ABSTRACT. Objective. Using data from a randomized, double-blind, placebo-controlled study, we assessed the capacity of clinical and nonsteroidal antiinflammatory drug (NSAID)-sparing endpoints, alone and in combination, to discriminate between treatment effects in axial spondyloarthritis (axSpA).

Methods. Patients with active NSAID-resistant axSpA received etanercept (ETN) 50 mg/week or placebo for 8 weeks and tapered/discontinued NSAID. In posthoc logistic regression analyses, OR were calculated that indicated the capacity of the following endpoints to discriminate between the effects of ETN and placebo at Week 8: Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) 50; BASDAI ≤ 3; Assessment of Spondylarthropathy international Society (ASAS) 20; ASAS40; Ankylosing Spondylitis Disease Activity Score (ASDAS) with C-reactive protein (CRP) < 1.3 and ASDAS-CRP < 2.1; ≥ 50% decrease from baseline in ASAS-NSAID score, score < 10, and score = 0; and each clinical and/or each NSAID measure.

Results. In 90 randomized patients (ETN, n = 42; placebo, n = 48), disease activity was similar between groups at baseline: mean (± SD) BASDAI (ETN vs placebo) 6.0 ± 1.6 versus 5.9 ± 1.5. NSAID intake was high: ASAS-NSAID score 98.2 ± 39.0 versus 93.0 ± 23.4. OR ranged from 1.6 (95% CI 0.5–5.4) for ASDAS-CRP < 1.3 to 5.8 (95% CI 1.2–29.1) for BASDAI50 and NSAID score of 0; most measures (34/45) reached statistical significance (α = 0.05) favoring ETN. Most combined outcome variables using OR were more discriminant than single outcome measures.

Conclusion. These findings suggest that changes in NSAID intake during treatment do not prevent demonstration of clinically relevant effects of biologic treatment, and combined (i.e., clinical with NSAID-sparing) endpoints were frequently more discriminant than single (i.e., clinical) endpoints. ClinicalTrials.gov (NCT01298531). (First Release November 15 2015; J Rheumatol 2015;42:2361–8; doi:10.3899/jrheum.150378)
cardiovascular system10,11. For diseases primarily treated with drugs that are clinically effective but pose a potentially serious safety risk, newer therapies are often introduced to provide equivalent or greater symptomatic improvement while “sparing” use of the potentially toxic conventional therapy. In rheumatic diseases, examples include NSAID-sparing symptomatic slow-acting drugs (e.g., chondroitin sulphate and glucosamine sulphate) in osteoarthritis12,13,14,15, corticosteroid-sparing methotrexate in polymyalgia rheumatica16, and corticosteroid-sparing biologics in rheumatoid arthritis17,18,19,20.

Treatment with antitumor necrosis factor (anti-TNF) biological agents is recommended in patients with definite axSpA who have high disease activity despite the use of at least 2 NSAID in the prior 4-week period21. Although these agents have an acceptable safety record, they are associated with a small but significant risk of serious infections22, particularly at high doses23. In clinical practice, beyond their clinical efficacy in NSAID-resistant patients, anti-TNF agents may reduce NSAID intake in patients with SpA and therefore decrease the risk of toxicity with longterm NSAID use. However, relatively few clinical trials have assessed the NSAID-sparing effects of such therapies24. In most clinical trials of anti-TNF therapy in axSpA, changes in disease activity measures [e.g., the Assessment of SpondyloArthritis international Society (ASAS) responder criteria] have been selected as primary endpoints, and the use of concomitant NSAID therapy has been either prohibited or required to remain stable during the blinded controlled phase of the study. Such a decision reflects the past and present fear that outcome measures used to assess disease activity would have less discriminant capacity if NSAID intake were substantially reduced in the active treatment arm versus the placebo arm. Moreover, the decision also likely reflects at least in part a lack of consensus on the appropriate methodology to use for the collection and reporting of concomitant therapy intake. Several different techniques have been proposed for the quantification and recording of NSAID intake, including the ASAS scoring system25.

In trials designed to evaluate treatments such as anti-TNF agents with possible NSAID-sparing effects, many questions remain about the discriminant capacity of outcome measures. Specifically, research has not yet shown whether clinical response criteria alone, NSAID response criteria alone, or a combination of such criteria might be more discriminant and therefore might result in a reduction of the number of patients required for inclusion in trials.

In an anti-TNF trial (the SPARSE study), the effect of treatment with the anti-TNF agent etanercept (ETN) on NSAID intake was examined using the latter ASAS-NSAID score25 in patients with axSpA who had been taking NSAID at baseline and were strongly advised to decrease and discontinue their NSAID use thereafter during an initial 8-week, double-blind, placebo-controlled treatment period. Primary and secondary outcomes of this study have been recently published elsewhere26. Using the SPARSE database, we performed posthoc, exploratory analyses to evaluate the capacity of conventional clinical outcome measures and NSAID-sparing outcome measures, assessed individually and in combination, to discriminate between the treatment effects of ETN and placebo.

MATERIALS AND METHODS

The following sections briefly summarize the methodology of the SPARSE study, which is described in detail in a previous publication26.

Study design. At screening, investigators instructed patients to discontinue their NSAID and restart the NSAID only in case of symptom flare, with treatment adjusted as needed to achieve optimal symptomatic control. Patients who experienced symptom flare after NSAID discontinuation and restarted NSAID treatment, and whose disease remained active, were randomized (1:1) to receive ETN 50 mg or placebo subcutaneously once weekly for 8 weeks, as well as their background NSAID as required. Investigators requested that patients taper and discontinue their NSAID intake during the randomized treatment period if clinically acceptable.

The SPARSE study was conducted in accordance with the International Conference on Harmonisation guidelines for good clinical practice and the Declaration of Helsinki. Study activities were not initiated until the Institutional Review Board approval and the patient-informed consent were obtained.

Patients. Eligible patients had axSpA according to the treating rheumatologist, with active axial involvement defined by a mini-Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) > 4 on a scale of 0 to 10, i.e., Q1 + Q2 + [(Q5 + Q6) / 2] + 3 267. An inadequate response was also required to at least 2 NSAID taken at maximum tolerated doses (based on medical history) for a total combined duration of at least 1 month. Patients were ineligible if they received prior treatment with a biologic agent or recent treatment with a corticosteroid or had uncontrolled inflammatory bowel disease or uveitis.

Clinical and NSAID-sparing outcome measures. The conventional clinical endpoints selected a priori for analysis, individually and in combination with the ASAS-NSAID score endpoints, included the BASDAI50 response and BASDAI ≤ 328. ASAS20 and ASAS40 responses29, and the Ankylosing Spondylitis Disease Activity Score (ASDAS) based on C-reactive protein (CRP) < 1.3 (inactive disease) and ASDAS-CRP ≥ 1.3 to < 2.1 (moderate disease activity state)30 at 8 weeks. The ASAS-NSAID score is based on the NSAID type, mean daily dose, and number of days with intake. The score was derived from data recorded on patient diary cards for the 7 days prior to the respective visit. Each daily dose of NSAID was converted to a percentage dose equivalent to 150 mg diclofenac. The daily doses were then totaled and the sum divided by the number of days in the period of interest. The minimum value was 0 (no NSAID intake) and a higher ASAS-NSAID value indicated greater NSAID consumption23. The binary ASAS-NSAID score endpoints chosen a priori as single outcomes and combined outcomes with clinical endpoints were ASAS-NSAID score of 0, ASAS-NSAID score < 10, and decrease in ASAS-NSAID score of ≥ 50% from baseline. Dichotomous (rather than continuous) outcomes were used in these analyses because they allowed for the combination of clinical and NSAID-sparing outcome measures and are more easily understood by clinicians.

Statistical analysis. All analyses were performed on the intent-to-treat (ITT) population, which included all randomized patients who received at least 1 dose of study medication. Baseline demographic and disease characteristics were summarized using descriptive statistics. The proportion of patients (95% CI) in the ETN 50 mg and placebo groups who achieved each conventional and NSAID-sparing endpoint at Week 8 was analyzed using logistic regression, with the corresponding baseline scores and treatment group included as covariates. For analyses of the NSAID-sparing endpoints,
when diary data were missing for a specific day, the missing data were counted as no intake. In addition, both a last observation carried forward method and a baseline observation carried forward approach (when no postbaseline diary data were available) were used. The estimated treatment difference (95% CI) was calculated as the difference between the proportion of patients treated with ETN who achieved the endpoint and the proportion of patients treated with placebo who achieved the endpoint.

To determine the capacity to discriminate between the treatment effects of ETN and placebo for each of the clinical and NSAID-sparing endpoints alone and in combination, OR (95% CI) were estimated from logistic regression models; the highest OR denoted the highest discriminant capacity in favor of ETN versus placebo.

RESULTS
Patients. Of 128 screened patients, 90 (ETN, n = 42; placebo, n = 48) were randomized into the 8-week, double-blind treatment period and included in the ITT population; 66 patients (ETN, n = 33; placebo, n = 33) completed the double-blind period. Patients in the ETN and placebo treatment groups had similar demographic and disease characteristics at baseline (Table 1). Of the 90 randomized patients, 51 (57%) had radiographic sacroilitis and 45 (50%) had sacroiliac joint inflammation on magnetic resonance imaging.

Clinical and NSAID-sparing effects. At Week 8, statistically significant between-group differences, favoring ETN over placebo, were found in the proportions of patients achieving BASDAI50, BASDAI ≤ 3, ASAS40, and ASDAS < 2.1, but not in the proportions achieving ASAS20 or ASDAS < 1.3 (Figure 1A). Significantly more patients receiving ETN than patients receiving placebo achieved each of the NSAID-sparing endpoints (i.e., an ASAS-NSAID score of 0, ASAS-NSAID score < 10, and a 50% reduction in ASAS-NSAID score). Statistically significant between-group differences were seen with 10 of 12 BASDAI combinations, 8 of 12 ASAS combinations, and 9 of 12 ASDAS combinations (Figure 1B–1D).

Discriminant capacity of clinical/NSAID-sparing outcome measures. The OR for the discriminant capacity of the clinical and NSAID-sparing outcome measures, alone and in combination, are shown in Figure 2. The majority of outcome measures (34 of 45) achieved statistical significance (α = 0.05) in discriminating between the treatment effects of ETN and placebo. Observed treatment effects for clinical and NSAID-sparing measures individually and in combination ranged from an OR of 1.6 (95% CI 0.5, 5.4) for ASDAS < 1.3 alone to 5.8 (95% CI 1.2, 29.1) for BASDAI50 and NSAID score of 0. Combined outcome variables (i.e., clinical + NSAID-sparing) were found to be at least as discriminant as single (i.e., clinical) outcome variables; the majority of combined outcome variables using OR were more discriminant than single outcome measures.

DISCUSSION
In our posthoc analysis of the SPARSE study, we found that clinical and NSAID-sparing outcome measures, individually

Table 1. Baseline demographics and disease characteristics. Safety population. Values are mean (SD) unless otherwise specified.

<table>
<thead>
<tr>
<th>Variables</th>
<th>ETN, 50 mg, n = 42</th>
<th>Placebo, n = 48</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient and disease characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, yrs</td>
<td>38.8 (12.3)</td>
<td>38.9 (11.4)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>18 (42.9)</td>
<td>16 (33.3)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>25.7 (4.8)</td>
<td>25.9 (4.9)</td>
</tr>
<tr>
<td>HLA-B27–positive, n (%)</td>
<td>28 (66.7)</td>
<td>31 (64.6)</td>
</tr>
<tr>
<td>Duration since diagnosis of axSpA, yrs</td>
<td>6.0 (9.0)</td>
<td>5.5 (7.4)</td>
</tr>
<tr>
<td>Positive pelvic radiograph*, n (%)</td>
<td>24 (57.1)</td>
<td>27 (56.3)</td>
</tr>
<tr>
<td>MRI sacroilitis-positive**, n (%)</td>
<td>21 (50.0)</td>
<td>24 (50.0)</td>
</tr>
<tr>
<td>ASAS axSpA criteria, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical arm</td>
<td>7 (16.7)</td>
<td>8 (16.7)</td>
</tr>
<tr>
<td>Imaging arm</td>
<td>32 (76.2)</td>
<td>37 (77.1)</td>
</tr>
<tr>
<td><strong>NSAID intake</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAS-NSAID score†</td>
<td>98.2 (39.0)</td>
<td>93.0 (23.4)</td>
</tr>
<tr>
<td><strong>Disease activity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASDAI, 0–100</td>
<td>6.0 (1.6)</td>
<td>5.9 (1.5)</td>
</tr>
<tr>
<td>ASDAS</td>
<td>3.4 (0.9)</td>
<td>3.2 (0.8)</td>
</tr>
<tr>
<td>ASDAS disease state, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inactive disease</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Moderate disease activity</td>
<td>5 (11.9)</td>
<td>3 (7.0)</td>
</tr>
<tr>
<td>High disease activity</td>
<td>19 (45.2)</td>
<td>23 (53.5)</td>
</tr>
<tr>
<td>Very high disease activity</td>
<td>18 (42.9)</td>
<td>17 (39.5)</td>
</tr>
<tr>
<td>CRP level, mg/dl</td>
<td>1.0 (1.3)</td>
<td>0.9 (1.4)</td>
</tr>
<tr>
<td>Abnormal CRP level, n (%)</td>
<td>21 (50.0)</td>
<td>15† (34.9)</td>
</tr>
</tbody>
</table>

* Grade ≥ 3 unilaterally or grade ≥ 2 bilaterally based on 1984 modified New York criteria for radiographic axSpA. ** According to local rheumatologist or radiologist. † Last observation carried forward method, with imputation, intent-to-treat population. ‡ Fifteen of 43 patients in placebo/ETN 50 mg group with CRP levels available at baseline. Abnormal CRP = > 1.25 × the upper limit of normal (4.9 mg/l). Inactive disease = ASDAS < 1.3, moderate disease activity = 1.3 ≤ ASDAS < 2.1, high disease activity = 2.1 ≤ ASDAS < 3.5, and very high disease activity = ASDAS ≥ 3.5. ETN: etanercept; BMI: body mass index; axSpA: axial spondyloarthritis; MRI: magnetic resonance imaging; ASAS: Assessment of SpondyloArthritis international Society; NSAID: nonsteroidal antiinflammatory drug; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; ASDAS: Ankylosing Spondylitis Disease Activity Score; CRP: C-reactive protein.

and in combination, were useful tools in assessing differences between the treatment effects of ETN and placebo in patients with axSpA. In the overall analysis, ASDAS < 2.1 combined with NSAID-sparing endpoints (i.e., ASDAS < 2.1 and/or 50% decrease in NSAID score, ASDAS < 2.1 and/or NSAID score of 0, and ASDAS < 2.1 or NSAID score < 10) showed the highest treatment discriminant capacities of the outcome measures as a group, whereas ASDAS < 1.3, individually and in combination with the NSAID-sparing endpoints, showed the lowest discriminant capacities. The duration of the randomized clinical trial was limited to 8 weeks, which might have been too short for a sufficient number of patients to achieve remission. Consequently, it cannot be excluded that this remission outcome would perform better with trials of longer duration. The results presented here for the dichotomous ASDAS < 2.1 endpoint confirm findings from the primary publication of the SPARSE study, which demon-
strated that ASDAS treated as a continuous variable had greater discriminant capacity than BASDAI in detecting differences in treatment effect. Differences between the discriminant capacities of other clinical measure combinations (i.e., ASAS and BASDAI measures) appeared to be relatively small (with the exception of BASDAI50 and an NSAID score of 0, which had the highest OR of all measures). Two NSAID intake measures (i.e., 50% reduction...
in NSAID intake and NSAID score of 0) were at least as discriminant as the conventional outcome measures; these results are of importance because when initiating anti-TNF therapy, clinicians also aim to decrease or discontinue NSAID use, particularly in patients at high risk of gastrointestinal, renal, or cardiovascular disease.

Our observation that outcome measures in combination were generally more discriminant than individual measures is also a clinically relevant finding that should be considered in the design and conduct of future clinical trials as well as the analysis of data collected in longitudinal observational studies (e.g., cohorts and registries). To allow for a more...
Figure 2. Discriminant capacity of (A) individual clinical and NSAID-sparing outcome measures, and (B) combined clinical and NSAID-sparing outcome measures in axSpA at Week 8. Based on logistic regression analyses with baseline scores and treatment group as covariates. For combined endpoints, baseline NSAID score was also in the model. Highest OR = highest discriminant capacity of treatment effects. LOCF method and ITT population. ASAS: Assessment of SpondyloArthritis international Society; ASDAS: Ankylosing Spondylitis Disease Activity Score; axSpA: axial spondyloarthritis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; ITT: intent-to-treat; LOCF: last observation carried forward; NSAID: nonsteroidal antiinflammatory drug.
robust analysis of these data, we calculated OR; the lower limits of the 95% CI for the OR may be particularly relevant to researchers when calculating sample size of new clinical trials.

To our knowledge, this double-blind, placebo-controlled study (i.e., the SPARSE study) was the first to assess the NSAID-sparing effect of an anti-TNF agent using the ASAS-NSAID score. The primary findings of this study support the NSAID-sparing effects and symptomatic benefits of such treatment. Estimated between-group differences in the proportions of patients achieving NSAID-sparing endpoints ranged from 23% (for ASAS-NSAID score < 10, p = 0.022) to 34% (for ≥ 50% decrease in ASAS-NSAID score, p = 0.003) after 8 weeks of treatment. Significant differences between the treatment groups in the proportions of patients achieving clinical endpoints at 8 weeks ranged from 21% (for BASDAI50, p = 0.034) to 32% (for ASDAS < 2.1, p = 0.006) in favor of the biologic agent. Interestingly, the treatment effects observed using the conventional outcome measures (e.g., ASAS responder criteria) were of a similar magnitude in our study, in which NSAID intake was tapered, as in other clinical trials, in which NSAID intake was maintained at stable levels during the control period.

Similar results have also been reported in clinical trials of other anti-TNF agents in radiographic and nonradiographic axSpA, although comparison of these trials is more challenging because of differences in patient selection, treatment duration, and statistical methods. The relative consistency of results across these trials suggests that changes in NSAID intake do not alter the discriminant capacity of the conventional outcome measures.

Important attributes of the SPARSE study include the prospective, randomized, double-blind, placebo-controlled design and use of many different disease assessment tools. Limitations include the 8-week duration of the study’s double-blind, placebo-controlled period, the relatively small number of patients evaluated (n = 90), and the amount of data found to be missing in patients’ paper diaries (summarized in the primary publication). Data collection through electronic patient diaries or physician interviews may be preferable options in future studies.

This study’s findings, which may influence future clinical trial design, indicate that reduced NSAID intake during treatment in patients with axSpA does not preclude demonstration of clinically relevant treatment efficacy with an anti-TNF agent. Although not surprising to clinicians after 15 years of experience with anti-TNF therapy in axSpA, confirmation of their observations in a clinical study setting is nonetheless meaningful. In addition, they suggest that clinical and NSAID-sparing outcome measures, individually and in combination, may be valid means of discriminating treatment effects in axSpA. However, further research is needed to examine more closely the validity of these combinations of “clinical” and “therapeutic” outcome measures, to determine their relevance in patients with radiographic and nonradiographic axSpA, to establish the ideal combination of clinical and NSAID-sparing outcomes with the greatest discriminant capacity, and to confirm that similar results are attainable with other treatments and over other treatment durations.

ACKNOWLEDGMENT


REFERENCES


