Nonsteroidal Antiinflammatory Drugs for Axial Spondyloarthritis: A Cochrane Review

Féline P.B. Kroon, Lennart R.A. van der Burg, Sofia Ramiro, Robert B.M. Landewé, Rachelle Buchbinder, Louise Falzon, and Désirée van der Heijde

ABSTRACT. Objective. To determine the benefits and harms of nonsteroidal antiinflammatory drugs (NSAID) in axial spondyloarthritis (axSpA).

Methods. Systematic review using Cochrane Collaboration methodology. Inclusion criteria: randomized controlled trials (RCT) and quasi-RCT (to June 2014), investigating NSAID versus any control for axSpA, and observational studies of longterm effects (≥ 6 mos) of NSAID on radiographic progression or adverse events. Main outcomes were pain, Bath Ankylosing Spondylitis Disease Activity Index, Bath Ankylosing Spondylitis Functional Index, Bath Ankylosing Spondylitis Metrology Index, radiographic progression, number of withdrawals because of adverse events, and number of serious adverse events. Risk of bias was assessed.

Results. Thirty-five RCT, 2 quasi-RCT, and 2 cohort studies were included. Twenty-nine RCT and 2 quasi-RCT (n = 4356) were included in pooled analyses [traditional NSAID vs placebo (n = 5), cyclooxygenase-2 (COX-2) vs placebo (n = 3), COX-2 vs traditional NSAID (n = 4), NSAID vs NSAID (n = 24), naproxen vs other NSAID (n = 3), and low- vs high-dose NSAID (n = 5)]. Compared with placebo, both traditional and COX-2 NSAID were consistently more efficacious at 6 weeks and equally safe after 12 weeks. No significant differences in benefits or harms between the 2 NSAID classes and no important differences in benefits or withdrawals because of adverse events between different NSAID were found, especially if studies with high risk of bias were excluded. Single studies suggest NSAID may retard radiographic progression, especially by continuous rather than on-demand NSAID use.

Conclusion. High-quality evidence indicates that both traditional and COX-2 NSAID are efficacious for treating axSpA, and harms are not different from placebo in the short term. Various NSAID are equally effective. (First Release February 1 2016; J Rheumatol 2016;43:607–17; doi:10.3899/jrheum.1507271)

Key Indexing Terms: NONSTEROIDAL ANTIINFLAMMATORY DRUGS ANKYLOSING SPONDYLITIS AXIAL SPONDYLOARTHRITIS SYSTEMATIC REVIEW METAANALYSIS

Spondyloarthritis (SpA) is an umbrella term consisting of ankylosing spondylitis (AS), psoriatic arthritis, arthritis/spondyloarthritis with inflammatory bowel disease, and reactive arthritis. Patients with typical SpA features that do not fulfill the criteria for 1 of these subgroups have been incorporated in the SpA concept as undifferentiated SpA. More recently, patients with SpA are distinguished according to their clinical presentation as patients with either predominantly peripheral (including peripheral arthritis, enthesitis, and/or dactylitis) or axial SpA (axSpA); involvement of the
sacroiliac joints and/or the spine). Classification criteria for axSpA distinguish between nonradiographic axSpA (nr-axSpA; i.e., not having established radiographic changes in the sacroiliac joint) and radiographic axSpA or AS (i.e., presence of radiographic changes in the sacroiliac joint)\cite{6,7}. The prevalence of axSpA in Western European countries is between 0.3\% and 2.5\% and the prevalence rate of AS is up to 0.53\% in Western countries\cite{8}.

Nonsteroidal antiinflammatory drugs (NSAID), including traditional NSAID and selective cyclooxygenase inhibitors (COX-2; COXIB), are recommended as first-line drug treatment for patients with axSpA with pain and stiffness\cite{9}. While continuous NSAID treatment is favored for patients with persistently active, symptomatic disease, certain cardiovascular (CV), gastrointestinal (GI), and renal risks should be taken into account\cite{9}. NSAID are associated with a variety of GI toxicities\cite{10,11,12,13,14,15,16}, increased risk of CV events\cite{17,18,19}, development of congestive heart failure\cite{20}, either reversible or permanent renal toxicity, and a variety of damage to electrolyte and water homeostasis\cite{21}. It is therefore crucial to know whether the benefits offset the risks, especially because the therapy is often given for extended periods of time.

We performed a Cochrane systematic review to synthesize the available evidence assessing the benefits and harms of NSAID in controlling symptoms, disease activity, and radiographic progression in patients with axSpA.

**MATERIALS AND METHODS**

This paper is a shortened co-publication of a Cochrane review\cite{22}. A more detailed description of the methodology can be found in the original publication.

**Inclusion and exclusion criteria.** We considered all randomized controlled trials (RCT) and quasi-RCT (i.e., where allocation was not truly random) without language restrictions that were available as a full trial report. We included trials of adults with axSpA, as determined by clinical diagnosis or fulfillment of the modified New York criteria or the Assessment of SpondyloArthritis international Society (ASAS) axSpA criteria, including nr-axSpA and AS. Studies containing patients with other diagnoses (for example, trials that included participants based upon fulfillment of the Amor or European Spondyloarthropathy Study Group criteria) were only eligible if the results from patients with axSpA were presented separately\cite{12,14}.

We included studies comparing NSAID in all possible variations (dosage, intensity, mode, duration or timing of delivery, traditional, and COX-2 selective) to placebo, no therapy, another NSAID, other pharmacological therapy, nonpharmacological therapy, combination therapy, different doses or modes of delivery, or frequency or duration.

Because radiographic progression and long-term safety are unlikely to be assessed in short-term RCT, we also included observational cohort studies to investigate the effect of NSAID on these specific outcomes. Cohort studies assessing radiographic progression had to have a minimum duration of 6 months to be included.

**Search strategy.** We searched the following databases up to June 2014: MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials, as well as additional resources including the Database of Abstracts of Review of Effects, Scopus for conference proceedings, and clinical trial registries for ongoing and recently finished studies (clinicaltrials.gov and the World Health Organization International Clinical Trials Registry Platform).

**Study selection and data collection.** Two review authors (FK, LvdB) independently screened retrieved titles and abstracts, and full-text papers if necessary to determine inclusion. In case of nonconsensus, a third review author (SR) served as adjudicator. Data extraction was performed by the same authors using a standardized data extraction form. Raw data (i.e., means and SD for continuous outcomes and number of events for dichotomous outcomes) were extracted for outcomes of interest.

**Assessment of risk of bias in included studies.** Two review authors (FK, LvdB) independently assessed risk of bias of each included RCT with regard to random sequence generation, allocation concealment, blinding (of participants, care provider, and outcome assessor), incomplete outcome data, selective outcome reporting, and other sources of bias according to the Cochrane risk of bias tool\cite{23}. Each criterion was judged as “low risk of bias,” “high risk of bias,” or “unclear” (either lack of information or uncertainty over the potential for bias). The same authors also independently assessed risk of bias of each included observational study regarding study participation, study attrition, prognostic factor measurement, outcome measurement, confounding measurement and account, and analysis\cite{24}.

**Outcome measures.** We included outcomes at the latest followup in each trial. Main efficacy outcomes were (1) pain [on a visual analog scale (VAS) or numerical rating scale; back pain was used, but if not present in a study, overall pain was used], (2) disease activity assessed by the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)\cite{25}, (3) physical function assessed by the Bath Ankylosing Spondylitis Functional Index (BASFI)\cite{26}, (4) spinal mobility assessed by the Bath Ankylosing Spondylitis Metrology Index (BASMI)\cite{27}, and (5) radiographic progression assessed by the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS)\cite{28}. Main safety outcomes were total number of withdrawals because of adverse events, and number of serious adverse events.

Secondary outcomes included disease activity, fulfillment of response criteria, spinal mobility, and proportion of patients reporting pain relief of 50\% or greater. Disease activity was assessed by the Ankylosing Spondylitis Disease Activity Index\cite{29}, patient’s global assessment of disease activity, duration and severity of morning stiffness, C-reactive protein, and erythrocyte sedimentation rate. Two response criteria were assessed: the ASAS20 response criteria\cite{30} and the ASAS partial remission criteria\cite{31}. Spinal mobility was assessed by lateral spinal flexion, chest expansion, tragus-to-wall distance, occiput-to-wall distance, intermalleolar distance, and modified Schober test\cite{32}.

Only radiographic and long-term safety outcomes were extracted from the observational studies.

**Data analysis.** Comparisons of traditional NSAID versus placebo and COX-2 NSAID versus placebo were deemed the most important comparisons. Additional comparisons were COX-2 versus traditional NSAID, 1 NSAID versus another, lower versus higher dose NSAID, and continuous versus on-demand use. Metaanalysis was only performed if the studies were clinically and statistically sufficiently homogeneous. Clinical homogeneity was assessed with respect to intervention and control groups, outcome measure, and timing of assessment. The I² statistic was used to test for statistical heterogeneity, interpreted in accordance with the Cochrane Handbook for Systematic Reviews\cite{33}. We used a random-effects model as the default option to be conservative, independent of the I². We calculated mean differences (MD) for continuous outcomes and risk ratios (RR) for dichotomous data, both with corresponding 95\% CI. For studies containing more than 2 intervention groups (e.g., group A, B, and C), we included the same group of participants only once in the metaanalysis (e.g., group A vs group C, or group B vs group C, or a combination of groups A and B vs group C). In case 2 comparisons were deemed necessary (e.g., group A vs group C and group B vs group C), we split the number of participants of the group with the “shared intervention” (group C) into 2 equally large groups. Whenever we had to decide between multiple dosages of an NSAID for studies containing more than 2 intervention groups, we used the proposed equivalent dose of 150 mg diclofenac as defined by ASAS\cite{34}.

Preplanned sensitivity analyses were performed to assess the effect of including trials with high or unclear risk of bias in all significant comparisons where sufficient studies existed.
Summary of findings tables. The main results of the review are presented in summary of findings tables, including an overall grading of the evidence using the Grading of Recommendations Assessment, Development, and Evaluation approach. In these tables, we provided the absolute percent difference, the relative percent change from baseline, and the number needed to treat (NNT; only when the outcome showed a statistically significant difference). The NNT for continuous outcomes was calculated using the Well’s calculator software.

RESULTS
Description of studies. A detailed description of the search results and characteristics of included studies can be found in the original publication. We initially identified 7883 records; 177 qualified for full review and 39 were finally included in this review (Figure 1).

Randomized and quasi-randomized studies. Thirty-five RCT and 2 quasi-RCT involving 4908 participants (range 14–611, mean 133), published between 1966 and 2006, were included. Twenty-four studies (65%) were published before 1990. The mean age of participants (reported in 26/37 trials) was 40.5 years (SD 11.1) and 81% were men (reported in 36 studies). Treatment duration ranged from 1 week to 2 years, with a median duration of 12 weeks.

Cohort studies. One retrospective cohort study published in 1976 (n = 40, variable followup up to 20 yrs) and 1 prospective cohort study published in 2012 (n = 174, followup 2 yrs) were included.

Risk of bias. A detailed description of the risk of bias for each of the included studies is presented in the original publication. Most trials (n = 29) were at unclear risk of selection bias, although blinding of participants and personnel was adequate in 24 trials. Twenty-five trials were at low risk of attrition bias and 29 trials had a low risk of reporting bias. Risk of bias in both cohort studies was judged high for study participation, and low or unclear for all other criteria.

Effects of interventions. Thirty-one trials (n = 4356 participants) contributed to the pooled analyses [traditional NSAID vs placebo (n = 5), COX-2 vs placebo (n = 3), COX-2 vs traditional NSAID (n = 4)].

Figure 1. Flow chart.
Table 1. Summary of findings: traditional NSAID compared with placebo for axial spondyloarthritis.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect, Risk Ratio (95% CI)</th>
<th>No. Participants, Studies</th>
<th>Quality of the Evidence, GRADE#</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain on VAS, scale 0–100 mm, higher is worse, followup: 2 to 6 weeks</td>
<td>The mean pain score in the control group was 61 points**</td>
<td>The mean pain scores in the intervention groups were 16.5 points lower (12.2 to 20.8)</td>
<td>850 (4 studies)</td>
<td>High</td>
<td>Absolute percent difference: 17% lower (12% to 21%). Relative percent change from baseline: 21% lower (16% to 27%)<em>. NNT: 4 (3 to 6)</em>.</td>
</tr>
<tr>
<td>Withdrawals because of adverse events, followup: 2 to 12 weeks</td>
<td>52 per 1000††</td>
<td>39 per 1000 (24 to 63)</td>
<td>1165 (5 studies)</td>
<td>High</td>
<td>Absolute percent difference: 0% more (3% less to 2% more). Relative percent difference from baseline: 54% decrease (54% decrease to 21% increase). NNT: 3 (2 to 4)*†.</td>
</tr>
<tr>
<td>BASDAI, scale 0–100, higher is worse, followup: 6 weeks</td>
<td>The mean BASDAI in the control group was 54.7 points</td>
<td>The mean BASDAI in the intervention group was 17.5 points lower (11.8 to 23.1)</td>
<td>190 (1 study)</td>
<td>Moderate†††</td>
<td>Absolute percent difference: 18% lower (12% to 23%). Relative percent change from baseline: 28% lower (19% to 37%). NNT: 3 (2 to 4)*†.</td>
</tr>
<tr>
<td>BASFI, scale 0–100, higher is worse, followup: 6 weeks</td>
<td>The mean BASFI in the control group was 50.0 points†††</td>
<td>The mean BASFI in the intervention group was 9.1 points lower (5.1 to 13.0)</td>
<td>356 (2 studies)</td>
<td>High</td>
<td>Absolute percent difference: 9% lower (5% to 13%). Relative percent change from baseline: 17% lower (9% to 24%)<em>†. NNT: 5 (3 to 8)</em>†.</td>
</tr>
<tr>
<td>BASMI, scale 0–10, higher is worse</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>None of the trials included in this comparison reported BASMI.</td>
</tr>
<tr>
<td>Radiographic progression, mean change in mSASSS, higher is worse</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>None of the trials included in this comparison reported mSASSS.</td>
</tr>
<tr>
<td>No. serious adverse events, followup: 6 to 12 weeks</td>
<td>2 per 1000§</td>
<td>3 per 1000 (1 to 16)</td>
<td>671 (3 studies)</td>
<td>Moderate§§</td>
<td>Absolute percent difference: 0% more (1% less to 2% more). Relative percent change from baseline: 69% increase (64% decrease to 697% increase).</td>
</tr>
</tbody>
</table>

* The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). # GRADE Working Group grades of evidence: high quality (further research is very unlikely to change our confidence in the estimate of effect), moderate quality (further research is likely to have an important effect on our confidence in the estimate of effect and may change the estimate), low quality (further research is very likely to have an important effect on our confidence in the estimate of effect and is likely to change the estimate), very low quality (we are very uncertain about the estimate). ** Assumed risk based on mean control group final values from Dougdos, et al80 and van der Heijde, et al40. *** Estimated relative changes based on mean (SD) pain on VAS in placebo group at baseline 77.22 (15.24) from van der Heijde, et al40. † Based on MCID of 15 points on a 0–100 point scale. †† Based on MCID of 15 points on a 0–100 point scale. ††† Potential imprecision because of data available only from a single study (n = 190). ‡ Estimated relative changes based on mean (SD) BASDAI in placebo group at baseline 61.78 (18.70) from van der Heijde, et al40. §§ Based on MCID of 10 points on a 0–100 point scale. §§§ Assumed risk based on the control group final values from van der Heijde, et al40. §§§§ Potential imprecision because the 95% CI includes “no effect” and the upper confidence limit also crosses “appreciable harm.” NSAID: nonsteroidal antiinflammatory drug; GRADE: Grading of Recommendations Assessment, Development, and Evaluation; VAS: visual analog scale; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; mSASSS: modified Stoke Ankylosing Spondylitis Spinal Score; MCID: minimal clinically important difference; NA: not applicable.
Table 2. Summary of findings: COX-2 NSAID compared with placebo for axial spondyloarthritis.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect</th>
<th>No. Participants, Studies</th>
<th>Quality of the Evidence, GRADE#</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain on VAS, scale 0–100 mm, higher is worse, followup: 6 weeks</td>
<td>The mean pain scores across control groups was 64 points**</td>
<td>2.14 (0.36–12.56)</td>
<td>669 (3 studies)</td>
<td>Low+++</td>
<td>Absolute percent difference: 22% lower (7% to 36%). Relative percent change from baseline: 28% lower (10% to 47%)**. NNT: 3 (2 to 24)1.</td>
</tr>
<tr>
<td>Withdrawals because of adverse events, followup: 6 to 12 weeks</td>
<td>11 per 1000 ††</td>
<td>24 per 1000 (4 to 142)</td>
<td>2.14 (0.36–12.56)</td>
<td>669 (3 studies)</td>
<td>Low+++</td>
</tr>
<tr>
<td>BASDAI, scale 0–100, higher is worse, followup: 6 weeks</td>
<td>The mean BASDAI in the control group was 54.7 points</td>
<td>193 (1 study)</td>
<td>Moderate²</td>
<td>**</td>
<td>Absolute percent difference: 22% lower (17% to 27%). Relative percent change from baseline: 36% lower (27% to 44%)². NNT: 2 (1 to 3)².</td>
</tr>
<tr>
<td>BASFI, scale 0–100, higher is worse, followup: 6 weeks</td>
<td>The mean BASFI in the control group was 50.0 points**</td>
<td>349 (2 studies)</td>
<td>High</td>
<td>**</td>
<td>Absolute percent difference: 13% lower (9% to 17%). Relative percent change from baseline: 25% lower (18% to 32%)³. NNT: 3 (2 to 4)³.</td>
</tr>
<tr>
<td>BASMI, scale 0–10, higher is worse</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>None of the trials included in this comparison reported BASMI.</td>
</tr>
<tr>
<td>Radiographic progression, mean change in mSASSS, higher is worse</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>None of the trials included in this comparison reported mSASSS.</td>
</tr>
<tr>
<td>No. serious adverse events, followup: 6 to 12 weeks</td>
<td>2 per 1000 ††</td>
<td>2 per 1000 (0 to 13)</td>
<td>0.92 (0.14–6.21)</td>
<td>669 (3 studies)</td>
<td>Moderate⁵</td>
</tr>
</tbody>
</table>

* The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). ☀ GRADE Working Group grades of evidence: high quality (further research is very unlikely to change our confidence in the estimate of effect), moderate quality (further research is likely to have an important effect on our confidence in the estimate of effect and may change the estimate), low quality (further research is very likely to have an important effect on our confidence in the estimate of effect and is likely to change the estimate), and very low quality (we are very uncertain about the estimate). ** Assumed risk based on the control group final values from van der Heijde, et al⁸⁰. *** Estimated relative changes based on mean (SD) pain on VAS in placebo group at baseline 77.22 (15.24) from van der Heijde, et al⁸⁰. † Based on MCID of 15 points on 0–100 point scale. †† Assumed risk based on the mean risk in the control groups. ††† Potential imprecision because of the 95% CI includes “no effect” and the upper confidence limit also crosses “appreciable harm,” as well as inconsistency in the results with large heterogeneity (I² = 84%). †‡ Potential imprecision because of data available only from a single study (n = 193). †‡‡ Estimated relative changes based on mean (SD) BASDAI in placebo group at baseline 61.78 (18.70) from van der Heijde, et al⁸⁰. †‡§ Based on MCID of 10 points on a 0–100 point scale. †‡‖ Estimated relative changes based on mean (SD) BASFI in placebo group at baseline 54.12 (26.99) from van der Heijde, et al⁸⁰. †‡¶ Based on MCID of 10 points on a 0–100 point scale. †‡‖‖ Estimated relative changes based on mean (SD) BASMI in placebo group at baseline 54.12 (26.99) from van der Heijde, et al⁸⁰. †‡‖‖‖ Based on MCID of 10 points on a 0–100 point scale. †‡‖‖‖‖ Potential imprecision because the 95% CI includes “no effect” and the upper confidence limit also crosses “appreciable harm.” COX-2: cyclooxygenase-2; NSAID: nonsteroidal antiinflammatory drug; VAS: visual analog scale; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; mSASSS: modified Stoke Ankylosing Spondylitis Spinal Score; NNT: number needed to treat; MCID: minimal clinically important difference; NA: not applicable; GRADE: Grading of Recommendations Assessment, Development, and Evaluation.

In the trials: pain on VAS (2 trials, n = 349, MD –21.68, 95% CI –35.94 to –7.42; Figure 3A), BASDAI (1 trial, n = 193, MD –22.00, 95% CI –27.44 to –16.56), and BASFI (2 trials, n = 349, MD –13.42, 95% CI –17.35 to –9.49). No studies reported data for BASMI or radiographic progression, our other main efficacy outcomes. There were no between-group differences in the number of withdrawals because of adverse events (Figure 3B) and number of (serious or any) adverse events. Similar to traditional NSAID, there were more GI adverse events in patients taking COXIB compared with placebo (3 studies, n = 669, RR 1.80, 95% CI 1.22–2.67). There were no between-group adverse events. However, 5 trials (n = 1289) found more GI adverse events in patients taking NSAID compared with placebo (RR 1.92, 95% CI 1.41–2.61), and 4 studies (n = 1144) found fewer neurological adverse events (including headache and dizziness) in the NSAID group compared with the placebo group (RR 0.44, 95% CI 0.24–0.82). There were no between-group differences in the number of respiratory, hematomatologic, or dermatological adverse events. COX-2 NSAID versus placebo. Three studies (n = 669, duration 6 to 12 weeks) provided data for this comparison⁴¹,⁴⁸,⁸⁰. Significant effects favoring COX-2 NSAID over placebo were found for all main efficacy outcomes measured.
differences in the number of respiratory, neurological, or dermatological adverse events.

**COX-2 NSAID versus traditional NSAID.** There were 4 studies that compared COX-2 with traditional NSAID (n = 995). We found no between-group differences in any of the reported main efficacy (pain, BASDAI, BASFI, and BASMI) or safety outcomes.

**One NSAID versus another.** There were 24 trials that compared 1 NSAID to another (n = 2076). None of the NSAID performed consistently better than any other for any of the reported main efficacy or safety outcomes. However, based upon 11 studies (n = 1135), the use of indomethacin resulted in significantly more adverse events (RR 1.25, 95% CI 1.06–1.48), and based upon 9 trials (n = 963), indomethacin was associated with more neurological adverse events (such as headache and dizziness) than other NSAID (RR 2.34, 95% CI 1.32–4.14). Adverse events in the other organ systems that were assessed were not more prevalent in 1 NSAID versus another.

There were 3 trials that compared naproxen with other NSAID (n = 646). Based upon 2 trials (n = 232), naproxen performed significantly worse than other NSAID with respect to improving pain (MD 6.80, 95% CI 3.72–9.88), although no difference was found for any of the other reported main efficacy (BASDAI and BASFI) or safety outcomes.

**Lower versus higher dose NSAID.** Five trials (n = 1136) and 1 posthoc analysis of a prospective cohort study (n = 164) compared low versus a higher dose of an NSAID. No clear dose-response effect on benefits or harms was found in any of the trials. In the posthoc analysis of the cohort study, fewer participants with AS with a higher NSAID intake showed worsening of mSASSS score by 2 units or more compared with those with a low NSAID intake (OR 0.15, 95% CI 0.02–0.96). No such effect was found for patients with nr-axSpA.

**Continuous versus on-demand use.** One trial with a posthoc analysis (n = 214) and 1 retrospective cohort study (n = 612...
40) compared continuous with on-demand NSAID use. These studies suggest that NSAID may be effective in retarding radiographic progression in the spine in axSpA, especially in certain subgroups of patients, e.g., patients with high CRP, and that this may be best achieved by continuous rather than on-demand use of NSAID.

Sensitivity analyses. Results of all efficacy variables remained unchanged when excluding trials with high or unclear risk of bias. However, there were no statistically significant differences in safety when trials with a high or unclear risk of bias were excluded from the safety analyses.

DISCUSSION

Based upon moderate- to high-quality evidence, both traditional and COX-2 NSAID are more efficacious than and as safe as placebo for patients with axSpA in the short term (up to 12 weeks). An increase in GI adverse events that was initially observed for traditional as well as COX-2 NSAID in comparison with placebo was no longer significant when studies with high or unclear risk of bias were excluded. The results of our review are in keeping with current recommendations that NSAID are appropriate first-line treatments of patients with axSpA with active disease before tumor necrosis factor inhibitor biologicals are applied9. They are effective and safe in the short term.

Previous systematic reviews that have investigated the effects of NSAID for SpA have reported broadly similar findings as our review, although these reviews limited inclusion to placebo-controlled trials83,84,85. Surprisingly, we could not confirm in our review the safety concerns associated with both traditional NSAID and COX-2 NSAID that have been reported in these reviews84,85 as well as in studies in other rheumatic diseases86,87,88,89. The safety concern reported in those studies was that NSAID cause an increased risk of GI toxicity, at a level lower with COXIB but still considerable also in this class of NSAID. We did not find statistically significant differences in safety between traditional NSAID, COXIB, and placebo, which could mean that short-term use of either class of NSAID in this population of patients is not associated with an increased risk of GI or other adverse events. But studies were small (likelihood of Type II error), and the duration of the studies does not preclude adverse events occurring at a later stage. Clear differences in short-term adverse events can also be absent because most patients with AS are younger and may be “healthier” (i.e., have fewer comorbidities) than patients with other rheumatic diseases [such as rheumatoid arthritis]
This is supported by the finding that biologicals also result in fewer adverse events in patients with AS than in patients with other rheumatic diseases. Previous systematic reviews, in axSpA and in other diseases, also indicate an increased risk of CV toxicity, most importantly in COX-2 NSAID, which we could not confirm in our review. As mentioned, it is technically still possible that lack of statistical power is at the basis of this, but we feel it is more likely that in the studied population and within the studied time frame (i.e., short-term), the risks of GI or CV toxicity are really not increased. We did not find sufficient data to draw conclusions on longterm safety, therefore one still depends on studies in other rheumatic diseases regarding longterm safety of NSAID in axSpA.

We considered the benefits and harms of naproxen in comparison with other NSAID because a metaanalysis of vascular and upper GI effects of NSAID in various patients (prescribed mostly for RA or OA, but also for prevention of colorectal adenomas or of Alzheimer disease) showed that naproxen was associated with less vascular (but increased upper GI) risk than other NSAID. In our review, we found no important differences in safety between naproxen and other NSAID, although naproxen appeared to be less effective in relieving pain. However, few studies (n = 3) could be included in this comparison, and therefore we could neither confirm nor reject the results of Bhala, et al concerning the safety of naproxen.

In general, we found no clear dose effect on benefits or harms, although 1 posthoc analysis of a prospective cohort study suggested that higher NSAID intake may retard radiographic progression. This finding, although derived from studies that only compared a few different doses of a few NSAID (celecoxib, etoricoxib, meloxicam, and ximopron), suggests that it might be preferable to choose a lower NSAID dosage to minimize the risk of adverse events. However, ASAS members who are experts in the field have agreed to use relatively high dosages of NSAID to treat patients with axSpA (150 mg diclofenac, or an equivalent dose of another NSAID) based upon their experience in clinical practice. Further robust data are needed to resolve this issue.

We found a suggestion from single studies that NSAID may be effective in retarding radiographic progression of the spine in axSpA, especially in certain subgroups of patients, e.g., those with high CRP, and this may be best achieved by continuous rather than on-demand use of NSAID. These findings are in keeping with a recent study that found that high disease activity leads to more structural damage in the spine. It has also been shown that radiographic damage is associated with impaired spinal mobility and function. These findings stress the importance of retarding the progression of structural damage in the spine, and taking NSAID for a longer period may be an effective way to do so. However, the risk/benefit of this strategy requires confirmation in further rigorous longterm studies that also consider safety, and until these data are available, the potential benefits of continuous NSAID use should be considered in comparison to the potential risks in individual patients.

Our study has several strengths. We performed a systematic review and metaanalysis using Cochrane methods, with predefined outcomes and a published protocol, and assessed outcomes of relevance as recommended by ASAS. The risk that bias was introduced by the methods used is low because all authors strictly followed the protocol outlined in our review. However, our review has several limitations, including that many trials were older (61% of the included studies were published before 1990). Consequently, many studies did not include some of the outcomes we had prespecified and they also did not include participants diagnosed with nr-axSpA. Although we expect that our results will also apply to patients with nr-axSpA, this requires confirmation. In addition, the RCT provided limited data regarding the longterm effects of NSAID because the median duration of NSAID treatment was 12 weeks, therefore we were unable to draw conclusions regarding their longterm benefits or harms. We attempted to address this by including observational studies, but the 2 included cohort studies did not include longterm safety as one of their outcomes and no studies on longterm safety could be included.

ACKNOWLEDGMENT

We thank Renea Johnston of the Cochrane Musculoskeletal Group for suggestions regarding the protocol.

REFERENCES


25. Kroon, et al. NSAID for axSpA


