A Randomized Trial of the Efficacy and Tolerability of the COX-2 Inhibitor Rofecoxib vs Ibuprofen in Patients With Osteoarthritis

Ric Day, MD; Briggs Morrison, MD; Armando Luza, MD; Oswaldo Castaneda, MD; Alberto Strusberg, MD; Menachem Nahir, MD; Knut Bjorn Helgetveit, MD; Barbara Kress, RN; Brian Daniels, MD; James Bolognese; Dave Krupa; Beth Seidenberg, MD; Elliot Ehrich, MD; for the Rofecoxib/Ibuprofen Comparator Study Group

Background: Nonsteroidal anti-inflammatory drugs (NSAIDs) inhibit both cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2). It is not known whether a specific inhibitor of COX-2 will provide efficacy in osteoarthritis (OA) comparable with NSAIDs. Therefore, we compared the efficacy and safety of the rofecoxib, which specifically inhibits COX-2, with those of the NSAID ibuprofen in patients with OA.

Objective: To compare the clinical efficacy and tolerability of rofecoxib (12.5 and 25 mg once daily) with ibuprofen (800 mg 3 times daily).

Methods: A randomized, double-blind trial of 809 adults with OA was conducted. Patients with OA in whom the knee or hip was the primary source of pain were randomized to 1 of 4 treatment groups on demonstration of disease activity: placebo; rofecoxib, 12.5 or 25 mg once daily; or ibuprofen, 800 mg 3 times daily. Clinical efficacy and safety were monitored during a 6-week treatment period.

Results: Both doses of rofecoxib demonstrated efficacy clinically comparable with ibuprofen as assessed by 3 primary end points (pain walking on a flat surface [Western Ontario and McMaster Universities Osteoarthritis Index], patient global assessment of response to therapy, and investigator global assessment of disease status) according to predefined comparability criteria. Both rofecoxib doses and the ibuprofen dose provided significantly (P<.001) greater efficacy than placebo on all primary end points. Results from secondary end points were consistent with those of the primary end points. All treatments were well tolerated; the overall incidence rates of clinical adverse experiences were not significantly different (P>.05) among the treatment groups.

Conclusion: Rofecoxib was well tolerated and provided clinical efficacy comparable with a high dose of the NSAID ibuprofen.

Arch Intern Med. 2000;160:1781-1787

©2000 American Medical Association. All rights reserved.
METHODS

The primary objective of this randomized, placebo-controlled clinical trial was to compare the clinical efficacy of rofecoxib (12.5 and 25 mg once daily) with ibuprofen (800 mg 3 times daily). All subjects gave written informed consent. The study protocol was approved by the institutional review boards or ethical review committees for all 49 investigative sites in 26 countries.

STUDY DESIGN

On confirmation of eligibility, patients were randomized to 1 of 4 treatment groups by a computer-generated allocation schedule: placebo, rofecoxib, 12.5 or 25 mg once daily, or ibuprofen, 800 mg 3 times daily. The primary purpose of this study was to compare the efficacy of rofecoxib with that of ibuprofen; a smaller placebo group was included to confirm that rofecoxib and ibuprofen had efficacy greater than that of the placebo. Thus, the allocation was 1:4:4:4 for placebo, both doses of rofecoxib, and ibuprofen. The masked allocation schedule was generated by an individual not otherwise involved with the study and kept concealed from all study participants. The allocation schedule was unblinded once all data had been entered, reviewed, and certified. Medication was provided in blister packages; study blinding was maintained by using a matching placebo for each study medication. Patients took 3 tablets each morning and 1 tablet at both midday and evening. Patients returned approximately every 2 weeks for 3 visits to assess both efficacy and safety. Patients were provided open-label acetaminophen for osteoarthritis pain not adequately controlled by the study medication; the maximum daily dose of acetaminophen allowed was 2600 mg, and the amount used was recorded. Patients returned 7 to 10 days after their last dose of study medication for posttherapy safety assessment.

ENTRY CRITERIA

The study included 2 groups of patients with OA. NSAID users: These patients discontinued their prior NSAID therapy on confirmation of eligibility. Following a washout period (longer than 5 plasma half-lives of prior NSAID use), patients’ pain walking on a flat surface was assessed using question 1 of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). A minimum value of 40 mm on the pain VAS (question 1 of the WOMAC) was required to be included in the study. The allocation schedule was generated by an individual not otherwise involved with the study and kept concealed from all study participants. The allocation schedule was unblinded once all data had been entered, reviewed, and certified. Medication was provided in blister packages; study blinding was maintained by using a matching placebo for each study medication. Patients took 3 tablets each morning and 1 tablet at both midday and evening. Patients returned approximately every 2 weeks for 3 visits to assess both efficacy and safety. Patients were provided open-label acetaminophen for osteoarthritis pain not adequately controlled by the study medication; the maximum daily dose of acetaminophen allowed was 2600 mg, and the amount used was recorded. Patients returned 7 to 10 days after their last dose of study medication for posttherapy safety assessment.

EFFICACY ASSESSMENTS

To obtain a comprehensive assessment of the effect of rofecoxib on the multiple clinical manifestations of OA, a variety of efficacy end points were included in the study that were derived from the assessments made by both the patient and the investigator.

At each visit, the patient completed the WOMAC and a global assessment of overall disease status (100-mm VAS, ranging from “very well” to “very poor”). At treatment visits, the patient also rated the overall response of his or her OA to study medication on a 0-to-4 Likert scale (“none” to “excellent”). The physician rated (1) overall assessments of disease status (0-to-4 Likert scale of “very poor” to “very well”), (2) overall response of the patient’s OA to study medication, and (3) study joint tenderness. Examination of study knee or hip joint for tenderness was performed with the patient in the supine position. Tenderness was defined as pain in response to passive motion or pressure; the hip was internally and externally rotated, and the knee was moved through the full available range to detect any end range pain and palpated around the medial and lateral joint lines while the knee was in the neutral position. Pain on palpation (knee only) or during passive range to 1000 mg. Rofecoxib is, therefore, a specific inhibitor of the COX-2 isozyme in humans.

Clinical evidence to support the hypothesis that rofecoxib has an improved GI tract safety profile compared with NSAIDs would consist of data demonstrating that rofecoxib provides improved GI tract safety compared with an NSAID at doses that provide comparable clinical efficacy. In 2 large multicenter
of motion (hip and knee) was graded according to the following scale: 0, no pain; 1, patient states there is pain; 2, patient states there is pain and winces; and 3, patient states there is pain, winces, and withdraws.

Other measurements of efficacy included amount of rescue acetaminophen consumed and discontinuation from the study because of lack of efficacy of the study medication. The primary end points were pain walking on a flat surface, patient response to therapy, and investigator global assessment of disease status. Secondary end points included the WOMAC subscales (pain, stiffness, and disability), patient global assessment of disease status, investigator assessment of response to therapy, patients discontinued from the study because of lack of efficacy, acetaminophen use, and study joint tenderness.

**TOLERABILITY ASSESSMENTS**

Spontaneously reported adverse experiences and vital signs were monitored at every visit. Laboratory investigations, including hematology (complete blood cell count with differential), chemistry (electrolyte, urea nitrogen, creatinine, total protein, albumin, calcium, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, and total bilirubin levels), and urinalysis (protein, glucose, pH, red blood cells, and white blood cells, with microscopic examination if there were any abnormal results) were performed at screening, randomization, 4 and 6 weeks of therapy, and the posttherapy visit. For all adverse experiences, the investigator recorded the intensity, the relation to test drug (“definitely not” and “probably not” related were scored as not drug-related adverse experiences; “possibly,” “probably,” and “definitely” related were scored as drug-related adverse experiences), the outcome, and any action taken.

**STATISTICAL ANALYSIS**

The primary measure for each efficacy end point (except discontinuation because of lack of efficacy) was the mean response (change from baseline) over all observation times in the 6-week treatment period. All data collected from discontinuation and unscheduled visits were included in this analysis; no missing values were imputed. For each end point, a patient had to have a baseline measurement and at least 1 measurement during the 6-week treatment period for the mean change from baseline to be computed. Only 14 of the 809 randomized patients were excluded from the analysis for one or more of the primary end points because of missing baseline or on-treatment data. Eighty-five percent of the randomized patients had a measurement recorded for all 3 primary end points at all of the planned observation times.

For each end point, analysis of covariance (ANCOVA) was used to model patient mean change from baseline as a function of the categorical predictors treatment, study center, and history of ulcer or upper GI tract bleeding, and a continuous covariate, the baseline measurement. Mean patient change from baseline and SEs resulting from the ANCOVA were used to compute 95% confidence intervals (CIs) for the between-treatment difference in mean response, tests to compare mean response with active treatments against the placebo, and posterior probabilities (based on Bayesian analyses with noninformative prior distributions) that the true mean differences in response to the active treatments were within the predefined clinical comparability bounds. All statistical tests for difference were 2-tailed with \( P = .05; P \leq .05 \) was considered statistically significant.

The primary hypothesis of this study was that rofecoxib would provide clinical efficacy comparable with ibuprofen as assessed by 3 primary end points: pain walking on a flat surface, patient response to therapy, and investigator global assessment of disease status. The following conditions had to be satisfied to conclude that the treatments were clinically comparable: in any 2 of the 3 primary end points, the 95% CIs of mean differences between treatment groups had to be within predefined comparability bounds (±10 mm on a 100-mm VAS and ±0.5 on a Likert scale), and all of the 3 posterior probabilities were required to be 0.930 or lower. These clinical comparability bounds are more conservative than those proposed by a consensus panel of rheumatologists and were derived from results of previous OA trials with rofecoxib.

Separate analyses were performed to evaluate effects on treatment differences of subgroup factors, including race, age, sex, study joint (knee vs hip), and prior OA medication use (NSAID vs acetaminophen). These were assessed individually by adding each subgroup factor and its interaction with treatment to the ANCOVA model for each of the 3 primary end points.

Safety was assessed by comparing incidence rates of adverse experiences and exceeding predefined limits of change in laboratory and vital sign variables between the treatment groups. These between-group comparisons were calculated using the Fisher exact test; a step-down approach (25 mg first, and if significant, followed by 12.5 mg) was used for the comparisons of rofecoxib doses vs placebo.

This study had greater than 99% power to demonstrate comparable efficacy (according to the criteria cited) between rofecoxib and ibuprofen if their true difference is 0. Power calculations were based on observed treatment effects in other placebo-controlled studies with rofecoxib. Since this study was designed using variability data from a pilot study (data on file, Merck Research Laboratories), provision was made for larger variability. If the underlying SDs were 25% larger than those observed in the pilot study, the power was approximately 94%.

**RESULTS**

Between April 30 and November 7, 1997, 1023 patients with OA were screened and 809 were enrolled in the study (Figure 1). Patients randomly assigned to the 4 treat-
ment groups had similar sociodemographic and clinical characteristics, including baseline values for efficacy end points (Table 1, Figure 1, and Figure 2; additional data not shown).

Of the 809 patients, 709 (88%) completed the study; the overall discontinuation rate was comparable among treatment groups (Figure 3; additional data not shown). There was a significantly higher discontinuation rate because of clinical adverse experiences in the ibuprofen group compared with the placebo group (P<.05), whereas both rofecoxib groups were not significantly different from placebo. There were significantly fewer discontinuations because of lack of efficacy (P=.009) in the active therapy groups compared with the placebo group (Figure 3). The number of patients who withdrew for other reasons was similar between all groups (P>.05).

Efficacy

Figure 2 presents the mean change from baseline during the 6-week treatment period for the 3 primary end points (pain walking on a flat surface, patient response to therapy, and investigator global assessment of disease status) and the secondary end point of the physical function subscale of the WOMAC; data for all primary and secondary end points are shown in Table 2. For all 4 end points, the treatment effect was similar among all active groups and was superior to the placebo group. Maximum treatment effects were seen (Figure 2) by the first visitation (2 weeks) and were sustained throughout the 6-week treatment period. The treatment effect of rofecoxib was consistently seen for all primary and secondary end points (Figure 2 and Table 2); for each end point the effect was similar among all active groups, and all active groups were superior to the placebo group.

The clinical efficacy of both rofecoxib doses was comparable with that of ibuprofen during 6 weeks of treatment using the prespecified comparability criteria (see the “Methods” section). For all 3 primary end points, the 95% CIs for the difference of mean response between each of the treatment pairs (rofecoxib, 25 mg, and ibuprofen; rofecoxib, 12.5 mg, and ibuprofen; and rofecoxib, 25-12.5 mg, and ibuprofen) were within the predefined comparability bounds, and the posterior probability that the true mean difference was within the predefined compara-

rability bounds was greater than 0.950. The effect of rofecoxib, 25 mg, was significantly superior to that of ibuprofen (P<.05) for 2 of the 3 primary end points (patient response to therapy [P=.005] and investigator global assessment of disease status [P=.005]).

Analyses were performed to determine if the treatment effects observed were consistent across various subgroups of patients using treatment-by-subgroup analyses for the 3 primary end points. Treatment effects were consistently observed whether the patients had knee vs hip as the primary study joint and whether they were an acetaminophen user vs an NSAID user at study entry. No statistically significant interactions were observed between treatment and study center, sex, race, age, or ARA functional class at study entry.

Tolerability

The incidence of any clinical adverse event was not significantly different among the treatment groups (41.9% in the placebo group vs 50.8% [rofecoxib, 12.5 mg], 53.3% [rofecoxib, 25 mg], and 51.8% [ibuprofen] in the active groups). Drug-related clinical adverse events were more common in all active therapy groups compared with placebo (10.8% vs 27.5%, 31.0%, and 30.5%, respectively; P=.003). Clinical adverse events that led to discontinuation from the study were most common in the ibuprofen group (1.4% vs 4.1%, 3.7%, and 8.4%, respectively; P=.03 vs placebo for ibuprofen only); this was mostly accounted for by adverse experiences related to the GI tract. Two asymptomatic gastric ulcers were observed in the study; both were in the ibuprofen treatment group. The most common clinical adverse experiences were epigastric discomfort (0% vs 5.7%, 5.8%, and 8.0%, respectively), diarrhea (4.1% vs 4.5%, 5.0%, and 5.2%, respectively), and nausea (1.4% vs 2.9%, 6.6%, and 3.6%, respectively). The incidence of any laboratory adverse event was not significantly different among the treatment groups (4.1% in the placebo group vs 10.7% [rofecoxib, 12.5 mg], 7.6% [rofecoxib, 25 mg], and 13.4% [ibuprofen] in the active groups). The mean changes in body weight and blood pressure were similar in all treatment groups. Adverse experiences of edema or hyper-
tension were reported at similar rates in all treatment groups.

The discovery of 2 isoforms of COX, the target enzyme inhibited by NSAIDs, has led to a number of questions concerning the role of inhibiting COX-1 vs COX-2 in terms of the efficacy and safety of this widely prescribed class of drugs. Previous work has demonstrated that specific inhibitors of COX-2 are efficacious in the treatment of OA, but left open the question of how that efficacy compares with NSAIDs, which inhibit both COX-1 and COX-2. In this report, we demonstrated that the efficacy of rofecoxib, which specifically inhibits COX-2, was comparable with that of ibuprofen. Importantly, we characterized the effect of rofecoxib on a variety of the clinical manifestations of OA, and in all cases, we found the efficacy of rofecoxib to be comparable with ibuprofen. These results were obtained in a large, diverse population of patients from 26 different countries, and the results were consistent across race, age, sex, study joint, and prior OA medication use (NSAID vs acetaminophen).

Our data also demonstrated that rofecoxib, 12.5 and 25 mg, provided comparable clinical efficacy. Based upon these and other data, it is recommended that 12.5 mg be used as the initial dose of rofecoxib for the treatment of OA.

The NSAIDs are associated with a number of toxic effects, the most important of which are related to the GI tract and the kidney. To firmly establish an improved safety profile of rofecoxib in contrast to NSAIDs, it is important that the safety profiles be compared using doses that provide equivalent efficacy. This study rigorously demonstrated that both once-daily doses (12.5 mg and 25 mg) of rofecoxib...
An important clinical adverse effect of NSAIDs is their propensity to lead to serious upper GI tract events, such as perforations, gastric and duodenal ulcers, and upper GI tract bleeding. In other clinical studies, assessed by endoscopy, the incidence of abnormalities of the GI mucosa associated with rofecoxib, 25 and 50 mg, was substantially less than that associated with ibuprofen, 800 mg 3 times daily. Our study was not intended to assess endoscopically diagnosed ulcers and was not large enough or long enough to compare the incidence of serious upper GI tract events. However, in an overview analysis of all clinical trials performed with rofecoxib, including our study, the incidence of serious upper GI tract events was found to be significantly less with rofecoxib compared with NSAIDs.

In summary, we have demonstrated that specific inhibition of COX-2 provides efficacy in the treatment of OA that is comparable with that of high doses of the NSAID ibuprofen. The safety of rofecoxib, 12.5 and 25 mg once daily, was not significantly different from placebo and ibuprofen, 800 mg 3 times daily.

Accepted for publication December 8, 1999.

This work was funded by grants from Merck & Co Inc, West Point, Pa.

Presented as an abstract at the European League Against Rheumatism 14th Congress in Glasgow, Scotland, June 6-12, 1999.

Other members of the Rofecoxib/Ibuprofen Comparator Study Group are as follows: A. Alves de Matos, MD, Lisbon, Portugal; Andre Beaudier, MD, Saint-Foy, Quebec; Jamie Beier, MD, Esbjerg, Denmark; William Bensen, MD, Hamilton, Ontario; Marco Broggini, MD, Varese, Italy; Ricardo Castro, MD, Sao Paulo, Brazil; H. Eckhardt, MD, Vogtareuth, Germany; Carlos Fuentealba, MD, Santiago, Chile; D. Geissler, MD, Klagenfurt, Austria; Mel Gomes, MD, Lisbon, Portugal; Nigel Gilchrist, MD, Christchurch, New Zealand; Francesc Gutierrez, MD, Santiago, Chile; Kurt Haas, MD, Stockholm, Sweden; Gabriel Herrera-Beaumont, MD, Madrid, Spain; Boulou Harroui, MD, Montreal, Quebec; Bruce Kirkham, MD, Kogarah, Australia; Peter Kotz, MD, Vienna, Austria; Michel

Table 2. Efficacy Results by End Point According to Treatment Assignment*

<table>
<thead>
<tr>
<th>Efficacy End Points</th>
<th>Placebo (n = 74)</th>
<th>Rofecoxib, 12.5 mg (n = 244)</th>
<th>Rofecoxib, 25 mg (n = 242)</th>
<th>Ibuprofen, 2400 mg (n = 249)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain walking on a flat surface†</td>
<td>-18.92 (−23.72 to −14.12)</td>
<td>-34.32‡ (−37.03 to −31.60)</td>
<td>-35.07‡ (−37.82 to −31.33)</td>
<td>-33.55‡ (−36.26 to −30.84)</td>
</tr>
<tr>
<td>Patient global assessment of response to therapy§</td>
<td>-1.56 (−1.77 to −1.36)</td>
<td>-2.28‡ (−2.39 to −2.16)</td>
<td>-2.44‡ (−2.56 to −2.33)</td>
<td>-2.22‡ (−2.34 to −2.11)</td>
</tr>
<tr>
<td>Investigator global assessment of disease status§</td>
<td>-1.00 (−1.17 to −0.83)</td>
<td>-1.47‡ (−1.56 to −1.37)</td>
<td>-1.59‡ (−1.68 to −1.49)</td>
<td>-1.40‡ (−1.50 to −1.31)</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical function subscale†</td>
<td>-8.76 (−12.72 to −4.79)</td>
<td>-18.73‡ (−20.98 to −16.49)</td>
<td>-20.64‡ (−22.91 to −18.37)</td>
<td>-18.06‡ (−20.30 to −15.82)</td>
</tr>
<tr>
<td>Pain subscale†</td>
<td>-11.89 (−15.98 to −7.80)</td>
<td>-23.37‡ (−25.68 to −21.05)</td>
<td>-24.78‡ (−27.13 to −22.44)</td>
<td>-22.89‡ (−25.21 to −20.58)</td>
</tr>
<tr>
<td>Stiffness subscale†</td>
<td>-8.88 (−13.38 to −4.38)</td>
<td>-21.24‡ (−23.77 to −18.70)</td>
<td>-20.79‡ (−23.33 to −18.24)</td>
<td>-20.17‡ (−22.69 to −17.65)</td>
</tr>
<tr>
<td>Patient global assessment of disease status†</td>
<td>-10.02 (−14.60 to −5.45)</td>
<td>-26.93‡ (−29.52 to −24.34)</td>
<td>-29.05‡ (−31.66 to −26.43)</td>
<td>-25.28‡ (−27.87 to −22.69)</td>
</tr>
<tr>
<td>Patients discontinued because of lack of efficacy, No. (%)</td>
<td>9 (12.2)</td>
<td>8 (3.3)</td>
<td>7 (2.9)</td>
<td>9 (3.6)</td>
</tr>
<tr>
<td>Investigator global assessment of response to therapy§</td>
<td>-1.70 (−1.91 to −1.49)</td>
<td>-2.44‡ (−2.56 to −2.32)</td>
<td>-2.56‡ (−2.68 to −2.44)</td>
<td>-2.40‡ (−2.52 to −2.28)</td>
</tr>
<tr>
<td>Tenderness in study joint†</td>
<td>-0.56 (−0.70 to −0.42)</td>
<td>-0.84‡ (−0.93 to −0.76)</td>
<td>-0.93‡ (−1.01 to −0.85)</td>
<td>-0.82‡ (−0.90 to −0.74)</td>
</tr>
<tr>
<td>Acetaminophen usage for rescue, tablets/d</td>
<td>1.36 (1.14 to 1.58)</td>
<td>0.88 (0.74 to 1.01)</td>
<td>0.82 (0.68 to 0.95)</td>
<td>0.99 (0.85 to 1.13)</td>
</tr>
</tbody>
</table>

*All values are least squares means (95% confidence intervals) except for the end point “patients discontinued because of lack of efficacy.”
†Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), a visual analog scale of 0 to 100 mm.
‡P<.009 compared with placebo.
§Likert scale (0-4).
||Scale of 0 to 3.


32. Hawkev C, Laine L, Mortensen E, et al. Treatment of osteoarthritis with rofe-

33. coxib, a cyclooxygenase-2 (COX-2) specific inhibitor, was associated with a lower incidence of gastrointestinal ulcers compared to ibuprofen and was comparable to placebo treatment [abstract]. In: Proceedings of the 14th Congress of the Eu-

34. ropean League Against Rheumatism; 14th Congress; June 6-12, 1999; Glasgow, Scotland.


40. Rofecoxib prescribing information. West Point, PA: Merck & Co Inc.