Efficacy of non-steroidal anti-inflammatory drugs for low back pain: a systematic review of randomised clinical trials

Bart W Koes, Rob J P M Scholten, Jan M A Mens, Lex M Bouter

Abstract

Purpose—To assess the efficacy of non-steroidal anti-inflammatory drugs (NSAIDs) for low back pain.

Data sources—Computer aided search of published randomised clinical trials and assessment of the methods of the studies.

Study selection—26 randomised clinical trials evaluating NSAIDs for low back pain were identified.

Data extraction—Score for quality (maximum = 100 points) of the methods based on four categories: study population; interventions; effect measurement; data presentation and analysis. Determination of success rate per study group and evaluation of different contrasts. Statistical pooling of placebo controlled trials in similar patient groups and using similar outcome measures.

Results—The methods scores of the trials ranged from 27 to 83 points. NSAIDs were compared with placebo treatment in 10 studies. The pooled odds ratio in four trials comparing NSAIDs with placebo after one week was 0.53 (95% confidence intervals 0.32 to 0.89) using the fixed effect model, indicating a significant effect in favour of NSAIDs compared with placebo. In nine studies NSAIDs were compared with other (drug) therapies. Of these, only two studies reported better results of NSAIDs compared with paracetamol with and without dextropropoxyphene. In the other trials NSAIDs were not better than the reference treatment. In 11 studies different NSAIDs were compared, of which seven studies reported no differences in effect.

Conclusions—There are flaws in the design of most studies. The pooled odds ratio must be interpreted with caution because the trials at issue, including the high quality trials, did not use identical outcome measures. The results of the 26 randomised trials that have been carried out to date, suggest that NSAIDs might be effective for short-term symptomatic relief in patients with uncomplicated low back pain, but are less effective or ineffective in patients with low back pain with sciatica and patients with sciatica with nerve root symptoms.

Low back pain is an important medical and socio-economical problem in western societies.1–3 A variety of therapeutic interventions are available, but, their efficacy often remains unknown.4–6 Consequently, decisions regarding optimal management strategies are not easy for physicians and therapists involved with the care for patients with low back pain. Possibly as a consequence of this situation the management of low back pain shows typically a large variation.6–9 The Quebec Task Force on Spinal Disorders reported in 1987 that the efficacy of most interventions had not been demonstrated by sound randomised clinical trials.7 In our recent series of review articles we assess the available randomised clinical trials to evaluate the scientific evidence of common interventions for low back pain. In earlier review articles we have reported on the efficacy of exercise therapy, spinal manipulation and mobilisation, bed rest and orthoses, back schools, traction therapy, and epidural corticosteroid injections.10–15 In this article we will focus on the efficacy of non-steroidal anti-inflammatory drugs (NSAIDs) for low back pain.

Worldwide NSAIDs seem to be the most commonly prescribed medications16 and they are also widely used for patients with rheumatic disorders, including low back pain.17 The US clinical guidelines for the management of acute low back pain state that there is fair to good evidence for the prescription of NSAIDs for symptom control when the patients’ response to non-prescription analgesics is inadequate.18 Their recommendation is based on four randomised clinical trials meeting their selection criteria only.19–22 The clinical guidelines from the UK, based on the same information, also recommend prescription of NSAIDs (and simple analgesics) in the early management strategy as symptom pain relief to prevent disability.21
The rationale of NSAIDs treatment for low back pain is based on both their analgesic potential and their anti-inflammatory action. To determine the current situation regarding the efficacy of NSAIDs for low back pain, we systematically assessed the evidence from published randomised clinical trials. As even randomised clinical trials may show biased outcomes related to methodological shortcomings in the design,12 strong emphasis is laid on the methodological quality of the trials.

Methods

SELECTION OF STUDIES

A MEDLINE literature search was carried out for the period 1966-1994 (keywords (MeSH): backache, low back pain, anti-inflammatory agents, non-steroidal (including all minor subheadings). An EMBASE (Drugs and Pharmacology) search was carried out for the period 1980-1994 (keywords: non-steroid anti-inflammatory agent, backache, low back pain). In addition, the references given in relevant publications were further examined. Abstracts and unpublished studies were not selected. Studies had to meet the following criteria: (1) concerned a randomised clinical trial; (2) one treatment regimen included an NSAID (additional interventions were permitted); (3) the study subjects suffered from low back pain (or at least a subgroup of which the results are presented separately); and (4) the article was written in English.

ASSESSMENT OF METHODOLOGICAL QUALITY

All eligible trials were scored according to the criteria listed in table 1. The criteria are based on generally accepted principles of intervention research. Similar criteria have previously been used to assess the methodological quality of trials evaluating other therapeutic interventions for low back pain.10-14 To each criterion a weight was attached indicating their putative relative importance. The maximum score for each study was 100 points. Items B, C, E, J, K, M, O, are relevant for assessing the internal validity of the trials.13 All trials were assessed by two reviewers (RPMS, JMAM) independently of each other. In a subsequent meeting they had to reach consensus on each criterion they initially disagreed upon. Where disagreement persisted, a third reviewer (BWK) made the final decision. The assessments resulted in a hierarchical list in which higher scores indicate studies of higher methodological quality. The outcome of the studies will be discussed in relation to their methodological scores.

OUTCOME OF THE STUDIES AND STATISTICAL POOLING

A study was judged to be positive if the authors concluded that the NSAID at issue was more effective than the reference treatments (for example, placebo capsules, other NSAIDs or other (drug)therapy). Usually this meant that the difference in effect for the primary outcome was statistically significant at the conventional 5% level. In a negative study the authors reported no differences between the study treatments, or even better results in favour of the reference treatment.

Pooling was to be limited to studies of which the characteristics (that is, NSAID treatment/reference treatment, patients, and outcome) were clinically sufficiently similar. After assessment of the trials we agreed that only the placebo controlled trials were sufficiently similar to permit statistical pooling. We attempted to pool data for acute and chronic low back pain patients separately, using the (forced) success rates determined one and two weeks after randomisation. The results of a subset of trials were pooled statistically using Peto’s ‘observed minus expected’ method. We included a test for homogeneity of the odds ratios (ORs) of the randomised controlled trials.29 If there was heterogeneity, we present ORs and 95% confidence intervals (CIs) using the fixed effects model as well as the more conservative random effects model.30 Results are presented as ORs with corresponding 95% CIs. Treatment failures were compared between the intervention groups: an OR below 1 indicates a better outcome of the NSAID at issue. Sensitivity analysis was carried out by performing separate meta-analyses on subsets of trials based on methodological quality (that is, those higher and lower than 50 points).

Results

A total of 26 trials met the inclusion criteria and were included in this review. Of these, four trials were published between 1960-1970, four between 1971-1980, 13 between 1981-1990, and five were published after 1991. Table 2 presents the trials in hierarchical order, according to their methodological quality.

Initially, there was disagreement between the two independent reviewers in 207 (20%) of the 1040 items scored. Disagreement mainly occurred because of reading and interpretation errors, and in the assessment of the two studies using a crossover design.22,29 Most of the disagreement was solved in a subsequent consensus meeting. The third reviewer had to
Table 2  Randomised trials on the efficacy of NSAIDs in order of methods score

<table>
<thead>
<tr>
<th>First author (reference)</th>
<th>Scores for methods criteria*</th>
<th>Total score</th>
<th>Indication†</th>
<th>Conclusion‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosie³¹</td>
<td>2 — 5 2 4 17 8 — 5 5 5 10 10 — 5 5 83</td>
<td>Acute low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Amlie²⁰</td>
<td>2 — 4 2 4 17 8 5 — 5 5 6 6 — — 5 5 73</td>
<td>Acute low back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Goldie³²</td>
<td>2 4 3 4 — 5 5 5 6 6 — — 5 5 62</td>
<td>Acute low back pain and sciatica</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Weber³³</td>
<td>2 — 3 — 4 8 10 5 — 5 5 6 6 — — 5 5 59</td>
<td>Acute sciatica with nerve root symptoms</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Hickey³⁴</td>
<td>2 4 1 2 4 — 8 — 5 5 8 10 10 — 5 5 61</td>
<td>Chronic low back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Blazek³⁵</td>
<td>2 — 3 3 4 — 8 — 5 5 5 8 8 — 5 5 61</td>
<td>Acute low back pain and lumboischialgie (herniated disc)</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Hosie³¹</td>
<td>2 — 3 1 2 17 8 5 — — 5 8 8 — — — 57</td>
<td>Acute back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Aoki³⁶</td>
<td>2 — — 2 — 4 17 8 — 5 — 5 8 8 — — 5 55</td>
<td>Acute and chronic back pain</td>
<td>Slightly positive</td>
<td></td>
</tr>
<tr>
<td>Szpalski³⁷</td>
<td>2 — 4 1 4 — 8 5 — 5 5 8 8 — — 5 5 49</td>
<td>Acute low back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Siegert³⁸</td>
<td>2 — — 2 4 — 8 — 5 5 10 10 — — 5 5 49</td>
<td>Spondylolisthesis</td>
<td>Not clear</td>
<td></td>
</tr>
<tr>
<td>Sweetman³⁹</td>
<td>2 — 4 — 2 — 6 — 5 — 5 10 10 — — 5 5 49</td>
<td>Acute low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Weber⁴⁰</td>
<td>2 — 1 2 4 — 8 5 — 5 5 8 8 — — — 48</td>
<td>Prolapsed discs</td>
<td>Not clear</td>
<td></td>
</tr>
<tr>
<td>Szpalski³⁷</td>
<td>2 — 3 2 4 — 8 — 5 — 5 6 6 — — 5 5 45</td>
<td>Chronic low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Matsumoto⁴⁴</td>
<td>1 — — — 4 8 6 — 5 5 5 4 2 — — — 5 45</td>
<td>Chronic low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Berry⁴⁵</td>
<td>2 — 1 2 4 — 8 5 5 5 4 4 — — — 44</td>
<td>Chronic back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Weber⁴⁰</td>
<td>2 — 2 — — — 8 5 — — 5 8 8 — — 5 5 43</td>
<td>(Sub)acute sciatica</td>
<td>Slightly positive</td>
<td></td>
</tr>
<tr>
<td>Wiesel⁴⁶</td>
<td>2 — 3 3 4 — 6 — 5 5 — 4 — — 5 5 42</td>
<td>Acute low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Waterworth⁴⁷</td>
<td>2 — 2 2 4 — 8 — 5 5 — 10 — — 5 5 38</td>
<td>Acute low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Brown⁴⁸</td>
<td>2 — 1 — 2 — 8 — 5 5 — 8 2 — — 5 5 38</td>
<td>Acute low back pain</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Evans³⁰</td>
<td>1 — 1 — 4 — 10 — 5 5 — 6 2 — — — 34</td>
<td>Acute low back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Jacobs⁴ⁱ</td>
<td>1 — — — 2 — 8 5 — — 5 6 6 — — — 33</td>
<td>Low back pain with or without nerve root pain</td>
<td>Positive (in subgroup)</td>
<td></td>
</tr>
<tr>
<td>Aghababian³¹</td>
<td>2 — 3 1 — — 8 — 5 — 8 — 8 — — 8 32</td>
<td>Acute low back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Boström³²</td>
<td>2 — — 1 4 — 6 5 — 5 2 2 — — 5 5 30</td>
<td>Acute back pain</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td>Postacher³³</td>
<td>2 — — 2 — 4 — 5 — — 6 5 — 5 27</td>
<td>Acute and chronic low back pain</td>
<td>Not clear</td>
<td></td>
</tr>
</tbody>
</table>

* See Appendix for details of criteria. † The labels ‘acute’ and ‘chronic’ are according to the authors of the study. ‡ Conclusion of the author(s) of the study; positive conclusion = NSAIDs better than the reference treatment; negative conclusion = NSAIDs worse than or equally effective as reference treatment.
make a final decision in 15 instances, mainly relating to criterion (C) ‘comparability of baseline characteristics’ in the case of the two crossover trials.

Table 2 shows the wide range in methodological scores (range 26-83). There were nine studies that scored more than 50 points (maximum score = 100). The median score was 48 points, indicating the overall moderate methodological quality of the trials. The most prevalent methodological shortcomings were (B) no description of randomisation procedure, (C) non-similarity regarding relevant baseline characteristics, (D) no adequate description of drop outs, (F) the small size of the study populations included, (H) no placebo control group, (M) no blinded outcome measurement (N) no long term (six months or longer) follow up (O) no intention to treat analysis, including a worst case analysis in cases with more than 10% loss to follow up.

If we consider the validity items only (items: B, C, E, J, K, M, and O from table 1) there seem to be no important changes in the hierarchy of the trials. In the top of the list Hosie31 remains the best study with 34 (71%) out of the maximum of 48 points for validity, followed by Goldie32 with 33 points. At the bottom of the list Postacinni33 remains with two points.

Overall, there were nine positive and 12 negative studies. In two studies positive results were reported for a subgroup only, and in three studies no conclusion was drawn. As the NSAIDs were compared with different reference treatments we present the results for comparisons with placebo (table 3), other (drug) therapy (table 4), and other NSAIDs (table 5), separately.

COMPARISONS WITH PLACEBO THERAPY
In five of 10 trials in which an NSAID was compared with a placebo the authors reported better results with the NSAID (table 3). Two trials reported positive results in a subgroup only, and in two other trials the authors reported no differences between the NSAID and the placebo. In one trial no conclusion was drawn. Of the five studies with methodological scores above 50 points, two reported a favourable outcome of NSAID in patients with acute low back pain. One reported favourable results of NSAID in a subgroup of acute low back pain (that is, those with initial moderate to severe pain) only. The two other studies reported no differences in effect between the NSAID and the placebo in patients with (a) acute low back pain and sciatica and (b) acute sciatica with nerve root symptoms.

In four of 10 trials patients were allowed to use rescue analgesics, usually paracetamol and codeine. In two of these the patients in the placebo group significantly used more rescue analgesics.20-21 In the two other trials there were no significant differences between the study groups regarding the use of additional analgesics.22 25 In three trials no rescue analgesics were permitted22 24 50 and in three other publications rescue analgesics are not mentioned at all.

COMPARISONS WITH OTHER (DRUG) THERAPIES
There were nine trials comparing NSAIDs with other (drug) therapies (table 4). In five trials NSAIDs were not better than the reference treatment in patients with acute low back pain (four studies) and in chronic low back pain (one study). In three trials NSAIDs were reported to be better than the reference treatment in acute low back pain (two studies) and in chronic low back pain (one study). In one study no conclusion was drawn. Unfortunately, only one study scored more than 50 points. In this study NSAIDs were found to be more effective than paracetamol in patients with chronic low back pain.29

COMPARISONS BETWEEN DIFFERENT NSAIDS
In 11 trials a comparison was made between different NSAIDs (table 5). In seven of these, there were no differences in effect between the NSAIDs for patients with acute or chronic low back pain. In three studies positive results were reported of one NSAID over the other(s) and in one study no conclusion was drawn. There were only three studies with more than 50 points. All three showed no difference in effect between the NSAIDs under study, although the authors of one study were more positive about one of the drugs.37

SIDE EFFECTS
Complications or side effects of NSAIDs were reported in most of the trials included in this review. The number of patients reporting side effects varied from 0% to 31%. The side effects usually concerned mild to moderately severe events, such as abdominal pain and diarrhoea, and other side effects such as oedema, dry mouth, rash, dizziness, headache, tiredness, etc. There seemed to be no clear difference in the reported number or severity of side effects, or both, between the different types of NSAIDs.

STATISTICAL POOLING
Only the placebo controlled studies were regarded to be sufficiently similar to permit statistical pooling of the data. In general, the methodological quality of the placebo controlled trials was higher than the trials investigating other contrasts. Most placebo controlled trials involved patients with acute low back pain (duration less than six weeks). Of the 10 trials, seven involved patients with acute low back pain, two with chronic low back pain, and in one study the duration was not described.36 We refrained from performing a pooling of the two trials on chronic low back pain, because in one of these success rates could not be extracted. Of the seven placebo controlled trials on acute low back pain, three trials presented insufficient data to extract success rates. Unfortunately, this concerned two studies with relatively high methods scores.32 33 Efforts to contact the authors to obtain additional information did not succeed.

Four studies were included in the meta-analysis in which the short-term results were pooled (fig 1).21 32 36 47 The χ² value for homogeneity of the ORs was 4.34 (3 df; p = 0.227). The pooled odds ratio for the success rate determined after one
<table>
<thead>
<tr>
<th>Author</th>
<th>Reference</th>
<th>NSAIDs dose/frequency/duration</th>
<th>Reference treatment(s)</th>
<th>Method scores</th>
<th>Results*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amlie</td>
<td>(20)</td>
<td>piroxicam 20mg capsules/2 twice a day first two days, one twice a day next 5 days/7d (140)</td>
<td>placebo capsules (142)</td>
<td>73</td>
<td>More pain relief than placebo measured with visual analog scale after 3 days. After 7 days no significant differences. Side effects similar.</td>
</tr>
<tr>
<td>Goldie</td>
<td>(32)</td>
<td>indomethacin 25mg capsules/3 times a day for 50 capsules (25)</td>
<td>placebo capsules (25)</td>
<td>62</td>
<td>No of patients with complete relief of pain after 7 and 14 days (i) 7, 14 (ii) 9, 16. No significant differences. Side effects comparable.</td>
</tr>
<tr>
<td>Weber</td>
<td>(33)</td>
<td>piroxicam 20mg capsules/40mg dfirst two days, 20mg d next 12 days/14d (120)</td>
<td>placebo capsules (94)</td>
<td>59</td>
<td>Reduction of pain in back and leg measured by visual analog scales after 4 weeks the same in the two groups (data in graphs). No significant differences. More side effects in (i) 22% (ii) 13%.</td>
</tr>
<tr>
<td>Lacey</td>
<td>(36)</td>
<td>piroxicam 10mg capsules/4 times a day first two days, two times a day next 12 days/14d (168)</td>
<td>placebo capsules (169)</td>
<td>57</td>
<td>Patient (%) improved after 1 week only in subgroup with initial moderate/severe pain (i) 82% (ii) 51% (iii) 36%. No differences for subgroup with mild initial pain. Results after 2 weeks not reported (no data presented on side effects for subgroup with back pain.</td>
</tr>
<tr>
<td>Szpalski</td>
<td>(38)</td>
<td>tenoxicam 20mg im injection on day one + 20mg capsules 1d for day 2 to 14 (−7 days bed rest) (37)</td>
<td>placebo injection + placebo capsules (36)</td>
<td>33</td>
<td>Mean pain intensity on VAS on day 1, 8 and 15 (i) 7.4, 1.9, 0.6 (ii) 7.1, 2.8, 0.8.</td>
</tr>
</tbody>
</table>
| Radin      | (43)      | phenylbutazone 100mg capsules/6 times a day for 2 days, 4 times a day next 8 days (6) | placebo capsules | 48         | Group (i) improvement in motor weakness and painful straight leg raising. No results are reported for group (ii). Side effects reported in group (i) 3%

*Results of the most important outcome measures according to the authors of the study. P values < 0.05 were taken as significant. †d = per day.
Table 4  Details of trials comparing the efficacy of NSAID with other (drug) therapies for low back pain

<table>
<thead>
<tr>
<th>First author (reference)</th>
<th>NSAIDs dose/frequency/duration (no of patients)</th>
<th>Reference treatment(s) (no of patients)</th>
<th>Methods score</th>
<th>Results*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hickey</td>
<td>(i) diflunisal 500 mg capsules/twice d/4 wk (16)</td>
<td>(ii) paracetamol 1000 mg/four times d/4 wk (13)</td>
<td>61</td>
<td>No of patients with none or mild low back pain after 2 and 4 weeks (i) 11, 13 (ii) 9, 7. Significantly more patient in (i) (10 out of 16) considered the therapy as good or excellent (ii) (4 out of 12). Side effects similar (i) (2) (ii) 1</td>
</tr>
<tr>
<td>Sweetman</td>
<td>(i) mefenamic acid 500 mg three times d + placebo twice d (40)</td>
<td>(ii) Chlormezanone 100 mg and paracetamol 450 mg two capsules three times d + placebo three times d (42) (ii) Ethoheptazine 75 mg and meprobamate 150 mg and aspirin 250 mg two capsules + placebo three times d (40) (ii) meptazinol 200 mg capsule/4 times d/3 wk (35)</td>
<td>49</td>
<td>No of patients reporting no pain after 1 and 7 days (i) 7, 21 (ii) 12, 23 (iii) 10, 20. No of patients with adverse events (i) 9 (ii) 10 (iii) 16</td>
</tr>
<tr>
<td>Videman</td>
<td>(i) diflunisal 250 mg capsule/4 times d/3 wk (35)</td>
<td>(ii) meptazinol 200 mg capsule/4 times d/3 wk (35)</td>
<td>49</td>
<td>Mean change in degree of pain on 100 mm VAS at three weeks (i) 45 (ii) 40. Similar improvement regarding capacity for daily tasks (data in graphs). No significant differences. (Side effects similar (i) 19 (ii) 23 patients)</td>
</tr>
<tr>
<td>Wiesel</td>
<td>(i) aspirin 625 mg capsules/4 times d/2 wk (~)</td>
<td>(ii) phenylbutazone 100 mg capsules/4 times d (first 5 days). No further information given (~)</td>
<td>42</td>
<td>Mean no of days before return to full activity (i) 5.7 (ii) 6.5 (ii) 5.7. No significant differences. No data on side effects given</td>
</tr>
<tr>
<td>Waterworth</td>
<td>(i) diflunisal 500 mg capsules, 1000 mg immediately, 500 mg twice d/10 d (36)</td>
<td>(ii) physiotherapy: local heat, ultrasound and exercises (5 × 45 min session weekly) (34) (iii) spinal manipulation and/or McKenzie therapy (5 × 45 min session weekly) (38)</td>
<td>39</td>
<td>Mean change in pain intensity on 4 points scale after 4 and 12 days: (i) −0.9, −1.7 (ii) −0.9, −1.6 (iii) −1.1, −1.7. No significant differences in pain and mobility</td>
</tr>
<tr>
<td>Brown</td>
<td>(i) diflunisal (capsules)/initial dose 1000 mg, 500 mg every 12 hours/15 d (19)</td>
<td>(ii) acemetacinophen 300 mg with codeine 50 mg/two capsules initially one capsule every 4 h/15 d (21)</td>
<td>38</td>
<td>Pain assessment by patient and investigator on 3 point ordinal scale show similar improvement curves (data in graphs). No of patient rating drugs as excellent or very good (i) 9 (ii) 9. No significant differences. Side effects: more side effects in (ii) 10 than in (i) 3. Mean daily pain index during intervention period (on point ordinal scale) (i) 1.4 (ii) 1.5 (iii) 1.4 (iv) 1.4 (v) 1.7 (vi) 1.7. (iii) significantly different from (v and vi). (i) significantly different from (v). Side effects: more side effects in (i) 20 (ii) 19 (v) 19 than in (iii) 12 (vi) 12 (v) 12 (iv) 4</td>
</tr>
<tr>
<td>Evans</td>
<td>(i) aspirin 300 mg capsule 4 times d/1 wk (30) (ii) indomethacin 50 mg3 times d/1 wk (30) (iii) mefenamic acid 250 mg2 capsules 3 times d/1 wk (30) (iv) paracetamol 500 mg capsules2 capsules 4 times d/1 wk (30) (v) phenylbutazone 100 mg 5 times d/1 wk (30)</td>
<td>(v) dextropropoxyphene 32.5 mg and paracetamol 325 mg capsules2 capsules 4 times d/1 wk (30) (vi) paracetamol 500 mg capsules2 capsules 4 times d/1 wk (30)</td>
<td>34</td>
<td>No of patients reporting marked improvement after 2, 4, and 7 days. (i) 10, 14, 24 (ii) 10, 14, 31 (iii) 8, 18, 30. Group (i) significantly better after 4 days (based on total improvement). No other significant differences were found. No significant side effects were found</td>
</tr>
<tr>
<td>Baenajian</td>
<td>(i) diflunisal capsules 500 mg/twice d (<del>) (ii) diflunisal capsules 500 mg + 5 mg cyclobenzaprine/twice d (</del>)</td>
<td>(iii) cyclobenzaprine capsules 5 mg/twice d (~)</td>
<td>30</td>
<td>Mean improvement on combined pain, disability, and spinal mobility score (5-32) after 3 wks, 2, and 6 mths. In subgroup with acute pain (i) 3.0, 10.7, 14.0 (ii) 7.5, 9.7, 12.3 (iii) 5.0, 8.4, 10.2 (iv) 5.4, 7.5, 7.3 (v) not included (vi) 1.8, 7.3, 11.0. Group (ii) significantly better than others after 3 wks; no other differences. In subgroup with chronic pain (i) 2.6, 2.2, 4.0 (ii) 2.2, 2.6, 4.3 (iii) 3.9, 4.2, 6.0 (iv) not included (v) 0.5, 4.6, 8.9 (vi) 0.7, 1.2, 2.0. Group (i) not significantly better. No data on side effects reported</td>
</tr>
<tr>
<td>Postacchini</td>
<td>(i) diclofenac 'full dosage'10-14 days (acute patients), 15-20 days (chronic patients) (81)</td>
<td>(ii) chiropractic manipulation (87) (iii) physiotherapy (78) (iv) bed rest (29) (v) back school (90) (vi) placebo (anti-oedema gel) (73)</td>
<td>27</td>
<td>Results of the most important outcome measures according to the authors of the study. p Values &lt; 0.05 were taken as significant.</td>
</tr>
</tbody>
</table>
Table 5 Details of trials comparing the efficacy of different NSAIDs for low back pain

<table>
<thead>
<tr>
<th>First author (reference)</th>
<th>NSAIDs dose/frequency/duration (no of patients)</th>
<th>Methods score</th>
<th>Results*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosie et al.</td>
<td>(i) ibuprofen capsules 400 mg three times d (+ placebo foam 3 times d)/14 d (147) (ii) felbinox (foam 3%) 3 times d (+ placebo capsules 3 times d)/14 d (140)</td>
<td>83</td>
<td>Patients (%) reporting none or mild severity after 1 and 2 weeks (i) 84, 92 (ii) 76, 88. No significant differences between the groups. No of side effects (i) 22 (ii) 26.</td>
</tr>
<tr>
<td>Blazek et al.</td>
<td>(i) diclofenac 25 mg capsules/four times d first 4 days and three times d next 8 days/12 d (14) (ii) ibuproxan 300 mg capsules/four times d first 4 days and three times d next 8 days/12 d (14)</td>
<td>61</td>
<td>Average improvement on ordinal 5 points scale (0= no response, 4=very good response) during and after the intervention period of 12 days according to physician and patient (i) 2.6 and 2.8 (ii) 2.8 and 3. No significant differences in recovery rate. Sid effects: mild side effects in three patients in each group.</td>
</tr>
<tr>
<td>Aoki</td>
<td>(i) piroxicam 20 mg capsules/once d/14 d (116) (ii) indomethacin 25 mg capsules/three times d/14 d (114)</td>
<td>57</td>
<td>No of patients (%) who are (very much) improved after 1 and 2 weeks assessed by a physician (i) 49% (ii) 46%. No significant differences. Side effects: similar (i) 11% (ii) 13%.</td>
</tr>
<tr>
<td>Siegmeth et al.</td>
<td>(i) ibuprofen/1200 mg d/14 d (15) (ii) diclofenac/75 mg d/14 d (15)</td>
<td>49</td>
<td>No of patients reporting to be improved after 1, 3, 4 weeks (i) 5, 10, 6 (ii) 5, 12, 11. No significant differences. Side effects similar: one in each group.</td>
</tr>
<tr>
<td>Orava et al.</td>
<td>(i) diflunisal 500 mg capsules/twice d/7 d (66) (ii) indomethacin 50 mg capsules/three times d/7 d (67)</td>
<td>48</td>
<td>No of patients (5) assessing therapy as good or excellent after 3 and 7 days (i) 45%, 64% (ii) 45%, 64%. No significant differences. More side effects in (i) than in (ii).</td>
</tr>
<tr>
<td>Videman et al.</td>
<td>(i) piroxicam 20 mg + placebo capsules one d + twice placebo/6 wk (14) (ii) indomethacin 25 mg capsules/three times d/6 wk (14)</td>
<td>45</td>
<td>Mean improvement on VAS (range 0-31) after 6 weeks (i) 8 (ii) 9. Similar improvement rates (data in graphs). Side effects similar (i) 13 (ii) 15.</td>
</tr>
<tr>
<td>Matsumo et al.</td>
<td>(i) ketoprofen 25 mg capsules/150 mg d/duration not given (77) (ii) diclofenac sodium 25 mg capsules/75 mg d/duration not given (78)</td>
<td>45</td>
<td>No of patients improved after 1, 2 weeks (i) 71%, 86% (ii) 62%, 79%. No significant differences. Side effects similar in both groups (i) 18% (ii) 21%.</td>
</tr>
<tr>
<td>Berry et al.</td>
<td>(i) naproxen sodium 275 mg capsules/2 times 2 capsules d/14 d (37) (ii) diflunisal 250 mg capsules/2 times d/14 d (37)</td>
<td>44</td>
<td>Decrease in pain by VAS (i) reduction of pain (ii) no change. Data in graphs. Group (i) somewhat better than (ii). Side effects similar in the two groups (i) 18 (ii) 18.</td>
</tr>
<tr>
<td>Wiesel et al.</td>
<td>(i) aspirin 625 mg capsules/4 times d/2 wk (−) (ii) phenylbutazone 100 mg capsules/4 times d/first 5 days.</td>
<td>42</td>
<td>Mean no of days before return to full activity (i) 5.7 (ii) 6.5. No significant differences. No data on side effects given.</td>
</tr>
<tr>
<td>Evans et al.</td>
<td>(i) aspirin 300 mg capsules 4 times d/1 wk (30 co) (ii) phenylbutazone 100 mg capsules/3 times d/1 wk (30 co)</td>
<td>34</td>
<td>Mean daily pain index during intervention period (on 4 point ordinal scale) (i) 1.4 (ii) 1.5 (iii) 1.4 (iv) 1.4 (v) 1.7 (vi) 1.7. (iii) significantly different from (iv and vi). (i) significantly different from (v). More side effects in (i) 20 (ii) 19 (v) 19 than in (iii) 12 (vi) 13 (iv) 4.</td>
</tr>
<tr>
<td>Aghababian et al.</td>
<td>(i) diflunisal capsules 1000 mg initially, 500 mg every 8-12 hrs/2 wks (16) (ii) naproxen capsules 500 mg initially, 250 mg every 6-8 h/2 wks (17)</td>
<td>32</td>
<td>No of patients (%) reporting no pain (on a ordinal 4 point scale) after 2 weeks (i) 81% (ii) 41%. No significance tests reported. No adverse experiences were reported by the patients.</td>
</tr>
</tbody>
</table>

* Results of the most important outcome measures according to the authors of the study. p Values < 0.05 were taken as significant.
Efficacy of NSAIDs for low back pain

This review shows some important methodological shortcomings in randomised trials evaluating the efficacy of NSAIDs in low back pain. The randomisation procedure was seldomly described, making it impossible for the reader of the article to discover if procedures were used that definitely excluded bias. Data on similarity of relevant baseline characteristics were often not presented, making it difficult to assess whether the study groups were sufficiently similar regarding their prognosis. Perhaps even more disturbing was the finding that number of drop outs and the reason for it were often not reported, while selective drop out of patients and loss to follow up may easily cause bias.

The small size of the study populations was also a commonly identified problem. For this reason, studies may lack the statistical power to detect clinically relevant differences in effects between the interventions under study, which of course only is a problem if pooling is not feasible. Another problem with smaller sample sizes is that important (un)known prognostic variables might not be in balance between the study groups after randomisation. Such situations may lead to biased outcomes if, by chance, patients in one group had a more favourable prognosis.

Another problem refers to the blinding of patients with respect to the interventions under study. Although it can be argued that patients will not be able to detect the content of the drug given, one should preferably evaluate whether the blinding was indeed successful by asking the patients to indicate which intervention they thought they had received.

The wide range of scores for methodological quality suggests that there is much room for improvement in future studies. It must be noted, however, that the reported methodological flaws are not unique for clinical trials evaluating the efficacy of NSAIDs. In general, the NSAIDs trials (median 48, range 27-83) seem to score somewhat higher than trials evaluating other interventions for low back pain. For example, trials evaluating the efficacy of spinal manipulation and mobilisation (median 35, range 20-56), exercise therapy (median 40, range 24-61) and back schools (median 36, range 16-70), traction therapy (median 36, range 23-66) all had a lower median methods score. It must be noted, however, that the methodological assessment was focused on the publication of the trial at issue. It might well be that the authors of a trial in fact conducted a high quality trial, meeting most of the criteria from table 1, but for some reason did not report the details in their article.

Another problem relating to the reporting of the trials is the clinical description of the study population. The descriptions and definitions of back pain (for example, acute and chronic, recurrence status, sciatica) varied widely among the studies included in this review. In some instances the definitions were not described at all. This situation hampers the interpretation of the study results. In future studies some standardisation of the description and classification of patients with low back pain (for example, the classification of the Quebec Task Force on Spinal Related Disorders) might be desirable.

Efficacy

The results of the 26 randomised trials that have been published to date, suggest that NSAIDs are effective for symptomatic short-term relief in patients with uncomplicated low back pain. The placebo controlled studies with methods scores above 50 points suggest that NSAIDs are effective in patients with (uncomplicated) low back pain, but are less effective or ineffective in patients with low back pain with sciatica and patients with sciatica with nerve root symptoms. The latter seems to be somewhat surprising because in patients with sciatica and nerve root symptoms some inflammation process is suggested to be part of the cause of the symptoms. One might have expected that NSAIDs would be effective in these patients because of the anti-inflammatory component of the drug. Whether NSAIDs are more effective than other (drug) therapies, including

![Figure 1: Pooled OR (Peto) in four placebo controlled trials evaluating NSAIDs for acute low back pain.](http://ard.bmj.com/)

Reference 38
Reference 21
Reference 47
Reference 32
Pooled OR

Favours NSAID
Favours placebo

| 0.1 | 1 | 10 |
---|---|---|

Discussion

METHODOLOGICAL QUALITY

This review shows some important methodological shortcomings in randomised trials evaluating the efficacy of NSAIDs in low back pain. The randomisation procedure was seldomly described, making it impossible for the reader of the article to discover if procedures were used that definitely excluded bias. Data on similarity of relevant baseline characteristics were often not presented, making it difficult to assess whether the study groups were sufficiently similar regarding their prognosis. Perhaps even more disturbing was the finding that number of drop outs and the reason for it were often not reported, while selective drop out of patients and loss to follow up may easily cause bias.

The small size of the study populations was also a commonly identified problem. For this reason, studies may lack the statistical power to detect clinically relevant differences in effects between the interventions under study, which of course only is a problem if pooling is not feasible. Another problem with smaller sample sizes is that important (un)known prognostic variables might not be in balance between the study groups after randomisation. Such situations may lead to biased outcomes if, by chance, patients in one group had a more favourable prognosis.

Another problem refers to the blinding of patients with respect to the interventions under study. Although it can be argued that patients will not be able to detect the content of the drug given, one should preferably evaluate whether the blinding was indeed successful by asking the patients to indicate which intervention they thought they had received.

The wide range of scores for methodological quality suggests that there is much room for improvement in future studies. It must be noted, however, that the reported methodological flaws are not unique for clinical trials evaluating the efficacy of NSAIDs. In general, the NSAIDs trials (median 48, range 27-83) seem to score somewhat higher than trials evaluating other interventions for low back pain. For example, trials evaluating the efficacy of spinal manipulation and mobilisation (median 35, range 20-56), exercise therapy (median 40, range 24-61) and back schools (median 36, range 16-70), traction therapy (median 36, range 23-66) all had a lower median methods score. It must be noted, however, that the methodological assessment was focused on the publication of the trial at issue. It might well be that the authors of a trial in fact conducted a high quality trial, meeting most of the criteria from table 1, but for some reason did not report the details in their article.

Another problem relating to the reporting of the trials is the clinical description of the study population. The descriptions and definitions of back pain (for example, acute and chronic, recurrence status, sciatica) varied widely among the studies included in this review. In some instances the definitions were not described at all. This situation hampers the interpretation of the study results. In future studies some standardisation of the description and classification of patients with low back pain (for example, the classification of the Quebec Task Force on Spinal Related Disorders) might be desirable.

Efficacy

The results of the 26 randomised trials that have been published to date, suggest that NSAIDs are effective for symptomatic short-term relief in patients with uncomplicated low back pain. The placebo controlled studies with methods scores above 50 points suggest that NSAIDs are effective in patients with (uncomplicated) low back pain, but are less effective or ineffective in patients with low back pain with sciatica and patients with sciatica with nerve root symptoms. The latter seems to be somewhat surprising because in patients with sciatica and nerve root symptoms some inflammation process is suggested to be part of the cause of the symptoms. One might have expected that NSAIDs would be effective in these patients because of the anti-inflammatory component of the drug. Whether NSAIDs are more effective than other (drug) therapies, including

![Figure 1: Pooled OR (Peto) in four placebo controlled trials evaluating NSAIDs for acute low back pain.](http://ard.bmj.com/)

Reference 38
Reference 21
Reference 47
Reference 32
Pooled OR

Favours NSAID
Favours placebo

| 0.1 | 1 | 10 |
---|---|---|
The pooling of results from individual trials was confined to the placebo controlled trials on acute low back pain only. The trials investigating other contrasts were considered to be too heterogeneous regarding methodological quality, patients' characteristics, contrasts under study, and outcome measurements, to allow pooling of their results. The meta-analyses of the results after one week indicated a pooled OR of 0.53 using the fixed effects model, indicating that NSAIDs were significantly more effective than placebo. The random effects model resulted in more or less the same point estimates with wider confidence intervals so that the point estimates reached borderline significance only. Given the last remark and because the sensitivity analysis indicated that the weaker trials reported larger effects the positive short-term effects of NSAIDs must be viewed with some caution. Caution is also indicated because the outcome measures in the pooled analysis were not identical. All four outcome measures consisted of a global (subjective) assessment of the clinical progress of the patient measured on an ordinal scale. In all four studies we were able to dichotomise the outcome measures into 'successes' (for example, complete relief of pain, noticeable improvement, definitive positive effect) and 'failures' (for example, slight improvement, no chance/improvement, worse). However, whether these outcome measures are similar enough to permit statistical pooling remains an arbitrarily judgement. The outcomes after two weeks were more or less similar to those after one week. Again significant positive results were found for NSAIDs compared with placebo with the fixed effect model, however, the more conservative random effect model resulted in non-significant findings.

SIDE EFFECTS
Numerous articles have reported on the side effects of NSAIDs, especially gastrointestinal events. In the studies presented in this review, side effects were also frequently reported, including abdominal pain, diarrhoea, oedema, dry mouth, rash, dizziness, headache, tiredness, etc. Most side effects were considered to be mild to moderately severe according to the authors of the studies. There seemed to be no clear difference in the reported number or severity, or both, of side effects between the different types of NSAIDs in the studies included in this review. However, the sample sizes of the studies, in general, were relatively low, permitting an inaccurate estimate of side effects only. Therefore, from the trials described in this review no clear conclusion can be drawn regarding the risks for gastrointestinal and other side effects when using NSAIDs.

LIMITATIONS
There are certain limitations to the methods used in this systematic review. Publication bias cannot be ruled out, so it is possible that trials that were not published because of their (negative) results were missed. As we, for practical reasons, included English language papers only, there might also be a possibility for language bias, in the sense that perhaps the results of trials published in other languages might systematically differ from trials published in the English literature. Furthermore, the two independent reviewers were not blinded with respect to the source and outcome of the trials. However, the methodological criteria were quite strict and easy to apply. These criteria have been used for a number of reviews on conservative interventions for low back pain. In addition, a recent meta-analysis of spinal manipulation for low back pain demonstrated that the results of our scoring method were similar to results obtained by the scoring method of Chalmers et al. One of the drawbacks of using this list of methodological criteria might be that trials showing a ‘fatal mistake’ (for example, irrelevant outcome measures, drop out rate exceeding 50%) may end up with a comparatively high score because they meet most of the other criteria. Studies with the highest methods scores should therefore be checked regarding such fatal flaws. No ‘fatal flaws’ were identified in the best studies (methods scores more than 60 points) in this review.

In conclusion, there are flaws in the design of most studies. The results of the 26 randomised trials that have been carried out to date, suggest that NSAIDs might be effective for short-term symptomatic relief in patients with uncomplicated low back pain, but are less effective or ineffective in patients with low back pain with sciatica and patients with sciatica with nerve root symptoms.

This study was supported by a grant from the Dutch Health Insurance Executive Board.

Appendix

Explanation of the criteria from table 1. Each criterion must be applied independently of the other criteria.

A Description of inclusion and exclusion criteria (1 point). Restriction to a homogeneous study population (1 point).
B Similarity for: duration of complaints, value of outcome measures, age, recurrence status, and radiating complaints (1 point each).
C Randomisation procedure described (2 points).
D Information from which group and with reason for withdrawal.
E Loss to follow up: all randomised patients minus the number of patients at main moment of effect measurement for the main outcome measure, divided by all randomised patients times 100.
F Smallest group immediately after randomisation.
G NSAID therapy explicitly described (5 points).
H Comparison with an existing treatment modality.
I Other medical interventions are avoided in the design of the study (except analgesics, advice on posture or use at home of heat, rest, or a routine exercise scheme).
J Comparison with a placebo therapy.
K Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
L Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
M Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
N Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
O Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
P Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
Q Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
R Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
S Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
T Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
U Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
V Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
W Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
X Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
Y Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
Z Placebo controlled: attempt of blinding (3 points), blinding evaluated and fully successful (2 points).
Efficacy of NSAIDs for low back pain

Efficacy of non-steroidal anti-inflammatory drugs for low back pain: a systematic review of randomised clinical trials
Bart W Koes, Rob J P M Scholten, Jan M A Mens and Lex M Bouter

doi: 10.1136/ard.56.4.214

Updated information and services can be found at:
http://ard.bmj.com/content/56/4/214

These include:

References
This article cites 41 articles, 3 of which you can access for free at:
http://ard.bmj.com/content/56/4/214#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/