2003); female = 50% (Schnitzer 2000); female = 62% (Uberall 2012); female: 50% (Vorsanger 2008)) and ethnicity (all involved a large number of Caucasian participants (> 85% of the randomized population)). Two trials included patients with previous low-back surgery if it was performed more than five years previously, but only if it was associated with complete pain relief (Schnitzer 2000; Peloso 2004). However, failed back surgery pain was an exclusion criteria across the other trials. All trials also excluded patients with pain in areas other than the low-back and individuals with a past history of substance abuse. Patient history prior to enrolment, including factors such as number of patients actively employed, status of compensation or the average duration of pain before entry into the trial, was not stated.

In two RCTs, tramadol was combined with acetaminophen (paracetamol) (Ruoff 2003; Peloso 2004). The average daily dose of tramadol was approximately 150 mg (Ruoff 2003; Peloso 2004), 242 mg (Schnitzer 2000), 200 mg (Uberall 2012) and 200 to 300 mg/day (Vorsanger 2008). Three trials had a double-blind phase duration of 90 days (Peloso 2004; Ruoff 2003; Vorsanger 2008); whereas two studies were just over four weeks in duration (Schnitzer 2000; Uberall 2012). The included trials did not allow initiation of other treatments during the follow-up periods, although two trials permitted continuation of physiotherapy started prior to inclusion in the trial (Schnitzer 2000; Peloso 2004). None of the included trials documented the number of people receiving concurrent treatments or the types of concurrent treatment they received. Two trials allowed the concomitant use of diclofenac (Uberall 2012) or acetaminophen (Vorsanger 2008).

Each of the included studies included pain intensity as the primary outcome. Three trials used a visual analogue scale (VAS) (Ruoff 2003; Peloso 2004; Vorsanger 2008). In Schnitzer 2000, trial authors used the primary efficacy outcome of "distribution of time to therapeutic failure" (which the trial authors defined as the time to discontinuation of therapy due to inadequate pain relief). They included pain measured with the VAS as a secondary outcome. Uberall 2012 used the change from baseline as the primary outcome. Four trials used the RMDQ to measure functional outcome and Uberall 2012 opted for the pain disability index (PDI).

Two out of five trials employed an enriched enrolment randomized withdrawal trial (Schnitzer 2000; Vorsanger 2008).

**Buprenorphine compared to placebo**

We included two RCTs that compared transdermal buprenorphine versus placebo for treatment of people with CLBP (Gordon 2010; Steiner 2011). Gordon 2010 used a crossover design, each period included a four-weeks follow-up; Steiner 2011 used a 15-week enrichment design including three weeks of open-label titration followed by 12-week randomized and double-blind fashion.

Both studies reported similar demographics including mean age (50.7 (Gordon 2010); 49.4 (Steiner 2011)); and sex (female (Steiner 2011): 55%; female: 60.3% (Gordon 2010)). In Steiner 2011, 70% of the participants were caucasian, while Gordon 2010 did not report the ethnicity of the participants. Both studies had substance abuse as an exclusion criteria. Additionally, Steiner 2011 monitored opioid abuse or diversion behaviour and listed radicular symptoms as an exclusion criteria. Gordon 2010 used concurrent physiotherapy as an exclusion criteria; in contrast, Steiner 2011 allowed physiotherapy if participants started it at least two weeks prior to study entry. The trial authors did not state working or compensation status in the study demographics.

Gordon 2010 used buprenorphine patches of 10 or 20 mcg/hour, up to a maximum dose of 40. Steiner 2011 titrated the dose of buprenorphine from 5 mcg/hour to 20 mcg/hour during the run-in period and maintained a maximum dose of 20 mcg/hour during the double-blind phase. Gordon 2010 allowed participants to use analgesic rescue with acetaminophen and Steiner 2011 allowed participants to use acetaminophen plus ibuprofen and oxycodone IR during the first six weeks of the double-blind phase. Gordon 2010 allowed participants to use non-opioid analgesics (antidepressants or anticonvulsants).

Steiner 2011 defined the primary outcome as the "average pain in the last 24 hours at week 12". Gordon 2010 used daily pain intensity as the primary outcome. The trials authors measured functional status using either the ODI (Steiner 2011) or the QBPD (Gordon 2010).

**Strong opioids (morphine, hydromorphone, oxymorphone, tapentadol or oxycodone) compared to placebo**

We included seven RCTs in this category: Buynak 2010 used tapentadol; two RCTs used morphine (Khoromi 2007; Chu 2012); two RCTs evaluated oxymorphone (Hale 2007; Katz 2007); one RCT assessed hydromorphone (Hale 2010); and one RCT focused on oxycodone (Webster 2006). Notably, three of the seven included RCTs were not designed with the primary objective of demonstrating the effectiveness of the opioid for the treatment of people with CLBP. Webster 2006 aimed to explore an opioid alternative (oxycodone combined with low-dose naltrexone) to avoid physical dependence after long-term treatment. Khoromi 2007 explored the effectiveness of morphine in chronic radicular LBP. Chu 2012 focused on the potential development of opioid tolerance versus opioid-induced hyperalgesia.

In the included trials, the mean age of participants ranged between 45 years (Chu 2012) and 53 years (Khoromi 2007). The proportion of women did not significantly differed across the trials (female = 59.3.1% (Buynak 2010); female = 43.9% (Chu 2012); female = 45.1% (Hale 2007); female = 50.4% (Hale 2010); female: 53.2% (Katz 2007); female: 50% (Khoromi 2007); female: 61.2% (Webster 2006)). The vast majority of the participants were Caucasian. Four trials excluded patients with any history of opioid abuse (Buynak 2010;

Chu 2012; Khoromi 2007; Webster 2006), but three included chronic opioid users (Hale 2007; Katz 2007; Hale 2010). History of failed back surgery pain or LBP that could have some benefit with spine surgery were also exclusion criteria. All RCTs excluded patients with radicular symptoms or neurological abnormalities in the lower extremities, except for Khoromi 2007, which focused on patients with sciatica. All RCTs allowed physiotherapy or physical exercise if participants started at least two weeks prior to the trial start. Only Khoromi 2007 described the work or compensation status of the participants.

The mean dose of opioids was 78 mg morphine (Chu 2012); 62 mg morphine (Khoromi 2007); 100 to 250 mg tapentadol (40 to 100 morphine equivalent) (Buynak 2010); 80.9 mg oxymorphone (243morphine equivalent) (Hale 2007); 39.2 mg oxymorphone (117.6 morphine equivalent) (Katz 2007); 37.8 mg hydromorphone (189 morphine equivalent) (Hale 2010), and 39 mg of oxycodone (58.5 morphine equivalent) (Webster 2006). The included trials that evaluated strong opioids for LBP did not allow participants to use concurrent analgesics (including antidepressants and anticonvulsants). All strong opioid trials used pain scores as a primary outcome. Additionally, one trial used quantitative sensory testing (Chu 2012). Trial authors measured functional status using SF-36 (Buynak 2010), RMDQ (Hale 2010; Chu 2012), ODI (Webster 2006; Khoromi 2007); but two oxymorphone trials did not report functional scores (Hale 2007; Katz 2007).

Most of the studies had a duration of 12 weeks (Webster 2006; Hale 2007; Katz 2007; Hale 2010). The tapentadol trial had the longest follow-up (15 weeks) (Buynak 2010). The morphine trials ran for four weeks (Chu 2012) and nine weeks (Khoromi 2007). Three out of seven RCTs used an enriched enrolment randomized withdrawal design (Hale 2007; Katz 2007; Hale 2010) and the pharmaceutical industry conducted and sponsored all three of these RCTs.

#### Opioids compared with other analgesics

O'Donnell 2009 reported two trials with identical methodology (randomized, double-blinded and parallel) that compared the effectiveness of tramadol (50 mg four times a day) versus celecoxib (200 mg twice a day) for treatment of CLBP. 798 participants (mean age 47.5; female 56.4%) received tramadol versus 785 participants (mean age 48; female 58.7%) treated with celecoxib. The rate of participants that completed the six weeks of follow-up was significantly higher in the celecoxib group (86% versus 71.8%). 62% of the participants were Caucasian. This is the only trial that used the rate of participants that had at least 30% improvement in pain ratings from baseline to week 6. The trial authors excluded patients with low back surgery within six months prior to study entry or those taking any kind of analgesic treatment.

Five of the studies that we already described had a second active arm: Buynak 2010 compared tapentadol to sustained-release (SR) oxycodone. Webster 2006 used oxycodone and included two additional arms of a tablet combining oxycodone and naltrexone; we did not use any of the data from the combination arms in the review. Vorsanger 2008 analyzed the dose response of tramadol and had two tramadol treatment arms; we included the data using the higher dose. Khoromi 2007 (crossover design) included one period of nortriptyline. Finally, Uberall 2012 included a control arm of tramadol and we used these data in the review.

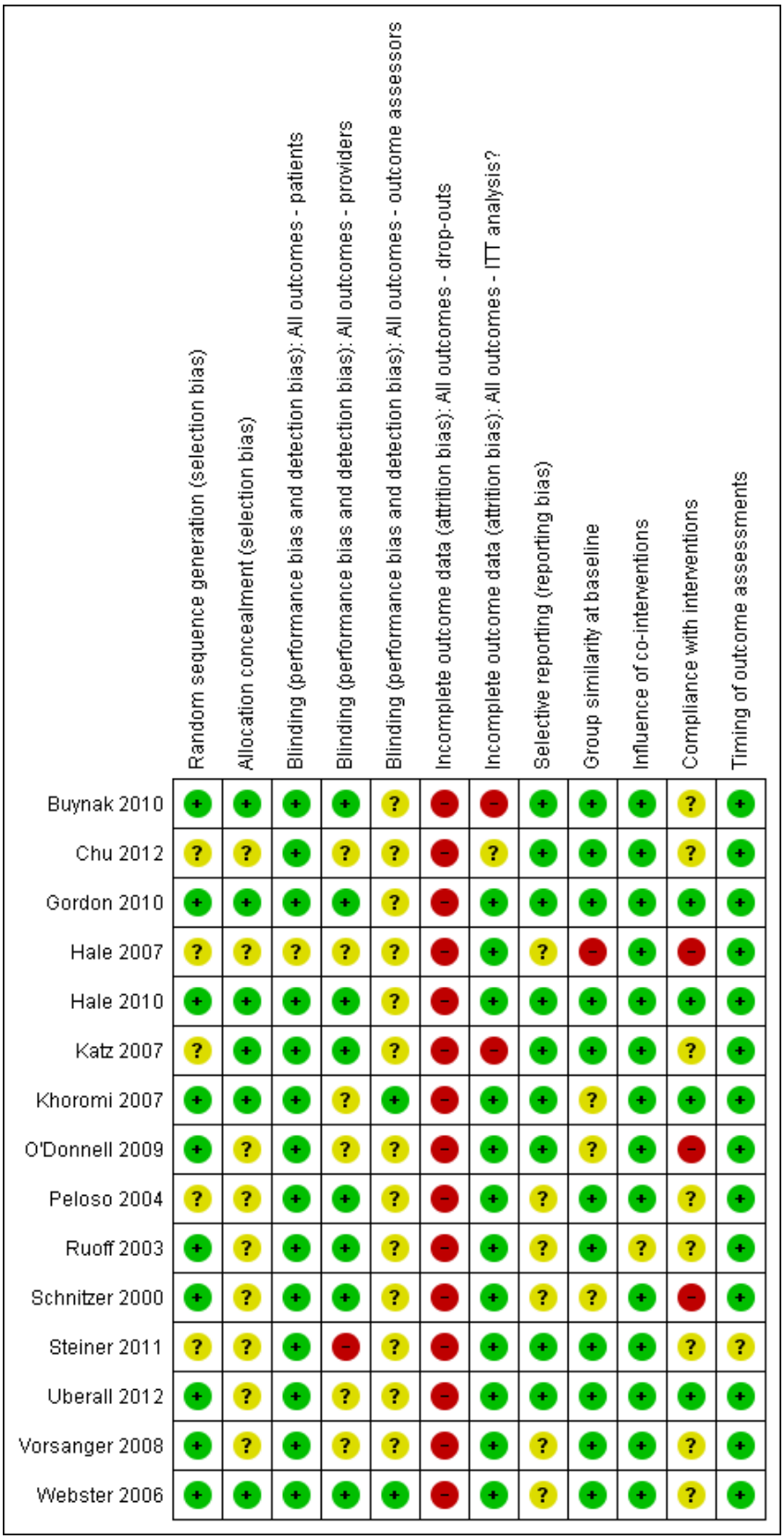
### Excluded studies

We excluded 36 studies from the review: six were developed in an open label fashion (Jamison 1998 (included in the original review); Adams 2006; Allan 2005; Gaertner 2006; Pascual 2007; Peniston 2009); 12 compared opioid versus opioid (Beaulieu 2007; Gostick 1989; Hale 1997; Hale 1999; Hale 2005; Hale 2009; Likar 2007; Nicholson 2006a; Perrot 2006; Rauck 2006; Rauck 2007; Salzman 1999); two used an opioid for analgesic rescue (Cloutier 2013; Vondrackova 2008); one did not meet our definition of CLBP (Gordon 2010a); three due to follow-up < four weeks (Kuntz 1996; Li 2008; Muller 1998); in two, the study population had < 50% of participants with a primary diagnosis of CLBP (Landau 2007; Moulin 1996); seven were secondary analyses (Gould 2009; Kalso 2007) or observational studies (Taylor 2007; Volinn 2009; Wallace 2007; Weinstein 2006; Wiesel 1980); one focused on the effectiveness of opioids for breakthrough LBP (Portenoy 2007); one focused on the timing or scheduling of the drug (Nicholson 2006), or pharmacokinetics issues (Sarbu 2008). See Characteristics of excluded studies.

## Risk of bias in included studies

We presented the results of the included articles in Figure 4.

**Figure 4. Summary of risk of bias of included studies.**



### Allocation

Only four trials described the method used for sequence generation and allocation concealment (Webster 2006; Khoromi 2007; Buynak 2010; Gordon 2010); five trials described adequately the sequence generation (Schnitzer 2000; Ruoff 2003; Vorsanger 2008; O'Donnell 2009; Uberall 2012) and six trials reported the allocation concealment (Schnitzer 2000; Ruoff 2003; Katz 2007; Vorsanger 2008; O'Donnell 2009; Uberall 2012).

### Blinding

Participants and medication providers were properly blinded in most of the studies through the use of physically identical capsules/tablets (Buynak 2010; Chu 2012; Gordon 2010; Hale 2010; Katz 2007; Khoromi 2007; Peloso 2004; Ruoff 2003; Schnitzer 2000; Steiner 2011; Uberall 2012; Vorsanger 2008; Webster 2006). However, a method to keep the outcome assessors blinded was generally flawed in all trials except for Khoromi 2007 (outcomes assessors could have guessed the allocation based on the side effects profile of the opioids).

### Incomplete outcome data

All included studies had a drop-out rate over 20% that qualified them for high risk of bias; however, ITT analysis played in favour of most of them. A Last-Observation-Carried-Forward analysis was qualified as high risk of bias (Figure 4).

### Selective reporting

Only nine out of 15 studies indicated pre-trial registration on a clinical trial registry (Buynak 2010; Chu 2012; Gordon 2010; Hale 2010; Katz 2007; Khoromi 2007; O'Donnell 2009; Steiner 2011; Uberall 2012); however, most of the trials reported outcomes that were clinically relevant.

### Other potential sources of bias

All studies showed that participants between groups were similar. Only one study assessed the count of tablets to verify the compliance of medication intake (Uberall 2012). More commonly, the number of drop-outs due to non-compliance to medications was reported. In several the studies (Hale 2007; Hale 2010; Katz 2007; Vorsanger 2008; Webster 2006) the use of analgesics was restricted to the study drugs. We constructed funnel plots but we could not identify any evidence of publication bias (Figure 5; Figure 6; Figure 1; Figure 2).

**Figure 5. Funnel plot of comparison: 1 Tramadol compared to placebo, outcome: 1.1 Pain intensity (higher score means worse pain levels).**

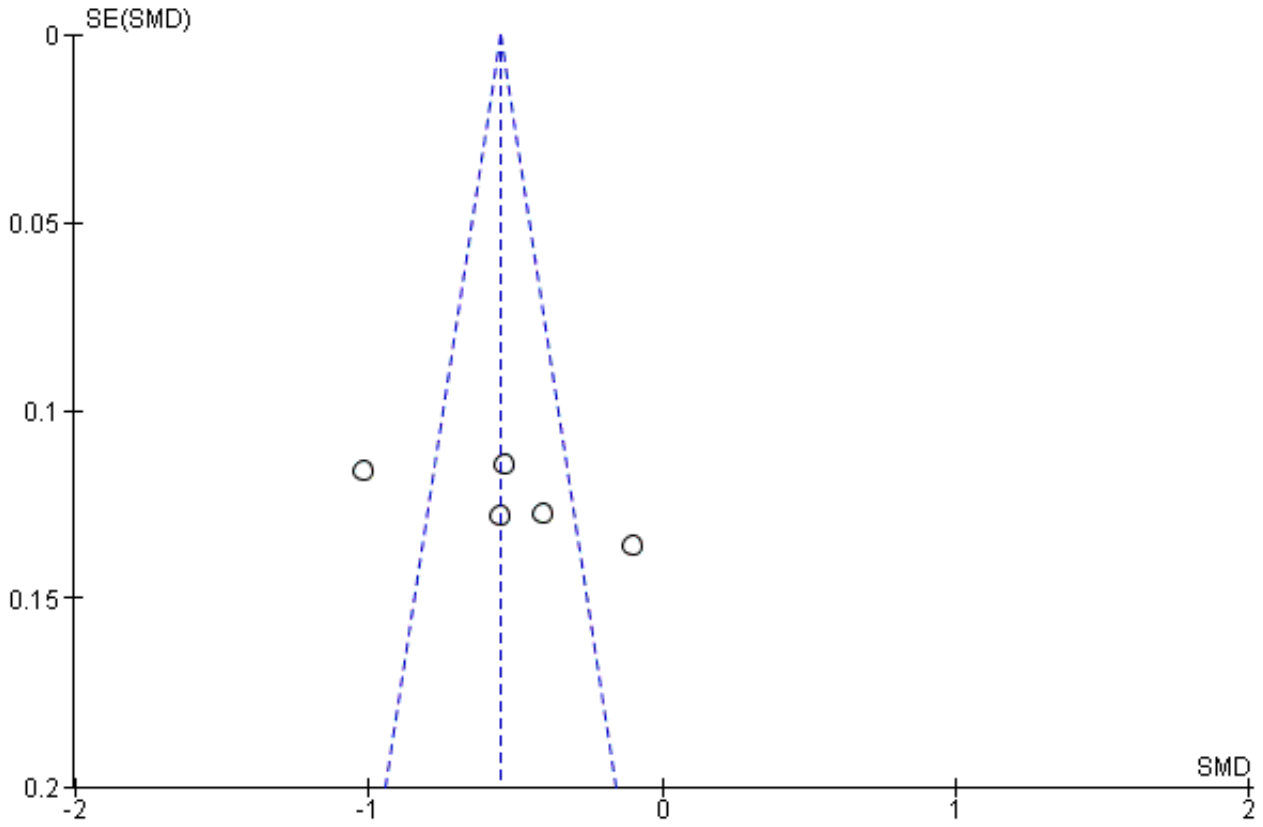
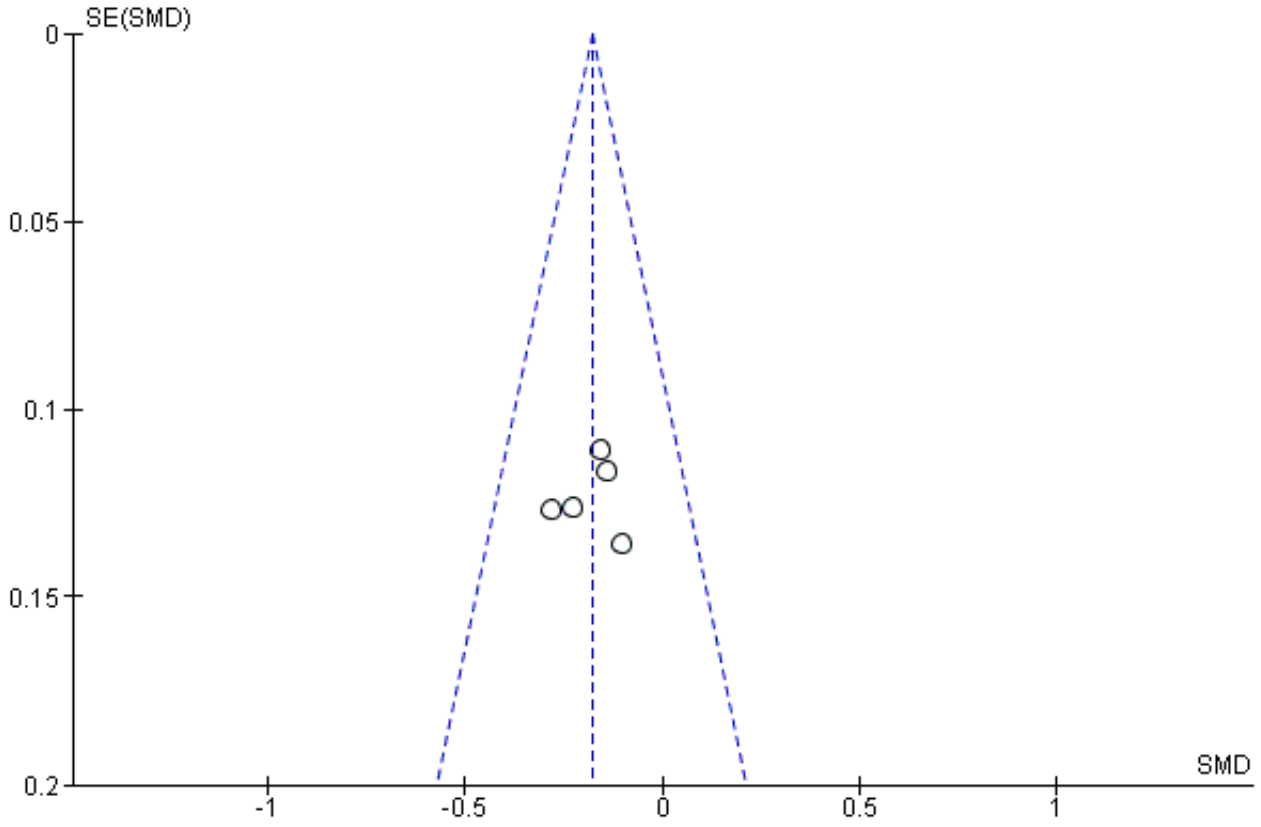


**Figure 6. Funnel plot of comparison: 1 Tramadol compared to placebo, outcome: 1.2 Disability (higher ratings mean greater disability).**




## Effects of interventions

See: **Summary of findings for the main comparison** Strong opioids compared to placebo for chronic low-back pain

### Efficacy of tramadol compared to placebo

A total of 1378 participants were included in five studies of tramadol compared to placebo (Peloso 2004, Ruoff 2003, Schnitzer 2000, Uberall 2012, Vorsanger 2008). Meta analysis (fixed effects) was used to combine the results of these studies. There is *low quality evidence* (Table 1) that tramadol is better than placebo in improving pain (SMD -0.55, 95% CI -0.66 to -0.44) (see Analysis 1.1); and *moderate quality evidence* that tramadol is better than placebo in improving functional outcomes (SMD -0.18, 95%CI -0.29 to -0.07) (see Analysis 1.2).

### Efficacy of buprenorphine compared to placebo

A total of 653 participants were included in two studies of transdermal buprenorphine compared to placebo (Gordon 2010 and Steiner 2011). Meta-analysis (fixed effects) was used to combine the results of these studies. There is *very low quality evidence* (Table 2) that transdermal buprenorphine is better than placebo in improving pain (SMD -2.47, 95%CI -2.69 to -2.25) (see Analysis 2.1); and *very low quality evidence* of no difference on functionality outcomes (SMD -0.14, 95%CI -0.53 to 0.25) (see Analysis 2.4).

### Efficacy of strong opioids compared to placebo

We identified seven RCTs for inclusion but we could only use six in the meta-analysis, as we could not obtain relevant data for the primary outcome from the authors of Hale 2007. A total of 1887 participants were included in six studies of strong opioids compared to placebo (Buynak 2010, Chu 2012, Hale 2010, Katz 2007, Khoromi 2007, Webster 2006). Meta-analysis (fixed effects) was used to combine the results of these studies. There is *moderate quality evidence* (Table 3) that strong opioids are better than placebo in reducing pain (SMD -0.43, 95%CI -0.52 to -0.33) (see Analysis 3.1); and *moderate quality evidence* that they are better than placebo in improving functional outcomes (SMD -0.26, 95% CI -0.37 to -0.15) (see Analysis 3.4).

### Adverse effects of opioids compared to placebo

Ten studies described one or more of 14 adverse events (Analysis 4.2). People treated with opioids had a statistically significant higher incidence of nausea (10%, 95% CI 7% to 14%), dizziness (8%, 95% CI 5% to 11%), constipation (7%, 95% CI 4% to 11%), vomiting (7%, 95% CI 4% to 9%), somnolence (6%, 95% CI 3% to 9%) and dry mouth (6%, 95% CI 2% to 10%) than people treated with placebo. People who received opioids had a < 5% higher incidence of headaches, pruritis, fatigue, anorexia, increased sweating and hot flushes compared to placebo. People treated with either opioids or placebo showed no differences regarding the number of people with upper respiratory tract infection (-2%, 95% CI -8% to 3%) or sinusitis (2%, 95% CI -3% to 6%).

### Effectiveness of opioids versus other drugs

We could not perform a meta-analysis of data comparing opioids (tramadol) and NSAIDs (such as celecoxib) as we only found one RCT with 1583 participants (O'Donnell 2009). There is *very low quality evidence* (Table 4) that tramadol is better than celecoxib in reducing pain (RR 0.82, 95% CI 0.76 to 0.90) (see Analysis 5.1). There was no information about functional status outcomes.

Two RCTs, including 272 participants in total, compared opioids to the antidepressants nortriptyline (Khoromi 2007) or flupirtine (Uberall 2012). Meta-analysis (fixed effects) was used to combine the results of these studies. There is *very low quality evidence* (Table 5) of no difference in pain outcomes (SMD 0.21, 95% CI -0.03 to 0.45) (see Analysis 6.1); and there is *very low quality evidence* of no difference for functional status outcomes (SMD -0.11, 95% -0.63 to 0.42) (see Analysis 6.2)

### Sensitivity analysis

We could not assess the secondary objectives of this review due to paucity of data. In particular, we could not perform subgroup analyses on the following categories:

- Route of opioid delivery (oral, intramuscular, transdermal);
- Type of opioid (morphine, codeine, oxycodone, hydromorphone, fentanyl);
- Duration of treatment (shorter than 12 months, 12 months or longer);
- CLBP non-surgical versus prior spine surgery (failed back surgery syndrome);
- CLBP with or without radiating symptoms;
- Pharmaceutical sponsored studies compared to non-sponsored trials;
- Enriched versus non-enriched enrolment randomized design.

## DISCUSSION

We included 15 RCTs in this review that assessed the use of opioids for longer than four weeks in the management of CLBP. Overall, the quality of the evidence ranged from very low to moderate regarding use of opioids compared to placebo for pain and functional outcomes. The magnitude of the effect sizes were small to medium. All trials suffered from attrition bias with a large number of drop-outs. Many trials employed an enriched enrolment design which is known to under-report adverse events (Furlan 2011). The duration of the included RCTs was longer than four weeks but shorter than 15 weeks. Also, there was poor generalizability to populations at high risk for complications. We identified very few active-controlled (non placebo-controlled) trials. We identified an insufficient number of trials that examined use of tramadol compared to NSAIDS (such as celecoxib) or compared use of opioids with use of antidepressants to treat people with CLBP.

### 1) Strict inclusion criteria and duration of treatment

In the included trials, CLBP was well-defined. However, these trials imposed limitations by excluding patients who presented with pain outside this area (even those with radicular symptoms), had previous unsuccessful lumbar surgery or a history of substance abuse. Given the heterogeneous nature of the CLBP population, narrowly defined criteria prevent extrapolation of results to a more diverse group commonly seen in clinical settings. Importantly, exclusion of failed back surgery syndrome is also significant since it may occur in 10% to 40% of lumbar spine operations and contributes to CLBP (Oaklander 2001).

Our review excluded trials of opioid use in CLBP that were shorter than four weeks. Only two trials followed the participants for more than three months (15 weeks) (Buynak 2010; Steiner 2011). While these trials lasted substantially longer than most involving opioids and CLBP, we consider these articles to have a 'short-term' time frame. This limited treatment duration, when in reality patients are often treated for years, leaves important unanswered questions including long-term efficacy, safety, tolerance and pain sensitivity (Ballantyne 2003). Only one study focused on the potential development of opioid tolerance (Chu 2012), but the participants were followed for only one month. The high drop-out rate in the included studies demonstrates the huge challenge of developing double-blinded and placebo-controlled studies for long-term follow-up. We recommend that future studies should compare opioids to other analgesics with the goal of obtaining long-term data on relative effectiveness and safety. These studies should also enroll patients commonly presenting with CLBP, including those with prior spine surgery and at variable risk for opioid misuse or abuse (for example, explicitly identifying risk using valid questionnaires).

### 2) Poorly-defined study population

In the included RCTs that compared opioids with placebo, the study authors did not report sufficient information regarding the history of study populations. Although study authors documented demographic data well, many studies neglected to report other parameters affecting outcomes, such as duration of pain prior to enrolment, employment or compensation status or poor response to previous treatment, including opioids (Sanders 1986; Greenough 1993; Andersson 1999). Thus we were unable to compare intervention and placebo arms based on potentially relevant factors other than age, sex and race. Finally, all studies permitted physiotherapy under certain circumstances, but none of the trials reported the number of patients who may have received concurrent treatment or the types of therapy these patients obtained.

### 3) Limited interpretation of functional improvement

Most of the studies used validated questionnaires to assess functional outcomes. As noted by our results, the pooled SMD favoured in a moderate grade the use of tramadol or strong opioids for improvement of the functional outcomes. Further information is required for any recommendation for transdermal buprenorphine.

An additional limitation regarding functional outcomes is the difficulty associated with the interpretation of these data in meaningful economic or social activities, such as return-to-work or improvement in ADLs. This issue is not specific to these trials, but highlights a problem present in the pain literature when attempting to interpret improvement registered in research-based tools alone.

### 4) High drop-out rates and ITT analysis

Most studies had significant drop-out rates (> 20%). Although the reasons were clearly documented, the implications on final outcomes could be significant. Experimental mortality (loss of patients during the trial) with greater loss in the control arm could enhance the effect seen in favour of treatment. In addition,

substantial drop-outs reduce the power of the study, compromising the ability to detect a significant difference. Overall, interpretation of the study outcomes with any level of confidence is questionable, given the significant number of drop-outs.

Several studies stated that efficacy analysis was performed on the ITT population. However, some of them failed to perform a proper ITT. The method of handling absent data for patients lost to follow-up was documented through the use of LOCF. This method has been criticized given the potential overestimation of the effect (Moore 2012).

### 5) Comparison to other reviews on opioids in CNCP and CLBP

Two recently published systematic reviews have addressed the issue of opioids in the pharmacological management of CNCP (Furlan 2011) and CLBP (Kuijpers 2011). Furlan 2011 concluded that opioids were more effective than placebo for improving both pain and function in the management of CNCP. The results were significant for both neuropathic and nociceptive pain. Subgroup analyses revealed that only strong opioids (oxycodone and morphine) were statistically more effective in reducing pain but not function when compared to naproxen and nortriptyline. Kuijpers 2011 evaluated opioids, antidepressants and NSAIDs for CLBP; however, they excluded patients with sciatica. Several studies that we included in our review were not considered in Kuijpers 2011 due to timing of publication. Their conclusion regarding the effectiveness of opioids does not differ from our conclusion.

Our review confirms the effectiveness of tramadol, buprenorphine and strong opioids in the management of CLBP in the short-term. Other systematic reviews on opioids in people with CNCP have included people with multiple pathologies. This factor and the predominance of short-term studies could limit any meaningful interpretation when considering opioids for the long-term management of CLBP.

The results of our review differ from another published systematic review (Martell 2007). The review (Martell 2007) included 15 studies in the literature assessing the efficacy of opioids in CLBP. Nine of the studies considered comparisons among different opioids, while another six compared opioids with placebo or other analgesics. Meta-analysis of this latter group was completed with four of the six studies. The review authors found that opioids were ineffective in the management of CLBP when compared to the pooled sample of placebo and other analgesics.

The existence of discordant reviews has been previously described in the literature (Jadad 1997; Furlan 2001). The only common clinical query between our review and that of Martell 2007 related to the efficacy of opioids in CLBP. Notably, the two reviews used different outcomes to define the efficacy of opioids. Our review considered pain and function as outcome measures. Also, there were differences between inclusion and exclusion criteria. Our review restricted original articles to opioid treatment that was longer than one month in duration to provide more meaningful clinical interpretation in the management of CLBP. We excluded comparisons among opioids to avoid issues with head-to-head trials or equivalency determinations. We also considered only articles published in peer-review journals. Taking these criteria into account, we excluded nine trials (all opioid comparators) found in Martell 2007 from our review. From the six trials comparing the efficacy of opioids to placebo or another analgesic, we excluded five from our review for the following reasons: two trials were published in abstract form (Tennant 1993; Richards 2002), two trials had a treatment duration of less than 30 days (Kuntz 1996; Muller 1998) and one trial had a lack of randomization when comparing opioids to placebo (Hale 2005; Characteristics of excluded studies). The three trials (all involving tramadol) we used to derive our meta-analyses were absent from the Martell 2007 review. Although not specifically stated in their inclusion and exclusion criteria, the review authors may have excluded tramadol as an opioid, given its atypical status. Finally, Martell 2007 combined studies involving placebo and other comparators to determine efficacy. In this case, conceptual homogeneity may not have existed due to differences in patient response to an active control compared with placebo. Statistical pooling of these studies may lead to questionable results.

Many people experience recurring episodes of LBP or never fully recover from their initial episode (Abenhaim 1988; Von Korff 1996). With direct and indirect costs estimated to exceed $100 billion annually in the United States alone, LBP continues to inflict a huge economic toll on society (Hashemi 1997; Katz 2006). Opioids have become a popular tool to help manage patients with CLBP. The prevalence of opioid prescribing in CLBP varies by treatment setting but has been found to be as low as 3% or as high as 66% (Martell 2007). Moreover, the same review (Martell 2007) identified prescription of opioids to be more common to patients with impaired functional status. Despite significant concerns surrounding the use of opioids, there is still little evidence in the literature for their efficacy and effectiveness in long-term treatment of CLBP. Although few systematic reviews suggest that opioids are effective in the management of CNCP in general, the extrapolation of this evidence to CLBP is cautioned. Further, the few original studies that do exist focusing on opioids for the management of CLBP are of limited value in clinical practice given their lack of long-term follow-up and description of long-term safety profile. As the pendulum has swung from an 'opiophobic' to an 'opiophilic' society, physicians should question whether the current trend is based on evidence or simply the outcries of well-intentioned patient advocates and aggressive marketing efforts by the pharmaceutical industry (Chinellato 2003).

## AUTHORS' CONCLUSIONS

### Implications for practice

There is evidence (multiple high quality RCTs) that the use of tramadol (a weak atypical opioid) or strong opioids results in improved pain and moderate changes in function in the short-term in people with CLBP when compared with placebo. However, the general applicability of this treatment to the clinical setting is questionable. Several factors, including the strict inclusion criteria of the original studies, high drop-out rates, and the poor description of the study population regarding duration of pain, concurrent treatments, work status, and compensation, limit the reported results. Notably, a number of important outcomes that capture patient function were absent (such as return-to-work). Finally, there is strong evidence that nausea is more common in patients with CLBP being treated with opioids when compared to placebo.

### Implications for research

CLBP is a prevalent condition with significant socioeconomic implications in the Western world. Given the escalating use of opioids in CNCP (a subset of which is CLBP) more quality research is needed to understand i) the long term benefits and risks of opioid therapy including the different subgroups of CLBP (for example, failed back surgery syndrome and CLBP with radicular symptoms); ii) opioid effectiveness relative to other conventional physical and medical treatments; iii) characteristics of patients who are most likely to respond to long-term opioid therapy; and iv) the predictors of opioid side effects, abuse and misuse in this population.

## ACKNOWLEDGEMENTS

We are grateful to Rachel Couban (Cochrane Back Review Group Trials Coordinator) for assisting with literature searches and Ute Bültmann for helping to translate the non-English language study. We would like to gratefully and sincerely thank Teresa Marin (Managing Editor, Cochrane Back Review Group) for her invaluable assistance on the manuscript preparation.

# REFERENCES

## References to studies included in this review

**Buynak 2010** *{published data only}*

Buynak R, Shapiro DY, Okamoto A, Van Hove I, Rauschkolb C, Steup A, et al. Efficacy and safety of tapentadol extended release for the management of chronic low back pain: results of a prospective, randomized, double-blind, placebo- and active-controlled Phase III study. *Expert Opinion on Pharmacotherapy* 2010;**11**(11):1787-804.

**Chu 2012** *{published data only}*

Chu LF, D'Arcy N, Brady C, Zamora AK, Young CA, Kim JE, et al. Analgesic tolerance without demonstrable opioid-induced hyperalgesia: A double-blinded, randomized, placebo-controlled trial of sustained-release morphine for treatment of chronic nonradicular low-back pain. *Pain* 2012;**153**(8):1583-92.

**Gordon 2010** *{published data only}*

Gordon A, Callaghan D, Spink D, Cloutier C, Dzongowski P, O'Mahony W, et al. Buprenorphine transdermal system in adults with chronic low back pain: a randomized, double-blind, placebo-controlled crossover study, followed by an open-label extension phase. *Clinical Therapeutics* 2010;**32**(5):844-60.

**Hale 2007** *{published data only}*

Hale ME, Ahdieh H, Ma T, Rauck R. Efficacy and safety of OPANA ER (Oxymorphone Extended Release) for relief of moderate to severe chronic low back pain in opioid-experienced patients: A 12-week, randomized, double-blind, placebo-controlled study. *Journal of Pain* 2007;**8**(2):175-84.

**Hale 2010** *{published data only}*

Hale M, Khan A, Kutch M, Li S. Once-daily OROS hydromorphone ER compared with placebo in opioid-tolerant patients with chronic low back pain. *Current Medical Research and Opinion* 2010;**26**(6):1505-18.

**Katz 2007** *{published data only}*

Katz N, Rauck R, Ahdieh H, Ma T, Gerritsen van der Hoop R, Kerwin R, et al. A 12-week, randomized, placebo-controlled trial assessing the safety and efficacy of oxymorphone extended release for opioid-naive patients with chronic low back pain. *Current Medical Research and Opinion* 2007;**23**(1):117-28.

**Khoromi 2007** *{published data only}*

Khoromi S, Cui L, Nackers L, Max MB. Morphine, nortriptyline and their combination vs. placebo in patients with chronic lumbar root pain. *Pain* 2007;**130**(1-2):66-75.

**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.

**Peloso 2004** *{published data only}*

Peloso PM, Fortin L, Beaulieu A, Kamin M, Rosenthal N, Protocol TRP-CAN-1 Study Group. Analgesic efficacy and safety of tramadol/acetaminophen combination tablets (Ultracet) in treatment of chronic low back pain: a multicenter, outpatient, randomized, double blind, placebo controlled trial. *Journal of Rheumatology* 2004;**31**(12):2454-63.

**Ruoff 2003** *{published data only}*

Ruoff GE, Rosenthal N, Jordan D, Karim R, Kamin M, Protocol CAPSS-112 Study Group. Tramadol/acetaminophen combination tablets for the treatment of chronic lower back pain: a multicenter, randomized, double-blind, placebo-controlled outpatient study. *Clinical Therapeutics* 2003;**25**(4):1123-41.

**Schnitzer 2000** *{published data only}*

Schnitzer TJ, Gray WL, Paster RZ, Kamin M. Efficacy of tramadol in treatment of chronic low back pain. *Journal of Rheumatology* 2000;**27**(3):772-8.

**Steiner 2011** *{published data only}*

Steiner DJ, Sitar S, Wen W, Sawyerr G, Munera C, Ripa SR, et al. Efficacy and safety of the seven-day buprenorphine transdermal system in opioid-naive patients with moderate to severe chronic low back pain: an enriched, randomized, double-blind, placebo-controlled study. *Journal of Pain and Symptom Management* 2011;**42**(6):903-17.

**Uberall 2012** *{published data only}*

Uberall MA, Mueller-Schwefe GH, Terhaag B. Efficacy and safety of flupirtine modified release for the management of moderate to severe chronic low back pain: results of SUPREME, a prospective randomized, double-blind, placebo- and active-controlled parallel-group phase IV study. *Current Medical Research and Opinion* 2012;**28**(10):1617-34.

**Vorsanger 2008** *{published data only}*

Vorsanger GJ, Xiang J, Gana TJ, Pascual ML, Fleming RR. Extended-release tramadol (tramadol ER) in the treatment of chronic low back pain. *Journal of Opioid Management* 2008;**4**(2):87-97.

**Webster 2006** *{published data only}*

Webster LR, Butera PG, Moran LV, Wu N, Burns LH, Friedmann N. Oxytrex minimizes physical dependence while providing effective analgesia: A randomized controlled trial in low back pain. *Journal of Pain* 2006;**7**(12):937-46.

## References to studies excluded from this review

**Adams 2006** *{published data only}*

Adams EH, Chwiecko P, Ace-Wagoner Y, Mangefrida B, Duerden ME, Perdikis GC, et al. A study of AVINZA (morphine sulfate extended-release capsules) for chronic moderate-to-severe noncancer pain conducted under real-world treatment conditions-the ACCPT study. *Pain Practice* 2006;**6**(4):254-64.

**Allan 2005** *{published data only}*

Allan L, Richarz U, Simpson K, Slappendel R. Transdermal fentanyl versus sustained release oral morphine in strong-opioid naive patients with chronic low back pain. *Spine* 2005;**30**(22):2484-90.

**Beaulieu 2007** *{published data only}*

Beaulieu AD, Peloso P, Bensen W, Clark AJ, Watson CP, Gardner-Nix J, et al. A randomized, double-blind, 8-week crossover study of once-daily controlled-release tramadol versus immediate-release tramadol taken as needed for chronic noncancer pain. *Clinical Therapeutics* 2007;**29**(1):49-60.

**Cloutier 2013** *{published data only}*

Cloutier C, Taliano J, O'Mahony W, Csanadi M, Cohen G, Sutton I, et al. Controlled-release oxycodone and naloxone in the treatment of chronic low back pain: A placebo-controlled, randomized study. *Pain Research and Management* 2013;**18**(2):75-82.

**Gaertner 2006** *{published data only}*

Gaertner J, Frank M, Bosse B, Sabatowski R, Eisner F, Giesecke T, et al. Oral controlled-release oxycodone for the treatment of chronic pain. Data from 4196 patients. *Schmerz* 2006;**20**(1):61-8.

**Gordon 2010a** *{published data only}*

Gordon A, Rashiq S, Moulin DE, Clark AJ, Beaulieu AD, Eisenhoffer J, et al. Buprenorphine transdermal system for opioid therapy in patients with chronic low back pain. *Pain Research and Management* 2010;**15**(3):169-78. [PUBMED: 20577660]

**Gostick 1989** *{published data only}*

Gostick N, Allen J, Cranfield R, Currie J, Grillage M, Hildebrand PJ. A comparison of the efficacy and adverse effects of controlled release dihydrocodeine and immediate release dihydrocodeine in the treatment of pain in osteoarthritis and chronic back pain. *Edinburgh Symposium on Pain Control and Medical Education* 1989;**1**:137-43.

**Gould 2009** *{published data only}*

Gould EM, Jensen MP, Victor TW, Gammaitoni AR, White RE, Galer BS. The pain quality response profile of oxymorphone extended release in the treatment of low back pain. *Clinical Journal of Pain* 2009;**25**(2):116-22.

**Hale 1997** *{published data only}*

Hale M, Speight K, Harsanyi Z, Iwan T, Slagle NS, Lacouture PG. Efficacy of 12 hourly controlled-release codeine compared with as required dosing of acetaminophen plus codeine in patients with chronic low back pain. *Pain Research and Management* 1997;**2**(1):33-8.

**Hale 1999** *{published data only}*

Hale ME, Fleischmann R, Salzman R, Wild J, Iwan T, Swanton RE, et al. Efficacy and safety of controlled-release versus immediate-release oxycodone: randomized, double-blind evaluation in patients with chronic back pain. *Clinical Journal of Pain* 1999;**15**(3):179-83.

**Hale 2005** *{published data only}*

Hale ME, Dvergsten C, Gimbel J. Efficacy and safety of oxymorphone extended release in chronic low back pain: results of a randomized, double-blind, placebo- and active-controlled phase III study. *Journal of Pain* 2005;**6**(1):21-8.

**Hale 2009** *{published data only}*

Hale M, Upmalis D, Okamoto A, Lange C, Rauschkolb C. Tolerability of tapentadol immediate release in patients with lower back pain or osteoarthritis of the hip or knee over 90 days: a randomized, double-blind study. *Current Medical Research and Opinion* 2009;**25**(5):1095-104.

**Jamison 1998** *{published data only}*

Jamison RN, Raymond SA, Slawsby EA, Nedeljkovic SS, Katz NP. Opioid therapy for chronic noncancer back pain. A randomized prospective study. *Spine* 1998;**23**(23):2591-600. [PUBMED: 9854758]

**Kalso 2007** *{published data only}*

Kalso E, Simpson KH, Slappendel R, Dejonckheere J, Richarz U. Predicting long-term response to strong opioids in patients with low back pain: findings from a randomized, controlled trial of transdermal fentanyl and morphine. *BMC Medicine* 2007;**5**:39.

**Kuntz 1996** *{published data only}*

Kuntz D, Brossel R. Analgesic effect and clinical tolerability of the combination of paracetamol 500 mg and caffeine 50 mg versus paracetamol 400 mg and dextropropoxyphene 30 mg in back pain. *Presse Medicale* 1996;**25**(25):1171-4.

**Landau 2007** *{published data only}*

Landau CJ, Carr WD, Razzetti AJ, Sessler NE, Munera C, Ripa SR. Buprenorphine transdermal delivery system in adults with persistent noncancer-related pain syndromes who require opioid therapy: a multicenter, 5-week run-in and randomized, double-blind maintenance-of-analgesia study. *Clinical Therapeutics* 2007;**29**(10):2179-93.

**Li 2008** *{published data only}*

Li C, Ni J, Wang Z, Li M, Gasparic M, Terhaag B, et al. Analgesic efficacy and tolerability of flupirtine vs. tramadol in patients with subacute low back pain: a double-blind multicentre trial*. *Current Medical Research and Opinion* 2008;**24**(12):3523-30.

**Likar 2007** *{published data only}*

Likar R, Lorenz V, Korak-Leiter M, Kager I, Sittl R. Transdermal buprenorphine patches applied in a 4-day regimen versus a 3-day regimen: a single-site, Phase III, randomized, open-label, crossover comparison. *Clinical Therapeutics* 2007;**29**(8):1591-606.

**Moulin 1996** *{published data only}*

Moulin DE, Iezzi A, Amireh R, Sharpe WK, Boyd D, Merskey H. Randomised trial of oral morphine for chronic non-cancer pain. *Lancet* 1996;**347**(8995):143-7.

**Muller 1998** *{published data only}*

Muller FO, Odendaal CL, Muller FR, Raubenheimer J, Middle MV, Kummer M. Comaprison of the efficacy and tolerability of a paracetamol/codeine fixed dose combination with

tramadol in patients with refractory chronic back pain. *Arzneimittelforschung* 1998;**48**(6):675-9.

**Nicholson 2006** *{published data only}*

Nicholson B, Ross E, Weil A, Sasaki J, Sacks G. Treatment of chronic moderate-to-severe non-malignant pain with polymer-coated extended-release morphine sulfate capsules. *Current Medical Research and Opinion* 2006;**22**(3):539-50.

**Nicholson 2006a** *{published data only}*

Nicholson B, Ross E, Sasaki J, Weil A. Randomized trial comparing polymer-coated extended-release morphine sulfate to controlled-release oxycodone HCl in moderate to severe nonmalignant pain. *Current Medical Research and Opinion* 2006;**22**(8):1503-14.

**Pascual 2007** *{published data only}*

Pascual ML, Fleming RR, Gana TJ, Vorsanger GJ. Open-label study of the safety and effectiveness of long-term therapy with extended-release tramadol in the management of chronic nonmalignant pain. *Current Medical Research and Opinion* 2007;**23**(10):2531-42.

**Peniston 2009** *{published data only}*

Peniston JH, Gould E. Oxymorphone extended release for the treatment of chronic low back pain: a retrospective pooled analysis of enriched-enrollment clinical trial data stratified according to age, sex, and prior opioid use. *Clinical Therapeutics* 2009;**31**(2):347-59.

**Perrot 2006** *{published data only}*

Perrot S, Krause D, Crozes P, Naim C, GRTF-ZAL-1 Study Group. Efficacy and tolerability of paracetamol/tramadol (325 mg/37.5 mg) combination treatment compared with tramadol (50 mg) monotherapy in patients with subacute low back pain: a multicenter, randomized, double-blind, parallel-group, 10-day treatment study. *Clinical Therapeutics* 2006;**28**(10):1592-606.

**Portenoy 2007** *{published data only}*

Portenoy RK, Messina J, Xie F, Peppin J. Fentanyl buccal tablet (FBT) for relief of breakthrough pain in opioid-treated patients with chronic low back pain: a randomized, placebo-controlled study. *Current Medical Research and Opinion* 2007;**23**(1):223-33.

**Rauck 2006** *{published data only}*

Rauck RL, Bookbinder SA, Bunker TR, Alftine CD, Ghalie R, Negro-Vilar A, et al. A randomized, open-label study of once-a-day AVINZA (morphine sulfate extended-release capsules) versus twice-a-day OxyContin (oxycodone hydrochloride controlled-release tablets) for chronic low back pain: the extension phase of the ACTION trial. *Journal of Opioid Management* 2006;**2**(6):325-3, 331-3.

**Rauck 2007** *{published data only}*

Rauck RL, Bookbinder SA, Bunker TR, Alftine CD, Gershon S, de Jong E, et al. A randomized, open-label, multicenter trial comparing once-a-day AVINZA (morphine sulfate extended-release capsules) versus twice-a-day OxyContin (oxycodone hydrochloride controlled-release tablets) for the treatment of chronic, moderate to severe low back pain: improved physical functioning in the ACTION trial. *Journal of Opioid Management* 2007;**3**(1):35-43.

**Salzman 1999** *{published data only}*

Salzman RT, Roberts MS, Wild J, Fabian C, Reder RF, Goldenheim PD. Can a controlled-release oral dose form of oxycodone be used as readily as an immediate-release form for the purpose of titrating to stable pain control?. *Journal of Pain and Symptom Management* 1999;**18**(4):271-9.

**Sarbu 2008** *{published data only}*

Sarbu A, Radulescu F, Robertson S, Bouchard S. Onset of analgesic effect and plasma levels of controlled-release tramadol (Tramadol Contramid once-a-day) 200-mg tablets in patients with acute low back pain. *Journal of Opioid Management* 2008;**4**(5):285-92.

**Taylor 2007** *{published data only}*

Taylor DR, Webster LR, Chun SY, Reinking J, Stegman M, Shoemaker S, et al. Impact of breakthrough pain on quality of life in patients with chronic, noncancer pain: patient perceptions and effect of treatment with oral transmucosal fentanyl citrate (OTFC, ACTIQ). *Pain Medicine* 2007;**8**(3):281-8.

**Volinn 2009** *{published data only}*

Volinn E, Fargo JD, Fine PG. Opioid therapy for nonspecific low back pain and the outcome of chronic work loss. *Pain* 2009;**142**(3):194-201.

**Vondrackova 2008** *{published and unpublished data}*

Vondrackova D, Leyendecker P, Meissner W, Hopp M, Szombati I, Hermanns K, et al. Analgesic efficacy and safety of oxycodone in combination with naloxone as prolonged release tablets in patients with moderate to severe chronic pain. *Journal of Pain* 2008;**9**(12):1144-54.

**Wallace 2007** *{published data only}*

Wallace M, Rauck RL, Moulin D, Thipphawong J, Khanna S, Tudor IC. Once-daily OROS hydromorphone for the management of chronic nonmalignant pain: a dose-conversion and titration study. *International Journal of Clinical Practice* 2007;**61**(10):1671-6.

**Weinstein 2006** *{published data only}*

Weinstein SM, Shi M, Buckley BJ, Kwarcinski MA. Multicenter, open-label, prospective evaluation of the conversion from previous opioid analgesics to extended-release hydromorphone hydrochloride administered every 24 hours to patients with persistent moderate to severe pain. *Clinical Therapeutics* 2006;**28**(1):86-98.

**Wiesel 1980** *{published data only}*

Wiesel SW, Cuckler JM, Deluca F, Jones F, Zeide MS, Rothman RH. Acute low-back pain. An objective analysis of conservative therapy. *Spine* 1980;**5**(4):324-30.

## References to ongoing studies

**NCT01081912** *{unpublished data only}*

A randomized double-blind, placebo-controlled Trial to evaluate the efficacy, tolerability and safety of hydrocodone

bitartrate controlled-release capsules in opioid-experienced subjects with moderate to severe CLBP.. Ongoing study March 4, 2010..

**NCT01358526** *{unpublished data only}*

A Randomized, double-blind, placebo-controlled, multicentre trial with an enriched study design to assess the efficacy and safety of oxycodone/naloxone controlled-release tablets (OXN) compared to placebo in opioid-experienced subjects with moderate to severe pain due to chronic low back pain who require around-the-clock opioid therapy. Ongoing study May 2011..

**NCT01455519** *{unpublished data only}*

True functional restoration and analgesia in non-radicular low back pain: a prospective double blind, placebo-controlled study of hydromorphone ER.. Ongoing study October 2011..

**NCT01571362** *{unpublished data only}*

A multicentre, 12 week, double-blind, placebo-controlled, randomized withdrawal study to determine the efficacy and safety of ALO-02 (oxycodone hydrochloride and naltrexone hydrochloride) extended-release capsules in subjects with moderate to severe CLBP.. Ongoing study June 2012..

**NCT01789970** *{unpublished data only}*

A 12 week, randomized, double-blind, placebo-controlled, randomized-withdrawal study to evaluate the efficacy and safety of hydrocodone bitartrate extended-release tablets (CEP-33237) at 30 to 90 mg every 12 hours for relief of CLBP who require opioid treatment for an extended period of time.. Ongoing study March 2013..

## Additional references

**Abenhaim 1988**

Abenhaim L, Suissa S, Rossignol M. Risk of recurrence of occupational back pain over three year follow up. *British Journal of Industrial Medicine* 1988;**45**(12):829-33.

**Andersson 1999**

Andersson GB. Epidemiological factors of chronic low-back pain. *Lancet* 1999;**354**(9178):581-5.

**Ballantyne 2003**

Ballantyne JC, Mao J. Opioid therapy for chronic pain. *New England Journal of Medicine* 2003;**349**(20):1943-53.

**Bohnert 2011**

Bohnert AS, Valenstein M, Bair MJ, Ganoczy D, McCarthy JF, Ilgen MA, et al. Association between opioid prescribing patterns and opioid overdose-related deaths. *Journal of the American Medical Association* 2011;**305**(13):1315-21. [PUBMED: 21467284]

**Boutron 2005**

Boutron I, Moher D, Tugwell P, Giraudeau B, Poiraudeau S, Nizard R, et al. A checklist to evaluate a report of a non pharmacological trial (CLEAR NPT) was developed using consensus. *Journal of Clinical Epidemiology* 2005;**58**(12):1233-40.

**Breckenridge 2003**

Breckenridge J, Clark JD. Patient characteristics associated with opioid v. non-steroidal anti-inflammatory drug management of chronic lower back pain. *Journal of Pain* 2003;**4**(6):344-50.

**Cassidy 1998**

Cassidy JD, Carroll LJ, Cote P. The Saskatchewan Health and Back Pain Survey: The prevalence of low back pain and related disability in Saskatchewan adults. *Spine* 1998;**23**(17):1860-66.

**Chinellato 2003**

Chinellato A, Terrazzani G, Walllye T, Giusti P. Opioids in Italy: is marketing more powerful than the law?. *Lancet* 2003;**362**(9377):78.

**Chou 2007**

Chou R, Huffman LH. Medications for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine* 2007;**147**(7):505-14.

**Deshpande 2007**

Deshpande A, Furlan A, Mailis-Gagnon A, Atlas S, Turk D. Opioids for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD004959.pub3]

**Deyo 2006**

Deyo RA, Mirza SK, Martin BL. Back pain prevalence and visit rates: Estimates from US national surveys, 2002. *Spine* 2006;**31**(23):2724-7.

**Edlund 2007**

Edlund MJ, Sullivan M, Steffick D, Harris KM, Wells KB. Do users of regularly prescribed opioids have higher rates of substance use problems than nonusers?. *Pain Medicine* 2007;**8**(8):647-56.

**Ferrell 2009**

Ferrell B, Argoff CE, Epplin J, Fine P, Gloth FM, Herr K, et al. Pharmacological management of persistent pain in older persons. *Journal of the American Geriatrics Society* 2009;**57**(8):1331-46. [PUBMED: 19573219]

**Fritsch 1996**

Fritsch EW, Heisel J, Rupp S. The failed back surgery syndrome-reasons, intraoperative findings and long-term results: a report of 182 operative treatments. *Spine* 1996;**21**(5):626-33.

**Furlan 2001**

Furlan AD, Clarke J, Esmail R, Sinclair S, Irvin E, Bombardier C. A critical review of reviews on the treatment of chronic low back pain. *Spine* 2001;**26**(7):E155-62.

**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.

**Furlan 2010**

Furlan AD, Reardon R, Weppler C. Opioids for chronic noncancer pain: a new Canadian practice guideline. *Canadian Medical Association journal* 2010;**182**(9):923-30. [PUBMED: 20439443]

**Furlan 2011**

Furlan A, Chaparro LE, Irvin E, Mailis-Gagnon A. A comparison between enriched and nonenriched enrollment randomized withdrawal trials of opioids for chronic noncancer pain. *Pain Research and Management* 2011;**16**(5):337-51. [PUBMED: 22059206]

**Gore 2012**

Gore M, Sadosky A, Stacey BR, Tai KS, Leslie D. The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings. *Spine* 2012;**37**(11):E668-77. [PUBMED: 22146287]

**Greenough 1993**

Greenough CG. Recovery from low back pain. 1-5 year follow-up of 287 injury-related cases. *Acta Orthopaedica Scandinavica. Supplementum* 1993;**254**:1-34.

**Hashemi 1997**

Hashemi L, Webster BS, Clancy EA, Volinn E. Length of disability and cost of workers' compensation low back claims. *Occupational and Environmental Medicine* 1997;**39**(10):937-45.

**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 2011.

**Jadad 1997**

Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *Canadian Medical Association Journal* 1997;**156**(10):1411-6.

**Katz 2006**

Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *Journal of Bone and Joint Surgery* 2006;**88**(Suppl 2):21-4.

**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.

**Martell 2007**

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

**Moore 2012**

Moore RA, Straube S, Eccleston C, Derry S, Aldington D, Wiffen P, et al. Estimate at your peril: imputation methods for patient withdrawal can bias efficacy outcomes in chronic pain trials using responder analyses. *Pain* 2012;**153**(2):265-8.

**Morley-Foster 2003**

Morley-Foster PK, Clark AJ, Speechley M, Moulin D. Attitudes toward opioid use for chronic pain: A Canadian physician survey. *Pain Research and Management* 2003;**8**(4):189-94.

**Mueller 2007**

Mueller PS, Montori VM, Bassler D, Koenig BA, Guyatt GH. Ethical issues in stopping randomized trials early because of apparent benefit. *Annals of Internal Medicine* 2007;**146**(12):878-81.

**Noble 2008**

Noble M, Tregear SJ, Treadwell JR, Schoelles K. Long-term opioid therapy for chronic noncancer pain: a systematic review and meta-analysis of efficacy and safety. *Journal of Pain and Symptom Management* 2008;**35**(2):214-28.

**Oaklander 2001**

Oaklander AL, North RB. Failed back surgery syndrome. In: Loeser JD, Butler SH, Chapman CR, Turk DC editor(s). Bonica's management of pain. 3rd Edition. Philadelphia PA: Lippincott Williams and Wilkins, 2001:1540-9.

**Ostelo 2003**

Ostelo RWJG, de Vet HCW, Waddell G, Kerckhoffs MR, Leffers P, van Tulder MW. Rehabilitation after lumbar disc surgery. *Cochrane Database of Systematic Reviews* 2002, Issue 2. [DOI: 10.1002/14651858.CD003007]

**Peul 2007**

Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, et al. Surgery versus prolonged conservative treatment for sciatica. *New England Journal of Medicine* 2007;**356**(22):2245-56.

**Review Manager (RevMan) [Computer program]**

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan) 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2013.

**Sanders 1986**

Sanders ME, Meyers JE. The relationship of disability to compensation status in railroad workers. *Spine* 1986;**11**(2):141-3.

**Shekelle 1995**

Shekelle PG, Markovich M, Louie R. An epidemiological study of episodes of back pain care. *Spine* 1995;**20**(15):1668-73.

**Spitzer 1987**

Spitzer WO, Leblanc FE, Depuis M. Scientific approach to the assessment and management of activity-related spinal disorders: a monograph for physicians. *Spine* 1987;**12**(Suppl 7):S1-S59.

**Thomas 1999**

Thomas E, Silman AJ, Croft PR, Papageorgiou AC, Jayson M, Macfarlane G. Predicting who develops chronic low back pain

in primary care: a prospective study. *British Medical Journal* 1999;**318**(7199):1662-7.

**Turk 1997**

Turk DC, Okifuji A. What factors affect physicians' decisions to prescribe opioids for chronic noncancer pain patients?. *Clinical Journal of Pain* 1997;**13**(4):330-6.

**Turk 2011**

Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. *Lancet* 2011;**377**(9784):2226-35. [PUBMED: 21704872]

**van Tulder 2003**

van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 2003;**28**(12):1290-9.

**Volinn 1997**

Volinn E. The epidemiology of low back pain in the rest of the world. A review of surveys in low- and middle-income countries. *Spine* 1997;**22**(15):1747-54.

**Von Korff 1996**

Von Korff M, Saunders K. The course of back pain in primary care. *Spine* 1996;**21**(24):2833-7.

**Von Korff 2004**

Von Korff M, Deyo RA. Potent opioids for chronic musculoskeletal pain: flying blind?. *Pain* 2004;**109**(3):207-9.

**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.

**Wenghofer 2011**

Wenghofer EF, Wilson L, Kahan M, Sheehan C, Srivastava A, Rubin A, et al. Survey of Ontario primary care physicians' experiences with opioid prescribing. *Canadian Family Physician* 2011;**57**(3):324-32. [PUBMED: 21402971]

**Wilson 2013**

Wilson HD, Dansie EJ, Kim MS, Moskovitz BL, Chow W, Turk DC. Clinicians' attitudes and beliefs about opioids survey (CAOS): Instrument development and results of a national physician survey. *Journal of Pain* 2013;**14**(6):613-27.

## CHARACTERISTICS OF STUDIES

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

**Buynak 2010**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, placebo-controlled, parallel, three arms design for 15 weeks (3 week titration and 12 week maintenance period). |
| Participants | Adults with LBP for ≥ 3 months were included. The study required participants to be dissatisfied with their current treatment, and to have a baseline pain intensity of ≥ 5/10. Daily doses of opioids had to be equivalent to ≤ 160 mg of oral morphine in the opioid users. |
| Interventions | Tapentadol ER Group: during the titration period (3 weeks) started at 50 mg BID and 3 days later 100 mg BID. Dosing was adjusted each 3 days as required up to a maximum dose of 250 mg BID.<br><br>Oxycodone HCL CR: during the titration period (3 weeks) started at 10 mg BID and 3 days later 20 mg BID. Dosing was adjusted at a minimum of 3 day intervals as required up to a maximum dose of 50 mg BID.<br><br>Placebo capsules and tablets were administered to maintain blinding to the intervention. |
| Outcomes | Two different primary efficacy endpoints: change from baseline in mean pain intensity at week 12 of the maintenance period (week 15 of the study; US primary endpoint) or change from baseline in mean pain intensity over the entire 12 week maintenance period (European Union and other regions' primary endpoint).<br><br>Secondary outcomes: Brief Pain Inventory (BPI), Short Form-36 health survey (SF-36), euroQol-5 Dimension health questionnaire, sleep questionnaire, patient's global impression of change and % patients who responded with 30% or 50% reduction in pain intensity. |
| Notes | A significant proportion of patients from all 3 groups dropped out of the study. We received additional (unpublished) information from the authors that we used in the meta-analysis. |

***Risk of bias***

**Buynak 2010** *(Continued)*

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Randomization of patients to treatment was based on a computer-generated randomization list, balanced by randomly permuted blocks, and stratified by study site". |
| Allocation concealment (selection bias) | Low risk | "Randomization was implemented through an interactive voice response system (IVRS) that assigned patients to blinded study medication". |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | The study authors used placebo and capsules (one for each active treatment) to maintain the blind in this double-blind, double-dummy design. |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | "Investigators were not provided with the randomization codes and the schedule was maintained with the Interactive Voice Response System". |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | It is not clear if the outcomes assessors for efficacy were the same ones for safety. The side effect profile of the medication could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 167/319 (52%) in the placebo group; 155/318 (48.7%) in the tapentadol group; 195/328 (59.4%) in the oxycodone group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | High risk | The study authors performed ITT analysis using last observation carried forward, which can increase the risk of bias. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00449176); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Low risk | Patients did not differ in the baseline characteristics based on the reported table 1. |
| Influence of co-interventions | Low risk | Only acetaminophen was allowed across the groups. |
| Compliance with interventions | Unclear risk | The authors reported a high rate of non-compliance (based on the flow chart). |
| Timing of outcome assessments | Low risk | Timing of outcomes assessment was identical in both groups. |

**Chu 2012**

| | |
|---|---|
| Methods | Single-centre, randomized, double blind, placebo-controlled, parallel design for 1 month. |
| Participants | Adults with moderate to severe CLBP. Eligible patients were between ages 18 and 70 years; diagnosed with chronic nonmalignant, nonradicular LBP of at least 6 months duration. Participants were not currently taking opioid pain medication in excess of 30 mg oral morphine equivalents per day, which the research group defined as low-dose opioid therapy. |

**Chu 2012** *(Continued)*

| | |
|---|---|
| Interventions | Titrated oral morphine starting at 15 mg twice per day, followed every 2 days by a dose increase of 1 capsule per day, if tolerated, until (1) adequate analgesia (as determined by the subjects) had been achieved, (2) side effects (severe sedation, nausea or vomiting, constipation, sleep disturbances) limited further titration, or (3) a total of 8 capsules (120 mg/d of oral morphine if on active treatment) had been reached. |
| Outcomes | The primary outcome measure was opioid-induced hyperalgesia using cold pain and heat pain tolerance. The study also measured analgesic tolerance, pain scores, RMDQ, BDI, as well as symptoms and signs of opioid withdrawal. |
| Notes | The study was performed to distinguish between two different long-term effects of opioids: hyperalgesia versus tolerance. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | "Patients were randomly assigned to receive either sustained acting morphine (15 mg MS-Contin; Purdue Pharma, Stamford, CT) or weight-matched placebo capsules". There was no description of the method of randomization. |
| Allocation concealment (selection bias) | Unclear risk | Not clearly stated: "Of the 139 randomized patients, 69 were allocated to the morphine group and 70 were allocated to the placebo group". |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | "Study drugs were encapsulated in an opaque blinding capsule (DBCaps; Capsugel, Peapack, NJ) to ensure adequate blinding of the study medications". |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | Unclear whether the researchers took any approach to blind the clinicians who monitored the opioid titration. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | Unclear whether the researchers took any approach to blind the outcomes assessors. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 21/69 (30.4%) in the morphine group; 15/70 (21.4%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Unclear risk | Unclear from the text if the pain scores reported came from only those who completed the trial or all participants. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00246532); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Low risk | Groups were comparable regarding the most important demographic characteristics. |
| Influence of co-interventions | Low risk | "no patients enrolled in this study were using anticonvulsant or antidepressant drugs" |
| Compliance with interventions | Unclear risk | A higher number of participants in the morphine group discontinued the treatment compare to the patients assigned to the placebo group. |

**Chu 2012** *(Continued)*

| | | |
|---|---|---|
| Timing of outcome assessments | Low risk | Timing of outcomes assessment was identical in both groups: 1 month. |

**Gordon 2010**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, placebo-controlled, cross-over study; 4 week period. Patients who completed the trial were eligible for a 6-month open-label phase. |
| Participants | Adults reporting LBP, at least moderate in intensity, for more than 3 months and requiring more than one tablet of opioids. Patients with pain refractory to opioids were excluded as well as those with previous surgical/invasive interventions, narcotics or alcohol abusers, or those with a significant cardiovascular, pulmonary, liver or gastrointestinal disease. |
| Interventions | Patients underwent a 2 to 7 day washout of opioid analgesia before receiving patches of buprenorphine 10 µg/h or matching placebo patches. All patches were to be worn for 6 to 8 days. The initial dose was titrated weekly to 20 mcg/h and a maximum of 40 mcg/h using 10- and 20-mcg/h patches based on pain relief and adverse events. |
| Outcomes | Pain diaries (unmarked 0-100 mm VAS) two times/day and a 5-point ordinal scale; sleep questionnaires; PDI; Quebec Back Pain Disability Scale (QBPDI); SF-36 health survey; 0 to 3 treatment effectiveness; period 1, 2, none, preference; 1 to 4 ordinal scale for benefit; Subjective Opioid Withdrawal Scale; 1 to 3 ordinal scale for side effects. |
| Notes | Some important demographics were not described such as work status, disability or low-back diagnosis. We received additional (unpublished) information from the authors that we used in the meta-analysis. |

***Risk of bias***

| **Bias** | **Authors' judgement** | **Support for judgement** |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "The randomization code was generated using PROC PLAN in SAS version 6.12 (SAS Institute Inc., Cary, North Carolina)". |
| Allocation concealment (selection bias) | Low risk | "A block-randomization procedure was used to generate the treatment allocations: for every 4 successive patients, 2 received BTDS in the first phase and 2 received BTDS in the second phase. Study monitors, investigators, coordinators, pharmacists, patients, and sponsor clinical research personnel remained blinded to treatment allocation throughout the conduct of the study". |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | The study used matching placebo patches. |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | Clinical personnel remained blinded to treatment allocation. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | Research personnel remained blinded to treatment allocation. Opioid withdrawal symptoms might induce bias, but investigators reported no symptoms. However, a significant difference in nausea was reported in the adverse effects table that could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20%: 29/78 (37.1%) participants did not complete the 8 weeks of treatment. |

**Gordon 2010** *(Continued)*

| | | |
|---|---|---|
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The authors performed ITT and per-protocol analyses. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at Current Controlled Trials: ISRCTN 06013881; the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Low risk | Not applicable (cross-over design). |
| Influence of co-interventions | Low risk | Not applicable (cross-over design). |
| Compliance with interventions | Low risk | The authors reported that only one patient was non-compliant with the study medication. |
| Timing of outcome assessments | Low risk | Timing of outcomes assessment was identical in both groups. |

**Hale 2007**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, placebo-controlled, parallel, enrichment design clinical trial. |
| Participants | 251 patients were screened and included for the titration stage; 47/101 discontinued the intervention due to adverse events and 10 due to lack of efficacy. 143 patients were randomized; 49 completed in the active group compared to 18 in the placebo group.<br><br>Inclusion criteria: 18 years old, moderate/severe CLBP for at least 3 hours/day for minimum of 3 months, receiving opioids (60mg/d morphine equivalent) for 2 weeks before screening. Patients excluded: pregnant or lactating women, secondary source of pain such as infection or tumour, back surgery within 6 months, suspected neoplasm, dysphagia, hypersensitivity to opioids, seizure history, colostomy. Average age: 46, 49.1 (placebo, open-label). Work status was not documented. Pain diagnosis: degenerative disc disease (DDD) (approx 40%), herniated disc (20%), osteoarthritis (OA) (20%), spinal stenosis (2-5%), trauma (17%), other (25%), NB assumed not mutually exclusive groups. Duration of LBP was not documented. Previous non-opioid treatments were not documented. Previous back surgeries were not documented. |
| Interventions | Oxymorphone PO BID, range of opioid dose: 20-260 mg after open-label phase. Median opioid dose: 60 mg. Duration of treatment: 12 weeks during the double-blind period. Placebos were administered in the same fashion; no description of placebo type was provided. |
| Outcomes | VAS Pain Scores. The increase from baseline (at randomization) to final visit was 31.6 mm for placebo versus 8.7 mm with OPANA ER (P <0.0001). During double-blind treatment, placebo patients were approximately 8-fold more likely than OPANA ER patients to discontinue because of lack of efficacy (P < 0.001). |
| Notes | Previous surgeries, employment status, non-opioid medications were not documented. Enrichment design that could underestimate the side effects profile of the interventions. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|

**Hale 2007** *(Continued)*

| | | |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | The authors did not report the method for sequence generation. |
| Allocation concealment (selection bias) | Unclear risk | The authors did not document the method for concealment of allocation. |
| Blinding (performance bias and detection bias) All outcomes - patients | Unclear risk | The authors did not report the physical characteristics of the placebos. |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | Other than participants, it was unclear who else was blinded. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | Described as 'double-blind' but not clearly stated. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 21/70 (30%) in the oxymorphone group and 55/73 (75.3%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The authors used conservative methods for the measurement of efficacy and safety in patients who discontinued the treatment. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | High risk | 52.8 % women in the open label; however, 40% were randomized to active treatment and 24% to placebo. The study authors observed higher pain ratings and higher mean daily dosage of analgesic in the placebo group compared with the opioid arm. |
| Influence of co-interventions | Low risk | "To prevent confounding of the study through the use of other analgesics, short-acting nonsteroidal anti-inflammatory drugs or other adjuvant analgesics were not permitted". |
| Compliance with interventions | High risk | 49/70 participants in the active group and 18/73 participants in the placebo group completed the study. |
| Timing of outcome assessments | Low risk | Twelve week follow-up in both groups. |

**Hale 2010**

| | |
|---|---|
| Methods | Enrichment design, placebo-controlled, clinical trial (2 to 4 weeks open label + 12 weeks double-blinded) developed in 66 centres in USA. |
| Participants | 459 patients met the inclusion criteria but only 268 were randomized. 134 patients were analysed for the primary outcome. Male and female - 18 to 75 years old, with moderate-to-severe CLBP, at least 3 hours per day, 20 days per month, for 6 months, with non-neuropathic or neuropathic characteristics based on the Quebec Task Force Classification of Spinal Disorders (QTFCSD). Opioid treatment with > 60 mg oral morphine equivalent (12 mg hydromorphone), but < 320 mg morphine (64 mg hydromor- |

**Hale 2010** *(Continued)*

| | |
|---|---|
| | phone) per day within 2 months prior to the screening visits, and on stable doses of all prior analgesics for at least 2 weeks prior to the screening visit. Exclusion criteria: Any other chronic pain condition that would have interfered with the assessment of LBP, surgical procedure for back pain within 6 months, fibromyalgia, CRPS, acute spinal cord compression, lower extremity weakness or numbness, cauda equina, diabetic amyotrophy, diskitis, back neoplasm GI dysfunction, psychiatric condition, MAOIs within 14 days, bowel obstruction within 60 days. |
| Interventions | 12-64 mg of hydromorphone (HM) CR once a day plus HM IR as rescue medication versus matching placebos. |
| Outcomes | The primary efficacy assessment was the mean change from baseline to week 12 or final visit of the double-blind phase in weekly pain intensity 0 to 10 NRS. Secondary measures: mean change from baseline to week 12 of the double-blind phase in weighted mean pain intensity NRS score (i.e., area under the pain intensity NRS score versus time curve [area under the curve, AUC]), based on pain intensity NRS scores recorded in the patient diaries; mean change from baseline to each visit in pain intensity during the 12 week double-blind phase based on the pain intensity NRS scores. |
| Notes | All patients were opioid users. Duration of LBP was not documented. We received additional (unpublished) information from the authors that we used in the meta-analysis. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "This randomization schedule was generated by ICON Clinical Research". |
| Allocation concealment (selection bias) | Low risk | A randomization number was assigned on day 1 of the double-blind phase via an interactive voice response system (IVRS) to encode the patient's assignment to one of the two treatment groups, according to the computer-generated randomization schedule. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | After randomization, the central laboratory blinded hydromorphone results to help protect the overall blind, even though the patients may or may not have been using the rescue medication provided. |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | After randomization, the central laboratory blinded hydromorphone results to help protect the overall blind, even though the patients may or may not have been using the rescue medication provided. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | All personnel involved with the double-blind phase of the study, including the sponsor and relevant investigational staff, were blinded to the medication codes. The side effect profile of the medication could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 68/134 (50%) in the Hydromorphone group and 90/134 (67%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed primary and secondary efficacy analyses on the ITT population, defined as all patients randomized to the double-blind phase who received at least one dose of study medication after randomization. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00549042); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |

**Hale 2010** *(Continued)*

| | | |
|---|---|---|
| Group similarity at baseline | Low risk | All baseline characteristics were similar between groups |
| Influence of co-interventions | Low risk | The study authors planned primary and rescue analgesia with the study medication. |
| Compliance with interventions | Low risk | Eleven patients per group dropped out from the study due to noncompliance to the medications. |
| Timing of outcome assessments | Low risk | Twelve week follow-up in both groups |

**Katz 2007**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, placebo-controlled, parallel, enrichment design, clinical trial. |
| Participants | 326 patients were screened and included in the open-label titration period; 59/120 discontinued the intervention due to adverse events compared to 4/120 secondary to lack of efficacy. 205 patients were randomized; 71/105 completed the study in the oxymorphone group versus 47/100 in the placebo group.<br><br>Inclusion criteria: adult opioid-naive patients, defined as those taking < 5 mg/day of oxycodone or equivalent in the 2 weeks before screening, experiencing at least moderate low-back pain (> 50/100 mm) on an everyday basis and lasting for > 3 months. The study excluded patients with reflex sympathetic distrophy, acute spinal cord compression, cauda equina compression, nerve root compression, meningitis and discitis. |
| Interventions | Open-label period: titration of Oxymorphone ER starting 5 mg PO q12h for 2 days and increasing 5 to 10 mg every 3 to 7 days until dose stabilization was achieved. This dose provided tolerability and efficacy (pain ≤ 40/100) for 5 consecutive days. Subsequently, patients were randomized into a 12 week double-blind treatment period in which they received their stabilized dose of oxymorphone ER or placebo every 12 hours. All patients were allowed oxymorphone immediate release (IR) as rescue medication for breakthrough pain. Rescue medication (NSAIDs) was restricted to a maximum of two doses each day. |
| Outcomes | Patients kept a daily diary record of the total oxymorphone ER (or placebo) and IR doses and provided safety and efficacy assessments at the site on days 0, 4, 7, 14, 21, 28, 42, 56, 70 and 84 (±1 to 3 days).<br><br>Primary outcome: VAS pain score. Secondary outcomes: early discontinuation due to lack of efficacy; patient and physician global rating of the medication. Adverse effects profile.<br><br>Placebo patients discontinued significantly sooner from lack of efficacy than those receiving oxymorphone ER (P < 0.0001). Pain intensity increased significantly more in the placebo group (least squares [LS] mean<br>change 26.9 ± 2.4 [median 28.0]) than in the oxymorphone ER group (LS mean change 10.0 ± 2.4 [median 2.0]; P < 0.0001). |
| Notes | Enrichment design that could underestimate the side effects profile of the interventions. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | The study authors did not report the method of sequence generation. "A randomization code was generated by the sponsor to ensure the appropriate number of patients was allocated to each treatment group at random". |

**Katz 2007** *(Continued)*

| | | |
|---|---|---|
| Allocation concealment (selection bias) | Low risk | "Patient medication kits were assigned unique 4-digit treatment numbers according to the randomization code". Central randomization by the sponsor. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | "The tablets of oxymorphone ER and placebo were over-encapsulated with gelatin to ensure that patients, investigator/study staff, and sponsor staff remained blind to study treatment" |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | "The tablets of oxymorphone ER and placebo were over-encapsulated with gelatin to ensure that patients, investigator/study staff, and sponsor staff remained blind to study treatment" |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | The side effect profile of the medication could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 34/105 (32%) in the oxymorphone group and 53/100 (53%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | High risk | The study authors performed Last Observation Carried Forward (LOCF) for non-completers and this could favour the active treatment arm. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00225797); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Low risk | Patients were similar at baseline for age, race, sex, etiology and average pain intensity. |
| Influence of co-interventions | Low risk | The study authors restricted all co-analgesics. |
| Compliance with interventions | Unclear risk | The study authors did not reported the rate of non-compliance with the interventions. |
| Timing of outcome assessments | Low risk | The authors recorded the primary outcome on a daily basis and the study lasted for 12 weeks. |

**Khoromi 2007**

| | |
|---|---|
| Methods | Single-centre, randomized, double-blind, active placebo-controlled, four period cross-over design for nine weeks each period. |
| Participants | Pain Model: chronic sciatica (>4/10 pain intensity). Median age: 52.5 (range: 30 to 64) and duration of pain 5 years (range:0.3 to 37). 61 patients were screened, 55 randomized and 28 completed all four treatment periods. 14/28 of the completers were female. Baseline pain score (average leg) was: 4.9 ± 2.43. |
| Interventions | During 5 weeks of dose escalation and 2 weeks of maintenance at the highest tolerated dose patients received BID ER morphine (15 to 90 mg; mean 62 mg), nortriptyline (25 to 100 mg; mean 84 mg), their combination (morphine 49 mg and NT 55 mg) or benztropine-active placebo (0.25 to 1 mg); subsequently, two weeks of tapering; next period started one pain score reached > 4/10. Opioids and antidepressants were not allowed. NSAIDs and acetaminophen were used as rescue medications. |

**Khoromi 2007** *(Continued)*

| | |
|---|---|
| Outcomes | Mean scores for average leg pain during the maintenance weeks. Pain diaries consigned 0 to 10 pain score at bedtime, average back, leg and overall pain, worst back, leg and overall. Secondary: Global pain relief scores, ODI, BDI, SF-36 and general health status instrument. |
| Notes | A carry-over effect was not noticed between treatments. 27/55 (49%) participants did not complete the four periods of treatment. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Patients were assigned by random numbers within blocks of four to one of four treatment sequences specified by a Latin square". |
| Allocation concealment (selection bias) | Low risk | "Randomization was performed by the NIH Pharmaceutical Development Service". |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | "During the MS Contin treatment period, each blue pill contained MS Contin 15 mg and each pink pill contained inert placebo" |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | Providers, apparently, were also blinded of the allocation of treatments. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Low risk | The rate of guessing by the nurses was above the rate for chance only (> 25%), but did not reach a high percentage. "Patients and research staff were blinded to the randomization order". |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-outs exceeded 20% in each group: 27/55 (49%) participants did not complete the 4 periods of treatment. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors included patients with at least two periods of treatment in the ITT analysis. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00009672); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Unclear risk | Not applicable (cross-over study). |
| Influence of co-interventions | Low risk | "None of the patients who opted to participate in the study were on opioids during the two months prior to study entry". |
| Compliance with interventions | Low risk | Only one patient dropped out due to non-compliance with the treatments. |
| Timing of outcome assessments | Low risk | Timing of outcomes assessment was identical in both groups. |

**O'Donnell 2009**

| | |
|---|---|
| Methods | Two identical trials published in a single paper that were multicentre, randomized, double-blind, parallel group and active-controlled (celecoxib) studies. One study was conducted in 56 centres and the other one in 59 centres located in the USA. |
| Participants | Adults > 18 years old, with a diagnosis of LBP > 12 weeks of duration, requiring regular use of analgesics, who experienced moderate to severe LBP at baseline visit.<br><br>Exclusion criteria: CLBP that was either neurological in aetiology, due to recent major trauma, or was due to a visceral disorder; rheumatoid arthritis; spondyloarthropathy; spinal stenosis; malignancy; fibromyalgia; a herniated disc associated with neurological impairment within the past 2 years; psoriasis; seizure disorder; alcohol/analgesic/narcotic or other substance abuse within the past 2 years; asthma; urticaria or allergic-type reactions after taking aspirin or NSAIDs; gastrointestinal (GI) perforations, obstructions or bleeding; failed back surgery pain; active/suspected oesophageal,gastric, pyloric channel or duodenal ulceration or bleeding within 90 days prior to the first dose of study medication; unstable cardiovascular (CV) disease. |
| Interventions | Eligible subjects were randomized in a 1:1 ratio to receive either celecoxib 200 mg twice a day (bid) or tramadol HCl 50 mg (no titration) four times a day (qid) for 6 weeks. |
| Outcomes | Efficacy assessments were performed at baseline, screening and weeks 1, 3 and 6. The primary efficacy evaluation was based on the 0 to 10 NRS pain scale. The primary efficacy endpoint was the proportion of subjects responding successfully to their respective treatments at week 6. Successful responders were defined as subjects completing 6 weeks of treatment and having a ≥ 30% improvement from baseline to week 6 on the NRS-pain scale. |
| Notes | Important clinical factors as previous surgeries, previous analgesics, and non pharmacological treatments were not described in the baseline assessment. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Randomization was carried out using a computer-generated schedule". |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | The participants were blinded but it was not clear who else was blinded in the trial. |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | The study authors did not test the success of blinding of the interventions. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | The rate of drop-out was higher and exceeded 20% in the tramadol group compared with the celecoxib group. |
| Incomplete outcome data (attrition bias) | Low risk | The study authors performed tests for superiority on the ITT population, defined as randomized subjects who received at least one dose of study medication, and a sensitivity analysis on the evaluated population. |

**O'Donnell 2009** *(Continued)*

| | | |
|---|---|---|
| All outcomes - ITT analysis? | | |
| Selective reporting (reporting bias) | Low risk | The trials were registered at clinicaltrials.gov (NCT00290901 and NCT00662558); the primary and the secondary outcomes were consistent in the protocols compared with the publication. |
| Group similarity at baseline | Unclear risk | We had insufficient information to permit judgement as the authors did not report some important baseline characteristics. |
| Influence of co-interventions | Low risk | Cointerventions were similar between groups. |
| Compliance with interventions | High risk | The study authors reported a higher dropout rate in the intervention group (tramadol) versus the celecoxib group. |
| Timing of outcome assessments | Low risk | Timing of outcomes assessment was identical in both groups. |

**Peloso 2004**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, and placebo-controlled trial. |
| Participants | Inclusion criteria: patients with CLBP severe enough to require medications for greater than 3 months.<br><br>Exclusion criteria: use of sedative hypnotics, short-acting analgesics, topical preparation/medications and anaesthetics or muscle relaxants for a period of less than 5 half-lives of the given medication prior to the double-blind phase; use of medication that could reduce the seizure threshold; use of opioids or initiation of nutraceuticals within 6 weeks of the double blind phase; history of seizure disorder or unstable medical disease, renal or hepatic dysfunction, substance abuse, inflammatory disease and more severe pain in a location other than the lower back or other disease states that may interfere with the interpretation of pain; neurological deficit in the lower extremities, tumour or infections of the spinal cord or meninges, symptomatic disk herniation, severe spinal stenosis, spondylolisthesis or instability of lumbar vertebrae, acute vertebral fractures; back surgery (except if procedure was > 5 years prior to study enrolment); intolerant to tramadol or acetaminophen.<br><br>This study enrolled 338 participants; 167 in the opioid group and 171 in the control group. |
| Interventions | Combination tablets of tramadol (37.5 mg) and acetaminophen (325 mg) compared with placebo. Patients were treated for 3 months. |
| Outcomes | Pain (VAS), Function (Roland Disability Questionnaire). |
| Notes | |

***Risk of bias***

| **Bias** | **Authors' judgement** | **Support for judgement** |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) | Low risk | "Identical appearing tablets containing either tramadol 37.5 mg/acetaminophen 325 mg or matching placebo". |

**Peloso 2004** *(Continued)*

| | | |
|---|---|---|
| All outcomes - patients | | |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | "After the initial titration phase, patients could adjust the daily dosage of study medication as needed up to a maximum of 2 tablets QID and a minimum of 3 tablets/day". |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | The active and the placebo group were indistinguishable for the outcomes assessors. The side effect profile of the medication could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 81/167 (48%) in the tramadol/acetaminophen group and 110/171 (64%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed an ITT analysis for safety and efficacy. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | Low risk | Demographic characteristics were similar between groups. |
| Influence of co-interventions | Low risk | The study authors did not allow any pain medication or treatment other than the study medication during the course of the study, except for rescue medication (acetaminophen 500 mg, up to 4 tablets daily) during the first 6 days of the double blind phase, provided the patient was taking no more than 6 tablets of study medication daily. |
| Compliance with interventions | Unclear risk | The study authors did not contemplate counting the tablets across the duration of the study. |
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

**Ruoff 2003**

| | |
|---|---|
| Methods | Multicentre randomized, double-blind, controlled trial. |
| Participants | Inclusion criteria: Patients with CLBP severe enough to require medications for greater than 3 months.<br><br>Exclusion criteria: Previously discontinued tramadol therapy due to adverse effects or tramadol within 30 days before entry; antidepressants, cyclobenzaprine or antiepileptic drug for pain or TENS, chiropractic or acupuncture within 3 weeks of double-blind phase; sedative hypnotics, short-acting analgesics, topical anaesthetics or muscle relaxants for a period of < 5 half-lives of the specific medication before the double-blind phase; corticosteroids: injections or systemic within 3 months before screening phase; severe pain in location other than lower back or neurologic deficits in the lower extremities; contraindications to opioids or acetaminophen; major psychiatric disorders; history of suicide or substance abuse<br><br>This study enrolled 322 participants; 162 in the opioid group and 160 in the control group. |
| Interventions | Combination tablets of tramadol (37.5 mg) and acetaminophen (325 mg) compared with placebo. Patients were treated for 91 days. |

**Ruoff 2003** *(Continued)*

| | |
|---|---|
| Outcomes | Pain (VAS), function (Roland disability questionnaire) and disability (Roland disability questionnaire). |
| Notes | |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Randomization was performed using SAS version 8 (SAS Institute Inc., Cary, North Carolina)". |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | "Patients, investigators, clinical staff, and study monitors remained blinded to treatment assignments until therapy was complete and the database was finalized". |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | "Patients, investigators, clinical staff, and study monitors remained blinded to treatment assignments until therapy was complete and the database was finalized". |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | The side effect profile of the medication could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 71/162 (43%) in the tramadol/APAP and 86/160 (53%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed ITT analysis for efficacy assessment. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | Low risk | Demographic characteristics and pain scores were similar at baseline. |
| Influence of co-interventions | Unclear risk | We had insufficient information to permit judgement. |
| Compliance with interventions | Unclear risk | The study authors did not count the tablets across the duration of the study. |
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

**Schnitzer 2000**

| | |
|---|---|
| Methods | Multicentre randomized, double-blind, control trial (enrichment design). |
| Participants | Inclusion criteria: patients with CLBP severe enough to require medications for greater than 3 months |

**Schnitzer 2000** *(Continued)*

| | |
|---|---|
| | Exclusion criteria: neurologic deficit in the lower extremities; tumours or infections; lesion amenable to surgery; more severe pain in a location other than the low-back, fibromyalgia, disk herniation, spondylolisthesis, spinal stenosis, instability of lumbar vertebrae; vertebral fracture; conditions such as tumour, infection, inflammatory disease, significant hepatic or renal disease, morbid obesity or borderline personality disorder; use of systemic corticosteroids or injections in the lower back within 3 months; use of TENS; history of narcotic or alcohol abuse; score at least 3 out of 5 in the Waddell's test<br><br>This study recruited 380 patients with 254 participants enrolled in the randomized group; 127 in the opioid group and 127 in the control group. |
| Interventions | Tramadol (50 mg) with maximum 8 tablets per day (200 to 400 mg/day) or placebo. |
| Outcomes | Time to therapeutic failure, pain (VAS) and function (RMDQ). |
| Notes | |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Computer generated random numbers were used to ensure that any given patient would be assigned randomly to one of the 2 treatment groups". |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | "Identical appearing capsules of tramadol HCI and placebo were prepared by the R.W. Johnson Pharmaceutical Research Institute". |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | "Patients were randomized either to continue treatment with tramadol or to receive placebo. Dosage adjustments were allowed but the daily dose was to be maintained within the range 200-400 mg of tramadol, or an equivalent amount of placebo capsules". |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | Outcomes assessors could be bias based on the side effects profile of tramadol. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 36/127 (28.3%) in the tramadol group and 72/127 (56.7%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed ITT analysis. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | Unclear risk | We had insufficient information to permit judgement. |
| Influence of co-interventions | Low risk | "Patients were told to maintain a constant level of exercise throughout the study. Physiotherapy (i.e., back exercises, therapy, hot/cold packs, and massages) started before entrance into the open label/run-in phase was continued throughout both the open label and double blind phases of the study. Physio- |

**Schnitzer 2000** *(Continued)*

| | | |
|---|---|---|
| | | therapy could not be initiated during the open label or double blind phases of the study". |
| Compliance with interventions | High risk | More patients discontinued in the placebo groups due to inadequate pain relief. The study authors did not count the tablets across the duration of the study. |
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

**Steiner 2011**

| | |
|---|---|
| Methods | Double-blind, placebo-controlled study with an enriched design. 27-day open-label with a TD, and those who tolerated and responded to treatment were randomized into a 12 week, double-blind, placebo-controlled phase. |
| Participants | Opioid naive patients. Men and women aged ≥ 18 years; moderate to severe LBP persisting for a minimum of 3 months prior to study entry. Subjects were naïve (< 5 mg of oxycodone/day in the last 14 days). Patients with spinal stenosis, spondylosis, spondylolisthesis and OA were eligible. |
| Interventions | At the investigator's discretion, patients receiving BTDS (Buprenorphine transdermal system) 20 or placebo TDS 20 who experienced unacceptable side effects were permitted to decrease the dosage of double-blind study medication to BTDS 10 or placebo TDS 10 and could remain at that level. If analgesia was deemed inadequate with BTDS 10 or placebo TDS 10, patients were allowed to retitrate their dosages up to BTDS 20 or placebo TDS 20, at the investigator's discretion. |
| Outcomes | The primary efficacy outcome was the "average pain over the last 24 hours" score at Week 12 on an 11-point numerical rating scale collected at each clinic visit during the double blind phase (Weeks 1, 2, 4, 8, and 12). Secondary efficacy variables were sleep disturbance, as measured by the sleep disturbance subscale of the 12-item Medical Outcomes Study (MOS) Sleep Scale (0 to 100 questionnaire where higher scores indicate greater sleep disturbance); the mean daily number of tablets of non-opioid supplemental analgesic medications used during Weeks 2 through 12 of the double-blind phase. |
| Notes | There were eight exploratory variables: percent reduction in average pain score; Patient Global Impression of Change (PGIC); ODI; BPI; MOS short-form health survey (SF-36), were all collected at the beginning of the run-in period, at the randomization visit, and at each clinic visit during the double-blind phase (Weeks 1, 2, 4, 8, and 12); daily "pain right now" for those who took supplemental analgesics. The use of oxycodone for supplemental analgesia for the first six days of the double-blind phase was recorded in patient diaries. The study authors calculated the time from randomization to discontinuation because of lack of therapeutic effect. |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | "Patients who tolerated BTDS and achieved the required analgesic response (described below) were randomized in a 1:1 ratio to BTDS 10 or matching placebo transdermal system (TDS) for the 12-week double-blind phase". The study authors did not report the method of randomization. |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | ..."matching placebo transdermal system (TDS)" |

**Steiner 2011** *(Continued)*

| | | |
|---|---|---|
| Blinding (performance bias and detection bias) All outcomes - providers | High risk | The provider was alert of efficacy and side effects of the medications and had the chance to increase or decrease the dosage of the drug provided. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | We had insufficient information to permit judgement. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 86/256 (34%) in the buprenorphine group and 84/283 (30%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed sensitivity analyses of the primary efficacy variable by applying different methods for imputing missing pain scores, including the retained drop-out ITT analysis. In the retained drop-out ITT analysis, the study authors included any "average pain over the last 24 hours" scores collected subsequent to the discontinuation of study drug and prior to completion or discontinuation from the study in the analysis, and attributed them to the randomized treatment. |
| Selective reporting (reporting bias) | Low risk | The trial was registered at clinicaltrials.gov (NCT00490919); the primary and the secondary outcomes were consistent in the protocol compared with the publication. |
| Group similarity at baseline | Low risk | All demographics reported were similar between groups. |
| Influence of co-interventions | Low risk | All patients were provided with IR oxycodone for supplementary analgesia during the first six days following randomization. During weeks 2 to 12 of the double-blind phase, patients were permitted to use sponsor-provided acetaminophen 500 mg every six hours up to a maximum of 2 g/day for supplemental analgesia. Alternatively, patients for whom acetaminophen was contraindicated could use ibuprofen 200 mg every six hours up to a maximum of 800 mg/day. Use of supplemental analgesic medication was to be suspended for 30 hours prior to assessing pain at study visits to reduce any confounding effect of this medication on analgesia provided by the patient's blinded treatment. |
| Compliance with interventions | Unclear risk | We had insufficient information to permit judgement. |
| Timing of outcome assessments | Unclear risk | Timing assessment was identical between groups. |

**Uberall 2012**

| | |
|---|---|
| Methods | Randomized, double-blind, placebo-controlled, and active-controlled multicenter study. Patients with at least moderate LBP the treatment over 6 weeks (1 week wash out, 4 weeks treatment period and 1 week follow-up phase). |
| Participants | Adults aged 18 to 75 years, with LBP > 3 months. Patients who were taking analgesics for LBP but the treatment was not satisfactory, reporting at least moderate pain (> 3/10).<br><br>Exclusion criteria: neurological etiology, recent low-back trauma, significant medical or psychiatric disease. Other reasons for exclusion: rheumatoid arthritis, psoriasis, spondyloarthropathies, metabolic |

**Uberall 2012** *(Continued)*

| | |
|---|---|
| | bone disease, spinal stenosis, spinal fractures, fibromyalgia, herniated disc, substance abuse, pregnancy, child-bearing potential, amongst others. |
| Interventions | Participants had a wash-out phase for one week. They were assigned to receive flupirtine 400 mg OD, Tramadol 200 mg OD or matching placebos during four weeks. |
| Outcomes | Participants recorded lowest, average and highest pain scores. They also recorded categorical pain scores too. SF-12 and a short version of the SF-36 were also collected. QLIP inventory was also used. Other assessments included the patient's global assessment of disease status scale and the patient investigator global assessment of response to therapy. Additionally, participants and investigators used a seven step global impression of change scale after completion of the 4-week treatment period. |
| Notes | |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | "Randomization of patients to treatment was based on a computer-generated allocation list with a block-size of six, and stratified by study site". |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | Study authors described the study as double blinded with a double-dummy methodology. |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | Outcomes were collected by pain diaries. However, further information was required to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | It was unclear if patients were aware of the allocation based on the potential side effects of each treatment. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 33/118 (27%) in the tramadol group, 28/123 (22%) in the flupirtine group and 26/122 (21%) in the placebo group. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The study authors performed ITT analysis. |
| Selective reporting (reporting bias) | Low risk | The study was registered (EudraCT 2009-013268-38) and the outcomes by protocol versus reported were equal. |
| Group similarity at baseline | Low risk | Demographics were similar between groups. |
| Influence of co-interventions | Low risk | Patients had access to diclofenac as rescue therapy. They were urged to discontinue its use 24 hours before clinic visits. No other medications were allowed. |
| Compliance with interventions | Low risk | "Rates of compliance with study medication, which were estimated by returned tablet count, were more than 95% in all three study groups". |

**Uberall 2012** *(Continued)*

| | | |
|---|---|---|
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

**Vorsanger 2008**

| | |
|---|---|
| Methods | Multicentre, 3 arms (2 of tramadol and one of placebo), randomized, enrichment design trial; 3 weeks of open trial plus 12 weeks of double-blinded phase. |
| Participants | 619 patients were included in the open-label period; 128/233 discontinued the treatment due to adverse effects and 41/233 due to lack of efficacy; 386 were randomized; 42/128 discontinued the treatment in the tramadol 300 mg group versus 42/129 in the tramadol 200 mg group and 61/129 in the placebo group.<br><br>Inclusion criteria: adults with VAS Score ≥ 40/100. CLBP> 6 months. Requiring at least 90 days of NSAIDs, COX-2, opioids or muscle relaxant.<br><br>Exclusion criteria: CRPS, inflammatory pain, fibromialgia, lumbar spine surgery, not well controlled medical condition, spinal manipulation for CLBP, under TENS therapy patients, medications not allowed: NSAIDs, steroids, opioids, neuroleptics, SSRIs, SNRIs, carbamazepine and quinidine. |
| Interventions | Eligible patients receiving tramadol F.R 300 mg once daily at the end of the run-in period were randomized in a 1:1:1 ratio to receive in a double-blinded fashion tramadol extended release 300 mg, tramadol extended release 200 mg or placebo. |
| Outcomes | Pain intensity VAS since the previous visit, current pain intensity VAS, global assessment of study medication. Following randomization, mean scores/or pain intensity VAS since the previous visit, averaged over the 12 week study period, increased more in the placebo group (12.2 mm) than in the tramadol ER 300 mg (5.2 mm, P = 0. 009) and 200 mg (7.8 mm, P = 0.052) groups. Secondary efficacy scores for current pain intensity, VAS, patient global assessment, Roland Disability Index, and overall sleep quality improved significantly (P ≤ 0.029 each) in the tramadol ER groups compared with placebo. |
| Notes | Enrichment design that can underestimate the side effects profile of the interventions. |

***Risk of bias***

| **Bias** | **Authors' judgement** | **Support for judgement** |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer generated randomization |
| Allocation concealment (selection bias) | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | The study authors used only tramadol (100 mg tablets) and placebo tablets which were identical in appearance and texture. |
| Blinding (performance bias and detection bias) All outcomes - providers | Unclear risk | We had insufficient information to permit judgement. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Unclear risk | Outcomes assessor can be biased by the side effects profile of the drugs. |

**Vorsanger 2008** *(Continued)*

| | | |
|---|---|---|
| Incomplete outcome data (attrition bias)<br>All outcomes - drop-outs | High risk | % drop-out exceeded 20% in each group: 42/128 (32%) in the high dose tramadol group and 61/129 (47%) in the placebo group. |
| Incomplete outcome data (attrition bias)<br>All outcomes - ITT analysis? | Low risk | Efficacy analyses were performed on the ITT population, including all patients who received at least one dose of study medication and had primary efficacy information recorded at randomization. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | Low risk | Demographics were similar between groups. |
| Influence of co-interventions | Low risk | Study authors did not allow patients to use NSAID corticoesteroids, opioids, or other analgesic during the study, with the exception of low-dose aspirin or acetaminophen as described earlier. They also excluded neuroleptic, SSRIs, SNRIs, carbamazepine, or quinidine medications. |
| Compliance with interventions | Unclear risk | The study authors did not counting the tablets across the duration of the study. |
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

**Webster 2006**

| | |
|---|---|
| Methods | Multicentre, randomized, double-blind, active and placebo-controlled, four arms trial for 12 weeks (short-term). |
| Participants | The study authors screened 1061 patients for eligibility and randomized 719 patients into four groups. For each patient randomized to the placebo group, they allocated 2 patients to each of the active groups.<br><br>Inclusion criteria: patients aged 18 to 70, with CLBP for at least 6 months, requiring daily analgesics. Baseline pain intensity (PI) score 5 at the screening visit, a mean daily PI score 5 recorded in a diary over the last 3 days of a 4 to 10 day washout period while off all analgesics except acetaminophen, and a confirmatory PI score 5 at the baseline visit at the conclusion of the washout period.<br><br>Exclusion criteria: CLBP secondary to malignancy, autoimmune disease, fibromyalgia, recent fracture, or infection; positive urine drug screens for any illicit substance at baseline; history of substance abuse within 5 years, or involvement in litigation regarding their lower back condition; pregnancy; allergy to study medications; severe hepatic, pulmonary, or renal impairment; unstable cardiac disease; corticosteroid therapy; intraspinal analgesic infusion or spinal cord stimulator in the preceding month; major surgery in the preceding 3 months; percutaneous or open procedure of the lumbosacral spine in the preceding 4 months; or high doses of central nervous system depressants or phenothiazines. |
| Interventions | Patients were randomized to receive placebo, oxycodone qid, or oxytrex qid or oxytrex bid. Each oxytrex tablet contained 1 µg naltrexone; oxytrex bid and qid treatments provide 2 and 4 µg naltrexone/day, respectively. Following a washout, patients with pain > 5 on a 0 to 10 scale were dose-escalated weekly from 10 up to 80 mg/day until reaching adequate pain relief (< 2) or a tolerable level of side effects. Following titration, the dose was fixed for 12 weeks. |
| Outcomes | The primary efficacy measure used was the 11-point numerical diary Pain Intensity Scale. Patients were asked to record a numerical score at bedtime each day for the overall pain intensity during the past 24 hours (0 = no pain and 10 = severe pain). Secondary efficacy measures included the SF-12, and |

**Webster 2006** *(Continued)*

| | |
|---|---|
| | ODI which were collected at baseline, monthly and at the end of treatment. Other secondary efficacy assessments, conducted at each clinic visit included: the Quality of Analgesia, for which patients rated pain relief as "poor", "fair", "good", "very good" or "excellent" and the Global Assessment of Study Drug, for which patients gave an overall rating as "poor", "fair", "good", "very good" or "excellent", taking into consideration the quality of pain relief, side effects, activity level, mood and sense of well-being in this evaluation. |
| Notes | |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer generated. |
| Allocation concealment (selection bias) | Low risk | ..."Patients then remained on their individual fixed doses randomized via a central call-in system to 1 of 4 treatments in a 1:2:2:2". |
| Blinding (performance bias and detection bias) All outcomes - patients | Low risk | All study medications were identical in appearance, and patients, site personnel, and study monitors were blinded to treatment assignments. |
| Blinding (performance bias and detection bias) All outcomes - providers | Low risk | All study medications were identical in appearance, and patients, site personnel, and study monitors were blinded to treatment assignments. |
| Blinding (performance bias and detection bias) All outcomes - outcome assessors | Low risk | We had insufficient information to permit judgement. The side effects of the drugs could induce bias. |
| Incomplete outcome data (attrition bias) All outcomes - drop-outs | High risk | All active groups and the placebo group had a drop-out rate > 20%. |
| Incomplete outcome data (attrition bias) All outcomes - ITT analysis? | Low risk | The primary analysis population for both efficacy and safety included the ITT population consisting of all randomized patients who took at least one dose of study medication and had at least one post-baseline PI assessment. |
| Selective reporting (reporting bias) | Unclear risk | We could not identify the protocol of this study in the databases for registration of clinical trials. |
| Group similarity at baseline | Low risk | Demographics were similar between groups. |
| Influence of co-interventions | Low risk | "No other analgesics were allowed during the treatment period". |
| Compliance with interventions | Unclear risk | The study authors did not count the tablets across the duration of the study. |
| Timing of outcome assessments | Low risk | Timing assessment was identical between groups. |

VAS - visual analog scale

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

| Study | Reason for exclusion |
|---|---|
| Adams 2006 | Single arm, open label, short term study (3 months only) in opioid-naive patients. |
| Allan 2005 | Open label, randomized, parallel groups, multicenter study in opioid naive patients. Opioid versus opioid comparison. |
| Beaulieu 2007 | Opioid versus opioid comparison. |
| Cloutier 2013 | All participants received an opioid as analgesic rescue. |
| Gaertner 2006 | Narrative review of 6 previously published open label studies of CR oxycodone. |
| Gordon 2010a | The study included patients with at least 6 weeks of LBP. The inclusion criteria of this review was restricted to studies that have evaluated at least 3 months of LBP. |
| Gostick 1989 | Opioid versus opioid comparison. |
| Gould 2009 | Secondary analysis based on pain descriptors of Hale 2007. |
| Hale 1997 | Opioid versus opioid comparison. |
| Hale 1999 | Oxycodone CR versus Oxycodone IR comparison. |
| Hale 2005 | Patients randomized to the oxymorphone versus oxycodone phase only. Opioid compared with placebo phase was not randomized. |
| Hale 2009 | Tapentadol versus oxycodone comparison. The number of patients with CLBP was unclear. |
| Jamison 1998 | The study was not blinded. |
| Kalso 2007 | This is a secondary analysis of Allan 2005 (see above). |
| Kuntz 1996 | Less than 4 week study duration. |
| Landau 2007 | The primary site of pain was back in the 49.1% of the population. |
| Li 2008 | The study was developed for 7 days only. |
| Likar 2007 | Open label and opioid versus opioid comparison. |
| Moulin 1996 | Less than 50% of study sample included diagnosis of CLBP and results were not reported by diagnostic condition. |
| Muller 1998 | Less than four week study duration. |
| Nicholson 2006 | A comparison of a single medication administered in the morning versus the evening. |
| Nicholson 2006a | Opioid versus opioid comparison. |
| Pascual 2007 | Open label study. |
| Peniston 2009 | Retrospective analysis. |
| Perrot 2006 | Opioid versus opioid comparison. |

| Study | Reason for exclusion |
|---|---|
| Portenoy 2007 | The trial was focused in the treatment of breakthrough low-back pain. |
| Rauck 2006 | Open label study and opioid versus opioid comparison. |
| Rauck 2007 | Open label study and opioid versus opioid comparison. |
| Salzman 1999 | Opioid versus opioid comparison. |
| Sarbu 2008 | Study about pharmacokinetics properties of tramadol. |
| Taylor 2007 | Multicentre patient reported survey. |
| Volinn 2009 | Observational study. |
| Vondrackova 2008 | After opioid taper and opioid titration during pre-randomization, patients were converted to an equivalent study medication with additional OxyNorm as rescue medication in the double-blind phase. The study therefore does not represent a "true" placebo-controlled design. |
| Wallace 2007 | Open label observational study. |
| Weinstein 2006 | Open label study. |
| Wiesel 1980 | Prospective observational study. |

## Characteristics of ongoing studies *[ordered by study ID]*

**NCT01081912**

| | |
|---|---|
| Trial name or title | A randomized double-blind, placebo-controlled Trial to evaluate the efficacy, tolerability and safety of hydrocodone bitartrate controlled-release capsules in opioid-experienced subjects with moderate to severe CLBP. |
| Methods | Randomized, double-blind trial. |
| Participants | Adult patients suffering from CLBP. Subjects must be classified as non-neuropathic, neuropathic, or symptomatic for more than 6 months after LBP surgery. |
| Interventions | The trial will consist of a screening phase (up to 14 days), an open-label conversion and titration phase (up to 6 weeks), a 12-week placebo-controlled treatment phase, and a 2 week follow-up phone call. Participants are randomized to receive 10 to 50 mg of hydrocodone or placebo capsules. |
| Outcomes | Primary outcome: change in average pain intensity as measured daily by a 0 to 10 Numerical Rating Scale (NRS). Secondary outcome: the change in pain intensity as measured in the clinic by a 0 to 10 NRS. |
| Starting date | March 4, 2010. |
| Contact information | John Ning, MD. |
| Notes | Study completion date: October 2011. |

**NCT01358526**

| | |
|---|---|
| Trial name or title | A Randomized, double-blind, placebo-controlled, multicentre trial with an enriched study design to assess the efficacy and safety of oxycodone/naloxone controlled-release tablets (OXN) compared to placebo in opioid-experienced subjects with moderate to severe pain due to chronic low back pain who require around-the-clock opioid therapy |
| Methods | Randomized, double-blind trial. |
| Participants | Adults with moderate to severe CLBP as their predominant pain condition for at least 3 months prior to screening period; the pain must be related; sciatica must be ruled out. Study will include patients with a stable regimen of opioid. |
| Interventions | Oxycodone/naloxone controlled-release tablets (10/5 mg, 20/10 mg, 30/15 mg, 40/20 mg) taken orally every 12 hours or placebo for 12 weeks. |
| Outcomes | The "average pain over the last 24 hours" at week 12 of the double-blind period; The Sleep Disturbance Subscale of the MOS Sleep Scale; Patient Global Impression of Change (PGIC). |
| Starting date | May 2011. |
| Contact information | Purdue Pharma LP. |
| Notes | Study completion date: October 2012. |

**NCT01455519**

| | |
|---|---|
| Trial name or title | True functional restoration and analgesia in non-radicular low back pain: a prospective double blind, placebo-controlled study of hydromorphone ER. |
| Methods | Randomized, double-blind trial. |
| Participants | Patients with CLBP. Non-radiating pain (below buttocks), no frank weakness or atrophy, no sensory or reflex changes. |
| Interventions | Total target dose of 32 mg/day. All subjects will have a lead in for 2 weeks; then begin a "forced" 2-week up-titration schedule as follows: 8 mg/d (1 pill, 5 days), 16 mgday (2 pills, 5 days), and 24 mg/day (3 pills, 5 days) then finally 32 mg/day (4 pills/day) for the "stable dose" phase of the study, or identical placebo pills. |
| Outcomes | The efficacy of hydromorphone extended release in chronic non-radicular low-back pain to improve pain, function and activity. |
| Starting date | October 2011. |
| Contact information | akirsling@ric.org |
| Notes | Estimated primary completion date: November 2013. |

**NCT01571362**

| | |
|---|---|
| Trial name or title | A multicentre, 12 week, double-blind, placebo-controlled, randomized withdrawal study to determine the efficacy and safety of ALO-02 (oxycodone hydrochloride and naltrexone hydrochloride) extended-release capsules in subjects with moderate to severe CLBP. |

**NCT01571362** *(Continued)*

| | |
|---|---|
| Methods | Randomized, double-blind trial. |
| Participants | The patient has had moderate to severe CLBP for at least 3 months duration before screening. |
| Interventions | 20 to 160mg total daily dose of oxycodone, divided into symmetric doses and administered twice daily versus placebo for 12 weeks. |
| Outcomes | Change in baseline in daily average pain numerical rating scale scores; % reduction in daily average pain numerical rating scale scores; changes in brief pain inventory-short form; change from baseline in brief pain inventory-short form; coanalgesia requirement; changes in patient's global assessment of low back pain; changes in Roland Morris disability questionnaire; change from baseline in healthcare resource use questionnaire; changes in EQ-5D health questionnaire. |
| Starting date | June 2012. |
| Contact information | Pfizer. |
| Notes | Ongoing trial. |

**NCT01789970**

| | |
|---|---|
| Trial name or title | A 12 week, randomized, double-blind, placebo-controlled, randomized-withdrawal study to evaluate the efficacy and safety of hydrocodone bitartrate extended-release tablets (CEP-33237) at 30 to 90 mg every 12 hours for relief of CLBP who require opioid treatment for an extended period of time. |
| Methods | Randomized, double-blind trial. |
| Participants | The patient has had moderate to severe CLBP for at least 3 months duration before screening. |
| Interventions | Hydrocodone bitartrate extended-release tablets or placebo tablets will be self-administered by patients at doses of 15, 30, 45, 60, or 90 mg, with each dose taken every 12 hours, during 12 weeks. |
| Outcomes | Weekly average of daily worst pain intensity (WPI) scores; weekly average of daily average pain intensity (API) scores; adverse effects profile. |
| Starting date | March 2013. |
| Contact information | Teva GCO. |
| Notes | Estimated primary completion date: January 2014. |

## DATA AND ANALYSES

## Comparison 1. Tramadol compared to placebo

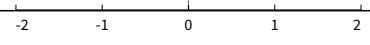
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Pain intensity (higher score means worse pain levels) | 5 | 1378 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.55 [-0.66, -0.44] |
| 2 Disability (higher ratings mean greater disability) | 5 | 1348 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.18 [-0.29, -0.07] |
| 3 Side effects | 5 | | Risk Difference (M-H, Random, 95% CI) | Subtotals only |
| 3.1 Nausea | 5 | 1401 | Risk Difference (M-H, Random, 95% CI) | 0.09 [0.05, 0.13] |
| 3.2 Constipation | 4 | 1147 | Risk Difference (M-H, Random, 95% CI) | 0.05 [0.02, 0.09] |
| 3.3 Somnolence | 3 | 911 | Risk Difference (M-H, Random, 95% CI) | 0.06 [-0.01, 0.13] |

### Analysis 1.1. Comparison 1 Tramadol compared to placebo, Outcome 1 Pain intensity (higher score means worse pain levels).

| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Std. Mean Difference Fixed, 95% CI | Weight | Std. Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Peloso 2004 | 167 | 47.4 (15) | 169 | 62.9 (15.5) | | 22.64% | -1.01[-1.24,-0.79] |
| Ruoff 2003 | 161 | 44.4 (14.5) | 157 | 52.3 (14.9) | | 23.36% | -0.54[-0.76,-0.31] |
| Schnitzer 2000 | 127 | 3.5 (2.8) | 127 | 5.1 (3) | | 18.62% | -0.55[-0.8,-0.3] |
| Uberall 2012 | 107 | 3.9 (2) | 110 | 4.1 (2) | | 16.5% | -0.1[-0.37,0.17] |
| Vorsanger 2008 | 127 | 30.5 (23) | 126 | 40.3 (25.2) | | 18.87% | -0.41[-0.65,-0.16] |
| **Total ***** | **689** | | **689** | | | **100%** | **-0.55[-0.66,-0.44]** |

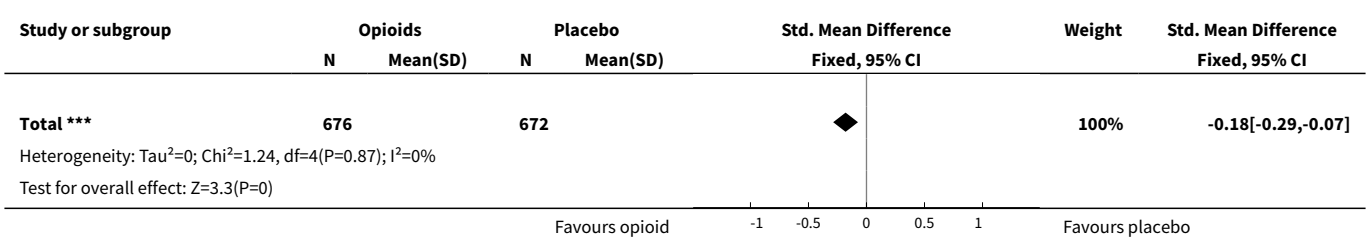
Heterogeneity: Tau²=0; Chi²=28.34, df=4(P<0.0001); I²=85.88%

Test for overall effect: Z=9.98(P<0.0001)

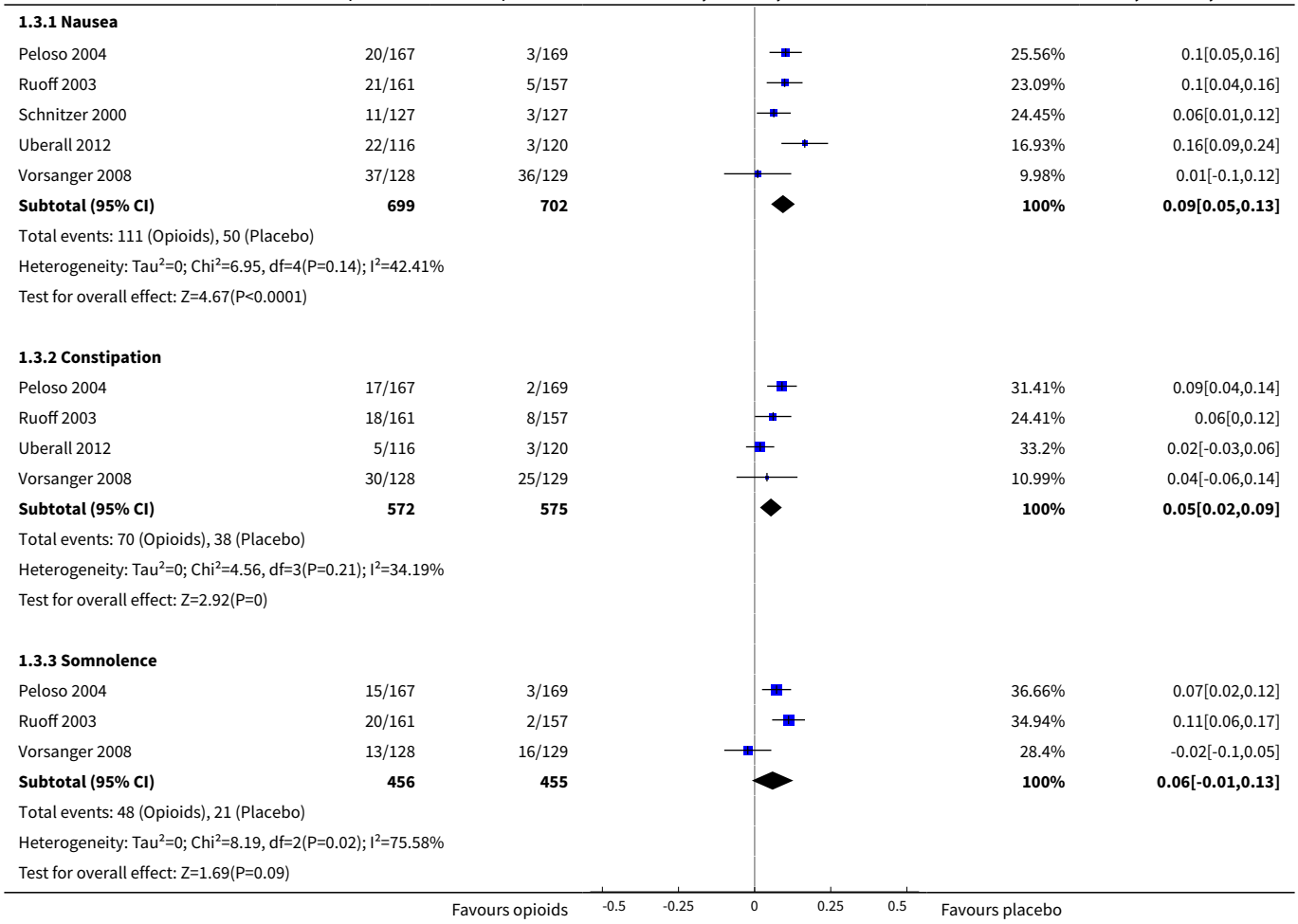



### Analysis 1.2. Comparison 1 Tramadol compared to placebo, Outcome 2 Disability (higher ratings mean greater disability).

| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Std. Mean Difference Fixed, 95% CI | Weight | Std. Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Peloso 2004 | 164 | 12.8 (5.9) | 163 | 13.7 (5.7) | | 24.3% | -0.15[-0.37,0.06] |
| Ruoff 2003 | 151 | 10.7 (6.3) | 146 | 11.6 (6.3) | | 22.07% | -0.14[-0.37,0.09] |
| Schnitzer 2000 | 127 | 8.8 (6.2) | 127 | 10.2 (6.2) | | 18.81% | -0.23[-0.47,0.02] |
| Uberall 2012 | 107 | 3.9 (1.9) | 110 | 4.1 (2) | | 16.15% | -0.1[-0.37,0.16] |
| Vorsanger 2008 | 127 | 8.2 (5.5) | 126 | 9.8 (5.9) | | 18.67% | -0.28[-0.53,-0.03] |




## Analysis 1.3. Comparison 1 Tramadol compared to placebo, Outcome 3 Side effects.

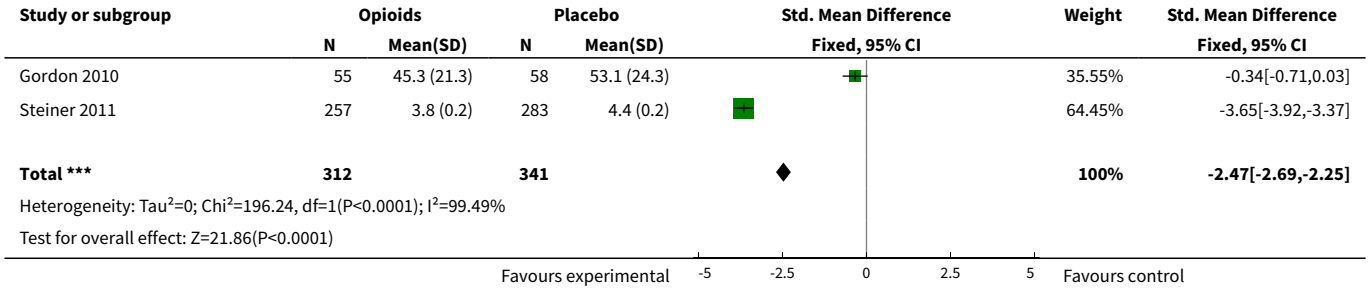



## Comparison 2. Buprenorphine compared to placebo

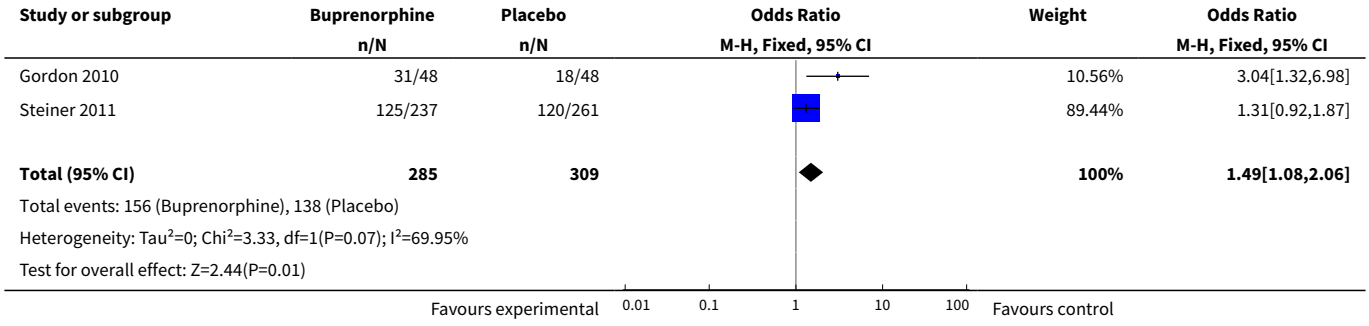
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Mean pain intensity | 2 | 653 | Std. Mean Difference (IV, Fixed, 95% CI) | -2.47 [-2.69, -2.25] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 2 At least 30% of pain relief or moderate improvement | 2 | 594 | Odds Ratio (M-H, Fixed, 95% CI) | 1.49 [1.08, 2.06] |
| 3 At least 50% of pain relief | 1 | 498 | Odds Ratio (M-H, Fixed, 95% CI) | 1.39 [0.97, 1.99] |
| 4 Disability (higher ratings mean greater disability) | 1 | 101 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.14 [-0.53, 0.25] |

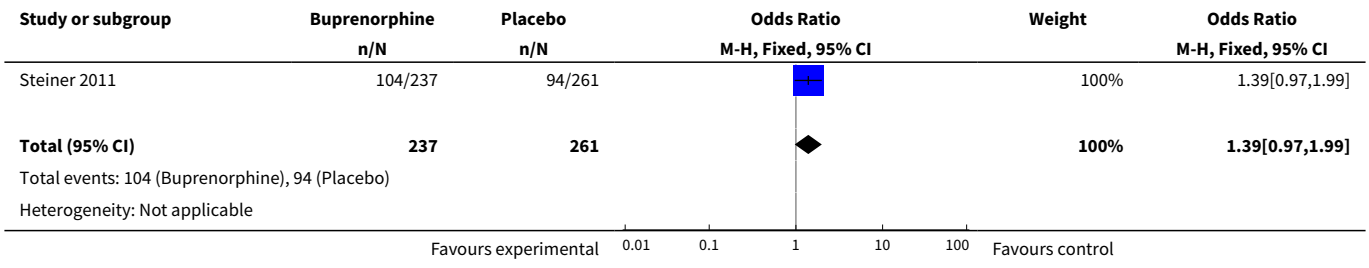
### Analysis 2.1. Comparison 2 Buprenorphine compared to placebo, Outcome 1 Mean pain intensity.

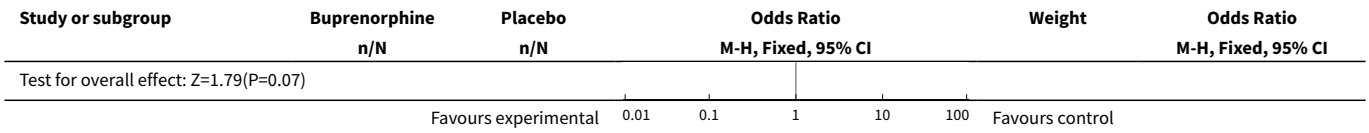



### Analysis 2.2. Comparison 2 Buprenorphine compared to placebo, Outcome 2 At least 30% of pain relief or moderate improvement.

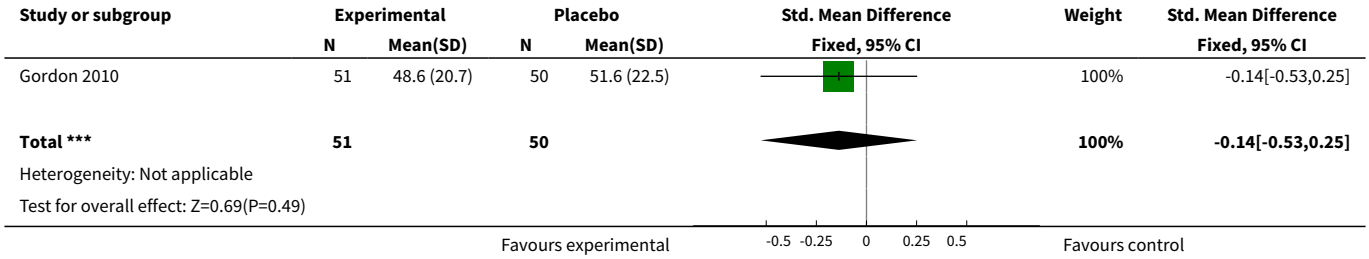



### Analysis 2.3. Comparison 2 Buprenorphine compared to placebo, Outcome 3 At least 50% of pain relief.





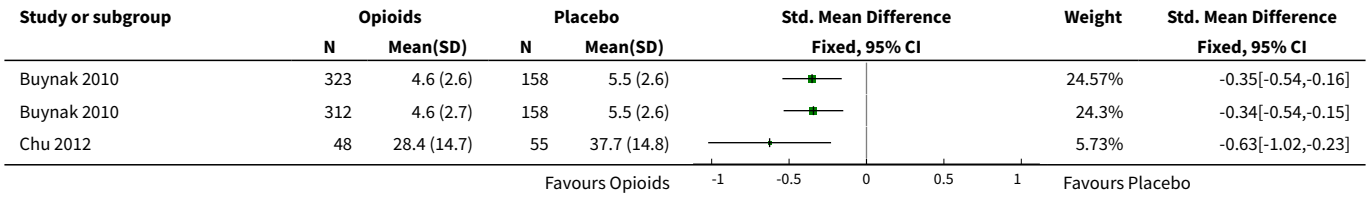

### Analysis 2.4. Comparison 2 Buprenorphine compared to placebo, Outcome 4 Disability (higher ratings mean greater disability).



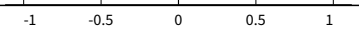

## Comparison 3. Strong opioids compared to placebo

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Mean pain intensity | 6 | 1887 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.43 [-0.52, -0.33] |
| 2 At least 30% of pain relief or moderate relief | 3 | 819 | Odds Ratio (M-H, Fixed, 95% CI) | 1.91 [1.41, 2.58] |
| 3 At least 50% of pain relief | 2 | 750 | Odds Ratio (M-H, Fixed, 95% CI) | 1.89 [1.34, 2.66] |
| 4 Disability | 4 | 1375 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.26 [-0.37, -0.15] |
| 5 Side effects | 5 | | Risk Difference (M-H, Random, 95% CI) | Subtotals only |
| 5.1 Nausea | 5 | 2346 | Risk Difference (M-H, Random, 95% CI) | 0.12 [0.05, 0.19] |
| 5.2 Constipation | 5 | 2346 | Risk Difference (M-H, Random, 95% CI) | 0.11 [0.04, 0.19] |
| 5.3 Somnolence | 5 | 2346 | Risk Difference (M-H, Random, 95% CI) | 0.06 [0.02, 0.10] |

### Analysis 3.1. Comparison 3 Strong opioids compared to placebo, Outcome 1 Mean pain intensity.

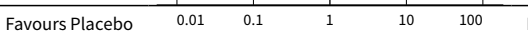


| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Std. Mean Difference Fixed, 95% CI | Weight | Std. Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Hale 2010 | 133 | 3.8 (1.8) | 133 | 4.8 (2) | | 15.09% | -0.52[-0.77,-0.28] |
| Katz 2007 | 105 | 29.5 (26.2) | 100 | 45.5 (26.9) | | 11.5% | -0.6[-0.88,-0.32] |
| Khoromi 2007 | 28 | 3.4 (2.5) | 28 | 3.8 (2.5) | | 3.28% | -0.16[-0.68,0.37] |
| Webster 2006 | 205 | 4 (2.5) | 101 | 5.2 (3.1) | | 15.54% | -0.44[-0.68,-0.2] |
| **Total *** ** | **1154** | | **733** | | | **100%** | **-0.43[-0.52,-0.33]** |
| Heterogeneity: Tau²=0; Chi²=5.39, df=6(P=0.49); I²=0% | | | | | | | |
| Test for overall effect: Z=8.82(P<0.0001) | | | | | | | |



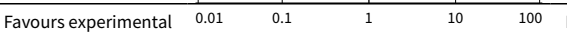

## Analysis 3.2. Comparison 3 Strong opioids compared to placebo, Outcome 2 At least 30% of pain relief or moderate relief.

| Study or subgroup | Opioids n/N | Placebo n/N | Odds Ratio M-H, Fixed, 95% CI | Weight | Odds Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| Buynak 2010 | 125/315 | 86/317 | | 85.26% | 1.77[1.26,2.47] |
| Katz 2007 | 66/71 | 34/47 | | 4.75% | 5.05[1.66,15.33] |
| Khoromi 2007 | 13/32 | 11/37 | | 9.99% | 1.62[0.6,4.38] |
| **Total (95% CI)** | **418** | **401** | | **100%** | **1.91[1.41,2.58]** |
| Total events: 204 (Opioids), 131 (Placebo) | | | | | |
| Heterogeneity: Tau²=0; Chi²=3.25, df=2(P=0.2); I²=38.47% | | | | | |
| Test for overall effect: Z=4.17(P<0.0001) | | | | | |




## Analysis 3.3. Comparison 3 Strong opioids compared to placebo, Outcome 3 At least 50% of pain relief.

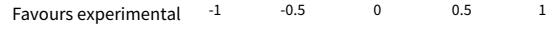
| Study or subgroup | Opioids n/N | Placebo n/N | Odds Ratio M-H, Fixed, 95% CI | Weight | Odds Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| Buynak 2010 | 85/315 | 60/317 | | 90.83% | 1.58[1.09,2.3] |
| Katz 2007 | 61/71 | 26/47 | | 9.17% | 4.93[2.04,11.9] |
| **Total (95% CI)** | **386** | **364** | | **100%** | **1.89[1.34,2.66]** |
| Total events: 146 (Opioids), 86 (Placebo) | | | | | |
| Heterogeneity: Tau²=0; Chi²=5.39, df=1(P=0.02); I²=81.45% | | | | | |
| Test for overall effect: Z=3.64(P=0) | | | | | |




## Analysis 3.4. Comparison 3 Strong opioids compared to placebo, Outcome 4 Disability.

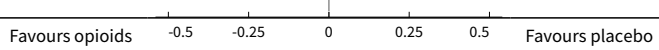
| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Std. Mean Difference Fixed, 95% CI | Weight | Std. Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Buynak 2010 | 323 | 4.3 (2.8) | 157 | 4.8 (2.8) | | 33.57% | -0.18[-0.37,0.01] |
| Buynak 2010 | 314 | 4.1 (2.8) | 158 | 4.8 (2.8) | | 33.28% | -0.25[-0.44,-0.06] |

| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Std. Mean Difference Fixed, 95% CI | Weight | Std. Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Chu 2012 | 48 | 5.1 (4.7) | 55 | 7.1 (4.6) | | 7.98% | -0.43[-0.82,-0.04] |
| Hale 2010 | 132 | 9.6 (6.3) | 132 | 11.7 (6.1) | | 20.75% | -0.34[-0.58,-0.09] |
| Khoromi 2007 | 28 | 25.7 (16.5) | 28 | 30.5 (15.9) | | 4.41% | -0.29[-0.82,0.23] |
| **Total *** ** | **845** | | **530** | | | **100%** | **-0.26[-0.37,-0.15]** |
| Heterogeneity: $Tau^2$=0; $Chi^2$=1.79, df=4(P=0.78); $I^2$=0% | | | | | | | |
| Test for overall effect: Z=4.62(P<0.0001) | | | | | | | |




## Analysis 3.5. Comparison 3 Strong opioids compared to placebo, Outcome 5 Side effects.

| Study or subgroup | Opioids n/N | Placebo n/N | Risk Difference M-H, Random, 95% CI | Weight | Risk Difference M-H, Random, 95% CI |
|---|---|---|---|---|---|
| **3.5.1 Nausea** | | | | | |
| Buynak 2010 | 64/318 | 29/319 | | 24.59% | 0.11[0.06,0.16] |
| Gordon 2010 | 39/73 | 12/68 | | 12.68% | 0.36[0.21,0.5] |
| Katz 2007 | 12/105 | 9/100 | | 20.49% | 0.02[-0.06,0.11] |
| Khoromi 2007 | 2/28 | 0/28 | | 16.42% | 0.07[-0.04,0.18] |
| Steiner 2011 | 240/1024 | 31/283 | | 25.83% | 0.12[0.08,0.17] |
| **Subtotal (95% CI)** | **1548** | **798** | | **100%** | **0.12[0.05,0.19]** |
| Total events: 357 (Opioids), 81 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=16.59, df=4(P=0); $I^2$=75.89% | | | | | |
| Test for overall effect: Z=3.43(P=0) | | | | | |
| **3.5.2 Constipation** | | | | | |
| Buynak 2010 | 44/318 | 16/319 | | 24.39% | 0.09[0.04,0.13] |
| Gordon 2010 | 12/73 | 4/68 | | 17.74% | 0.11[0,0.21] |
| Katz 2007 | 7/105 | 4/100 | | 22.62% | 0.03[-0.03,0.09] |
| Khoromi 2007 | 18/28 | 2/28 | | 8.92% | 0.57[0.37,0.77] |
| Steiner 2011 | 67/1024 | 3/283 | | 26.33% | 0.05[0.04,0.07] |
| **Subtotal (95% CI)** | **1548** | **798** | | **100%** | **0.11[0.04,0.19]** |
| Total events: 148 (Opioids), 29 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0.01; $Chi^2$=33.5, df=4(P<0.0001); $I^2$=88.06% | | | | | |
| Test for overall effect: Z=2.97(P=0) | | | | | |
| **3.5.3 Somnolence** | | | | | |
| Buynak 2010 | 26/318 | 8/319 | | 27.54% | 0.06[0.02,0.09] |
| Gordon 2010 | 16/73 | 5/68 | | 8.36% | 0.15[0.03,0.26] |
| Katz 2007 | 2/105 | 0/100 | | 28.54% | 0.02[-0.01,0.05] |
| Khoromi 2007 | 7/28 | 1/28 | | 4.08% | 0.21[0.04,0.39] |
| Steiner 2011 | 84/1024 | 6/283 | | 31.48% | 0.06[0.04,0.08] |
| **Subtotal (95% CI)** | **1548** | **798** | | **100%** | **0.06[0.02,0.1]** |
| Total events: 135 (Opioids), 20 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=13.29, df=4(P=0.01); $I^2$=69.9% | | | | | |
| Test for overall effect: Z=3.2(P=0) | | | | | |



**Comparison 4. Opioids (all types) compared to placebo**

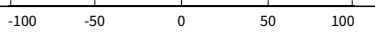
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Mean change in pain intensity in enriched design trials (higher changes indicate less favourable scores) | 3 | 382 | Mean Difference (IV, Fixed, 95% CI) | -21.34 [-22.77, -19.91] |
| 2 Side effects | 10 | | Risk Difference (M-H, Random, 95% CI) | Subtotals only |
| 2.1 Nausea | 10 | 3747 | Risk Difference (M-H, Random, 95% CI) | 0.10 [0.07, 0.14] |
| 2.2 Somnolence | 8 | 3257 | Risk Difference (M-H, Random, 95% CI) | 0.06 [0.03, 0.09] |
| 2.3 Constipation | 9 | 3493 | Risk Difference (M-H, Random, 95% CI) | 0.07 [0.04, 0.11] |
| 2.4 Headaches | 10 | 3747 | Risk Difference (M-H, Random, 95% CI) | 0.03 [0.01, 0.05] |
| 2.5 Dry mouth | 6 | 1724 | Risk Difference (M-H, Random, 95% CI) | 0.06 [0.02, 0.10] |
| 2.6 Dizziness | 9 | 3493 | Risk Difference (M-H, Random, 95% CI) | 0.08 [0.05, 0.11] |
| 2.7 Pruritis | 6 | 2865 | Risk Difference (M-H, Random, 95% CI) | 0.04 [0.02, 0.05] |
| 2.8 Fatigue | 6 | 1645 | Risk Difference (M-H, Random, 95% CI) | 0.03 [0.01, 0.05] |
| 2.9 Upper respiratory tract infection | 1 | 318 | Risk Difference (M-H, Random, 95% CI) | -0.02 [-0.08, 0.03] |
| 2.10 Sinusitis | 1 | 318 | Risk Difference (M-H, Random, 95% CI) | 0.02 [-0.03, 0.06] |
| 2.11 Vomiting | 7 | 3119 | Risk Difference (M-H, Random, 95% CI) | 0.07 [0.04, 0.09] |
| 2.12 Anorexia | 2 | 386 | Risk Difference (M-H, Random, 95% CI) | 0.04 [0.01, 0.07] |
| 2.13 Increased sweating | 4 | 1350 | Risk Difference (M-H, Random, 95% CI) | 0.04 [0.02, 0.05] |
| 2.14 Hot flushes | 2 | 593 | Risk Difference (M-H, Random, 95% CI) | 0.03 [0.00, 0.05] |

**Analysis 4.1. Comparison 4 Opioids (all types) compared to placebo, Outcome 1 Mean change in pain intensity in enriched design trials (higher changes indicate less favourable scores).**

| Study or subgroup | Opioids N | Opioids Mean(SD) | Placebo N | Placebo Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Hale 2007 | 49 | 8.7 (3) | 18 | 31.6 (2.9) | | 81.96% | -22.9[-24.48,-21.32] |
| Hale 2010 | 66 | 2 (10) | 44 | 16 (10) | | 14.08% | -14[-17.81,-10.19] |
| Katz 2007 | 105 | 10.9 (24.5) | 100 | 26 (27.9) | | 3.95% | -15.1[-22.3,-7.9] |
| **Total ***** | **220** | | **162** | | | **100%** | **-21.34[-22.77,-19.91]** |

Heterogeneity: $Tau^2$=0; $Chi^2$=20.85, df=2(P<0.0001); $I^2$=90.41%

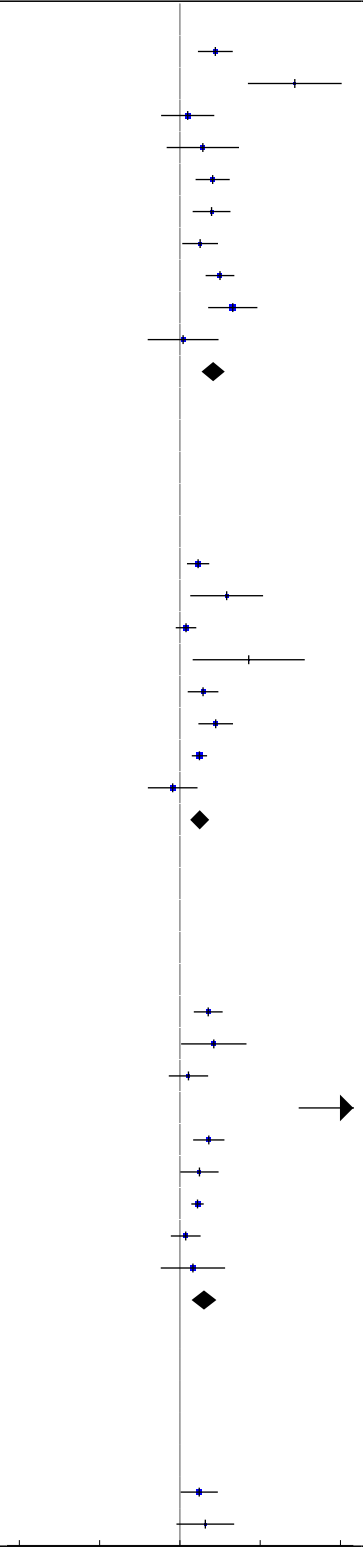
Test for overall effect: Z=29.21(P<0.0001)



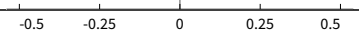

Favours Opioids Favours Placebo

## Analysis 4.2. Comparison 4 Opioids (all types) compared to placebo, Outcome 2 Side effects.

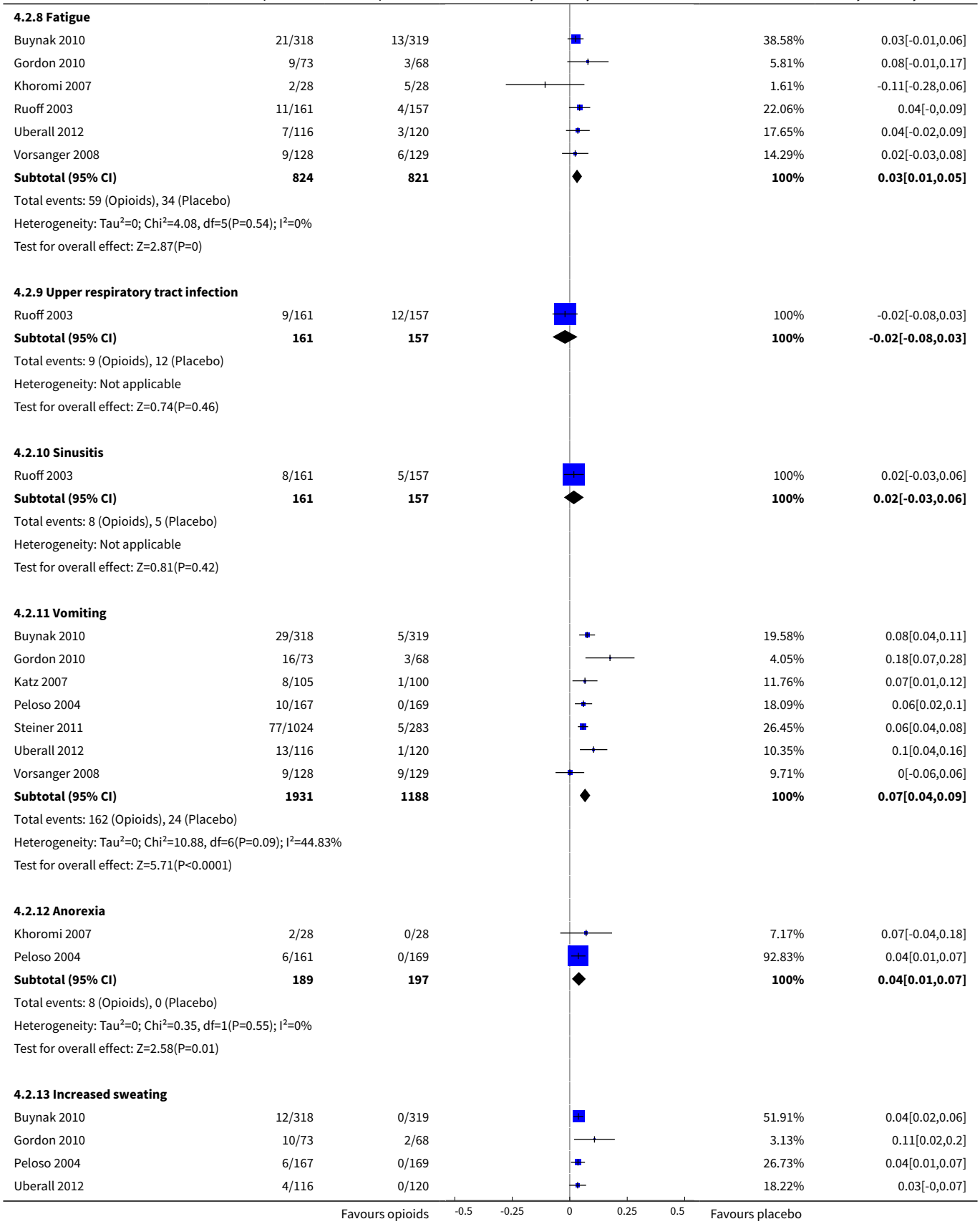
| Study or subgroup | Opioids n/N | Placebo n/N | Risk Difference M-H, Random, 95% CI | Weight | Risk Difference M-H, Random, 95% CI |
|---|---|---|---|---|---|
| **4.2.1 Nausea** | | | | | |
| Buynak 2010 | 64/318 | 29/319 | | 12.55% | 0.11[0.06,0.16] |
| Gordon 2010 | 39/73 | 12/68 | | 4.55% | 0.36[0.21,0.5] |
| Katz 2007 | 12/105 | 9/100 | | 9.13% | 0.02[-0.06,0.11] |
| Khoromi 2007 | 2/28 | 0/28 | | 6.49% | 0.07[-0.04,0.18] |
| Peloso 2004 | 20/167 | 3/169 | | 12.7% | 0.1[0.05,0.16] |
| Ruoff 2003 | 21/161 | 5/157 | | 11.95% | 0.1[0.04,0.16] |
| Schnitzer 2000 | 11/127 | 3/127 | | 12.37% | 0.06[0.01,0.12] |
| Steiner 2011 | 240/1024 | 31/283 | | 13.8% | 0.12[0.08,0.17] |
| Uberall 2012 | 22/116 | 3/120 | | 9.8% | 0.16[0.09,0.24] |
| Vorsanger 2008 | 37/128 | 36/129 | | 6.66% | 0.01[-0.1,0.12] |
| **Subtotal (95% CI)** | **2247** | **1500** | | **100%** | **0.1[0.07,0.14]** |
| Total events: 468 (Opioids), 131 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=24.07, df=9(P=0); $I^2$=62.6% | | | | | |
| Test for overall effect: Z=5.61(P<0.0001) | | | | | |
| **4.2.2 Somnolence** | | | | | |
| Buynak 2010 | 26/318 | 8/319 | | 17.51% | 0.06[0.02,0.09] |
| Gordon 2010 | 16/73 | 5/68 | | 5.26% | 0.15[0.03,0.26] |
| Katz 2007 | 2/105 | 0/100 | | 18.16% | 0.02[-0.01,0.05] |
| Khoromi 2007 | 7/28 | 1/28 | | 2.56% | 0.21[0.04,0.39] |
| Peloso 2004 | 15/167 | 3/169 | | 14.41% | 0.07[0.02,0.12] |
| Ruoff 2003 | 20/161 | 2/157 | | 13.07% | 0.11[0.06,0.17] |
| Steiner 2011 | 84/1024 | 6/283 | | 20.07% | 0.06[0.04,0.08] |
| Vorsanger 2008 | 13/128 | 16/129 | | 8.96% | -0.02[-0.1,0.05] |
| **Subtotal (95% CI)** | **2004** | **1253** | | **100%** | **0.06[0.03,0.09]** |
| Total events: 183 (Opioids), 41 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=20.46, df=7(P=0); $I^2$=65.78% | | | | | |
| Test for overall effect: Z=4.08(P<0.0001) | | | | | |
| **4.2.3 Constipation** | | | | | |
| Buynak 2010 | 44/318 | 16/319 | | 13.83% | 0.09[0.04,0.13] |
| Gordon 2010 | 12/73 | 4/68 | | 7.8% | 0.11[0,0.21] |
| Katz 2007 | 7/105 | 4/100 | | 11.92% | 0.03[-0.03,0.09] |
| Khoromi 2007 | 18/28 | 2/28 | | 3.02% | 0.57[0.37,0.77] |
| Peloso 2004 | 17/167 | 2/169 | | 13.39% | 0.09[0.04,0.14] |
| Ruoff 2003 | 18/161 | 8/157 | | 12.11% | 0.06[0,0.12] |
| Steiner 2011 | 67/1024 | 3/283 | | 16.32% | 0.05[0.04,0.07] |
| Uberall 2012 | 5/116 | 3/120 | | 13.67% | 0.02[-0.03,0.06] |
| Vorsanger 2008 | 30/128 | 25/129 | | 7.94% | 0.04[-0.06,0.14] |
| **Subtotal (95% CI)** | **2120** | **1373** | | **100%** | **0.07[0.04,0.11]** |
| Total events: 218 (Opioids), 67 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=35.94, df=8(P<0.0001); $I^2$=77.74% | | | | | |
| Test for overall effect: Z=3.8(P=0) | | | | | |
| **4.2.4 Headaches** | | | | | |
| Buynak 2010 | 63/318 | 44/319 | | 9.09% | 0.06[0,0.12] |
| Gordon 2010 | 9/73 | 3/68 | | 4.49% | 0.08[-0.01,0.17] |

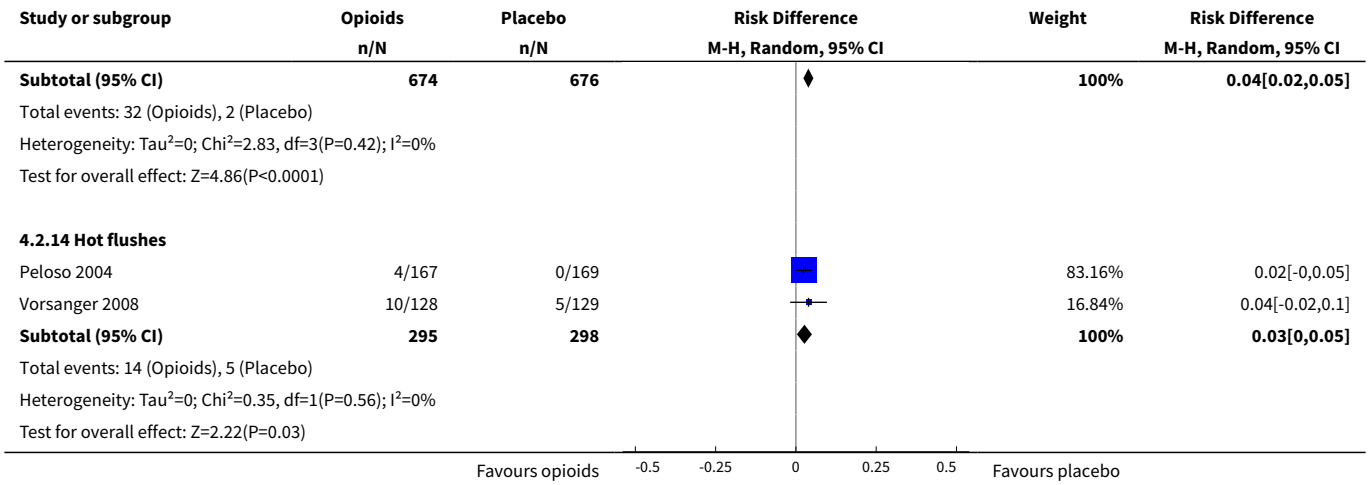


| Study or subgroup | Opioids n/N | Placebo n/N | Risk Difference M-H, Random, 95% CI | Weight | Risk Difference M-H, Random, 95% CI |
|---|---|---|---|---|---|
| Katz 2007 | 4/105 | 2/100 | | 12.49% | 0.02[-0.03,0.06] |
| Khoromi 2007 | 4/28 | 4/28 | | 1.2% | 0[-0.18,0.18] |
| Peloso 2004 | 11/167 | 7/169 | | 11.71% | 0.02[-0.02,0.07] |
| Ruoff 2003 | 14/161 | 6/157 | | 10.33% | 0.05[-0,0.1] |
| Schnitzer 2000 | 6/127 | 4/127 | | 11.82% | 0.02[-0.03,0.06] |
| Steiner 2011 | 100/1024 | 14/283 | | 18.77% | 0.05[0.02,0.08] |
| Uberall 2012 | 4/116 | 2/120 | | 14.5% | 0.02[-0.02,0.06] |
| Vorsanger 2008 | 10/128 | 21/129 | | 5.6% | -0.08[-0.16,-0.01] |
| **Subtotal (95% CI)** | **2247** | **1500** | | **100%** | **0.03[0.01,0.05]** |
| Total events: 225 (Opioids), 107 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=13.28, df=9(P=0.15); $I^2$=32.22% | | | | | |
| Test for overall effect: Z=2.67(P=0.01) | | | | | |
| **4.2.5 Dry mouth** | | | | | |
| Buynak 2010 | 26/318 | 7/319 | | 21.95% | 0.06[0.03,0.09] |
| Gordon 2010 | 13/73 | 0/68 | | 11.62% | 0.18[0.09,0.27] |
| Khoromi 2007 | 6/28 | 6/28 | | 3.26% | 0[-0.21,0.21] |
| Peloso 2004 | 11/167 | 0/169 | | 20.99% | 0.07[0.03,0.1] |
| Ruoff 2003 | 13/161 | 1/157 | | 20.01% | 0.07[0.03,0.12] |
| Uberall 2012 | 2/116 | 2/120 | | 22.17% | 0[-0.03,0.03] |
| **Subtotal (95% CI)** | **863** | **861** | | **100%** | **0.06[0.02,0.1]** |
| Total events: 71 (Opioids), 16 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=20.77, df=5(P=0); $I^2$=75.92% | | | | | |
| Test for overall effect: Z=2.96(P=0) | | | | | |
| **4.2.6 Dizziness** | | | | | |
| Buynak 2010 | 38/318 | 18/319 | | 14.32% | 0.06[0.02,0.11] |
| Gordon 2010 | 24/73 | 3/68 | | 5.37% | 0.28[0.17,0.4] |
| Katz 2007 | 5/105 | 3/100 | | 12.79% | 0.02[-0.04,0.07] |
| Khoromi 2007 | 4/28 | 1/28 | | 3.87% | 0.11[-0.04,0.25] |
| Peloso 2004 | 18/167 | 1/169 | | 13.51% | 0.1[0.05,0.15] |
| Ruoff 2003 | 12/161 | 2/157 | | 14.24% | 0.06[0.02,0.11] |
| Steiner 2011 | 102/1024 | 3/283 | | 17.84% | 0.09[0.07,0.11] |
| Uberall 2012 | 15/116 | 4/120 | | 10.28% | 0.1[0.03,0.16] |
| Vorsanger 2008 | 19/128 | 22/129 | | 7.78% | -0.02[-0.11,0.07] |
| **Subtotal (95% CI)** | **2120** | **1373** | | **100%** | **0.08[0.05,0.11]** |
| Total events: 237 (Opioids), 57 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=24.85, df=8(P=0); $I^2$=67.81% | | | | | |
| Test for overall effect: Z=4.7(P<0.0001) | | | | | |
| **4.2.7 Pruritis** | | | | | |
| Buynak 2010 | 23/318 | 6/319 | | 28.82% | 0.05[0.02,0.09] |
| Gordon 2010 | 17/73 | 14/68 | | 1.6% | 0.03[-0.11,0.16] |
| Katz 2007 | 3/105 | 1/100 | | 21.33% | 0.02[-0.02,0.06] |
| Ruoff 2003 | 11/161 | 2/157 | | 16.3% | 0.06[0.01,0.1] |
| Steiner 2011 | 87/1024 | 19/283 | | 26.07% | 0.02[-0.02,0.05] |
| Vorsanger 2008 | 14/128 | 10/129 | | 5.89% | 0.03[-0.04,0.1] |
| **Subtotal (95% CI)** | **1809** | **1056** | | **100%** | **0.04[0.02,0.05]** |
| Total events: 155 (Opioids), 52 (Placebo) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=3.94, df=5(P=0.56); $I^2$=0% | | | | | |
| Test for overall effect: Z=4.02(P<0.0001) | | | | | |




Favours opioids Favours placebo

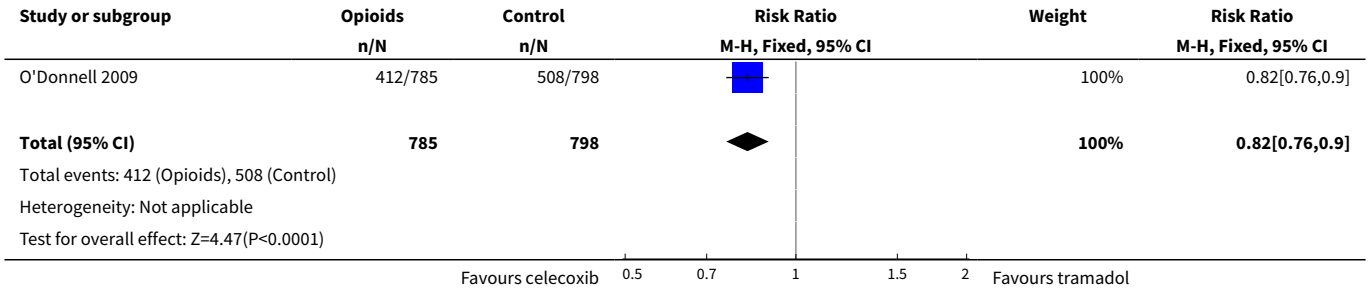





## Comparison 5. Tramadol compared to celecoxib

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 At least 30% of pain relief or moderate improvement | 1 | 1583 | Risk Ratio (M-H, Fixed, 95% CI) | 0.82 [0.76, 0.90] |

### Analysis 5.1. Comparison 5 Tramadol compared to celecoxib, Outcome 1 At least 30% of pain relief or moderate improvement.

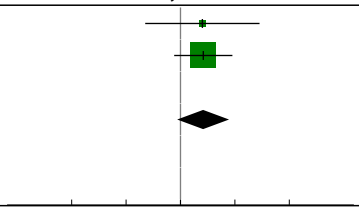



## Comparison 6. Opioids (all types) compared to antidepressants

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Pain (higher score means worse pain level) | 2 | 272 | Std. Mean Difference (IV, Fixed, 95% CI) | 0.21 [-0.03, 0.45] |
| 2 Disability (higher ratings mean greater disability) | 1 | 56 | Std. Mean Difference (IV, Fixed, 95% CI) | -0.11 [-0.63, 0.42] |

## Analysis 6.1. Comparison 6 Opioids (all types) compared to antidepressants, Outcome 1 Pain (higher score means worse pain level).

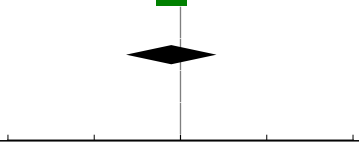
| Study or subgroup | Opioids | | Other analgesic | | Std. Mean Difference | Weight | Std. Mean Difference |
|---|---|---|---|---|---|---|---|
| | N | Mean(SD) | N | Mean(SD) | Fixed, 95% CI | | Fixed, 95% CI |
| Khoromi 2007 | 28 | 3.4 (2.5) | 28 | 2.9 (2.4) | | 20.59% | 0.2[-0.32,0.73] |
| Uberall 2012 | 107 | 3.9 (2) | 109 | 3.5 (1.8) | | 79.41% | 0.21[-0.06,0.48] |
| **Total ***** | **135** | | **137** | | | **100%** | **0.21[-0.03,0.45]** |
| Heterogeneity: $Tau^2$=0; $Chi^2$=0, df=1(P=0.98); $I^2$=0% | | | | | | | |
| Test for overall effect: Z=1.71(P=0.09) | | | | | | | |




Favours opioid — Favours other analge

## Analysis 6.2. Comparison 6 Opioids (all types) compared to antidepressants, Outcome 2 Disability (higher ratings mean greater disability).

| Study or subgroup | Opioids | | Other analgesics | | Std. Mean Difference | Weight | Std. Mean Difference |
|---|---|---|---|---|---|---|---|
| | N | Mean(SD) | N | Mean(SD) | Fixed, 95% CI | | Fixed, 95% CI |
| Khoromi 2007 | 28 | 25.7 (16.5) | 28 | 27.5 (16.7) | | 100% | -0.11[-0.63,0.42] |
| **Total ***** | **28** | | **28** | | | **100%** | **-0.11[-0.63,0.42]** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Z=0.4(P=0.69) | | | | | | | |




Favours opioid — Favours other analgesic

## ADDITIONAL TABLES

**Table 1. GRADE Table: Tramadol compared to placebo**

| Quality assessment | | | | | | | No of participants | | Effect | | Quality |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **No of studies** | **Design** | **Risk of bias** | **Inconsistency** | **Indirectness** | **Imprecision** | **Other considerations** | **Weak opioids compared to placebo** | **Control** | **Relative (95% CI)** | **Absolute** | |
| **Pain intensity (higher score means worse pain levels)** | | | | | | | | | | | |
| 5[1] | RCT | serious[2] | serious[3] | no serious indirectness | serious[4] | none[5] | 689 | 689 | - | SMD 0.55 lower (0.66 to 0.44 lower) | ⊕⊕○○ LOW |
| **Disability (higher ratings mean greater disability)** | | | | | | | | | | | |
| 5[1] | RCT | serious[2] | no serious inconsistency[6] | no serious indirectness | no serious imprecision[7] | none[5] | 676 | 672 | - | SMD 0.18 lower (0.29 to 0.07 lower) | ⊕⊕⊕○ MODERATE |

[1] Peloso 2004, Ruoff 2003, Schnitzer 2000, Uberall 2012 and Vorsanger 2008.

[2] High risk of attrition bias. Difficult to judge other biases because the information was insufficiently reported. It is likely that there was also performance and detection bias due to difficulty in blinding.

[3] Heterogeneity: $Chi^2 = 28.34$, $df = 4$ ($P < 0.0001$); $I^2 = 86\%$.

[4] SMD (effect size crosses 0.5. SMD is -0.55 [-0.66, -0.44].

[5] See Figure 3 (for pain) and Figure 4 (for function).

[6] Heterogeneity: $Chi^2 = 1.24$, $df = 4$ ($P = 0.87$); $I^2 = 0\%$.

[7] SMD (effect size) is -0.18 [-0.29, -0.07].

**Table 2. GRADE Table: Buprenorphine compared to placebo**

| Quality assessment | | | | | | | No of participants | | Effect | | Quality |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **No of studies** | **Design** | **Risk of bias** | **Inconsistency** | **Indirectness** | **Imprecision** | **Other considerations** | **Buprenorphine compared to placebo** | **Control** | **Relative (95% CI)** | **Absolute** | |
| **Mean pain intensity (higher score means worse pain levels)** | | | | | | | | | | | |

**Table 2. GRADE Table: Buprenorphine compared to placebo** *(Continued)*

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | RCT | very serious[1] | serious[2] | no serious indirectness | no serious imprecision[3] | none | 312 | 341 | - | MD 0.58 lower (0.61 to 0.55 lower) | ⊕◯◯◯ VERY LOW |
| **Disability (higher ratings mean greater disability)** | | | | | | | | | | | |
| 1 | RCT | serious[4] | no serious inconsistency | no serious indirectness | very serious[5] | none | 51 | 50 | - | MD 3 lower (11.44 lower to 5.44 higher) | ⊕◯◯◯ VERY LOW |

[1] One trial had a high risk of selection, performance and detection bias. The two trials included in this comparison had high risk of attrition bias.
[2] Heterogeneity: $Chi^2 = 196.24$, df = 1 ($P < 0.00001$); $I^2 = 99\%$.
[3] SMD -2.47 [-2.69, -2.25].
[4] The trial included in this comparison had a high risk of attrition bias.
[5] There is only one study. The SMD is -0.14 [-0.53, 0.25].

**Table 3. GRADE Table: Strong opioids compared to placebo**

| Quality assessment | | | | | | | No of participants | | Effect | | Quality |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **No of studies** | **Design** | **Risk of bias** | **Inconsistency** | **Indirectness** | **Imprecision** | **Other considerations** | **Strong opioids compared to placebo** | **Control** | **Relative (95% CI)** | **Absolute** | |
| **Mean pain intensity (higher score means worse pain levels)** | | | | | | | | | | | |
| 6 | RCT | serious[1] | no serious inconsistency[2] | no serious indirectness | no serious imprecision | none[3] | 1154 | 733 | - | SMD 0.43 lower (0.52 to 0.33 lower) | ⊕⊕⊕◯ MODERATE |
| **Disability (higher ratings mean greater disability)** | | | | | | | | | | | |
| 4 | RCT | serious[4] | no serious inconsistency[2] | no serious indirectness | no serious imprecision | none[5] | 845 | 530 | - | SMD 0.26 lower (0.37 to 0.15 lower) | ⊕⊕⊕◯ MODERATE |
| **At least 30% of pain relief or moderate relief** | | | | | | | | | | | |

**Table 3. GRADE Table: Strong opioids compared to placebo** *(Continued)*

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | RCT | serious[6] | no serious inconsistency | no serious indirectness | no serious imprecision[7] | none | 204/418 (48.8%) | 131/401 (32.7%) | OR 1.91 (1.41 to 2.58) | 154 more per 1000 (from 80 more to 229 more)<br>150 more per 1000 (from 76 more to 225 more) | ⊕⊕⊕○ MODERATE |
| | | | | | | | | 29.7% | | | |
| **At least 50% of pain relief** | | | | | | | | | | | |
| 2 | RCT | serious[8] | serious[9] | no serious indirectness | serious[10] | none | 146/386 (37.8%) | 86/364 (23.6%) | OR 1.89 (1.34 to 2.66) | 133 more per 1000 (from 57 more to 215 more)<br>156 more per 1000 (from 70 more to 240 more) | ⊕○○○ VERY LOW |
| | | | | | | | | 37.1% | | | |
| **Side effects - Somnolence** | | | | | | | | | | | |
| 5 | RCT | very serious[11] | no serious inconsistency | no serious indirectness | serious[10] | none | 135/1548 (8.7%) | 20/798 (2.5%) | See comment | 61 more per 1000 (from 20 more to 100 more)<br>61 more per 1000 (from 20 more to 100 more) | ⊕○○○ VERY LOW |
| | | | | | | | | 2.5% | | | |
| **Side effects - Nausea** | | | | | | | | | | | |
| 5 | RCT | very serious[11] | no serious inconsistency | no serious indirectness | no serious imprecision | none | 357/1548 (23.1%) | 81/798 (10.2%) | See comment | 122 more per 1000 (from 50 more to 190 more)<br>109 more per 1000 (from 45 more to 170 more) | ⊕⊕○○ LOW |
| | | | | | | | | 9.1% | | | |
| **Side effects - Constipation** | | | | | | | | | | | |
| 5 | RCT | very serious[11] | serious[12] | no serious indirectness | serious[10] | none | 148/1548 (9.6%) | 29/798 (3.6%) | See comment | 112 more per 1000 (from 40 more to 190 more)<br>154 more per 1000 (from 55 more to 262 more) | ⊕○○○ VERY LOW |
| | | | | | | | | 5% | | | |

[1] Four trials had low risk of selection bias and one trial (Chu 2012) was unclear. All five trials had low risk of performance bias, and low risk of reporting bias. However, all five trials suffered from high risk of attrition bias, and some trials also had high risk of detection bias becasue it was unclear if the outcome assessor were blinded.

[2] $I^2 = 0\%$

[3] See Figure 1.

[4] Selection bias: three trials low risk of bias, and one trial unclear (Chu 2012). All four trials had low risk of performance bias. Detection bias was unclear in three trials, except Khoromi 2007. Attrition bias was judged high in all four trials. Reporting bias was not a problem in any trial.
[5] See Figure 2.
[6] All trials had risk of attrition bias and performance bias. One trial was unclear about randomization method.
[7] Total number of events was 335.
[8] Both trials had high risk of attritiion bias. Both trials were unclear about performance bias. One trial was not clear about method of randomization.
[9] Heterogeneity: $Chi^2 = 5.39$, df = 1, P = 0.02; $I^2 = 81\%$.
[10] Total number of events < 300
[11] All trials had high risk of attrition bias. Most trials had a problem with performance bias, and one trial was not clear about method of randomization.
[12] Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 33.50$, df = 4, P < 0.00001; $I^2 = 88\%$.

**Table 4. GRADE Table: Tramadol compared to celecoxib**

| Quality assessment | | | | | | | No of participants | | Effect | | Quality |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **No of studies** | **Design** | **Risk of bias** | **Inconsistency** | **Indirectness** | **Imprecision** | **Other considerations** | **Opioids (all types) compared to NSAIDs** | **Control** | **Relative (95% CI)** | **Absolute** | |
| **Pain (higher score means worse pain level) (Better indicated by lower values)** | | | | | | | | | | | |
| 0 | No evidence available | | | | | none | 0 | - | - | not pooled | |
| **Disability (higher ratings mean greater disability) (Better indicated by lower values)** | | | | | | | | | | | |
| 0 | No evidence available | | | | | none | 0 | - | - | not pooled | |
| **At least 30% of pain relief or moderate improvement** | | | | | | | | | | | |
| 1 | RCT | serious[1] | no serious inconsistency | serious[2] | serious[3] | none | 412/785 (52.5%) | 508/798 (63.7%) | OR 0.63 (0.52 to 0.77) | 112 fewer per 1000 (from 62 fewer to 160 fewer) | ⊕○○○ VERY LOW |
| | | | | | | | | 63.7% | | 112 fewer per 1000 (from 62 fewer to 160 fewer) | |

[1] One study was included. There was uncertainty about allocation conceallment (selection bias), and blinding (performance and measurement bias). There was problem with drop-outs (attrition bias).
[2] Indirectness in the outcome measure. This trial used "at least 30% pain relief OR moderate improvement". There is not report of mean pain scores.
[3] Imprecision because there is only one study in this category.

**Table 5. GRADE Table: Opioids compared to antidepressants**

| Quality assessment | | | | | | | No of participant | | Effect | | Quality |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **No of studies** | **Design** | **Risk of bias** | **Inconsistency** | **Indirectness** | **Imprecision** | **Other considerations** | **Opioids (all types) compared to antidepressants** | **Control** | **Relative (95% CI)** | **Absolute** | |
| **Pain (higher score means worse pain level) (Better indicated by lower values)** | | | | | | | | | | | |
| 2 | RCTs | serious[1] | no serious inconsistency[2] | no serious indirectness | very serious[3] | none | 135 | 137 | - | SMD 0.21 higher (0.03 lower to 0.45 higher) | ⊕○○○ VERY LOW |
| **Disability (higher ratings mean greater disability)** | | | | | | | | | | | |
| 1 | RCT | serious[4] | no serious inconsistency | no serious indirectness | very serious[5] | none | 28 | 28 | - | SMD 0.11 lower (0.63 lower to 0.42 higher) | ⊕○○○ VERY LOW |

[1] Two studies were included. All studies had problems with drop-outs (attrition bias), One study did not have a clear description of concealment of allocation (selection bias). Both studies might have some issues with blinding of provider and outcome assessor (performance and measurement bias).

[2] Heterogeneity: $Chi^2 = 0.00$, $df = 1$ ($P = 0.98$); $I^2 = 0\%$.

[3] Two studies. Total population is less than 400. The 95% CI includes the no effect.

[4] Only one study was included in this comparison. This study had issues with drop-outs (attrition bias), and potentially blinding of providers (performance bias).

[5] There was only one study, and the 95% CI included the no effect.

# APPENDICES

## Appendix 1. MEDLINE search strategy

MEDLINE (OVID)

1 randomized controlled trial.pt.
2 controlled clinical trial.pt.
3 Randomized Controlled Trials/
4 Random Allocation/
5 Double-Blind Method/
6 Single-Blind Method/
7 or/1-6
8 Animal/ not Human/
9 7 not 8
10 clinical trial.pt.
11 explode Clinical Trials/
12 (clinical$ adj 25 trial$).tw.
13 ((singl$ or doubl$ or trebl$ or tripl$) adj(mask$ or blind$)).tw.
14 Placebos/
15 placebo$.tw.
16 random$.tw.
17 Research Design/
18 (latin adj square).tw.
19 or/10-18
20 19 not 8
21 20 not 9
22 Comparative Study/
23 explode Evaluation Studies/
24 Follow-Up Studies/
25 Prospective Studies/
26 (control$ or prospective$ or volunteer$).tw.
27 Cross-Over Studies/
28 or/22-27
29 28 not 8
30 29 not (9 or 21)
31 9 or 21 or 30
32 PAIN/pc, dt, rh, th [Prevention & Control, Drug Therapy, Rehabilitation, Therapy]
33 Chronic Disease/dt, pc, rh, th [Drug Therapy, Prevention & Control, Rehabilitation, Therapy]
34 (chronic adj3 pain).mp
35 Low Back Pain/
36 (low adj back adj pain).mp
37 or/ 32-36
38 exp Analgesics, opioid/
39 codeine.mp.
40 fentanyl.mp.
41 hydrocodone.mp.
42 hydromorphone.mp.
43 levorphanol.mp.
44 meperidine.mp.
45 morphine.mp.
46 oxycodone.mp.
47 oxymorphone.mp.
48 pentazocine.mp.
49 propoxyphene.mp.
50 sufentanil.mp.
51 tramadol.mp.
52 or/ 38-51
53 31 and 37 and 52

## Appendix 2. Other search strategies

EMBASE (OVID)

1 exp Clinical Trial/
2 exp randomization/
3 Double Blind Procedure/
4 Single Blind Procedure/
5 or/1-4
6 exp animal/
7 Nonhuman/
8 6 or 7
9 exp human/
10 8 not 9
11 5 not 10
12 (clinical$ adj25 trial$).tw.
13 ((singl$ or doubl$ or trebl$ or tripl$) adj (mask$ or blind$)).tw.
14 exp Placebo/
15 placebo$.tw.
16 random$.tw.
17 methodology/ or latin square design/
18 (latin adj square).tw.
19 or/12-18
20 19 not 10
21 20 not 11
22 comparative study/
23 evaluation/
24 Follow Up/
25 Prospective Study/
26 (control$ or prospective$ or volunteer$).tw.
27 Crossover Procedure/
28 or/22-27
29 28 not 10
30 29 not (11 or 21)
31 30 or 21 or 11
32 exp Chronic Pain/
33 exp PAIN/pc, rh, dt, th [Prevention, Rehabilitation, Drug Therapy, Therapy]
34 exp Chronic Disease/pc, rh, dt, th [Prevention, Rehabilitation, Drug Therapy, Therapy]
35 33 and 34
36 32 or 35
37 (chronic adj3 pain$).tw.
38 exp Low Back Pain/
39 (low adj back adj pain$).tw.
40 or/36-39
41 exp Narcotic Analgesic Agent/
42 codeine.mp.
43 fentanyl.mp.
44 hydrocodone.mp.
45 hydromorphone.mp.
46 levorphanol.mp.
47 meperidine.mp.
48 morphine.mp.
49 oxycodone.mp.
50 oxymorphone.mp.
51 pentazocine.mp.
52 propoxyphene.mp.
53 sufentanil.mp.
54 tramadol.mp.
55 or/41-54
56 31 and 40 and 55

CINAHL (Ebsco)

S69 S53 and S68
S68 S54 or S55 or S56 or S57 or S58 or S59 or S60 or S61 or S62 or S63 or S64 or S65 or S66 or S67
S67 (MH "Tramadol") OR "tramadol"
S66 (MH "Sufentanil") OR "sufentanil"
S65 (MH "Propoxyphene") OR "propoxyphene"
S64 (MH "Pentazocine") OR "pentazocine"
S63 "oxymorphone"
S62 (MH "Oxycodone") OR "oxycodone"
S61 (MH "Morphine+") OR "morphine"
S60 (MH "Meperidine") OR "meperidine"
S59 "levorphanol"
S58 "hydromorphone"
S57 "hydrocodone"
S56 (MH "Fentanyl+") OR "fentanyl"
S55 (MH "Codeine+") OR "codeine"
S54 (MH "Analgesics, Opioid+")
S53 S28 and S52
S52 S48 or S51
S51 S49 or S50
S50 (MM "Chronic Disease/DT/PC/RH/TH")
S49 (MM "Pain/PC/DT/RH/TH")
S48 S35 or S43 or S47
S47 S44 or S45 or S46
S46 "lumbago"
S45 (MH "Spondylolisthesis") OR (MH "Spondylolysis")
S44 (MH "Thoracic Vertebrae")
S43 S36 or S37 or S38 or S39 or S40 or S41 or S42
S42 lumbar N2 vertebra
S41 (MH "Lumbar Vertebrae")
S40 "coccydynia"
S39 "coccyx"
S38 "sciatica"
S37 (MH "Sciatica")
S36 (MH "Coccyx")
S35 S29 or S30 or S31 or S32 or S33 or S34
S34 lumbar N5 pain
S33 lumbar W1 pain
S32 "backache"
S31 (MH "Low Back Pain")
S30 (MH "Back Pain+")
S29 "dorsalgia"
S28 S26 NOT S27
S27 (MH "Animals")
S26 S7 or S12 or S19 or S25
S25 S20 or S21 or S22 or S23 or S24
S24 volunteer*
S23 prospectiv*
S22 control*
S21 followup stud*
S20 follow-up stud*
S19 S13 or S14 or S15 or S16 or S17 or S18
S18 (MH "Prospective Studies+")
S17 (MH "Evaluation Research+")
S16 (MH "Comparative Studies")
S15 latin square
S14 (MH "Study Design+")
S13 (MH "Random Sample")
S12 S8 or S9 or S10 or S11
S11 random*
S10 placebo*
S9 (MH "Placebos")
S8 (MH "Placebo Effect")

S7 S1 or S2 or S3 or S4 or S5 or S6
S6 triple-blind
S5 single-blind
S4 double-blind
S3 clinical W3 trial
S2 "randomi?ed controlled trial*"
S1 (MH "Clinical Trials+")

CENTRAL (Wiley)

#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
#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 MeSH descriptor: [Analgesics, Opioid] explode all trees
#30 opiate
#31 opioid
#32 codeine
#33 fentanyl
#34 hydrocodone
#35 hydromorphone
#36 levorphanol
#37 meperidine
#38 morphine
#39 oxycodone
#40 oxymorphone
#41 pentazocine
#42 propoxyphene
#43 tramadol
#44 tapentadol
#45 buprenorphine
#46 #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
#47 #28 and #46 in Trials

PsycInfo (OVID)

1 clinical trials/
2 controlled trial.mp.

3 RCT.mp.
4 (Random* adj3 trial).mp.
5 (clin* adj3 trial).mp
6 (sing* adj2 blind*).mp.
7 (doub* adj2 blind*).mp.
8 placebo.mp. or exp Placebo/
9 latin square.mp.
10 (random* adj2 assign*).mp.
11 prospective studies/
12 (prospective adj stud*).mp.
13 (comparative adj stud*).mp.
14 treatment effectiveness evaluation/
15 treatment effectiveness evaluation/
16 (evaluation adj stud*).mp.
17 exp Posttreatment Followup/
18 follow?up stud*.mp.
19 or/1-18
20 back pain/
21 lumbar spinal cord/
22 (low adj back adj pain).mp.
23 (back adj pain).mp.
24 spinal column/
25 (lumbar adj2 vertebra*).mp.
26 coccyx.mp.
27 sciatica.mp.
28 lumbago.mp.
29 dorsalgia.mp.
30 back disorder*.mp.
31 "back (anatomy)"/
32 ((disc or disk) adj degenerat*).mp.
33 ((disc or disk) adj herniat*).mp.
34 ((disc or disk) adj prolapse*).mp.
35 (failed adj back).mp.
36 or/20-35
37 exp opiates/
38 exp analgesic drugs/
39 codeine.mp. or exp Codeine/
40 fentanyl.mp. or exp Fentanyl/
41 hydrocodone.mp.
42 hydromorphone.mp.
43 levorphanol.mp.
44 exp Meperidine/ or meperidine.mp.
45 morphine.mp. or exp Morphine/
46 oxycodone.mp.
47 oxymorphone.mp.
48 pentazocine.mp. or exp Pentazocine/
49 propoxyphene.mp.
50 tramadol.mp. or exp Tramadol/
51 tapentadol.mp.
52 buprenorphine.mp.
53 or/37-52
54 36 and 53
55 19 and 54

## Appendix 3. Criteria for risk of bias assessment for RCTs

### 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 or 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 bias 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 one of the following, or an equivalent method, was used to conceal allocation: central allocation (including telephone, web-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 (for example, a list of random numbers); assignment envelopes were used without appropriate safeguards (for example, 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 blinding of participants was ensured 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 not likely 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 study***

There is a low risk of performance bias if blinding of personnel was ensured 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 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 the blinding of the outcome assessment was ensured 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 not likely to be influenced by lack of blinding, or:

- for patient-reported outcomes in which the patient was the outcome assessor (for example, 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 (for example, 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 not available 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 (for example, 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 (for example surgery), this item is irrelevant (van Tulder 2003).

**ITT analysis**

There is low risk of bias if all randomized patients were reported or 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 (for example, study funding).

## Appendix 4. Questions for clinical relevance assessment

1. Are the patients described in detail so that you can decide whether they are comparable to those that you see in your practice?
2. Are the interventions and treatment settings described well enough so that you can provide the same for your patients?
3. Were all clinically relevant outcomes measured and reported?
4. Is the size of the effect clinically important?
5. Are the likely treatment benefits worth the potential harms?

## WHAT'S NEW

| Date | Event | Description |
|---|---|---|
| 27 May 2014 | Amended | Correction made to description of risk of bias assessment. |

## HISTORY

Protocol first published: Issue 4, 2004
Review first published: Issue 3, 2007

| Date | Event | Description |
|---|---|---|
| 12 June 2013 | New search has been performed | We updated the inclusion criteria for study design. In 2007, we included randomized (RCT) and quasi-RCT trials. For the 2013 update, we restricted the inclusion criteria to RCTs that were double-blinded. We included 12 new trials in the update. We excluded one trial of the original review (not blinded and quasi-RCT). |
| 12 June 2013 | New citation required but conclusions have not changed | In the original review we concluded that tramadol improved pain and function. In this review update we found that all the evaluated opioids were effective for pain relief. We also found that all opioids, except buprenorphine, were effective for improvement of function. |
| 27 May 2008 | Amended | We converted to the new review format. |
| 18 May 2007 | New search has been performed | This review only includes studies identified up until March 2006. The literature search was updated in May 2007 and studies are currently awaiting assessment. |

## CONTRIBUTIONS OF AUTHORS

S Atlas contributed to study selection, reviewed and edited the protocol and review.

LE Chaparro contributed to study selection, risk of bias assessment, data extraction, data analysis and drafting of the review.

A Deshpande contributed to study selection, risk of bias assessment, data extraction, data analysis and drafted both the protocol and review.

A Furlan - contributed to study selection, risk of bias assessment, data extraction, data analysis, assisted with writing and editing the protocol and review.

A Mailis-Gagnon reviewed and edited the protocol and review.

D Turk reviewed and edited the protocol and review.

## DECLARATIONS OF INTEREST

None

## SOURCES OF SUPPORT

### Internal sources

- Canadian Institutes for Health Research, Canada.

  Andrea Furlan received a New Investigator Award from CiHR (2012-2017)

### External sources

- No sources of support supplied

## INDEX TERMS

### Medical Subject Headings (MeSH)

Analgesics, Opioid [adverse effects] [*therapeutic use]; Anti-Inflammatory Agents, Non-Steroidal [therapeutic use]; Chronic Pain [*drug therapy]; Low Back Pain [*drug therapy]; Randomized Controlled Trials as Topic

### MeSH check words

Adult; Female; Humans; Male; Middle Aged