Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials

L. A. C. Machado¹, S. J. Kamper¹, R. D. Herbert¹, C. G. Maher¹ and J. H. McAuley²

Objective. Estimates of treatment effects reported in placebo-controlled randomized trials are less subject to bias than those estimates provided by other study designs. The objective of this meta-analysis was to estimate the analgesic effects of treatments for non-specific low back pain reported in placebo-controlled randomized trials.

Methods. Medline, Embase, Cinahl, PsychInfo and Cochrane Central Register of Controlled Trials databases were searched for eligible trials from earliest records to November 2006. Continuous pain outcomes were converted to a common 0–100 scale and pooled using a random effects model.

Results. A total of 76 trials reporting on 34 treatments were included. Fifty percent of the investigated treatments had statistically significant effects, but for most the effects were small or moderate: 47% had point estimates of effects of <10 points on the 100-point scale, 38% had point estimates from 10 to 20 points and 15% had point estimates of >20 points. Treatments reported to have large effects (>20 points) had been investigated only in a single trial.

Conclusions. This meta-analysis revealed that the analgesic effects of many treatments for non-specific low back pain are small and that they do not differ in populations with acute or chronic symptoms.

Key words: Meta-analysis, Randomized-controlled trial, Treatment efficacy, Low back pain, Placebo effect.

Introduction

Low back pain is a highly prevalent health problem that is associated with enormous costs worldwide [1–3]. In developed countries, episodes of back pain are a leading cause of work absence, accounting for over 25% of all conditions involving days away from work [4, 5]. About 90% of the patients with low back pain will receive the diagnosis ‘non-specific low back pain’ (NSLBP), a term that signifies that no specific pathology or disease process has been identified by the clinician. Although pain improves rapidly in the first month with a typical episode of NSLBP, low levels of pain may continue for many months [6].

The number of studies investigating the effects of treatments for patients with NSLBP has increased dramatically in the past decade. Some of these studies compare outcomes in a treated group with outcomes of a group that is given placebo treatment or sham treatment. The use of a placebo is generally considered to be a good design feature because it controls for placebo effects and, more generally, for changes in patient behaviour caused by knowledge of allocation [7]. The provision of a placebo may also enable better control of other sources of bias in clinical trials, such as measurement bias, treatment non-compliance and loss to follow-up [7, 8].

The ability of placebos to control for bias in clinical research is closely linked to the facilitation of blinding [9]. In a recent meta-epidemiological study, the lack of blinding was associated with a 25% over-estimation of treatment effects when these effects were measured in terms of subjective outcomes, such as pain [odds ratio (OR) 0.75; 95% CI 0.61, 0.93] [10]. Thus, at least from an explanatory perspective, placebo-controlled trials may provide the least biased estimates of the analgesic effects of treatments.

To our knowledge there have not been any systematic reviews focusing on the analgesic effects of treatments estimated by placebo-controlled trials on NSLBP. Thus, we performed a systematic review and meta-analysis of placebo-controlled randomized trials investigating the effects of treatments for NSLBP.

Methods

Selection of studies

The electronic databases Medline, Embase, Cinahl, PsychInfo and the Cochrane Central Register of Controlled Trials were searched from the earliest record to November 2006 for placebo-controlled randomized trials of treatments for NSLBP. Our search strategy followed the recommendations of the Cochrane Back Review Group [11]. The results were combined with the terms ‘placebo’, ‘sham’, ‘attention-control’ or ‘minimal intervention’. We also searched cited references of relevant trial reports and reviews for potentially eligible studies.

Eligible studies were randomized controlled trials comparing treatments for NSLBP against placebo. To be included, they must have reported a continuous measure of pain. Studies in which participants presented with radicular syndrome, cauda equina syndrome, infection, neoplasm, fracture, inflammatory disease, pregnancy or spinal surgery in the past 12 months were excluded, as were primary prevention studies. Trials in which the placebo intervention was a contemporary treatment (e.g. an educational booklet) were excluded.

Data extraction

Two independent reviewers extracted data using a standard form. A third reviewer extracted data for non-English studies. Trial quality was assessed using the PEDro scale [12], an 11-item quality checklist. The full scale criteria can be viewed at http://www.pedro.fhs.usyd.edu.au/scale_item.html. Disagreements were resolved by discussion and consensus. Trials were included in the analysis regardless of their quality ratings. Data on continuous pain outcomes are often reported at several time points. We chose to extract data from the first assessment after the end of the therapy. This timing was decided a priori because it was considered the time-point where the largest analgesic effects would be observed.

For simplicity, trials comparing multiple treatments of diverse nature against the same placebo had each of their comparisons treated as an individual trial. However, when a single trial...
Data analysis

Where necessary, pain scores were re-scaled to a 0- to 100-point scale. For each trial, wherever possible, the size of the treatment effect was estimated by subtracting the mean pain in the treatment group from the mean pain in the control group. Methods described in the Cochrane Handbook [13] were used to calculate the variance of the estimate. Where there were insufficient data, the s.d. for assessments at baseline or the pooled s.d. of trials reporting on the same intervention was used [14]. The same procedure was used in one trial reporting implausible s.d. [15].

Where more than one trial estimated the effect of a particular treatment, a random effects model was used to obtain a pooled estimate of the effect (weighted mean difference, WMD) of that treatment. We used MIX (version 1.6) for the analyses [16].

A pre-specified secondary analysis was performed to evaluate the efficacy of treatments in populations with distinct duration of symptoms. Acute symptoms were defined as those present for <6 weeks, sub-acute symptoms as those present from 6 weeks to 3 months, and chronic symptoms as those present for over 3 months [17]. Trials not reporting the duration of the symptoms, or having a mix of patients with acute and chronic symptoms, were not included in the secondary analysis. To judge the magnitude of treatment effects (represented by the absolute differences between experimental and placebo groups at follow-up), we used the definitions of the American College of Physicians and the American Pain Society, as follows: large treatment effect (>20 points), moderate treatment effect (10-20 points) and small treatment effect (<10 points) [18].

Results

Selection of studies

Figure 1 describes the process of study selection. A total of 1031 papers were identified by the search strategy and screened for eligibility, of which 946 failed to meet the inclusion criteria. In 21 studies the reason for ineligibility was that the trial employed a placebo that is a contemporary treatment for NSLBP [19–39]. Table 1 lists the treatments used as placebos in these trials. Of the 85 eligible trials, 9 were excluded from the analysis because they provided insufficient data to estimate treatment effects [40–48]. Thus, 76 trials reporting on 34 different treatments were included in the analysis [15, 49–123].

Characteristics of studies

Table 2 describes the characteristics of the included trials. From the 76 trials included, 81 comparisons against placebo were considered. Muscle relaxants were tested in the largest number of trials (nine trials), while NSAIDs were tested on the largest number of participants (1349 participants). Trial quality was highly variable; individual items of the quality checklist are described in the Appendix (see supplementary data available at Rheumatology Online). Two trials investigating the effects of exercise [68] and spinal manipulative therapy (SMT) [115] scored ≤3 points on the PEDro scale, a score that has been considered the cut-off for low-quality trials in previous reviews [124, 125].

In 10 trials, the treatment under investigation was delivered in addition to baseline care provided to both the experimental and the placebo groups. Baseline treatments included exercise programmes [82, 103, 104, 107, 121], heat therapy [90], NSAIDs [85], pragmatic physiotherapy [62, 64] and the provision of educational material [52]. Concurrent therapies were allowed in 36 trials, which were mostly rescue medication or the continuation of previous treatment regimens. Because of the different duration of treatments, we had to use various time-points for data extraction, which ranged from 5 min after the intervention in a trial of neureflexotherapy to 52 weeks after intervention in a trial of intradiscal steroid injections (ISIs) (Table 2).

Analgesic efficacy

Seventeen of the investigated treatments (50%) had statistically significant effects compared with placebo (Fig. 2). Point estimates of the effects were small for 16 treatments (colchicine,
N-methyl-D-aspartate antagonists, shortwave, ISIs, percutaneous thermocoagulation intradiscal techniques, radiotherapy, traction, physiotherapy, prolotherapy, exercise, anti-depressants, behavioural, adenosine triphosphate, SMT, NSAIDs and magnets), moderate for 13 treatments (analgesics, radiofrequency denervation, herbal medicines, facet injections, laser, massage, muscle relaxants, anti-convulsants, back school, nerve blocks, transcutaneous electrical nerve stimulation, heat wrap therapy and acupuncture) and large for five treatments (neuroreflexotherapy, vitamin B12, immunoglobulins and electroacupuncture). However, with the possible exception of heat wrap therapy, the CIs about moderate estimates were not narrow enough to rule out small effects. Additionally, all five large estimates were based on just one small- or moderate-sized study. A post hoc analysis excluding the two low-quality trials [68, 115] produced even smaller analgesic effects for both exercise (pooled effect size −0.8, 95% CI −1.4; 95% CI −1.8, 9.0) and SMT (pooled effect −0.8, 95% CI −1.4; 9.0, 6.6).

We sought to determine if the effects of treatment, compared with placebo, varied with the duration of symptoms (acute, sub-acute or chronic). There were no studies reporting exclusively on sub-acute NSLBP, so trials having a mix of patients with acute and sub-acute symptoms (<3 months) or with sub-acute and chronic symptoms (≥6 weeks) were treated in our secondary analysis as acute or chronic NSLBP, respectively (Table 2). Figure 3 shows the analgesic efficacy, compared with placebo, of four treatments investigated in both acute and chronic populations. There was no evidence of substantial differences in effects between these populations.

**Discussion**

A meta-analysis of 76 placebo-controlled randomized trials revealed that the analgesic effects of many treatments for NSLBP are small. Large effects were only observed in single small trials. Additionally, treatment effects do not differ in populations with acute or chronic NSLBP.

Interestingly, treatment recommendations from recent clinical guidelines do not align with the results of this meta-analysis. For example, five of the treatments recommended in the 2007 guideline of the American Pain Society (anti-depressants, SMT, exercise, acupuncture, behavioural therapies) [18] were shown in this review to be not more effective than placebo. Inconsistencies like these are not surprising because, unlike our approach, guideline committees also consider the results from trials with a no-treatment control and trials comparing two active treatments when providing their recommendations. The findings of these latter types of trials (known as pragmatic trials) are generally considered more useful for clinicians because their design replicates more closely what happens in everyday clinical practice. However, in some pragmatic trials the interpretation of findings may be more difficult than in placebo-controlled trials. For example, a null result in a trial comparing two active treatments of unknown efficacy, often observed in the NSLBP literature, may mean that the treatments are equally effective or equally ineffective since they may not be superior to a placebo.

During the conduct of the present meta-analysis, another study with a similar aim was published by Keller et al. [126]. Despite the similar aims of the two meta-analyses, their methods and execution were fundamentally different. First, the search in the Keller review was less comprehensive than ours; the review did not include trials reporting on 27 treatments included in our review, including some commonly prescribed treatments for NSLBP, such as analgesics and anti-depressants. Second, the Keller review included trials with a no-treatment control, rather than restricting the analysis to trials with a placebo control. Trials with a no-treatment control have a higher risk of bias and so may...
provide overly optimistic estimates of treatment effects. Given these differences, we believe the present meta-analysis provides a more robust evaluation of the analgesic effects of treatments for NSLBP.

Despite the greater control of bias provided by the use of a placebo control in clinical trials, the use of placebos in trials evaluating non-pharmaceutical treatments for NSLBP has been contentious. Much of the debate does not relate to ethical issues, but to problems encountered during the design of proper placebos for these trials. The distinction between placebo effects and specific treatment effects may be ill-defined in trials of non-pharmaceutical treatments. This problem arises, in part, because there is often not a clear understanding of the mechanisms underlying some non-pharmaceutical treatments [127]. Thus, the selection of a placebo for these trials generally requires considerable thought to ensure that the placebo intervention does not share some of the specific therapeutic components of the experimental intervention. This issue is more of a concern when placebos are designed to resemble the experimental intervention [124].

In some placebo-controlled trials, the placebo treatment is actually used in clinical practice as a treatment. Examples are educational booklets [23, 28], massage [26] and exercises [25, 27, 36]. In this meta-analysis, we excluded trials using a placebo consisting of a contemporary treatment. We took this approach to minimize the possibility of under-estimation of treatment effects. The opposite problem can also arise in placebo-controlled trials: some placebos may lack credibility, which could cause an over-estimation of treatment effects. Unfortunately, trial reports usually contain insufficient information to judge whether this is a problem [124], so we could not exclude trials using placebos that are not credible. As a consequence it is possible that our estimates of the effects of treatments were exaggerated. It would seem unlikely, therefore, that our finding of small effects of treatments for NSLBP is due to the inadequate design of placebo in placebo-controlled trials.

Some authors have argued that the small effects of treatments for acute NSLBP are a consequence of the favourable natural history of acute NSLBP. The theory is that, at the conclusion of treatment in trials, control groups have improved substantially and so there is not ‘room’ for large treatment effects. To evaluate this argument we examined the baseline and follow-up scores from the acute trials included in the present meta-analysis. Three trials [64, 73, 85] did not report sufficient baseline data and were not considered. In 18 trials [49, 72, 78, 84, 91, 95, 97, 114, 115, 117, 120, 121] mean pain levels at baseline were 62.1 (S.D. 16.5) points in the treatment group and 61.5 (S.D. 15.9) points in the placebo group. At follow-up mean pain levels were 29.5 (S.D. 13.2) points in the treatment group and 39.6 (S.D. 17.1) points in the placebo group. This indicates that there is scope for treatment effects (i.e. mean between-group differences) as large as 40 points to be

---

**FIG. 2. Analgesic efficacy of treatments for NSLBP of any duration. Squares represent pooled estimates of random effects (multiple trials) or means (single trials).** Error bars are 95% CIs. Negative values favour treatment. In parentheses: number of trials; total number of participants. The dotted lines define the magnitude of effects: large (>20U); moderate (10–20U); small (<20U). ATP: adenosine triphosphate; ISIs: Intradiscal steroid injections; NMDA: N-methyl-D-aspartate. PTIT: percutaneous thermocoagulation intradiscal techniques; RF: radiofrequency; SMT: Spinal manipulative therapy; TENS: transcutaneous electrical nerve stimulation;
demonstrated in trials of NSLBP. Thus, the theory that there is no ‘room’ for trials to show large effects of treatments for acute NSLBP does not seem consistent with the data. Given the mean baseline pain observed in the present meta-analysis, on average, a 10-point difference in pain between treatment and placebo groups is equivalent to a 16% difference between groups in improvement of pain from baseline.

Another argument used to explain the small treatment effects found in the NSLBP literature is that most trial samples are conducted on samples from clinically heterogeneous populations. It is possible that specific treatments have large treatment effects on specific subgroups of patients with NSLBP [128, 129]. However, the evidence of a differential response of identifiable subgroups in the NSLBP literature is contradictory, as some authors report a differential response of subgroups [130–132] whereas others do not [133, 134].

The small effects found in this meta-analysis might also be attributed to the choice of outcome measure; i.e. reduction in pain. It could be argued that pain is not the most appropriate outcome to make a judgement on the efficacy of treatments that are designed to improve other outcomes, such as function or quality of life. However, we feel it is unlikely that an examination of other outcomes would produce meaningfully different conclusions to those in the current meta-analysis because, in previous reviews, pain has consistently shown larger responses to treatment than other outcomes for NSLBP [126, 135]. For example, in a previous meta-analysis on the effects of exercise therapy [135], the pooled effect of exercise for chronic NSLBP was, at short-term follow-up, 7.3 points on a 100-point scale (95% CI 3.7, 10.9) for pain outcomes, and only 2.5 points on a 100-point scale (95% CI 1.0, 3.9) for functional outcomes. A similar pattern was observed for intermediate- and long-term follow-ups [135].

Our meta-analysis has a number of strengths. First, this is the first meta-analysis to provide the estimates of true treatment effects of all treatments for NSLBP that have been tested against placebo. We used a comprehensive search strategy to identify potentially eligible trials, in contrast to other reviews that used previously published systematic reviews as the primary source of data [126, 136, 137]. Additionally, we excluded placebo-controlled trials in which the choice of placebo was inappropriate. One potential limitation of the present meta-analysis is the investigation of just one outcome. The outcome of pain was chosen because pain relief is ranked by patients as one of the most important components for the satisfactory management of low back pain and it is often the original motivation for seeking care from a health practitioner [138], and because most interventions appear to produce consistently greater reductions in pain than in other outcomes.

The available evidence from placebo-controlled trials shows only small to moderate treatment effects, over and above placebo, for many interventions that are currently used in the management of NSLBP. There seems to be a considerable scope for treatments for NSLBP to show large treatment effects but how this can be achieved is at present unclear.

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at Rheumatology Online.


