

A systematic review on the effectiveness of pharmacological interventions for chronic non-specific low-back pain

T. Kuijpers · M. van Middelkoop · S. M. Rubinstein ·
R. Ostelo · A. Verhagen · B. W. Koes ·
M. W. van Tulder

Received: 13 November 2009/Revised: 11 May 2010/Accepted: 19 July 2010/Published online: 31 July 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The objective of this review was to determine the effectiveness of pharmacological interventions [i.e., non-steroid anti-inflammatory drugs (NSAIDs), muscle relaxants, antidepressants, and opioids] for non-specific chronic low-back pain (LBP). Existing Cochrane reviews for the four interventions were screened for studies fulfilling the inclusion criteria. Then, the literature searches were updated. Only randomized controlled trials on adults (≥ 18 years) with chronic (≥ 12 weeks) non-specific LBP and evaluation of at least one of the main clinically relevant outcome measures (pain, functional status, perceived recovery, or return to work) were included. The GRADE approach was used to determine the quality of evidence. A total of 17 randomized controlled trials was included:

NSAIDs ($n = 4$), antidepressants ($n = 5$), and opioids ($n = 8$). No studies were found for muscle relaxants; 14 studies had a low risk of bias. The studies only reported effects on the short term (< 3 months). The overall quality of the evidence was low. NSAIDs and opioids seem to lead to a somewhat higher relief in pain on the short term, as compared to placebo, in patients with non-specific chronic low back pain; opioids seem to have a small effect in improving function for a selection of patients who responded with an exacerbation of their symptoms after stopping their medication. However, both types of medication show more adverse effects than placebo. There seems to be no difference in effect between antidepressants and placebo in patients with non-specific chronic LBP.

Electronic supplementary material The online version of this article (doi:10.1007/s00586-010-1541-4) contains supplementary material, which is available to authorized users.

Keywords Systematic review · Medication · Pharmacological interventions · Low-back pain · Interventions

T. Kuijpers (✉)
Dutch Institute for Health Care Improvement CBO,
Churchillaan 11, P.O. Box 20064, 3502 LB Utrecht,
The Netherlands
e-mail: t.kuijpers@cbo.nl

M. van Middelkoop · A. Verhagen · B. W. Koes
Department of General Practice, Erasmus MC,
University Medical Center Rotterdam,
Rotterdam, The Netherlands

S. M. Rubinstein · R. Ostelo · M. W. van Tulder
EMGO-Institute for Health and Care Research,
VU University Medical Center, Amsterdam, The Netherlands

M. W. van Tulder
Department of Health Economics and Health Technology
Assessment, Institute of Health Sciences,
Faculty of Earth and Life Sciences, VU University,
Amsterdam, The Netherlands

Introduction

Low-back pain (LBP) is related to disability and work absence and accounts for high economical costs in Western societies [1]. The management of LBP comprises a range of different intervention strategies including surgery, non-medical interventions such as exercise, behavioral therapy, and alternative therapies. Pharmacological interventions are the most frequently recommended intervention for back pain [2, 3]. Over the last years, a substantial number of randomized clinical trials and systematic reviews have been published. Based on this literature, this overview presents the current evidence on pharmacological interventions for non-specific chronic low back pain.

Objectives

The objective of this review was to determine the effectiveness of pharmacological interventions [i.e., non-steroid anti-inflammatory drugs (NSAIDs), muscle relaxants, antidepressants, and opioids] for non-specific chronic LBP.

Criteria for considering studies for this review

Types of studies

Only randomized controlled trials with at least 1 day of follow-up were considered in this systematic review.

Types of participants

In order to be included in this review, participants of the RCTs must fulfill the following inclusion criteria: adult subjects (≥ 18 years of age) with chronic (>12 weeks) non-specific LBP (including discopathy or any other non-specific degenerative pathology such as osteoarthritis).

The exclusion criteria were: (1) trials including subjects with specific LBP caused by pathologies such as vertebral spinal stenosis, ankylosing spondylitis, scoliosis, and coccydynia the diagnosis of which had to be confirmed by an MRI or another diagnostic imaging tool; (2) post-partum LBP or pelvic pain due to pregnancy; (3) post-operative studies;¹ (4) prevention studies; and (5) abstracts or non-published studies.

Types of interventions

The RCTs studying the following interventions were included in this overview: NSAIDs, muscle relaxants, antidepressants, and opioids. All types of NSAIDs (including COX-2), antidepressants (i.e., tricyclic and heterocyclic antidepressants, selective serotonin reuptake inhibitors, mono-amine oxidase inhibitors and ‘atypical’ antidepressants), muscle relaxants and opioids [given by oral, transdermal, mucosal (nasal or rectal), or intramuscular routes] were included.

Additional interventions were allowed in all studies if there was a contrast for the pharmacological intervention in the study.

Types of outcome measures

For inclusion, at least one of the following outcome measures should have been measured in the RCT: pain

intensity [e.g., visual analog scale (VAS), numerical rating scale (NRS), McGill pain questionnaire], back-specific functional status (e.g., Roland–Morris Disability Questionnaire, Oswestry Disability Index), perceived recovery (e.g., overall improvement), and return to work (e.g., return to work status, sick leave days). The primary outcomes for this review were pain and functional status.

Search methods for identification of studies

Existing Cochrane reviews for the four interventions were screened for studies fulfilling the inclusion criteria [4–7]. Then, the literature searches were updated from the last date onward for each of the interventions.

The search was conducted in MEDLINE, EMBASE, CINAHL, CENTRAL, and PEDro up to December 22, 2008. References of relevant studies were screened, and experts were approached in order to identify additional primary studies not identified in the previous steps. The language was limited to English, Dutch, and German, because these are the languages that the authors are able to read and understand. The search strategy outlined by the Cochrane Back Review Group (CBRG) was followed and is available upon request from the primary author [8].

Methods of the review

Study selection

Three authors (TK, SMR, and MM) independently screened the abstracts and titles retrieved by the search strategy and applied the inclusion criteria to all these abstracts. The first author screened all abstracts (TK), the others both half of the references. Full text of the article was obtained if the title and the abstract seemed to fulfill the inclusion criteria or if eligibility of the study was unclear. All full text articles from the existing Cochrane reviews were compiled and screened on inclusion criteria by the authors, independently. Any disagreements between the authors were resolved by discussion and consensus.

Risk-of-bias assessment

Two reviewers (TK and SMR) assessed the risk of bias (RoB) independently, using the criteria list advised by the CBRG, which consists of 11 items [8]. Items were scored as ‘positive’ if they fulfilled the criteria, as ‘negative’ when there was a clear RoB, and as ‘inconclusive’ if there was insufficient information. Differences in assessment were discussed during a consensus meeting. A total score was computed by adding the number of positive scores, and

¹ For opioids only studies specifically studying post-operative patients were excluded.

high quality was defined as fulfilling 6 or more (more than 50%) of the 11 internal validity criteria [9].

Data extraction

A standardized form was used for data extraction consisting of both descriptive data on the study population, the type of intervention examined, and quantitative data regarding the outcome measures. Data on the characteristics of the study population (gender, age), type and dose of medication, control treatment, and study results were also described.

Data analysis

If studies were clinically homogeneous regarding study population, types of treatment, reference treatment, outcomes, and measurement instruments, a meta-analysis was performed. If possible, we calculated the weighted mean difference (WMD) because this improves the interpretability of the results. If a WMD was not possible the standardized mean difference (SMD) was calculated. If trials reported outcomes as graphs, the mean scores and standard deviations (SDs) were estimated from these graphs. If SDs were not reported, they were calculated using the reported values of the confidence intervals, if possible. If the SD of the baseline score was reported, we used the ratio between the baseline score and SD to calculate the SD for other follow-up moments. Finally, if none of these data were reported, an estimation of the SD was based on study data (population and score) of other studies. In order to correct for error introduced by “double-counting” of subjects of “shared” interventions (i.e., two comparisons within one study that used the same control group as contrast) in the meta-analyses, the number of subjects in similar contrasts was divided by the number of comparisons that this one study added in the meta-analyses. For the comparisons where studies were clinically too heterogeneous, no meta-analysis was performed. We chose for random effect model when inspection of the forest plots showed heterogeneity represented by different directions of the effects or if $I^2 > 20\%$ (arbitrary cutoff point). When no new studies were added to the results of the original reviews we followed the presentation of the original results. Heterogeneity was tested with Chi-square and I^2 . For the statistical analyses the software package ‘Review Manager 5’ was used.

Quality of the evidence

Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) was used to evaluate the overall quality of the evidence and the strength of the

recommendations [10]. Quality of the evidence for a specific outcome was based upon five principal factors: (1) limitations (e.g., due to study design), (2) inconsistency of results, (3) indirectness (e.g., generalizability of the findings), (4) imprecision (e.g., sufficient data), and (5) other considerations, such as reporting bias. The overall quality was considered to be high when multiple RCTs with a low RoB provide consistent, generalizable, precise data for a particular outcome. The quality of the evidence was downgraded by one level when one of the factors described above was not met [10]. Single studies were considered imprecise (i.e., sparse data) and provide “low quality evidence”, which could be further downgraded to “very low quality evidence” if there were also limitations in design or indirectness. The following grades of quality of the evidence were applied:

<i>High quality</i>	Further research is very unlikely to change our confidence in the estimate of effect.
<i>Moderate quality</i>	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
<i>Low quality</i>	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.
<i>Very low quality</i>	We are very uncertain about the estimate.

To improve the readability of this review, a GRADE table was only completed when we completed a meta-analysis. If only one study was present for a given comparison, the results are described in the text and in the table with characteristics of included studies.

Adverse events were reported using relative risk (RR) for the total number of any adverse event.

Results

Description of studies

Study selection

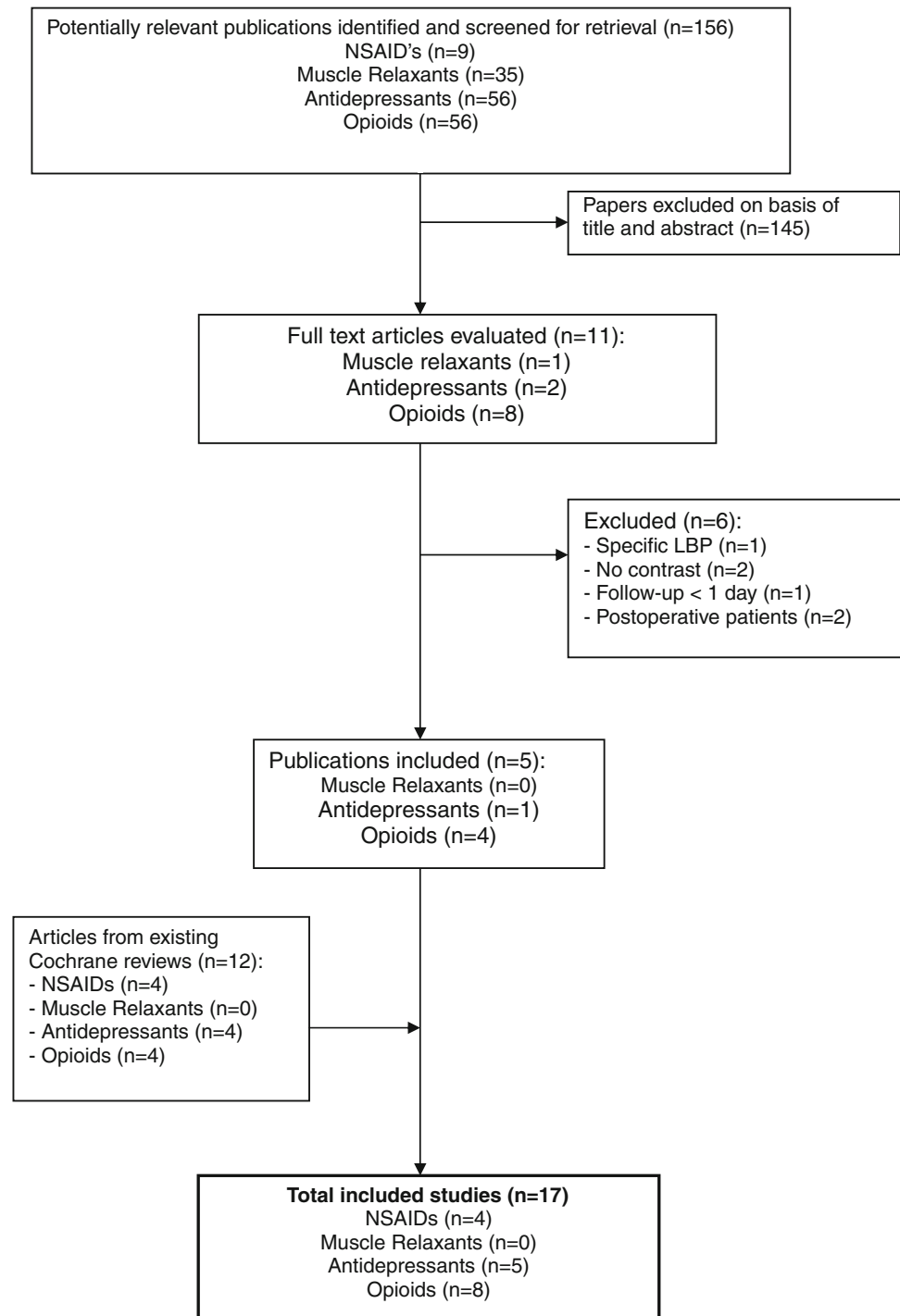
A total of 17 studies was included in this review (Fig. 1). From the four existing Cochrane reviews a total of 109 references was screened for eligibility. Of these 109 articles, 12 studies fulfilled the inclusion criteria and were included in this overview (Fig. 1). The most important reason for exclusion was inclusion of acute patients in the study. Additionally, 156 potentially new relevant titles and abstracts were identified for the four pharmacological

interventions and screened for potential inclusion. Of these abstracts, 11 full text articles were evaluated. A total of five studies fulfilled the inclusion criteria and were therefore included. The study characteristics of all included and excluded studies are summarized in Appendix 1 in Electronic Supplementary Material.

NSAIDs Four studies compared one or more types of NSAIDs with a placebo [11–14]. NSAIDs that were used in

the studies were naproxen, etoricoxib, rofecoxib, and valdecoxib. In three [12–14] of the four studies selective COX-2 NSAIDs were studied. In three [12–14] of the four studies assessing NSAIDs for chronic LBP a so called “flare design” was used, in which patients who were already responding well to NSAIDs were only included when they showed a large worsening in LBP complaints during a wash-out period.

Fig. 1 Flow diagram of inclusion and exclusion of articles for pharmacological interventions for chronic low back pain



Antidepressants Five studies were identified that compared antidepressant medication with placebo [15–19]. Two studies by Atkinson [16, 19], each included several comparisons, including different types of antidepressants with placebo and two antidepressants with each other. In the current review, we included two placebo comparisons from Atkinson [16] of maprotiline, and paroxetine, which we respectively refer to in our analysis as Atkinson [16]. We included three placebo comparisons from Atkinson [19] of desipramine <60 ng/mL (analysis: Atkinson [19]), desipramine >60 ng/mL (analysis: Atkinson [19]) and fluoxetine (analysis: Atkinson [19]).

All studies included patients because they had back pain, but only one study explicitly selected low-back patients with significant depressive symptoms [17]. Three studies excluded patients with major depression [15, 16, 19], and in one study it was unclear if patients were depressed or not [18].

Several types of antidepressants were investigated: tricyclic antidepressants (TCAs), serotonin selective reuptake inhibitors, and ‘atypical’ antidepressants (e.g., aminoketone antidepressants) [20, 21]. Most studies evaluated the effectiveness of TCAs with noradrenergic (maprotiline, nortriptyline, desipramine) [15, 16, 19]. Three studies evaluated a selective serotonin reuptake inhibitor (paroxetine, fluoxetine) [16, 17, 19] and one study investigated ‘atypical’ antidepressants, bupropion (aminoketone antidepressant) [18]. While two trials used an active placebo (i.e., preparation containing substances that mimic the adverse effects of antidepressants) [16, 19], the other trials used an inert placebo [15, 17, 18]. In all studies patients were allowed to continue regular medication (e.g., aspirin, NSAIDs), although opioids were usually not allowed. One study reported little or no information on additional medication [18].

Opioids Seven double-blinded RCTs compared opioids to inactive placebo in the management of chronic LBP [22–28]. Peloso [24] and Schnitzer [22] excluded patients with spinal stenosis, spondylolisthesis (Grade 2 or higher for Peloso [24]), and symptomatic disc herniation. Three trials were sponsored by Ortho-McNeil Pharmaceutical [22–24] and two by Endo Pharmaceuticals [25, 26].

Only one trial excluded patients with a history of lumbar spine surgery [28]. Two trials allowed patients with previous low-back surgery if it was performed more than 5 years previously and only if it was associated with complete pain relief [22, 24]. Three trials excluded patients with major surgery in the preceding 2 months [25], 3 months [27] or 6 months [26]. One study [23] did not report information about the surgery history of participants.

Tramadol (an atypical weak opioid) was the active agent in four studies [22–24, 28] with two studies using a

combination tablet containing acetaminophen (paracetamol) [23, 24]. Vorsanger [28] also compared two different doses of tramadol 200 mg (Vorsanger [28] in analysis) and 300 mg (Vorsanger [28] in analysis). The average daily dose of tramadol in the other studies was approximately 150 mg [23, 24] or 242 mg [22]. In two studies [26, 27] the opioid oxycodone extended release (ER) was studied. In one study the opioid oxytrex (oxycodone plus ultra low-dose naltrexone) was studied [27]. This study had three intervention arms: oxycodone qid (Webster [27]); oxytrex qid (Webster [27]); and oxytrex bid (Webster [27]).

Only one trial [29] included in this review assessed the effectiveness of opioids as compared to another analgesic (i.e., naproxen).

In five of the eight studies a so called “flare design” was used, in which patients who were already responding well to opioids were only included when they showed a large worsening in LBP complaints during a wash-out period.

Risk of bias in included studies

The results of the risk-of-bias assessment are shown in Table 1. Fourteen studies (82%) had a low RoB. All studies were described as randomized; however, only seven studies (41%) used an adequate randomization procedure in combination with an adequate concealment of treatment allocation. In only seven studies (41%) co-interventions were avoided or similar. Especially in the studies regarding NSAIDs and antidepressants this item was very poorly reported. Many studies had unacceptable compliance (10 studies; 59%) or unacceptable drop-out rates (11 studies; 67%) or both (8 studies; 41%).

Effects of intervention

A summary of the effect estimates can be found in Table 2. In addition, figures pertaining to the meta-analyses are available for all the studied interventions in Appendix 3 in Electronic Supplementary Material.

NSAIDs

NSAIDs versus placebo: pain intensity All four studies, which compared NSAIDs with placebo for chronic LBP reported sufficient data on pain intensity to enable statistical pooling. The Chi-square value for homogeneity of the WMD was 1.82 ($P > 0.5$), indicating statistical homogeneity among these studies.

There is low-quality evidence (four RCTs; $n = 1,020$; limitations in design; indirectness) that NSAIDs are more effective than placebo (WMD -12.40 ; 95% CI -15.53 to -9.26) (Appendix 3, graph 01.01 in Electronic Supplementary Material).

Table 1 Risk-of-bias assessment pharmacological studies

	Randomisation adequate?	Allocation concealed?	Groups similar at baseline?	Patient blinded?	Care provider blinded?	Outcome assessor blinded?	Co interventions avoided or similar?	Compliance acceptable?	Drop-out rate described and acceptable?	Timing outcome assessment similar?	Intention-to-treat analysis?	Total score
NSAID's												
Berry 1982	-	-	+	+	+	+	-	-	-	+	+	6
Birbara 2003	+	+	+	+	+	+	-	-	-	+	-	7
Coats 2004	+	?	+	+	?	?	?	+	-	+	-	6
Katz 2003	+	+	+	+	+	+	?	?	+	+	+	9
Antidepressants												
Atkinson 1998	+	+	+	+	+	+	+	+	-	+	+	10
Atkinson 1999	+	+	+	+	+	+	-	-	+	+	+	9
Dickens 2000	+	+	+	+	+	+	-	-	?	+	+	8
Katz 2005	+	?	-	+	+	+	-	?	+	+	-	6
Atkinson 2007	+	+	?	+	+	+	?	?	+	+	+	8
Opioids												
Jamison 1998	?	+	?	-	-	-	+	+	?	+	+	5
Schnitzer 2003	+	?	?	+	+	+	?	-	?	+	+	6
Ruoff 2003	+	?	+	+	+	+	?	-	?	+	-	6
Peloso 2004	?	?	+	+	+	+	+	-	?	+	-	6
Webster 2006	+	?	+	+	+	+	+	-	-	+	+	8
Hale 2007	?	?	-	+	+	+	+	-	-	+	-	5
Katz 2007	+	+	+	+	+	+	?	?	-	+	-	7
Vorsanger 2008	?	?	+	+	?	+	+	?	-	+	+	5

Table 2 Summary of effect estimates for pharmacological interventions

Outcome title	No. of studies	No. of participants	Statistical method	Effect size
NSAID's				
Comparison 01. NSAIDs versus placebo for chronic non-specific LBP, follow up ≤ 12 weeks				
Change in Pain Intensity from baseline on 100 mm VAS	4	1,020	Weighted mean difference (fixed) 95% CI	-12.40 [-15.53, -9.26]
Total adverse events	4	1,034	Relative risk (fixed) 95% CI	1.24 [1.07, 1.43]
Antidepressants				
Comparison 02. Antidepressants versus placebo for chronic non-specific LBP				
Pain	4	292	Standardised mean difference (random) 95% CI	-0.02 [-0.26, 0.22]
Adverse events	2	157	Relative risk (fixed) 95% CI	0.93 [0.84, 1.04]
Comparison 03. SSRIs versus placebo				
Pain	3	199	Standardised mean difference (random) 95% CI	0.11 [-0.17, 0.39]
Comparison 04. TCAs versus placebo				
Pain	2	104	Standardised mean difference (random) 95% CI	-0.11 [-0.72, 0.51]
Opioids				
Comparison 05. Opioids (all types) compared with placebo for chronic low back pain				
Pain	7	2,350	Standardised mean difference (fixed) 95% CI	0.54 [-0.72, -0.36]
Disability	4	1,258	Standardised mean difference (fixed) 95% CI	-0.19 [-0.31, -0.08]
Adverse events	4	1,176	Relative risk (random) 95% CI	1.28 [1.14, 1.44]

The four studies also reported data on adverse effects, the Chi-square value for homogeneity of the RR for adverse effects was 1.01 ($df = 3$; $P > 0.5$), indicating homogeneity among the studies. There is low-quality evidence (four RCTs; $n = 1,034$; limitations in design; indirectness) that there are statistically significant more adverse effects in the NSAIDs group (RR 1.24; 95% CI 1.07–1.43) (Appendix 3, graph 01.02 in Electronic Supplementary Material).

In a sensitivity analysis, we assessed the influence on the results if we leave out the study of Berry (1982), which studied traditional NSAIDs instead of COX-2 NSAIDs. We found no differences in effect estimates as compared with our main analysis (data available on request).

Antidepressants

Antidepressants versus placebo: pain intensity Of the five low-risk-of-bias studies comparing antidepressants with placebo, three trials reported no differences in pain between these treatments [16–18], while a treatment arm of Atkinson's 1999 study and two more low-risk-of-bias studies reported a greater reduction in pain with the use of antidepressants [15, 16, 19]. A meta-analysis of four small

placebo-controlled trials was performed [16–19], which included two studies by Atkinson [16, 19] with two and three intervention arms, respectively. One trial was excluded in the meta-analysis as follow-up means and SDs were not reported [15]. The Chi-square value for homogeneity of the SMD was 4.57 ($df = 6$; $P = 0.60$), indicating statistical homogeneity among the studies. There is moderate quality evidence (four RCTs; $n = 292$; limitations in design) that there is no difference in pain relief between antidepressants and placebo for patients with chronic non-specific LBP (SMD -0.02; 95% CI -0.26 to 0.22) (Appendix 3, graph 02.01 in Electronic Supplementary Material).

Antidepressants versus placebo: depression Four trials with a low RoB included depression as an outcome, which was measured by the Beck Depression Inventory [15, 16, 18], Hamilton Depression Scale [15, 16], and Montgomery–Asberg Depression Rating Scale [17]. These studies compared antidepressants with placebo and reported no statistically significant differences in depression [15–18]. Due to lack of data in three studies, only Dickens [17] reported data on depression, a meta-analysis could not be performed. Overall, these results suggest that there is very low quality evidence (inconsistent, imprecise, and reporting

bias) that antidepressants do not seem to reduce depression in patients with chronic LBP.

Antidepressants versus placebo: functional status One study with a low RoB included functional status as an outcome measure [17]. There is low-quality evidence (one RCT; $n = 92$; inconsistency; imprecision) that there is no difference in functional status with the use of antidepressants compared to placebo in patients with LBP.

Antidepressant type versus placebo: pain intensity Two pooled analyses were performed to evaluate the effect of two types of antidepressants on pain intensity (Appendix 3, graphs 03.01; 04.01 in Electronic Supplementary Material). Two trials with a low RoB [16, 19] were pooled comparing TCAs with placebo, including one trial with two intervention arms [19]. Three trials with a low RoB [16, 17, 19] were pooled to compare SSRIs with placebo. There is moderate evidence that SSRIs (three RCTs; $n = 199$; limitations in design; SMD 0.11; 95% CI -0.17 to 0.39) and TCAs (two RCTs; $n = 104$; limitations in design; SMD -0.11 ; 95% CI -0.72 to 0.51) are not more effective than placebo in reducing pain.

Adverse events Only two studies [15, 16] reported data about any adverse event during the study. The pooled results of these studies show that there is moderate evidence (two RCTs; $n = 157$; limitations in design) that there is no statistically significant difference between antidepressants and placebo in the occurrence of any adverse event during the study (RR 0.93; 95% CI 0.84–1.04) (Appendix 3, graph 02.02 in Electronic Supplementary Material). Adverse events that were frequently reported in both groups were dry mouth, insomnia, sedation, orthostatic symptoms and constipation.

In the study of Atkinson [19], adverse effects were reported that interfered at least ‘mildly’ with everyday function. Statistically significantly ($P < 0.001$) more adverse effects were reported in the experimental arms, desipramine $n = 19$ (63.3%) and fluoxetine $n = 16$ (51.6%), as compared to placebo $n = 3$ (13.6%).

Sensitivity analysis In the Cochrane review, the influence on the results was examined of the inclusion of the ‘positive’ trial by Atkinson [15], which did not report follow-up means and SDs, by calculating the follow-up means for pain intensity and depression from the baseline and change scores with different methods. The inclusion of Atkinson [15] in the meta-analyses for both pain and depression did not change the conclusions, that there is no difference in effect between antidepressants and placebo, and demonstrates the robustness of the findings.

Similarly, the addition of three intervention arms from the study by Atkinson [19] was examined in the meta-analyses. This study did not report statistically significant

results for the intention-to-treat analysis, only for the ‘complete cases analysis’. The Cochrane meta-analyses showed that the inclusion of Atkinson [19] did not change the conclusions for pain and antidepressant type.

Opioids

Opioids versus placebo: pain intensity A meta-analysis was performed to combine the results of seven trials [22–28]. Webster [27] and Vorsanger [28] included more than one intervention arm. There is low quality evidence (seven RCTs; $n = 2,350$; limitations in design; indirectness) that those who received opioids reported greater pain relief than those who received placebo (SMD -0.54 ; 95% CI -0.72 to -0.36) (Appendix 3, graph 05.01 in Electronic Supplementary Material).

Opioids versus placebo: functional status There is low evidence (four RCTs; $n = 1,258$; limitations in design; indirectness) that opioids (tramadol) are more efficacious than placebo for improving function as measured by the Roland Disability Questionnaire (RDQ, score 0–24, 0 = no disability) (SMD -0.19 (95% CI -0.31 to 0.08)) (Appendix 3, graph 05.02 in Electronic Supplementary Material).

Adverse events Four studies reported totals about adverse events. There is low evidence (four RCTs; $n = 1,176$; limitations in design; indirectness) that there are statistically significantly more adverse events in patients using opioids compared to placebo (RR 1.28; 95% CI 1.14–1.44) (Appendix 3, graph 05.03 in Electronic Supplementary Material). Adverse events most frequently reported were headache and nausea.

In a sensitivity analysis the influence of the inclusion in the analysis of patients with prior surgery was studied. We compared studies which included patients with prior surgery ≤ 1 year before entering the study or studies which reported no information [23, 25–27] with studies which included patients with prior surgery > 1 year before inclusion or studies which excluded patients with prior surgery [22, 24, 28]. We found no differences in effect estimates compared with our main analysis (data available on request).

In another sensitivity analysis we studied the influence of leaving studies out studies not studying tramadol as the active agent. We found no differences in effect estimates compared with our main analysis (data also available on request).

Opioids versus other drugs Only one study [29] (high RoB) compared opioids to another analgesic, i.e. naproxen. There is very low quality evidence (one RCT; $n = 23$; limitations in design; inconsistency; indirectness; imprecision)

that there is no difference in pain intensity (SMD -0.58 ; 95% CI -1.42 to 0.26) and function (SMD -0.06 ; 95% CI -0.88 to 0.76) between opioids compared to other drugs. This was likely due to the small sample size. Jamison 1998 found no improvement in function for opioids compared with naproxen.

Discussion

In this review, 17 RCTs were included that evaluated the effectiveness of pharmacological interventions for non-specific chronic low back pain.

The effectiveness of the different pharmacological interventions

No studies were found for muscle relaxants. In this review we found low quality evidence for effects on pain intensity for NSAIDs and opioids, and a small effect on function for opioids compared to placebo on the short term (<3 months) for patients with chronic low back pain. No effects were found for the use of antidepressants compared to placebo on any of the primary outcomes. NSAIDs and opioids seem to be associated with more adverse effects compared with placebo.

Methodological considerations

Despite the fact that the RoB of the studies was generally low, many studies showed flaws regarding concealment of treatment allocation, compliance, and drop-out rates. We feel the quality of future RCTs in the field of low back pain regarding these issues could be improved to reduce bias in future systematic reviews and overviews.

In three of the four studies assessing NSAIDs for chronic LBP and five of the eight studies on opioids, a so called “flare design” was used, in which patients who were already responding well to NSAIDs or opioids were only included when they showed a large worsening in LBP complaints during a wash-out period. This may have caused favourable results of the investigated NSAIDs and opioids, expressed in an overestimation of the effects and an underestimation of the adverse effects due to the selection of the study population, and certainly decreases the external validity for daily practice. It is uncertain if the results also apply to other patients with low back pain (who have not yet received NSAIDs or opioids for their LBP episode).

Adverse effects

In the studies presented in this review adverse effects were reported, although we would like to emphasize the need for

a complete and better report of adverse effects in clinical trials. Clearly, smaller randomised trials are unlikely to detect rare adverse events. Better reporting of adverse events in larger trials or prospective cohort studies is required.

According to the authors of the studies on NSAIDs, most adverse effects, including abdominal pain, diarrhea, edema, dry mouth, rash, dizziness, headache, and tiredness, were considered to be mild to moderately severe. However, the sample sizes of most of the studies were relatively small, and therefore no clear conclusion can be drawn from these studies regarding the risks of gastrointestinal and other adverse effects of NSAIDs. The statistical pooling of all adverse effects of NSAIDs compared to placebo for acute LBP indeed showed an increased RR, indicating the additional risk of using NSAIDs.

For antidepressants adverse effects, such as dry mouth, constipation, tachycardia, sedation, orthostatic hypotension, and tremor, were commonly reported, but no serious adverse effects were documented. However, the trials were also very small and not designed to evaluate adverse effects.

For opioids adverse effects were reported extensively and seemed to occur more in the opioid group compared to placebo, although here as well the numbers are small.

Overall it is difficult to draw firm conclusions regarding the risks for adverse effects of NSAIDs, antidepressants and opioids. Prospective studies with larger sample sizes are necessary to evaluate the incidence of both minor and major adverse effects.

Strengths and limitations

Several biases can be introduced in systematic reviews by literature search and selection procedure. We might have missed relevant unpublished trials, which are more likely to be small studies without positive results, leading to publication bias. Screening references of identified trials and systematic reviews may result in an over representation of positive studies in the review, because trials with a positive result are more likely to be referred to in other publications, leading to reference bias. Studies not published in English, Dutch or German were not included in this review. It is not clear whether a language restriction is associated with bias [30].

Another important limitation was the poor reporting of co-interventions, especially in the studies regarding NSAIDs and antidepressants, which hampered the study of the potential bias caused by this issue.

Implications for research

To conclude, we identified 17 RCTs that evaluated pharmacological treatment effects for patients with chronic

non-specific LBP. Most of the studies included in this review had a low RoB, although there were methodological weaknesses, especially regarding concealment of allocation, compliance, and drop-out rates. There is a need for future high-quality RCTs with special emphasis on these subjects.

Implications for practice

NSAIDs and opioids might be useful for short-term pain relief in patients with chronic LBP, who responded with an exacerbation of their symptoms after stopping their medication. However, possible adverse effects should be weighed in deciding which medication to prescribe.

Acknowledgments A grant was received from the Dutch Health Insurance Council.

Conflict of interest None.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Andersson GB (1999) Epidemiological features of chronic low-back pain. *Lancet* 354(9178):581–585
- Hart LG, Deyo RA, Cherkov DC (1995) Physician office visits for low back pain frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 20:11–19
- Vogt MT, Kwok CK, Cope DK, Osial TA, Culyba M, Starz TW (2005) Analgesic usage for low back pain: impact on health care costs and service use. *Spine* 30:1075–1081
- Roelofs PDDM, Deyo RA, Koes BW, Scholten RJPM, van Tulder MW (2008) Non-steroidal anti-inflammatory drugs for low back pain. *Cochrane Database Syst Rev* (1):CD000396. doi:10.1002/14651858.CD000396.pub3
- Van Tulder M, Touray T, Furlan AD, Solway S, Bouter LM (2003) Muscle relaxants for non-specific low-back pain. *Cochrane Database Syst Rev* (4):CD004252. doi:10.1002/14651858.CD004252
- Urquhart DM, Hoving JL, Assendelft WWJJ, Roland M, van Tulder MW (2008) Antidepressants for non-specific low back pain. *Cochrane Database Syst Rev* (1):CD001703. doi:10.1002/14651858.CD001703.pub3
- Deshpande A, Furlan A, Mailis-Gagnon A, Atlas S, Turk D (2007) Opioids for chronic low-back pain. *Cochrane Database Syst Rev* (3):CD004959. doi:10.1002/14651858.CD004959.pub3
- Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 34(18):1929–1941
- Van Tulder M, Suttrop M, Morton S, Bouter L, Shekelle P (2009) Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine* 34(16):1685–1692
- Atkins D, Best D, Briss PA et al (2004) Grading quality of evidence and strength of recommendations. *BMJ* 328:1490
- Berry H, Bloom B, Hamilton EBD, Swinson DR (1982) Naproxen sodium, diflunisal, and placebo in the treatment of chronic back pain. *Ann Rheum Dis* 41:129–132
- Birbara CA, Puopolo AD, Munoz DR, Sheldon EA, Mangione A, Bohidar NR et al (2003) Treatment of chronic low back pain with etoricoxib, a new cyclo-oxygenase-2 selective inhibitor: improvement in pain and disability—a randomized, placebo-controlled, 3-month trial. *J Pain* 4(6):307–315
- Coats TL, Borenstein DG, Nangia NK, Brown MT (2004) Effects of valdecoxib in the treatment of chronic low back pain: results of a randomized, placebo-controlled trial. *Clin Ther* 26(8):1249–1260
- Katz N, Ju WD, Krupa DA, Sperling RS, Bozalis Rodgers D, Gertz BJ et al (2003) Efficacy and safety of rofecoxib in patients with chronic low back pain: results from two 4-week, randomized, placebo-controlled, parallel-group, double-blind trials. *Spine* 28(9):851–858
- Atkinson HJ, Slater MA, Williams RA, Zisook S, Patterson TL, Grant I et al (1998) A placebo-controlled randomized clinical trial of nortriptyline for chronic low back pain. *Pain* 76:287–296
- Atkinson JH, Slater MA, Wahlgren DR, Williams RA, Zisook S, Pruitt SD et al (1999) Effects of noradrenergic and serotonergic antidepressants on chronic low back pain intensity. *Pain* 83:137–145
- Dickens C, Jayson M, Sutton C, Creed F (2000) The relationship between pain and depression in a trial using paroxetine in sufferers of chronic low back pain. *Psychosomatics* 41:490–499
- Katz J, Pennella-Vaughan J, Hetzel RD, Kanazi GE, Dworkin RH (2005) A randomised, placebo-controlled trial of bupropion sustained release in chronic low back pain. *J Pain* 6(10):656–661
- Atkinson JH, Slater MA, Capparelli EV, Wallace MS, Zisook S, Abramson I et al (2007) Efficacy of noradrenergic and serotonergic antidepressants in chronic back pain: a preliminary concentration-controlled trial. *J Clin Psychopharmacol* 27(2):135–142
- Mir S, Taylor D (1997) The adverse effects of antidepressants. *Curr Opin Psychiatry* 10:88–94
- Stahl SM (1998) Basic psychopharmacology of antidepressants: antidepressants have seven distinct mechanisms of action. *J Clin Psychiatry* 59(Suppl. 4):5–14
- Schnitzer TJ, Gray WL, Paster RZ, Kamin M (2000) Efficacy of tramadol in treatment of chronic low back pain. *J Rheumatol* 27(3):772–778
- Ruoff GE, Rosenthal N, Jordan D, Karim R, Kamin M (2003) 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. *Clin Ther* 25(4):1123–1141
- Peloso PM, Fortin L, Beaulieu A, Kamin M, Rosenthal N, Protocol TRP-CAN-1 Study Group (2004) 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. *J Rheumatol* 31(12):2454–2463
- Katz N, Rauck R, Ahdieh H, Ma T, Van Der Hoop RG, Kerwin R et al (2007) A 12-week, randomized, placebo-controlled trial assessing the safety and efficacy of oxymorphone extended release for opioid-naïve patients with chronic low back pain. *Curr Med Res Opin* 23(1):117–128
- Hale ME, Ahdieh H, Ma T, Rauck R (2007) 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. *J Pain* 8(2):175–184
- Webster LR, Butera PG, Moran LV, Wu N, Burns LH, Friedmann N (2006) Oxytrex minimizes physical dependence while

- providing effective analgesia: a randomized controlled trial in low back pain. *J Pain* 7(12):937–946
28. Vorsanger GJ, Xiang J, Gana TJ, Pascual ML, Fleming RR (2008) Extended-release tramadol (tramadol ER) in the treatment of chronic low back pain. *J Opioid Manag* 4(2):87–97
 29. Jamison RN, Raymond S, Slawsby EA, Nedeljkovi SS, Katz N (1998) Opioid therapy for chronic noncancer back pain: a randomized prospective study. *Spine* 23(23):2591–2600
 30. Egger M, Smith GD (1998) Bias in location and selection of studies. *BMJ* 316(7124):61–66