Evaluation of Health-Related Quality of Life of Rheumatoid Arthritis Patients Treated with Celecoxib

Sean Z. Zhao, Justus I. Fiechtner, Elizabeth A. Tindall, Seema D. Dedhiya, William W. Zhao, Jane T. Osterhaus, and Shawn S. Yu

Objective. *To study the functional status and health-related quality of life (HRQOL) of patients with rheumatoid arthritis (RA) after treatment with celecoxib, compared with placebo and naproxen.*

Methods. *This was a prospective, randomized, double-blind, parallel group trial conducted at 79 sites in the United States and Canada over a 12-week treatment period. Patients were randomly assigned to 5 groups: placebo, 100 mg twice a day of celecoxib, 200 mg twice a day of celecoxib, 400 mg twice a day of celecoxib, and 500 mg twice a day of naproxen. The Health Assessment Questionnaire (HAQ) disability index was used to measure functional status. The Medical Outcomes Study Short Form 36 (SF-36) was used to measure general HRQOL.*

Results. *Enrollees were 1,149 patients with diagnosed and active RA. At the end of the treatment period, patients in the 4 active treatment groups had significant improvement in both functional status and overall HRQOL in comparison with the placebo group. Patients in the twice-daily 100 mg celecoxib group significantly differed from placebo at weeks 2 and 6 on HAQ scores and at week 12 on 5 domains and both summary scores of the SF-36. Patients treated with twice-daily 200 mg celecoxib had significantly better functional status than placebo at all times of testing with the HAQ, and also had significantly better function than those treated with naproxen after 2 and 12 weeks of treatment. Patients in the twice-daily 200 mg and 400 mg celecoxib groups showed similar improvement in HRQOL as determined by the 8 domain scores and 2 summary scores of the SF-36.*

Conclusion. *Celecoxib was better than placebo and comparable with naproxen in improving functional status and overall HRQOL among RA patients.*

Key words. *Rheumatoid arthritis; Functional status; Health-related quality of life; Stanford Health Assessment Questionnaire; Medical Outcomes Study Short Form 36.*

INTRODUCTION

Rheumatoid arthritis (RA), a common connective tissue disease (1), is characterized by alternating periods of active inflammation and absence of symptoms, both of variable duration (2). It is a disease that causes great physical, social, and economic burdens.

Impairments, both as a direct result of RA and due to the pain and immobility associated with it, impose a substantial economic burden on patients and health care payers (3). Along with the economic

Supported by G.D. Searle & Co.

Sean Z. Zhao, MD, PhD (current address: Pharmacia and Upjohn, Peapack, New Jersey), Seema D. Dedhiya, MS, William W. Zhao, PhD, Jane T. Osterhaus, PhD, and Shawn S. Yu, PhD, G.D. Searle & Co., Skokie, Illinois; Justus I. Fiechtner, MD, MPH, Good Clinical Practice Michigan Corp., East Lansing; and Elizabeth A. Tindall, MD, Portland Medical Associates, Portland, Oregon.

Address correspondence to Seema D. Dedhiya, MS, Global Health Outcomes, G.D. Searle & Co., 5200 Old Orchard Road, Skokie, IL 60077.

Submitted for publication April 19, 1999; accepted in revised form December 17, 1999.

0893-7524/00/$5.00

burden, patients' pain and forced immobility gradually lead to difficulty with physical functioning, such as walking and ascending or descending stairs (4). Measuring functional status in RA patients is important because as patients become unable to perform activities of daily living, their condition interferes with their social lives and personal relationships, reduces their level of independence, and increases their need for assistance.

Traditional treatment for RA, consisting mainly of the use of nonsteroidal anti-inflammatory drugs (NSAIDs) and disease-modifying antirheumatic drugs, is often suboptimal, requiring a balance between tolerance and efficacy (2). While efficacious at relieving pain, NSAIDs are frequently associated with a number of side effects, some of which are significant and potentially life threatening, including gastrointestinal (GI) symptoms, alteration of platelet function, impairment of renal function, and interaction with other medications commonly used in the aging population (5–14).

The mechanism of action for currently available NSAIDs is the prevention of the formation of prostaglandins by inhibition of the enzyme cyclooxygenase (COX) (15). COX exists in at least two forms: COX-1 and COX-2. COX-1 is a constitutive form of COX, while COX-2 is an inducible form. Preclinical data suggest that inhibition of COX-2 may be responsible for the anti-inflammatory effects of COX enzyme inhibition, whereas inhibition of COX-1 may be responsible for the adverse effects in the upper GI tract and platelets (16–19). Most traditional NSAIDs are nonselective inhibitors of both COX-1 and COX-2. In contrast, celecoxib, a COX-2 specific inhibitor, has a greater affinity for COX-2 than COX-1 at therapeutic doses (20). This selectivity is predicted to enable celecoxib to have a superior therapeutic index compared with mixed COX-1/COX-2 inhibitors, yielding additional clinical benefit to patients who currently depend on NSAIDs for inflammation and pain relief. Studies in RA patients have demonstrated that celecoxib has equivalent efficacy to NSAIDs but with a much better adverse events profile (21). The overall advantage of using celecoxib to treat RA is expected to be in the improvement of patients' functional status with fewer side effects. In addition, the pain associated with the condition is expected to decrease, and physical functioning is expected to improve compared with placebo.

The Stanford Health Assessment Questionnaire (HAQ) is a reliable, validated, disease-specific instrument that has been used extensively to measure functional status in patients with RA (22–32). The HAQ disability index, used in this study, serves as the 20-item core instrument of the HAQ. It can be easily self-administered in less than 5 minutes and evaluates the following 8 categories of daily functioning: dressing, arising, eating, walking, hygiene, reach, grip, and activities. The 20 items are divided between these 8 categories, with each category having at least 2 items (33).

The Medical Outcomes Study Short Form 36 (SF-36) is the most frequently used general health measure in the United States (34). It is a reliable, validated, generic instrument that has been used extensively to measure health-related quality of life (HRQOL) in diverse patient groups, including arthritis patients (31,34–40). The SF-36 measures health across 3 dimensions (functional status, well-being, and overall evaluation of health) using 8 separate domains (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health). In addition to the 35 items that make up these 8 domains, 1 additional item measures health transition. This 36-item survey can be self-administered within 5 to 10 minutes.

This study evaluated the functional status and general HRQOL of RA patients after treatment with celecoxib, compared with naproxen and placebo treatments, specifically with respect to their disability and physical functioning. The hypothesis of the study was that patients receiving 100, 200, and 400 mg twice a day of celecoxib would experience significant improvements in their functional status and overall HRQOL compared with placebo, and equivalent to 500 mg twice a day of naproxen, a full therapeutic dose for treating RA.

PATIENTS AND METHODS

Data collection and study design. This was a prospective, randomized, double-blind, parallel group, multicenter study conducted at 79 sites in the US and Canada over a treatment period of 12 weeks. The study was conducted in accordance with the principles of good clinical practice and the Helsinki Declaration of 1975. The institutional review board at each clinical site approved the protocol, and all patients were required to provide written informed consent.

Male and female outpatients, 18 years of age or older, were eligible to participate in this study if they fulfilled the American College of Rheumatology (formerly the American Rheumatism Association) clinical criteria for a diagnosis of RA evident for a period of 3 months or longer (41), and were in a functional class of I–III (42). Following a 2- to 7-day washout

period of NSAIDs or any analgesic medication, symptomatic arthritis ("flare") was confirmed at a baseline visit according to predefined criteria for changes in the results of the arthritis assessments performed at screening. In this study, baseline flare was determined according to physicians' and patients' global assessments (43). Excluded from the study were patients who presented with serious concomitant GI, renal, hepatic, or coagulation disorders; malignancy; or ulcerations. Additionally, patients diagnosed with another inflammatory type of arthritis or with a secondary noninflammatory type of arthritis (e.g., osteoarthritis or fibromylagia) were excluded from participation.

Patients with diagnosed and active RA were enrolled and randomly assigned to 1 of the 5 treatment groups: placebo, twice-daily 100 mg celecoxib, twice-daily 200 mg celecoxib, twice-daily 400 mg celecoxib, and twice-daily 500 mg naproxen. Oral glucocorticoids (up to 10 mg prednisone/day) or disease-modifying drugs, low-dose aspirin (≤325 mg/day), and acetaminophen (up to 2 g/day for periods no greater than 3 consecutive days) were allowed as concomitant medications. Use of other NSAIDs, injectable corticosteroids, anticoagulants, or anti-ulcer drugs was prohibited.

Data were collected on patient demographic characteristics, functional status, and HRQOL. The HAQ was administered to study patients prior to receiving any drug or placebo (baseline), at the end of weeks 2 and 6, and at the end of treatment (week 12). The SF-36 was administered at baseline and at the end of the 12-week treatment period. Additionally, prior to enrollment, patients completed a medical history, physical examination, and clinical laboratory testing including a serologic test for *Helicobacter pylori.* Baseline and treatment period assessments of RA at weeks 2, 6, and 12 included a complete count of tender/painful joints, a complete count of swollen joints, patient's and physician's global assessment of arthritis, duration of morning stiffness, patient's assessment of arthritis pain on a visual analog scale, and plasma levels of C-reactive protein. An upper GI endoscopy evaluation was also conducted at baseline and at the end of the treatment. Safety was evaluated according to the incidence and type of adverse reactions, and the rate of withdrawal because of adverse reactions. Patients were also monitored for clinical laboratory abnormalities at each treatment visit. Further details of the study design, study population, and results of clinical efficacy and safety are reported elsewhere (43).

HAQ data were scored according to the developer's specifications, with a low score representing a better functional status. The items in the HAQ are rated on a 4-point scale with 0 representing "no difficulty," 1 representing "some difficulty," 2 representing "much difficulty," and 3 representing "unable to do." The highest score recorded on an item in a particular category represents the score for that category, unless aids or assistance are used, in which case a lower score is adjusted to 2 to represent a more limited functional status. The 8 individual category scores are totaled and divided by 8 to yield the functional disability index (FDI) that can be expressed on a scale of 0 to 3. The FDI is dependent on the number of categories answered and cannot be computed if fewer than 6 categories are completed (44). A lower FDI score represents a better functional status.

The SF-36 data were also scored according to developers' specifications. The questionnaire provides 8 domain scores and two summary scores (physical and mental component scores, PCS and MCS), but does not provide a total score. The 8 domain scores and the 2 summary scores range from 0 to 100, where higher scores indicate a better HRQOL. The 2 summary scores are standardized with a mean of 50 and a standard deviation equal to 10 (34,36).

Statistical methods. All randomized patients were included in the analyses of pretreatment variables. All statistical analyses on treatment period variables were based on the intent-to-treat population, defined as all patients who were randomized and took at least one dose of study medication. The last observation carried forward approach was used for imputing missing values (45). All statistical testing was 2-sided at a 5% level of significance, and changes in HAQ scores and SF-36 scores from baseline were reported as least square means of change. The statistical package SAS for Windows version 6.03 (SAS Institute, Cary, NC) was used to conduct all analyses. For statistical analyses with the center as factor, any center with no patients in at least 1 of the 5 treatment groups was pooled with another center geographically.

Homogeneity of treatment groups for categorical variables was analyzed by chi-square tests. A 2-way analysis of variance with treatment and center included in the model was used to examine the homogeneity among treatment groups for continuous variables and also to analyze baseline HAQ and SF-36 scores. Treatment comparisons were assessed using change from baseline scores for each of the categories and the FDI of the HAQ, and for the 8 domain scores and 2 summary scores of the SF-36. Statistical significance was tested using a 2-way analysis of

Table 1. Demographic and baseline characteristics of patients*

Characteristic	Placebo (n = 231)	100 mg bid celecoxib (n = 240)	200 mg bid celecoxib (n = 235)	400 mg bid celecoxib (n = 218)	500 mg bid naproxen (n = 225)	*P* value
Mean age in years (range)	54 (27–79)	54 (22–85)	55 (20–90)	54 (22–85)	55 (28–81)	NS
Female, %	73	74	73	72	72	NS
Weight (kg), mean ± SD	77 ± 19	79 ± 18	77 ± 19	78 ± 18	79 ± 19	NS
Body mass index (kg/m^2), mean ± SD	27.69 ± 6.27	28.31 ± 6.10	28.11 ± 6.07	28.42 ± 6.53	28.36 ± 6.16	NS
Duration of disease (years), mean ± SD	11 ± 11	11 ± 10	11 ± 10	10 ± 9	10 ± 9	NS
Patients global assessment, %†						NS
Fair	35	38	37	44	46	
Poor	53	51	48	41	43	
Very poor	11	12	14	15	11	
Physicians global assessment, %†‡						NS
Fair	43	45	37	43	46	
Poor	49	46	52	48	48	
Very poor	8	8	11	9	5	
Arthritis pain—VAS (mm), mean ± SD§	69 ± 19	67 ± 20	68 ± 20	66 ± 21	67 ± 18	NS
Duration of morning stiffness (min), mean ± SD	276.5 ± 350.5	279.4 ± 388.5	305.3 ± 409.8	310.9 ± 418.7	312.6 ± 407.6	NS

* bid = twice a day; NS = not significant; VAS = visual analog scale.
† Scale ranged from 1 (very good) to 5 (very poor).
‡ One patient in the 100 mg bid celecoxib treatment group was rated "good," and the score for one 500 mg bid naproxen group patient was missing.
§ Scale ranged from 0 (no pain) to 100 (very severe pain).

covariance with treatment and centers as factors, and baseline scores included in the model as a covariate. To adjust multiple treatment comparisons and multiple endpoints for HAQ and SF-36 data, two methods were applied. First, scores of FDI of HAQ, PCS, and MCS were chosen as primary HRQOL endpoints in the analysis. The conclusion regarding the treatment effects were specific to these 3 endpoints. The conclusion could be generalized to HRQOL only when all 3 endpoints showed statistical significance. Second, Hochberg's step-up procedure (46) was used to interpret the results of pairwise comparisons for twice-daily celecoxib 100 mg, 200 mg, and 400 mg, and twice-daily naproxen 500 mg treatment groups versus placebo and for twice-daily celecoxib 100 mg, 200 mg, and 400 mg treatment groups versus the twice-daily naproxen 500 mg treatment group.

RESULTS

Descriptive statistics. A total of 1,149 patients were enrolled in the trial. The demographic characteristics and disease activity of the patients at study entry are summarized in Table 1. There were no statistically significant differences between the 5 treatment groups with respect to demographic characteristics or disease activity. Of the 1,149 patients who enrolled, 688 patients (60%) completed the study. Reasons for early discontinuation included treatment failure (30%), adverse events (6%), protocol noncompliance (3%), pre-existing protocol violation (1%), and loss to followup (1%) (Table 2).

Functional status evaluation. Mean baseline HAQ scores for each of the 8 categories and the FDI were similar across treatment groups, and no statistically significant differences between treatment groups were noted (Table 3). During the treatment period, mean HAQ scores for patients in all treatment groups, including placebo, showed improvement from baseline (Table 4). At the end of week 2, significant differences were found in 7 categories in the twice-daily 100 mg celecoxib group, in all 8 categories in the twice-daily 200 mg and 400 mg celecoxib groups, and in 5 categories in the twice-daily 500 mg naproxen group as compared with the placebo group. At week 6, compared with placebo, significant improvements were evident in 3 categories of the twice-daily 100 mg celecoxib group, in 6 categories of the twice-daily 200 mg and 400 mg celecoxib groups, and in 4 categories of the twice-daily 500 mg naproxen group. At the end of the 12-week treatment period, significant differences were seen in 2 categories in the twice-daily 100 mg

Table 2. Reasons for study termination in all randomized patients, no. (%)*

	Placebo (n = 231)	100 mg bid celecoxib (n = 240)	200 mg bid celecoxib (n = 235)	400 mg bid celecoxib (n = 218)	500 mg bid naproxen (n = 225)
Completed study	101 (44)	154 (64)	158 (67)	137 (63)	138 (61)
Withdrawn	130 (56)	86 (36)†	77 (33)†	81 (37)†	87 (39)†
Reason for withdrawal‡					
Loss to followup	3 (1)	1 (<1)	3 (1)	1 (<1)	1 (<1)
Pre-existing violation	2 (<1)	1 (<1)	3 (1)	2 (<1)	0 (0)
Protocol noncompliance	10 (4)	4 (2)	4 (2)	7 (3)	9 (4)
Treatment failure	104 (45)	67 (28)†	50 (21)†	59 (27)†	65 (29)†
Adverse events	11 (5)	13 (5)	17 (7)	12 (6)	12 (5)

* bid = twice a day.
† Statistically significantly different from placebo ($P < 0.05$).
‡ Mutually exclusive and exhaustive categories.

celecoxib group, in 6 categories in the twice-daily 200 mg celecoxib group, in 5 categories in the twice-daily 400 mg celecoxib group, and in 3 categories in the twice-daily 500 mg naproxen group. Also, at the end of the 12 weeks, the FDI showed significant improvement of 0.29 (19.33%) in the twice-daily 200 mg celecoxib group, 0.28 (20.00%) in the twice-daily 400 mg celecoxib group, and 0.22 (14.67%) in the naproxen group. The twice-daily 100 mg celecoxib group showed significant changes from placebo in the FDI at the end of weeks 2 (0.24 or 17.14%) and 6 (0.21 or 15.00%), but the improvement at the end of week 12 (0.17 or 12.14%) was not statistically significant compared with placebo (Figure 1). Overall, all doses of celecoxib showed improvements in functional status that were better than in the placebo

Table 3. Health Assessment Questionnaire (HAQ) and Medical Outcomes Study Short Form 36 (SF-36) scores at baseline*

	Placebo (n = 231)		100 mg bid celecoxib (n = 240)		200 mg bid celecoxib (n = 235)		400 mg bid celecoxib (n = 217)		500 mg bid naproxen (n = 225)		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*P* value
HAQ categories											
Activities	1.55	0.91	1.52	0.84	1.57	0.93	1.52	0.98	1.61	0.87	NS
Arising	1.22	0.87	1.24	0.83	1.24	0.87	1.31	0.89	1.27	0.82	NS
Dressing and grooming	1.27	0.83	1.15	0.84	1.35	0.87	1.26	0.86	1.26	0.84	NS
Eating	1.30	0.92	1.22	0.91	1.36	0.95	1.33	0.91	1.40	0.90	NS
Gripping and grasp	1.67	0.82	1.59	0.80	1.75	0.82	1.63	0.85	1.63	0.79	NS
Personal hygiene	1.57	1.04	1.57	1.13	1.58	1.10	1.52	1.10	1.68	1.09	NS
Reach	1.60	0.92	1.62	0.86	1.64	0.99	1.52	1.00	1.62	0.92	NS
Walking	1.01	0.88	1.11	0.82	1.18	0.85	1.03	0.88	1.20	0.87	NS
Total functional disability index	1.40	0.66	1.40	0.65	1.50	0.73	1.40	0.72	1.50	0.67	NS
SF-36 domains											
Physical component score†	29.1	8.6	29.7	8.0	29.5	7.9	29.5	8.3	29.9	8.9	NS
Physical functioning	38.1	24.8	39.2	21.8	36.4	24.1	38.6	24.5	36.9	24.1	NS
Role physical	18.9	29.6	21.8	33.2	23.1	33.3	22.7	32.0	23.4	33.0	NS
Bodily pain	32.9	17.7	34.9	16.6	32.4	17.1	33.9	17.4	34.4	17.0	NS
General health	51.5	22.0	52.6	21.1	50.6	23.0	52.0	20.5	52.5	20.8	NS
Mental component score†	46.9	10.8	47.6	11.1	45.3	12.3	47.5	11.6	46.2	11.6	NS
Vitality	34.7	19.5	34.4	19.7	34.7	21.9	34.0	20.4	35.2	21.4	NS
Social functioning	57.3	28.1	62.0	26.4	55.4	27.2	60.0	28.9	60.2	27.5	NS
Role emotional	48.1	44.4	51.3	43.6	44.2	44.4	52.9	45.3	46.4	42.4	NS
Mental health	69.5	17.8	69.4	18.0	65.6	22.5	69.1	20.2	66.7	19.3	NS

* bid = twice a day; NS = not significant.
† Standardized with mean 50 and SD 10.

Table 4. Average change in Health Assessment Questionnaire (HAQ) scores at weeks 2, 6, and 12 after controlling for study centers and baseline scores*

HAQ categories	Placebo		100 mg bid celecoxib		200 mg bid celecoxib		400 mg bid celecoxib		500 mg bid naproxen		Overall *P* value†
	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%	
Week 2											
Dressing and grooming	−0.13	−10.24	−0.24	−20.87	−0.36‡	−26.67	−0.35‡	−27.78	−0.29‡	−23.02	<0.001
Arising	−0.07	−5.74	−0.29‡	−23.39	−0.32‡	−25.81	−0.40‡	−30.53	−0.30‡	−23.62	<0.001
Eating	−0.03	−2.31	−0.22‡	−18.03	−0.26‡	−19.12	−0.41‡§	−30.83	−0.19‡	−13.57	<0.001
Walking	0.04	3.96	−0.16‡	−14.41	−0.24‡	−20.34	−0.24‡	−23.30	−0.17‡	−14.17	<0.001
Reach	−0.08	−5.00	−0.33‡	−20.37	−0.23‡	−14.02	−0.23‡	−15.13	−0.15	−9.26	<0.01
Personal hygiene	−0.02	−1.27	−0.21‡	−13.38	−0.18‡	−11.39	−0.27‡§	−17.76	−0.09	−5.36	<0.01
Gripping and grasp	−0.03	−1.80	−0.20‡	−12.58	−0.24‡	−13.71	−0.29‡§	−17.79	−0.12	−7.36	<0.001
Activities	−0.02	−1.29	−0.25‡	−16.45	−0.34‡	−21.66	−0.32‡	−21.05	−0.23‡	−14.29	<0.001
Total functional disability index	−0.04	−2.86	−0.24‡	−17.14	−0.29‡§	−19.33	−0.32‡§	−22.86	−0.21‡	−14.00	<0.001
Week 6											
Dressing and grooming	−0.15	−11.81	−0.20	−17.39	−0.34‡	−25.19	−0.28‡	−22.22	−0.27	−21.43	<0.05
Arising	−0.11	−9.02	−0.30‡	−24.19	−0.40‡	−32.26	−0.34‡	−25.95	−0.31‡	−24.41	<0.001
Eating	−0.08	−6.15	−0.22‡	−18.03	−0.29‡	−21.32	−0.32‡	−24.06	−0.25‡	−17.86	<0.01
Walking	−0.02	−1.98	−0.11	−9.91	−0.22‡	−18.64	−0.23‡	−22.33	−0.18‡	−15.00	<0.01
Reach	−0.15	−9.38	−0.24	−14.81	−0.28	−17.07	−0.28	−18.42	−0.25	−15.43	NS
Personal hygiene	−0.10	−6.37	−0.13	−8.28	−0.17	−10.76	−0.23	−15.13	−0.10	−5.95	NS
Gripping and grasp	−0.15	−8.98	−0.22	−13.84	−0.30‡	−17.14	−0.30‡	−18.40	−0.20	−12.27	<0.05
Activities	−0.07	−4.52	−0.22‡	−14.47	−0.30‡	−19.11	−0.31‡	−20.39	−0.27‡	−16.77	<0.01
Total functional disability index	−0.11	−7.86	−0.21‡	−15.00	−0.30‡	−20.00	−0.29‡	−20.71	−0.25‡	−16.67	<0.001
Week 12											
Dressing and grooming	−0.17	−13.39	−0.17	−14.78	−0.29	−21.48	−0.22	−17.46	−0.22	−17.46	NS
Arising	−0.11	−9.02	−0.27‡	−21.77	−0.33‡	−26.61	−0.36‡	−27.48	−0.31‡	−24.41	<0.001
Eating	−0.06	−4.62	−0.15	−12.30	−0.26‡	−19.12	−0.31‡	−23.31	−0.20	−14.29	<0.01
Walking	0.00	0.00	−0.19‡	−17.12	−0.22‡	−18.64	−0.28‡	−27.18	−0.23‡	−19.17	<0.001
Reach	−0.14	−8.75	−0.15	−9.26	−0.34‡§	−20.73	−0.25	−16.45	−0.14	−8.64	<0.05
Personal hygiene	−0.07	−4.46	−0.09	−5.73	−0.15	−9.49	−0.21	−13.82	−0.11	−6.55	<0.05
Gripping and grasp	−0.10	−6.00	−0.18	−11.32	−0.28‡	−16.00	−0.30‡	−18.40	−0.17	−10.43	<0.05
Activities	−0.08	−5.16	−0.19	−12.50	−0.33‡	−21.02	−0.28‡	−18.42	−0.24‡	−14.91	<0.05
Total functional disability index	−0.10	−7.14	−0.17	−12.14	−0.29‡§	−19.33	−0.28‡	−20.00	−0.22‡	−14.67	<0.001

* Least square mean change values are reported. Negative values represent an improvement in functional status. Number of patients who completed the HAQ at week 2: 161 (placebo), 214 (100 mg twice a day [bid] celecoxib), 210 (200 mg bid celecoxib), 182 (400 mg bid celecoxib), and 197 (500 mg bid naproxen); number of patients who completed the HAQ at week 6: 112 (placebo), 171 (100 mg bid celecoxib), 176 (200 mg bid celecoxib), 156 (400 mg bid celecoxib), and 155 (500 mg bid naproxen); number of patients who completed the HAQ at week 12: 101 (placebo), 152 (100 mg bid celecoxib), 158 (200 mg bid celecoxib), 137 (400 mg bid celecoxib), and 138 (500 mg bid naproxen). NS = not significant.
† Overall *P* values are based on analysis of covariance testing. If *P* value > 0.05, pairwise analysis was not conducted.
‡ Statistically significantly different from placebo.
§ Statistically significantly different from 500 mg bid naproxen.

group, with the twice-daily 200 mg and 400 mg celecoxib groups showing improvements better than naproxen in some cases at weeks 2 and 12.

General HRQOL evaluation. Mean SF-36 scores at baseline, for each of the 8 domains, and the 2 summary scores (PCS and MCS) were similar across all treatment groups (Table 3). The results of mean change analyses for changes in SF-36 scores from baseline to week 12 are presented in Table 5.

Physical functioning domains. The twice-daily 200 mg celecoxib, 400 mg celecoxib, and 500 mg naproxen groups were significantly different from placebo at almost all instances on the 4 domains related to physical functioning ($P < 0.01$). An exception was seen for the twice-daily 500 mg naproxen group that did not show a significant difference on the role physical domain. In addition, the twice-daily 100 mg celecoxib group was significantly different from placebo on the bodily pain and general health domains, but not on the physical functioning and role physical domains.

Mental health domains. The 4 domains that represent mental health had significant improvements

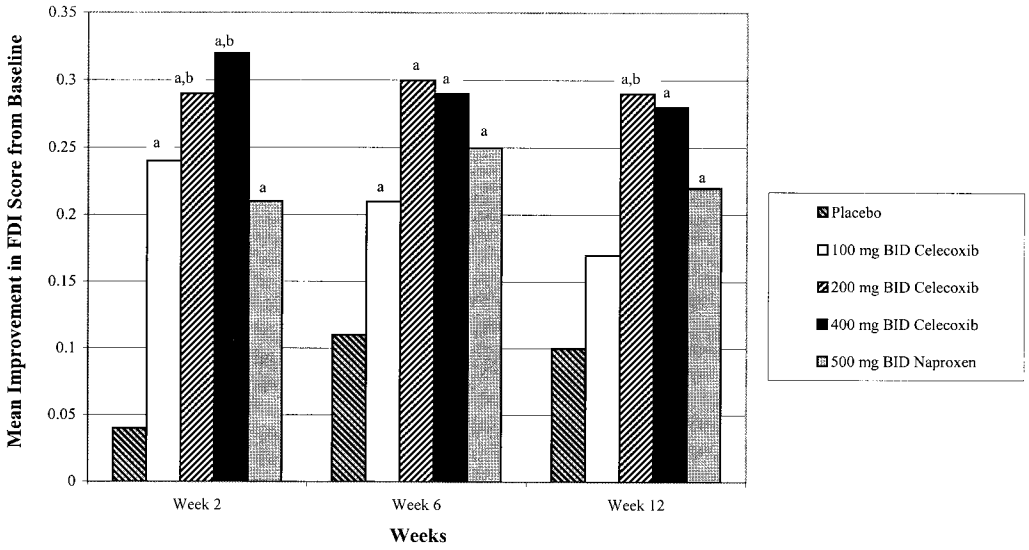

Figure 1. Mean improvement in Health Assessment Questionnaire (HAQ) functional disability index (FDI) scores from baseline to the end of treatment (week 12). During the 12-week treatment period, mean HAQ scores for patients in all treatment groups, including placebo, showed improvement from baseline. The 200 mg twice a day (BID) celecoxib group had significantly better functional status than placebo at all times of testing, and was significantly better than naproxen at the end of week 12 ($P < 0.001$). a = statistically significantly different from placebo; b = statistically significantly different from 500 mg BID naproxen.

from placebo in the twice-daily 200 mg celecoxib, 400 mg celecoxib, and 500 mg naproxen groups ($P < 0.05$). The twice-daily 100 mg celecoxib group also showed significant improvements in 3 of the 4 domains (i.e., significant improvement was not seen in the mental health domain). The placebo treatment group had negative scores for all but the role emotional domain.

Standardized summary scores. The mean change scores show a significant improvement in the PCS in all treatment groups compared with placebo ($P < 0.001$). The PCS results for the twice-daily 200 mg

Table 5. Average change in Medical Outcomes Study Short Form 36 (SF-36) scores for rheumatoid arthritis patients at week 12 after controlling for study centers and baseline scores*

	Placebo		100 mg bid celecoxib		200 mg bid celecoxib		400 mg bid celecoxib		500 mg bid naproxen		
SF-36	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%	*P* value
Physical component score†	0.9	3.0	2.5‡	8.2	4.3‡	14.5	4.4‡	14.9	2.7‡	9.1	<0.001
Physical functioning	0.6	1.5	3.9	10.0	9.1‡	25.1	9.1‡	23.6	5.5‡	14.9	<0.001
Role physical	4.3	22.9	10.9	49.9	15.6‡	67.3	14.0‡	61.8	10.1	43.1	<0.01
Bodily pain	2.1	6.3	7.5‡	21.6	12.1‡	37.5	9.9‡	29.1	9.0‡	26.1	<0.001
General health	−1.1	−2.0	2.6‡	4.9	3.8‡	7.4	4.9‡	9.4	3.1‡	5.9	<0.001
Mental component score†	−0.7	−1.5	1.8‡	3.8	2.8‡	6.2	1.8‡	3.8	2.1‡	4.5	<0.01
Vitality	−0.4	−1.2	5.1‡	14.8	9.5‡§	27.2	8.6‡	25.3	5.2‡	14.7	<0.001
Social functioning	−3.2	−5.5	4.7‡	7.6	9.7‡§	17.5	7.6‡	12.7	4.1‡	6.7	<0.0001
Role emotional	1.0	2.0	9.1‡	17.8	11.3‡	25.5	8.5‡	16.0	9.4‡	20.2	<0.05
Mental health	−0.8	−1.1	2.0	2.8	3.8‡	5.7	2.8‡	4.0	3.5‡	5.2	<0.01

* A negative value represents a decrease in health-related quality of life. Number of patients who completed the SF-36 at week 12: 100 (placebo), 152 (100 mg twice a day [bid] celecoxib), 156 (200 mg bid celecoxib), 134 (400 mg bid celecoxib), and 135 (500 mg bid naproxen). Mean = least square means; % = percentage change from baseline.
† Standardized with mean 50 and SD 10.
‡ Statistically significantly different from placebo.
§ Statistically significantly different from 500 mg bid naproxen.

and 400 mg celecoxib groups were also statistically significantly different from the PCS for the twice-daily 100 mg celecoxib group. The results of the MCS were similar to those of the PCS, with statistically significant differences from placebo occurring in all the other treatment groups ($P < 0.01$). There was a decline in the overall mental health of patients in the placebo group over the entire treatment period, as is indicated by the negative MCS.

DISCUSSION

The results of this study show the effect of celecoxib, a COX-2 specific inhibitor, on functional status and HRQOL in patients with chronic pain resulting from RA. Treatment with celecoxib for a period of 12 weeks was associated with a functional improvement as measured by the HAQ, and an overall improvement in HRQOL as shown by the SF-36. The magnitude of improvement observed with doses of twice-daily 100, 200, and 400 mg celecoxib, especially 200 and 400 mg doses, was comparable with that of twice-daily 500 mg naproxen. Naproxen is a nonselective inhibitor of both COX-1 and COX-2, and at its full therapeutic dose of 500 mg twice a day has been an efficacious NSAID for the treatment of RA. The improvement in FDI with a twice-daily 200 mg dose of celecoxib, the recommended dose for the treatment of RA (21), was statistically different from a twice-daily 500 mg dose of naproxen at weeks 2 and 12. At the 12-week assessment, the twice-daily 100 mg celecoxib treatment group did not show statistically significant improvement in HAQ disability index and 3 SF-36 domains compared with placebo; however, the improvement observed was not different than that in the naproxen group. The twice-daily 100 mg celecoxib dose has been previously established as efficacious and been proven to be similar to naproxen in its effects (43,47).

These findings are consistent with the discontinuation pattern among patients who dropped out due to treatment failure. Over one-half of the patients assigned to placebo treatment withdrew from the study at some point during the treatment period, in contrast to 33–37% of the celecoxib patients, and 39% of naproxen-treated patients. In general, adverse event incidences among the celecoxib-treated groups were higher than in the placebo group, but no statistically significant differences were found.

A lower HRQOL has been shown to be associated with increased health care use among osteoarthritis patients (48). The decrements that arthritis patients experience in HRQOL have been found to be larger with more severe arthritis, and an additive burden of diminished HRQOL has also been noted when other serious diseases (e.g., asthma, bronchitis, heart condition, hypertension, etc.) accompany it. These decrements not only concern patients and providers because of the burden placed on patients, but they also concern payers because of the relationship between decrements in HRQOL and the increased use of medical resources. Treatments that provide improved HRQOL and relief from patients' symptoms may reduce the economic burden placed on payers, while improving the level of patients' physical activity and overall HRQOL.

HRQOL assessments are an integral part of evaluating patients with arthritis, but specific dimensions of HRQOL, such as functional status, should also be measured to determine the effect arthritis has on patients' pain, stiffness, and physical functioning. Even though the SF-36 measures some aspects of functional status, disease-specific instruments such as the HAQ are able to provide more detailed information on the patient population. Results from the administration of the HAQ confirmed that celecoxib improved functional status among RA patients. An additional dimension that could be measured in a study is the specific effect of reduced GI problems achieved with the use of celecoxib as compared with naproxen. The HAQ or the SF-36 alone, however, cannot capture that information. Such GI-related information could be very valuable for future policy decisions regarding the favorability of the COX-2 specific inhibitors over traditional NSAIDs.

Treatments that provide improved functional status and relief from patients' symptoms may reduce the economic burden placed on health care payers while improving the level of patients' physical activity. The cost-effectiveness of celecoxib in relation to traditional NSAIDs needs to be further explored, taking into account the observed benefits in patient functioning and the reduction in adverse effects. The importance of pain relief and improvement in physical functioning is highlighted through the improvements in the patients' ability to care for themselves. Using celecoxib helped patients reduce their disability and perform daily physical activities, such as dressing themselves, shampooing their hair, reaching, and bending, with more ease. Celecoxib may even be able to provide sufficient relief to enable patients to increase their level of exercise, which may in turn help slow the progress of the disease.

The clinical results show that all doses of celecoxib were safe and generally well tolerated. In this study, significant anti-inflammatory, analgesic efficacy and a reduction in morning stiffness were associated with

the celecoxib treatment and led to significant functional improvement based on the HAQ FDI, and in HRQOL as demonstrated by the two summary scores and individual domain scores on the SF-36. In general, the efficacy of all 3 celecoxib groups was similar to twice-daily 500 mg naproxen as determined by the HRQOL measures, although statistically significant differences when compared with the naproxen group were detected only in the HAQ FDI for the twice-daily 200 mg celecoxib group, in the PCS for the twice-daily 400 mg celecoxib group, and in the vitality and social functioning domain scores for the twice-daily 200 mg celecoxib group.

Naproxen, a nonselective inhibitor of both COX-1 and COX-2, has been demonstrated to be an efficacious NSAID at its full therapeutic dose of 500 mg twice a day. Furthermore, the efficacy of celecoxib, combined with the observation that the drug is well tolerated and has superior upper GI safety when compared with naproxen, indicates that celecoxib may represent a major advance over current RA therapy (43). In conclusion, the results show that celecoxib was better than placebo and comparable with naproxen in improving functional status and overall HRQOL among RA patients.

We thank Somali Misra Burgess, PhD, of Strategic Outcomes Services, Inc., for her assistance in preparing the manuscript.

REFERENCES

1. Binder AI. Current management of rheumatoid arthritis. Br J Hosp Med 1992;48:106–14.
2. Rakel RE. Conn's current therapy. Philadelphia: WB Saunders; 1996.
3. Gabriel SE, Crowson CS, Campion ME, O'Fallon WM. Direct medical costs unique to people with arthritis. J Rheumatol 1997;24:719–25.
4. Hudgins TH, Brander VA, Chang RW. Rehabilitation advances in the treatment of arthritis and musculoskeletal disease. Curr Opin Rheumatol 1997;9:112–7.
5. Garcia Rodriguez LA, Jick H. Risk of upper gastrointestinal bleeding and perforation associated with individual non-steroidal anti-inflammatory drugs. Lancet 1994;343:769–72.
6. Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. Ann Intern Med 1991;115:787–96.
7. Langman MJS, Weil J, Wainwright P, Lawson DH, Rawlins MD, Logan RFA, et al. Risks of bleeding peptic ulcer associated with individual non-steroidal anti-inflammatory drugs. Lancet 1994;343:1075–8.
8. Carmichael J, Shankel SW. Effects of nonsteroidal anti-inflammatory drugs on prostaglandins and renal function. Am J Med 1985;78:992–1000.
9. Whelton A, Stout RL, Spilman PS, Klassen DK. Renal effects of ibuprofen, piroxicam, and sulindac in patients with asymptomatic renal failure. Ann Intern Med 1990;112:568–76.
10. Schlondorff D. Renal complications of nonsteroidal anti-inflammatory drugs. Kidney Int 1993;44:643–53.
11. Houston MC. Nonsteroidal anti-inflammatory drugs and antihypertensives. Am J Med 1991;90:42S–7S.
12. Burch JW, Stanford N, Majerus PW. Inhibition of platelet prostaglandin synthetase by oral aspirin. J Clin Invest 1978;61:314–9.
13. Vane JR. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. Nat New Biol 1971;231:232–5.
14. Todesco S. Special considerations in the use of NSAIDs in the elderly. Eur J Rheumatol Inflamm 1994; 14:7–13.
15. Mizuno K, Yamamoto S, Lands WEM. Effects of nonsteroidal anti-inflammatory drugs on fatty acid cyclooxygenase and prostaglandin hydroperoxidase activities. Prostaglandins 1982;23:743–57.
16. Masferrer JL, Zweifel BS, Manning PT, Hauser SD, Leahy KM, Smith WG, et al. Selective inhibition of inducible cyclooxygenase 2 in vivo is antiinflammatory and nonulcerogenic. Proc Natl Acad Sci U S A 1994;91:3228–32.
17. Siebert K, Zhang Y, Leahy K, Hauser S, Masferrer J, Perkins W, et al. Pharmacological and biochemical demonstration of the role of cyclooxygenase 2 in inflammation and pain. Proc Natl Acad Sci U S A 1994; 91:12013–7.
18. DeWitt DL, Meade EA, Smith WL. PGH synthase isoenzyme selectivity: the potential for safer nonsteroidal antiinflammatory drugs. Am J Med 1993;95: 40S–4S.
19. Mitchell JA, Akarasereenont P, Thiemermann C, Flower RJ, Vane JR. Selectivity of nonsteroidal antiinflammatory drugs as inhibitors of constitutive and inducible cyclooxygenase. Proc Natl Acad Sci U S A 1993;90:11693–7.
20. Penning TD, Talley JJ, Bertenshaw SR, Carter JS, Collins PW, Docter S, et al. Synthesis and biological evaluation of the 1,5-diary pyrazole class of cyclooxygenase-2 inhibitors: identification of 4-[5-(4-methylphenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl]benze nesulfonamide (SC-58635, celecoxib). J Med Chem 1997;40:1347–65.
21. G.D. Searle & Co. Celebrex (celecoxib capsules) package insert. Skokie (IL): G.D. Searle & Co.; 1998.
22. Bombardier C, Buchbinder R, Tugwell P. Efficacy of cyclosporin A in rheumatoid arthritis: long-term follow-up data and the effect on quality of life. Scand J Rheumatol 1992;95:29–33.
23. Wolfe F, Hawley DJ. The relationship between clinical activity and depression in rheumatoid arthritis. J Rheumatol 1993;20:2032–7.
24. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ,

Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire functional disability index in patients with rheumatoid arthritis. J Rheumatol 1988;15:1480–8.
25. Thompson PW, Pegley FS. A comparison of disability measured by the Stanford Health Assessment Questionnaire disability scales (HAQ) in male and female rheumatoid outpatients. Br J Rheumatol 1991;30:298–300.
26. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.
27. Guillemin F, Briançon S, Pourel J. Validity and discriminant ability of the HAQ functional index in early rheumatoid arthritis. Disabil Rehabil 1992;14:71–7.
28. Ziebland S, Fitzpatrick R, Jenkison C, Mowat A, Mowat A. Comparison of two approaches to measuring change in health status in rheumatoid arthritis: the Health Assessment Questionnaire (HAQ) and modified HAQ. Ann Rheum Dis 1992;51:1202–5.
29. Tennant A, Hillman M, Fear J, Pickering A, Chamberlain MA. Are we making the most of the Stanford Health Assessment Questionnaire? Br J Rheumatol 1996;35:574–8.
30. Hawley DJ, Wolfe F. Sensitivity to change of the Health Assessment Questionnaire (HAQ) and other clinical and health status measures in rheumatoid arthritis: results of short-term clinical trials and observational studies versus long-term observational studies. Arthritis Care Res 1992;5:130–6.
31. Talamo J, Frater A, Gallivan S, Young A. Use of the Short Form 36 (SF36) for health status measurement in rheumatoid arthritis. Br J Rheumatol 1997;36:463–9.
32. Brown JH, Kazis LE, Spitz PW, Gertman P, Fries JF, Meenan RF. The dimensions of health outcomes: a cross-validated examination of health status measurement. Am J Public Health 1984;74:159–61.
33. Fitzpatrick R. The measurement of health status and quality of life in rheumatological disorders. Baillieres Clin Rheumatol 1993;7:297–317.
34. Martin DP, Engelberg R, Agel J, Swiontkowski MF. Comparison of the Musculoskeletal Function Assessment questionnaire with the Short Form-36, the Western Ontario and McMaster Universities Osteoarthritis Index, and the Sickness Impact Profile health-status measures. J Bone Joint Surg Am 1997;79-A:1323–35.
35. McHorney CA, Ware JE, Rachel Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40–66.
36. Ruta DA, Hurst NP, Kind P, Hunter M, Stubbings A. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the Short Form 36-item health survey (SF-36). Br J Rheumatol 1998;37:425–36.
37. Kreibich DN, Vaz M, Bourne RB, Rorabeck CH, Kim P, Hardie R, et al. What is the best way of assessing outcome after total knee replacement? Clin Orthop 1996;331:221–5.
38. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992; 30:917–25.
39. Kosinski M, Keller SD, Ware JE, Hatoum HT, Kong SX. The SF-36 health survey (SF-36) as a generic outcome measure in clinical trials of patients with osteoarthritis and rheumatoid arthritis: relative validity of scales in relation to clinical measures of arthritis severity. Med Care 1999;37:MS23–39.
40. Hagen KB, Smedstad LM, Uhlig T, Kvien TK. The responsiveness of health status measures in patients with rheumatoid arthritis: comparison of disease-specific and generic instruments. J Rheumatol 1999;26: 1474–80.
41. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
42. Hochberg MC, Chang RW, Dwosh I, Lindsey S, Pincus T, Wolfe F. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. Arthritis Rheum 1992;35:498–502.
43. Simon LS, Weaver AL, Graham DY, Kivitz AJ, Lipsky PE, Hubbard RC, et al. The anti-inflammatory and upper gastrointestinal effects of celecoxib in rheumatoid arthritis: a randomized, controlled trial. JAMA 1999;282:1921–8.
44. Ramey DR, Fries JF, Singh G. The Health Assessment Questionnaire 1995—status and review. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia: Lippincott-Raven; 1996. p. 227–37.
45. Siddiqui O, Ali MW. A comparison of the random-effects pattern mixture model with last-observation-carried-forward (LOCF) analysis in longitudinal clinical trials with dropouts. J Biopharm Stat 1998;8:545–63.
46. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988;75:800–2.
47. Zhao SZ, McMillen JI, Markenson JA, Dedhiya SD, Zhao WW, Osterhaus JT, et al. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis treated with celecoxib. Pharmacotherapy 1999;19:1269–78.
48. Cronan TA, Shaw WS, Gallagher RA, Weisman M. Predicting health care use among older osteoarthritis patients in an HMO. Arthritis Care Res 1995;8:66–72.