

Rofecoxib Shows Consistent Efficacy in Osteoarthritis Clinical Trials, Regardless of Specific Patient Demographic and Disease Factors

LISA M. DETORA, DAVID KRUPA, JAMES BOLOGNESE, RHODA S. SPERLING, and ELLIOT W. EHRICH

ABSTRACT. Objective. To evaluate the efficacy of rofecoxib (Vioxx™) in subpopulations of patients with osteoarthritis (OA) identified by demographic or baseline disease characteristics, or varied OA involvement.

Methods. Data were combined from three 6-week double blind trials in patients with OA of the knee or hip. All trials contained placebo, 12.5 mg rofecoxib, and 25 mg rofecoxib arms (the only trials to date containing all 3 treatments). Analyses were performed on subgroups categorized according to the following baseline demographics and disease characteristics [age, sex, height, weight, body mass index, American Rheumatism Association (ARA) functional class, joint tenderness, joint stiffness, Western Ontario-McMaster University OA Index (WOMAC) functional subscale, unilateral/bilateral joint involvement, number of joint groups involved]. Three primary endpoints — Pain Walking on Flat Surface (WOMAC), Patient Global Assessment of Response to Therapy, and Investigator Global Assessment of Disease Status — were analyzed. The global assessments, which provided data on overall aspects of OA, regardless of affected joint, were used to assess effects among patients with one, 2, 3, or 4 joint groups affected (from among the following: interphalangeal/first carpal-metacarpal joint, spine, hip, or knee).

Results. Data from 1501 patients were included. No consistent treatment-by-subgroup interaction was observed with all 3 primary endpoints for patients taking placebo or 12.5 or 25 mg rofecoxib. Rofecoxib showed generally consistent efficacy across subgroups of patients identified by sex, race, age, OA location(s), prior OA therapy, baseline study joint tenderness or swelling (patients with knee OA only), and ARA functional class level.

Conclusion. In this combined analysis, no specific factor predicted a differential treatment effect to rofecoxib. (J Rheumatol 2001;28:2494–503)

Key Indexing Terms:
OSTEOARTHRITIS
TREATMENT EFFECT

ROFECOXIB

CYCLOOXYGENASE
SUBGROUP ANALYSIS

American College of Rheumatology (ACR) guidelines for the medical management of osteoarthritis (OA) of the hip and knee recommend the use of analgesics such as acetaminophen, or cyclooxygenase-2 (COX-2) inhibitors, or nonsteroidal antiinflammatory drugs (NSAID) in combination with exercise, education, and social support¹. No specific factors have been shown to predict pharmacologic treatment responses, hence recommendations are generally similar regardless of the type or number of joints affected¹⁻⁶ or the OA disease classification (primary or idiopathic, secondary, localized, generalized, or of specific joints). However, if specific factors that predicted enhanced or

reduced pharmacologic treatment efficacy could be identified, physicians could more easily choose appropriate therapies from among available options⁷⁻⁹.

Rofecoxib (4-[4-(methylsulfonyl) phenyl]-3-phenyl-2(5H)-furanone), Vioxx™, MK-0966, Merck and Co., Inc.), a selective COX-2 inhibitor, at doses of 12.5 or 25 mg once daily has been found in multiple clinical trials to be effective in the longterm treatment of OA¹⁰⁻¹³. Published data on rofecoxib also indicate an improved gastrointestinal (GI) safety profile compared to nonselective NSAID^{14,15}. No data have yet been published about any variable treatment effects among subgroups of OA patients taking rofecoxib. We performed a series of analyses on data from 3 different rofecoxib treatment trials of OA of the knee or hip to identify whether a specific patient subgroup(s) had enhanced or reduced treatment efficacy, based on demographic characteristics, disease criteria, and location of OA.

MATERIALS AND METHODS

The study consisted of a series of analyses conducted on a combined database of patient information derived from studies including rofecoxib 12.5

From Pulmonary-Immunology and Clinical Biostatistics Merck Research Laboratories, Merck & Co., Inc. Rahway, New Jersey, USA.

Supported by a grant from Merck & Co., Inc.

L.M. DeTora, PhD, Pulmonary-Immunology; D. Krupa, MS, Clinical Biostatistics; J.A. Bolognese, MS, Clinical Biostatistics; R.S. Sperling, MD, Pulmonary-Immunology; E.W. Ehrich, MD, Pulmonary-Immunology.

Address reprint requests to Dr. L. DeTora, RY32-305, Box 2000, Rahway, NJ 07065.

Submitted August 14, 2000; revision accepted June 15, 2001.

mg, 25 mg, and placebo treatment groups to determine if any particular subgroup of patients had different levels of efficacy, as assessed by the primary endpoints for the study. Published findings indicate that subgroup analyses should be conducted according to a defined plan, that data from all patients should be used, and that data from smaller trials should be combined to enhance the precision of analyses^{16,17}.

Analysis of a small number of subgroups was prespecified for the rofecoxib OA development program. Prespecified subgroups were characterized by: location of OA, history of serious GI adverse events in relation to prior NSAID therapy, prior therapy, age, sex, and race; analyses were not to be performed if patient numbers were insufficient. In planning for the current analysis, these data were reviewed and additional subgroups were chosen to more fully assess areas for further investigation of differences in efficacy in the existing database, particularly in terms of specific OA disease characteristics.

Studies and patients included in these analyses. Data were combined from each available study in patients with OA of the knee or hip, each containing placebo, 12.5 mg, and 25 mg rofecoxib arms. Three studies met the criteria and were included in the analysis. One trial was a placebo controlled dose-ranging trial (Phase IIb) conducted in the US¹⁰, comparing rofecoxib 5 mg (n = 149), 12.5 mg (n = 144), 25 mg (n = 137), 50 mg (n = 97), and placebo (n = 145). Data were combined with data from 2 identical placebo and active-comparator controlled Phase III clinical trials, one in the US¹¹ and the other in 26 countries¹². These studies compared rofecoxib 12.5 mg (n = 219 US, n = 244 international) and 25 mg (n = 227 US, n = 242 international) to ibuprofen 2400 mg (n = 221 US, n = 249 international) and placebo (n = 69 US, n = 74 international). All 3 trials shared entry criteria and primary and secondary endpoints. Each trial was double blinded and had a 6-week treatment period. The methods for each of the 3 individual trials are published¹⁰⁻¹². Each trial was conducted in accord with ethical considerations regarding the welfare of patients as specified in the Declaration of Helsinki at the time of study initiation (1996, 1996, and 1997, respectively).

Results from individual trials were generally consistent. In this analysis, their data were combined to maximize statistical precision. Only patients who received placebo or rofecoxib 12.5 or 25 mg once daily were included due to a combination of differences in study design and sparseness of data for other doses of rofecoxib or active comparators.

Primary efficacy measurements in the studies. The 3 primary endpoints prespecified in each trial were: Pain Walking on Flat Surface [Western Ontario-McMaster University OA Index (WOMAC), 0–100 mm on visual analog scale (VAS)]¹⁷; Patient Global Assessment of Response to Therapy (0 to 4 point Likert scale); and Investigator Global Assessment of Disease Status (0 to 4 point Likert scale). In each study, Pain Walking on a Flat Surface referred only to the primary study joint, and the Patient and Investigator Global Assessments referred to the patient's overall arthritis symptoms. Secondary endpoints and their results were published¹⁰⁻¹² and included a Patient Global Assessment of Disease Status (0 to 100 mm VAS) and Investigator Global Assessment of Response to Therapy (0 to 4 point Likert scale). The choice of the primary endpoints in these studies was determined by the recommendations of an international rheumatology consensus group (OMERACT) and the Osteoarthritis Research Society (OARS)¹⁸⁻²⁰. In addition, the primary endpoints were highly correlated^{21,22}, and it was expected that no specific endpoint should provide discrepant results, compared with other endpoints.

Patient subgroups in each study and the combined analysis. Subgroups to analyze were chosen based on available data recorded during the general physical examination (such as age or prior therapy), regarding the specific diagnosis of OA (such as tenderness or stiffness in the affected joint), and criteria measured specifically for inclusion in clinical trials (such as baseline WOMAC data). In each trial, patients had OA in a knee or hip, which was specified as the primary study joint. However, baseline data were obtained about the presence or absence of patients' OA symptoms in the interphalangeal, carpometacarpal (hand), or spine joints, and whether these

symptoms were unilateral or bilateral (for all joints). These subgroups thus reflected risk factors for and classification of OA, as well as measures identified by OMERACT^{1,3,18-20,23,24}.

Patient subgroups were categorized according to baseline data for (1) demographic variables of age, sex, weight, body mass index (BMI, calculated as kg/m²), prior OA therapy (NSAID or acetaminophen), and race; and (2) disease characteristics of American Rheumatism Association (ARA) functional class (I, II, or III), duration of OA, study joint tenderness (all patients), study joint swelling (knee only), WOMAC functional subscale, unilateral/bilateral joint involvement, number of joint groups involved, and primary study joint.

The prespecified analysis of patients with a history of GI bleeding or ulcer associated with NSAID use was not included due to the small number of patients who reported such a history (0, 34, and 45 patients in the placebo, 12.5 mg, and 25 mg rofecoxib groups, respectively).

Statistical methods for the combined analysis. In the OA clinical development program for rofecoxib, statistical methods were predefined to assess differences between treatments across various patient subgroups and as an exploratory exercise to identify areas requiring further investigation.

All patients with a baseline measurement and at least one on-treatment value were included. The primary measure of the effect of treatment on the clinical efficacy endpoints was the mean of all data collected over the 6-week treatment period. For endpoints measured at baseline, on-treatment values were adjusted by subtracting baseline values.

An analysis of covariance (ANCOVA) was used to model treatment response as a function of protocol, study center within protocol, treatment, baseline response as a continuous covariate, subgroup factor, and treatment-by-factor interaction. Inclusion of the treatment-by-factor interaction term in the ANCOVA model allows statistical testing of the consistency of differences between treatments across levels of the factor. The significance of the interaction effects was assessed with a formal statistical test based on the F distribution. In these analyses, $p < 0.05$ was considered to be statistically significant.

Additional analyses were performed to assess differences in treatment effect in patients with or without interphalangeal, carpometacarpal, or spine OA in addition to the primary study joint. These analyses were carried out for 2 primary endpoints: Patient Global Assessment of Response to Therapy (0 to 4 Likert scale) and Investigator Global Assessment of Disease Status (0 to 4 Likert scale). Analyses of the secondary endpoints, Patient Global Assessment of Disease Status (0–100 mm VAS) and Investigator Global Assessments of Response to Therapy (0 to 4 Likert scale), were performed to assess consistency.

For reference purposes, an analysis was performed comparing placebo, 12.5 mg, 25 mg rofecoxib and ibuprofen for the 2 active comparator controlled studies for the prespecified subgroups mentioned above, plus ARA functional class, baseline joint tenderness and swelling, and with or without thumb, hand or spine OA.

Since multiple statistical tests were performed, it was expected that one or more significant interactions (i.e., $p \leq 0.050$) would be detected due to chance alone. In addition, when sample sizes for subgroups were small, treatment effect comparisons were not estimated precisely and the ability to detect true differences was low.

In light of these caveats, an additional analysis was performed where the magnitude of differences in treatment effect across subgroups was estimated in a manner that enabled the comparison against clinical comparability criteria. To maximize precision, response was the mean across the 2 rofecoxib dose groups previously shown to have comparable effects¹⁰⁻¹². When the drug effect is expressed as $[(\text{mean}_{12.5\text{mg}} + \text{mean}_{25\text{mg}})/2 - \text{mean}_{\text{PBO}}]$, the difference in drug effect is $[(\text{mean}_{12.5\text{mg}} + \text{mean}_{25\text{mg}})/2 - \text{mean}_{\text{PBO}}]_{\text{Group 1}} - [(\text{mean}_{12.5\text{mg}} + \text{mean}_{25\text{mg}})/2 - \text{mean}_{\text{PBO}}]_{\text{Group 2}}$ [PBO: placebo] If the drug effects in groups 1 and 2 are consistent, this value will be close to 0. A 95% confidence interval is computed for the between-subgroup difference in drug effect to enable assessment of the closeness to 0 and to compare the potential magnitude of differences in treatment effect relative to clinical comparability criteria. This analysis was performed for

all subgroups assessed by 3 primary endpoints (age, race, BMI, sex, prior therapy, ARA class, duration of OA, unilateral vs bilateral involvement, baseline study joint tenderness and swelling, number of joint groups involved, primary study joint, and WOMAC functional subscale), and the results were plotted.

Clinical comparability can be assessed using various bounds or criteria. A difference of 0.5 units or less on the Likert scale and 10 units or less on the visual analog scale was considered clinically equivalent, as in the primary analyses of these trials^{10-12,22}.

RESULTS

Patients included in the analyses. All 3 trials were conducted in men and women between the ages of 38 and 92 years with OA of the hip or knee¹⁰⁻¹². A total of 1501 patients were enrolled into the 3 treatment groups across 3 trials; 1491 patients had a baseline and at least one treatment value for at least one of the primary endpoints. The patient populations studied were predominantly white (69 to 90%), female (71 to 80%), and obese (mean weight 76 to 89 kg), with mean ages varying from 61 to 64 years. The majority

of patients had knee OA (72 to 81%) of long-standing duration (mean duration, 9 to 11 yrs) and were ARA functional class II (62%). Baseline values of the primary endpoints, Pain Walking on a Flat Surface (WOMAC) and Investigator Global Assessment of Disease Status, indicate a similar mean degree of OA symptoms across all treatment groups at the time of randomization (Table 1). The third primary endpoint, Patient Global Assessment of Response to Therapy, was not measured at baseline. In the 3 trials 35, 22, and 37% of patients had OA of the hand, thumb, or spine, respectively, in addition to the primary study joint. Overall, 1003 (67%) of 1491 patients included in these analyses had multiple nonstudy joints affected by OA: 479 (32%) had OA in only the primary study joint group, 538 (36%) in 2 joint groups, 343 (23%) in 3 joint groups, and 122 (8%) in all 4 joint groups. As well, 368 (25%) patients had unilateral joint involvement and 1114 (75%) had bilateral involvement.

Efficacy of rofecoxib in individual patient subgroups. In the analyses described above, treatment effects with rofecoxib

Table 1. Summary of patient baseline characteristics.

	Phase IIb US N = 426 n (%)	Phase III US N = 515 n (%)	Phase III Multinational N = 560 N (%)
12.5 mg rofecoxib group	145 (34)	69 (13)	74 (13)
25 mg rofecoxib group	144 (34)	219 (43)	244 (44)
Placebo group	137 (32)	227 (44)	242 (43)
Sex			
Female	306 (72)	385 (75)	452 (81)
Male	117 (27)	130 (25)	108 (19)
Race			
Other	51 (12)	63 (12)	176 (31)
White	375 (88)	452 (88)	384 (69)
Age, yrs, mean (min to max)	61.9 (38 to 92)	61.0 (39 to 91)	63.4 (40 to 86)
Prior OA therapy			
NSAID	426 (100)	465 (90)	499 (89)
Acetaminophen	0 (0)	50 (10)	61 (11)
Primary study joint			
Hip	138 (32)	124 (24)	118 (21)
Knee	288 (68)	391 (76)	442 (79)
Study joint tenderness (0 to 3 scale), mean	1.98	1.81	1.91
Study joint swelling (knee only)			
Present	168 (39)	244 (47)	212 (38)
Absent	120 (28)	147 (29)	230 (41)
ARA Class*			
I	66 (15)	75 (15)	72 (13)
II	282 (66)	321 (62)	334 (60)
III	78 (18)	119 (23)	154 (27)
Mean duration, yrs	10.4	10.1	8.5
Efficacy endpoints**			
Pain Walking on a Flat Surface (WOMAC), mean (SD)	73.93 (15.82)	74.79 (15.78)	72.45 (14.52)
Min to max	6 to 100	23 to 100	40 to 100
Investigator Global Assessment of Disease Status, mean (SD)	2.93 (0.67)	2.86 (0.66)	2.96 (0.62)
Min to max	1 to 4	1 to 4	2 to 4

* Patients with ARA functional class IV were not permitted to enter the studies. ** No baseline values were obtained for Investigator or Patient Global Assessment of Response to Therapy.

were generally consistent for all 3 primary endpoints across the subgroups of patients categorized by the following criteria: (1) demographic factors: age, sex, race (white vs other), prior therapy (NSAID vs acetaminophen), weight in kg, BMI; (2) disease characteristics: baseline joint tenderness (all patients) or swelling (knee OA only), ARA functional class I, II or III, duration of OA, or primary study joint (hip or knee), WOMAC functional subscale, unilateral or bilateral joint involvement, or the number of joint groups involved.

Other than the few exceptions discussed below, the tests of statistical interaction did not detect any significant dependencies between the treatment effect and patient subgroups. Table 2 shows the results of analyses for Patient Global Assessment of Response to Therapy and Pain Walking on a Flat Surface. Figure 1 indicates that differences in treatment effect were generally within predefined clinical comparability bounds. The between-subgroup difference in drug effect is indicated by the distance to the left or right of zero.

An improved response in patients in the first subgroup listed is indicated by values to the left of zero, and vice versa.

The assessment of patients categorized by the presence or absence of baseline joint swelling (among patients with knee OA only) showed a statistically significant test of interaction ($p < 0.05$), favoring patients without swelling at baseline, for the Patient Global Assessment of Response to Therapy. Although differences in treatment effect tended in the same direction, neither of the 2 remaining endpoints showed a significant interaction.

Patients with 2, 3, or 4 joint groups (from among interphalangeal/carpometacarpal, spine, knee, or hip) affected by OA showed similar treatment effects as assessed by Patient Global Assessments of Response to Therapy and Disease Status. A further analysis of Investigator Global Assessment of Response to Therapy and Investigator Global Assessment of Disease Status yielded results consistent with patient assessments. These results were consistent with the results of an analysis of patients with or without OA of the inter-

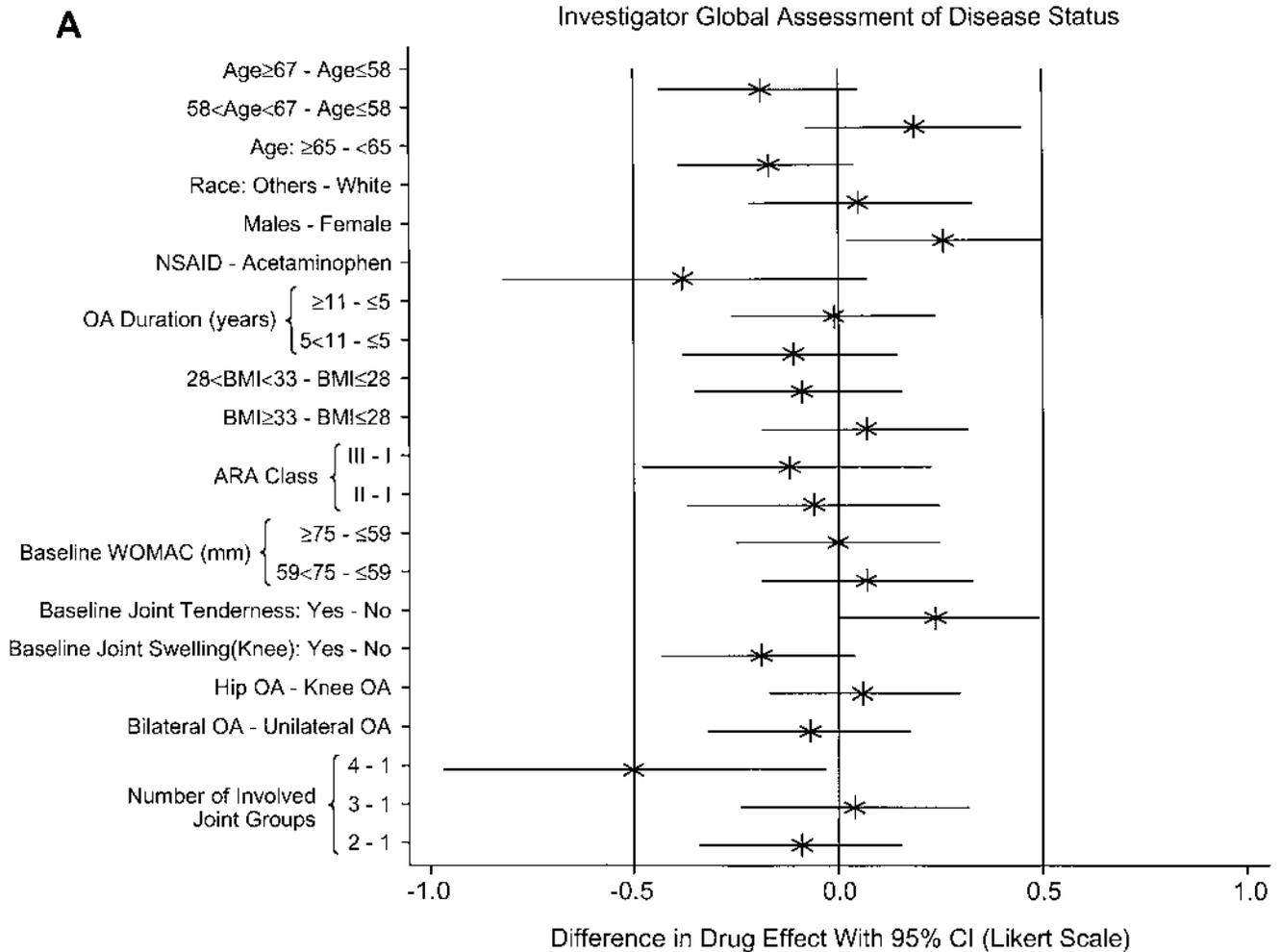


Figure 1A. Summary of results with the primary endpoint "Investigator Global Assessment of Disease Status." Error bars indicate the 95% confidence interval. If error bars do not cross 0, this is an indication of statistically significant differential drug effect across levels of the subgroup factor, and that further investigation is warranted.

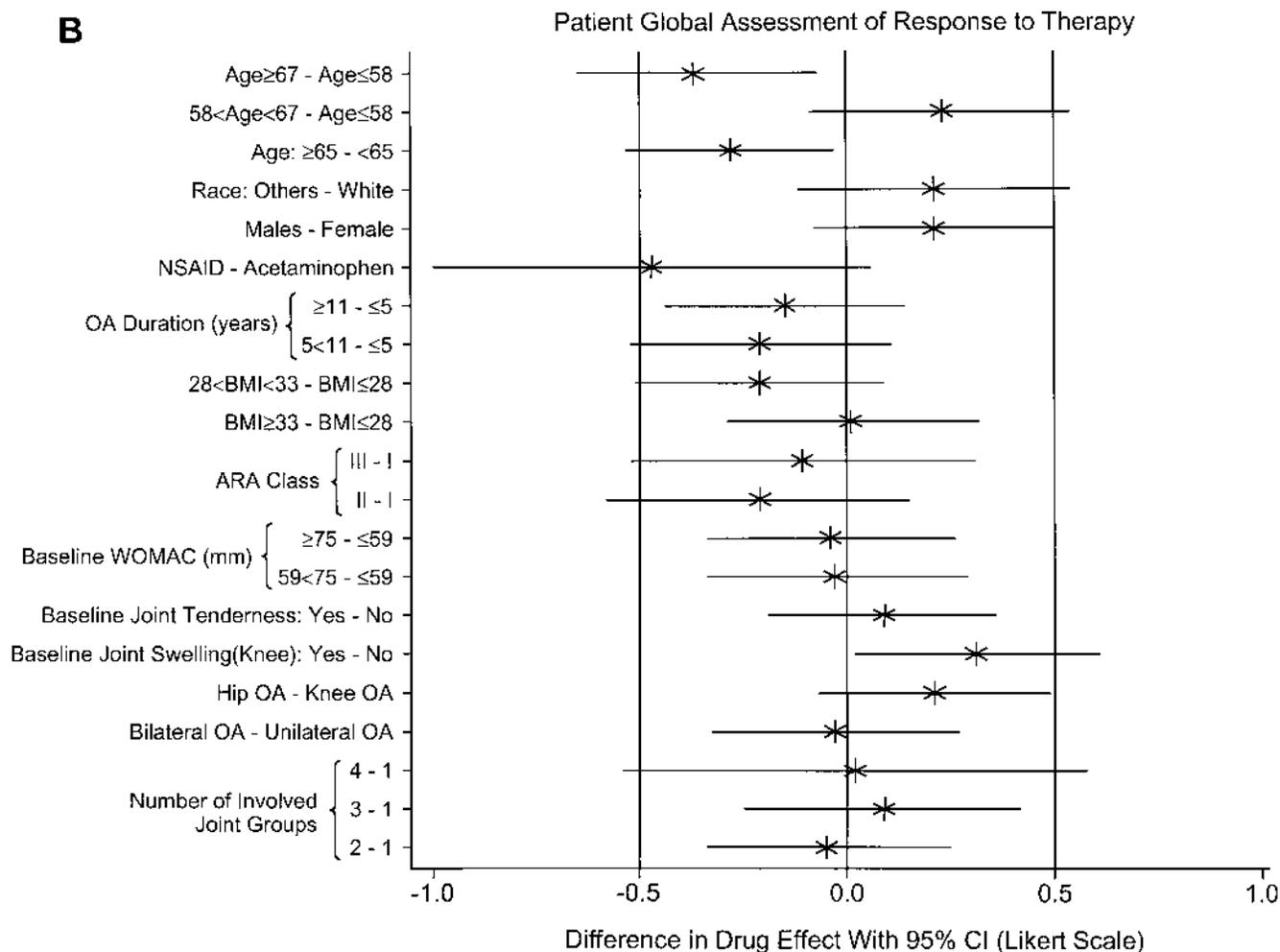
B

Figure 1B. Summary of results with the primary endpoint “Patient Global Assessment of Response to Therapy.” Error bars indicate the 95% confidence interval. If error bars do not cross 0, this is an indication of statistically significant differential drug effect across levels of the subgroup factor, and that further investigation is warranted.

phalangeal or carpometacarpal joints or spine in addition to the primary joint, and of patients with unilateral vs bilateral OA (Table 3).

In these analyses, patients receiving rofecoxib 12.5 mg or 25 mg showed a similar magnitude of response (Table 2). The summary of the results with each of the primary endpoints — Pain Walking on a Flat Surface (WOMAC), Patient Global Assessment of Response to Therapy, and Investigator Global Assessment of Disease Status — were generally similar (Table 2, Figure 1). A small difference in treatment effect was noted, favoring older patients, for 2 endpoints: Patient Global Assessment of Response to Therapy and Pain Walking on a Flat Surface (Figure 1). Further examination of the Patient Global Assessment showed that there was an increased placebo response rate in patients under age 65 years (1.45 Likert units) compared with patients aged 65 years and older (1.13 Likert units). For Pain Walking on a Flat Surface, a difference was noted between patients over age 67 compared with those under

age 58 years; no difference was noted for those under age 65 years compared with those 65 years and older.

For reference, an analysis of the 2 active-comparator controlled studies showed no clinically significant treatment by factor interactions for patients receiving placebo or 12.5 mg or 25 mg rofecoxib or ibuprofen for any of the subgroups studied. As in the analyses for all 3 studies, only one interaction p value ≤ 0.05 was observed, favoring patients in the middle third of the age analysis by thirds, for a single endpoint (Patient Global Assessment of Response to Therapy). This finding was inconsistent with those for age above or below 65 years for the same endpoint or for the other 2 primary endpoints for either analysis.

DISCUSSION

In an analysis designed to assess the consistency of treatment effects in patients with OA, rofecoxib 12.5 mg and 25 mg showed generally consistent efficacy across subgroups of patients categorized by either demographic factors or

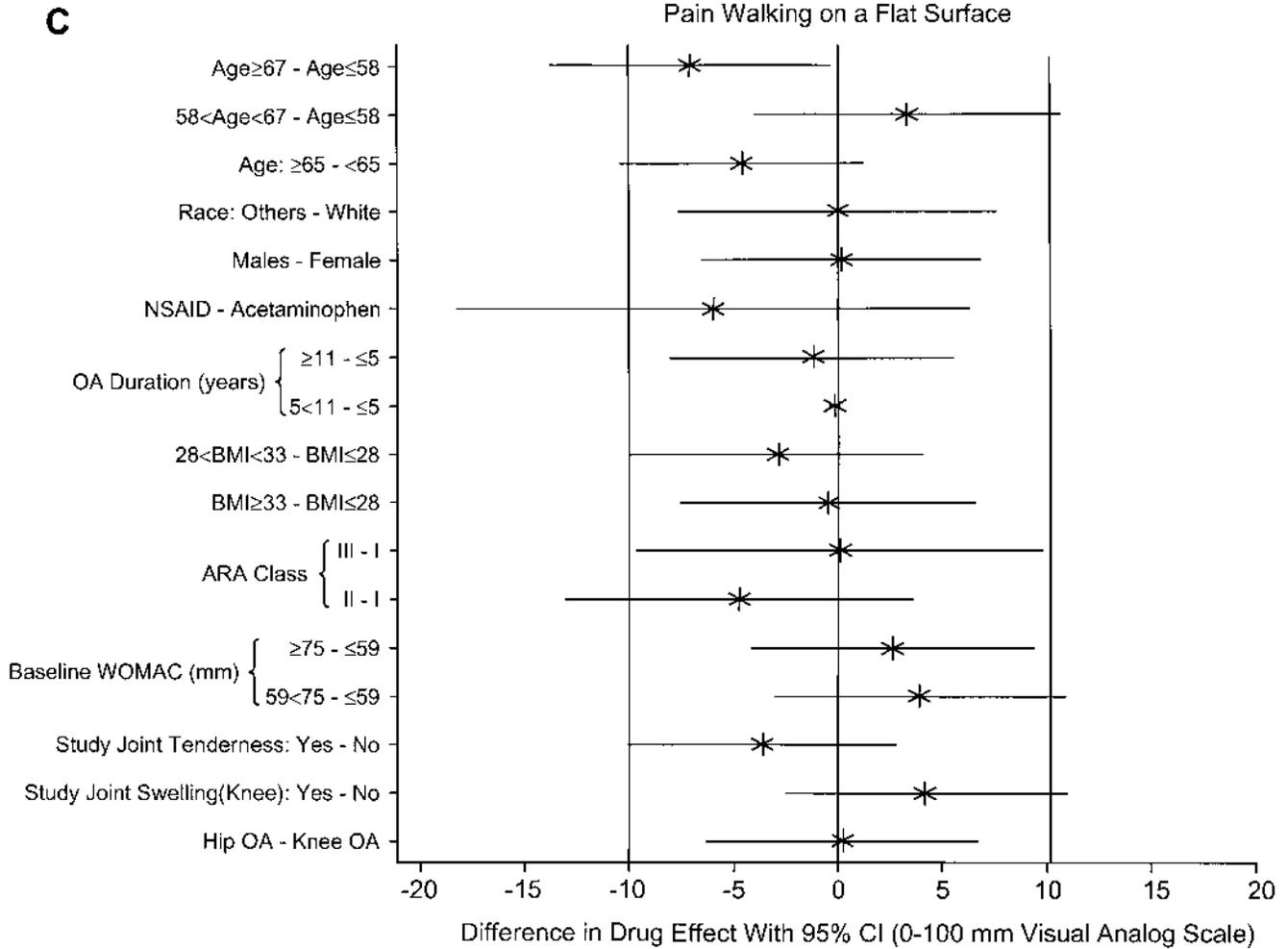


Figure 1C. Summary of results with the primary endpoint “Pain Walking on a Flat Surface.” Error bars indicate the 95% confidence interval. If error bars do not cross 0, this is an indication of statistically significant differential drug effect across levels of the subgroup factor, and that further investigation is warranted.

disease characteristics and location. The results of this analysis showed an overall consistency and correlation between the endpoints for each subgroup factor. Rofecoxib was similarly effective in patients with knee or hip OA, with unilateral or bilateral joint involvement, and in patients with one, 2, 3, or 4 joint groups involved in addition to the primary study joint.

In biostatistical analyses²⁵ of patient subgroups defined by baseline characteristics, emphasis should be placed on exploratory findings via estimation techniques, and p values should not be overinterpreted. Statistical tests of interaction are recommended to assess differential treatment effects across subgroups. However, interaction tests may lack power, and the likelihood of finding statistically significant differences due to chance alone can be high when many subgroups are tested. Thus, effects should be clinically important and consistent across analyses and endpoints to be believed. As noted above, any statistical differences were

isolated and comparison of these differences relative to comparability criteria suggest that they were not clinically meaningful, and thus consistent with the expectation that at least one significant difference might be detected due to chance.

Although the patient populations in each trial were intended to be generally similar, the clinical profiles of the patients in individual trials included in the current analysis showed some differences¹⁰⁻¹². Patients in the international trial¹² weighed 10 to 13 kg less and had OA one to 2 fewer years than patients in the US trials¹¹, and were more likely to be ARA functional class III or Hispanic. These differences, however, did not correspond with differences among the treatment groups for baseline efficacy values, results from the primary studies, or in these analyses. Rather, these differences enhance the applicability of the subgroup analysis results across a wider patient population.

The analyses presented above have certain limitations. A

Table 2A. Patient Global Assessment of Response to Therapy (Likert) — Results from the primary endpoints, stratified by subgroups. Effect difference from placebo for 12.5 and 25 mg rofecoxib. Results should be interpreted with caution due to small sample sizes in some subgroups.

Subgroup	Sample Size			Difference from Placebo (95% CI)		Interaction Test p
	Placebo	12.5 mg	25 mg	12.5 mg–Placebo	25 mg–Placebo	
Baseline demographics						
Age, yrs						
< 65	156	353	316	-0.7 (-0.9, -0.6)	-0.9 (-1.1, -0.7)	0.081
≥ 65	127	250	277	-1.0 (-1.3, -0.8)	-1.1 (-1.3, -0.9)	
Race						
White	236	478	477	-0.9 (-1.1, -0.8)	-1.0 (-1.2, -0.9)	0.459
Other	47	125	116	-0.7 (-1.0, -0.4)	-0.8 (-1.1, -0.5)	
Thirds of Body Mass Index						
≤ 28	101	206	233	-0.7 (-0.9, -0.5)	-1.0 (-1.3, -0.8)	0.138
28–33	88	199	192	-0.9 (-1.1, -0.6)	-0.9 (-1.1, -0.6)	
≥ 33	94	193	167	-1.1 (-1.3, -0.8)	-1.1 (-1.3, -0.9)	
Sex						
Female	214	465	445	-1.0 (-1.1, -0.8)	-1.0 (-1.2, -0.9)	0.068
Male	69	138	148	-0.6 (-0.9, -0.3)	-0.9 (-1.2, -0.7)	
Disease characteristics						
Prestudy analgesic therapy						
Acetaminophen	16	44	49	-0.4 (-0.9, 0.2)	-0.6 (-1.2, -0.1)	0.185
NSAID	267	559	544	-0.9 (-1.1, -0.8)	-1.0 (-1.2, -0.9)	
ARA functional class						
I	42	83	86	-0.7 (-1.0, -0.3)	-0.9 (-1.3, -0.5)	0.691
II	178	369	375	-0.9 (-1.1, -0.8)	-1.0 (-1.2, -0.9)	
III	63	151	132	-0.8 (-1.1, -0.5)	-0.9 (-1.2, -0.7)	
Duration of OA, yrs						
≤ 5	103	229	228	-0.7 (-1.0, -0.5)	-0.9 (-1.1, -0.7)	0.374
5–11	81	160	154	-1.1 (-1.3, -0.8)	-1.0 (-1.3, -0.8)	
≥ 11	97	210	204	-0.9 (-1.1, -0.7)	-1.1 (-1.3, -0.8)	
Side of body						
Unilateral	64	145	160	-0.8 (-1.1, -0.6)	-1.0 (-1.3, -0.7)	0.956
Bilateral	219	458	433	-0.9 (-1.0, -0.7)	-1.0 (-1.2, -0.8)	
Baseline swelling of primary study joint (knee OA only)*						
No	82	222	189	-1.2 (-1.4, -0.9)	-1.2 (-1.4, -0.9)	0.035
Yes	127	237	254	-0.8 (-1.0, -0.6)	-1.0 (-1.2, -0.7)	
Baseline tenderness of primary study joint						
No	79	175	155	-1.0 (-1.2, -0.7)	-1.0 (-1.3, -0.8)	0.758
Yes	202	428	438	-0.8 (-1.0, -0.7)	-1.0 (-1.1, -0.8)	
Number of joint groups involved						
1	92	193	194	-0.8 (-1.1, -0.6)	-1.1 (-1.3, -0.8)	0.390
2	106	220	211	-1.0 (-1.2, -0.7)	-1.0 (-1.2, -0.8)	
3	70	134	136	-0.9 (-1.1, -0.6)	-0.9 (-1.1, -0.6)	
4	15	56	52	-0.7 (-1.3, -0.2)	-1.1 (-1.7, -0.6)	
Knee or hip as primary study joint						
Knee	208	458	445	-0.9 (-1.1, -0.8)	-1.0 (-1.2, -0.9)	0.274
Hip	75	145	148	-0.7 (-1.0, -0.4)	-0.9 (-1.1, -0.6)	
Thirds of baseline WOMAC functional subscale						
≤ 59	97	213	183	-0.9 (-1.1, -0.7)	-0.9 (-1.2, -0.7)	0.670
59–75	83	201	202	-0.9 (-1.1, -0.6)	-1.0 (-1.2, -0.8)	
≥ 75	101	188	208	-0.8 (-1.1, -0.6)	-1.1 (-1.3, -0.8)	

* For interaction test, $p < 0.05$ for patients without baseline swelling vs those with baseline swelling.

planned analysis of patients with or without prior GI ulceration was not completed due to small sample size. Detailed analysis of patients by individual race categories could not be done as very few patients of Asian, Native American, or mixed race groups were enrolled in these trials; these groups were combined into the “other” category with patients of Black or Hispanic race. Further study will be required to

obtain detailed information about pharmacodynamic differences in these groups. Some analyses included groups that were widely disparate in size, as in the prior therapy analysis, where only about 10% of patients enrolled were prior acetaminophen users (Table 1). This analysis should therefore be interpreted cautiously.

Although, as noted above, no clinically significant differ-

Table 2B. Pain Walking on a Flat Surface (WOMAC VAS) — Results from the primary endpoints, stratified by subgroups. Effect difference from placebo for 12.5 and 25 mg rofecoxib. Results should be interpreted with caution due to small sample sizes in some subgroups.

Subgroup	Placebo	Sample Size		Difference from Placebo (95% CI)		Interaction Test p
		12.5 mg	25 mg	12.5 mg–Placebo	25 mg–Placebo	
Baseline demographics						
Age, yrs						
< 65	156	353	317	-11.5 (-15.7, -7.2)	-14.3 (-18.6, -10.0)	0.250
≥ 65	126	251	277	-16.9 (-21.7, -12.1)	-18.1 (-22.8, -13.4)	
Race						
White	235	479	478	-13.6 (-17.1, -10.1)	-16.3 (-19.8, -12.8)	0.554
Other	47	125	116	-15.3 (-22.8, -7.8)	-14.6 (-22.1, -7.0)	
Thirds of Body Mass Index						
≤ 28	100	207	233	-10.8 (-16.2, -5.5)	-16.9 (-22.1, -11.6)	0.565
28–33	88	199	192	-14.8 (-20.4, -9.2)	-13.9 (-19.5, -8.3)	
≥ 33	94	193	168	-16.8 (-22.3, -11.2)	-16.8 (-22.4, -11.1)	
Sex						
Female	214	466	446	-14.8 (-18.4, -11.1)	-15.2 (-18.8, -11.6)	0.079
Male	68	138	148	-11.2 (-17.7, -4.7)	-18.4 (-24.9, -12.0)	
Disease characteristics						
Prestudy analgesic therapy						
Acetaminophen	16	44	49	-8.0 (-20.9, 4.8)	-10.6 (-23.2, 2.0)	0.634
NSAID	266	560	545	-14.3 (-17.5, -11.0)	-16.3 (-19.6, -13.0)	
ARA functional class						
I	42	83	86	-10.3 (-18.6, -2.0)	-13.7 (-21.9, -5.5)	0.487
II	177	370	375	-15.5 (-19.5, -11.5)	-18.0 (-22.0, -14.0)	
III	63	151	133	-12.1 (-18.7, -5.5)	-11.8 (-18.5, -5.0)	
Duration of OA, yrs						
≤ 5	103	230	229	-12.3 (-17.4, -7.1)	-14.2 (-19.4, -9.0)	0.558
5–11	81	160	154	-17.7 (-23.7, -11.8)	-19.1 (-25.1, -13.1)	
≥ 11	96	210	204	-12.9 (-18.3, -7.5)	-15.9 (-21.3, -10.5)	
Side of Body						
Unilateral	63	146	160	-12.3 (-18.9, -5.8)	-15.0 (-21.5, -8.5)	0.868
Bilateral	219	458	434	-14.4 (-18.0, -10.8)	-16.3 (-20.0, -12.6)	
Baseline swelling of primary study joint (knee OA only)						
No	82	222	188	-17.9 (-23.5, -12.3)	-16.5 (-22.3, -10.8)	0.081
Yes	127	237	255	-10.9 (-15.7, -6.2)	-15.1 (-19.8, -10.4)	
Baseline tenderness of primary study joint						
No	79	175	154	-11.3 (-17.3, -5.4)	-13.3 (-19.4, -7.2)	0.537
Yes	202	429	440	-14.9 (-18.6, -11.1)	-17.1 (-20.8, -13.4)	
Number of joint groups involved						
1	91	193	193	-16.0 (-21.6, -10.5)	-19.2 (-24.8, -13.6)	0.629
2	106	220	212	-14.6 (-19.8, -9.5)	-15.0 (-20.2, -9.9)	
3	70	135	137	-11.2 (-17.7, -4.7)	-12.5 (-19.0, -6.0)	
4	15	56	52	-9.5 (-22.2, 3.2)	-16.2 (-29.1, -3.3)	
Knee or hip as primary study joint						
Knee	208	458	445	-14.1 (-17.8, -10.5)	-15.9 (-19.6, -12.2)	0.912
Hip	74	146	149	-13.3 (-19.5, -7.1)	-16.3 (-22.5, -10.1)	
Thirds of Baseline WOMAC functional subscale						
≤ 59	97	214	183	-17.3 (-22.5, -12.1)	-16.7 (-22.0, -11.3)	0.522
59–75	83	201	203	-12.1 (-17.6, -6.5)	-14.1 (-19.6, -8.5)	
≥ 75	101	188	208	-11.6 (-16.8, -6.3)	-17.1 (-22.3, -12.0)	

For interaction test, $p < 0.05$ for patients without baseline swelling vs those with baseline swelling.

ences in treatment effects were noted for patients based on age, further study of the effects of rofecoxib in older patients might be warranted. This is especially important considering that OA progressively affects the majority of the population over age 65 years and worsens with age, and that no remittive therapies exist. A study comparing rofecoxib 12.5 mg,

25 mg, placebo, and naproxen was conducted for 6 weeks in 341 patients age 80 and over²⁶.

The recent availability of selective COX-2 inhibitors (Coxibs) like rofecoxib, has 2 implications when considered in the context of the known variable treatment responses of OA patients to different NSAID. First, selective COX-2

Table 3. Results of analysis of patients with and without OA of the spine, interphalangeal, or thumb carpal joint. Results should be interpreted with caution due to small sample sizes in some subgroups.

Investigator Global Assessment of Disease Status (0 to 4 point Likert scale)						
Category	Sample Sizes			Difference from Placebo (95% CI)		Interaction p
	Placebo	12.5 mg	25 mg	12.5 mg-Placebo	25 mg-Placebo	
All patients	281	603	598	-0.56 (-0.68, -0.45)	-0.64 (-0.76, -0.53)	—
Presence of spine OA						
No	178	385	366	-0.55 (-0.69, -0.40)	-0.63 (-0.77, -0.48)	0.972
Yes	103	218	232	-0.58 (-0.78, -0.38)	-0.63 (-0.83, -0.43)	
Presence of interphalangeal OA						
No	200	382	396	-0.51 (-0.65, -0.37)	-0.63 (-0.77, -0.49)	0.341
Yes	81	221	202	-0.67 (-0.88, -0.46)	-0.68 (-0.89, -0.46)	
Presence of thumb carpal OA						
No	222	465	469	-0.52 (-0.66, -0.39)	-0.66 (-0.79, -0.53)	0.152
Yes	59	138	129	-0.71 (-0.96, -0.45)	-0.64 (-0.89, -0.38)	

Patient Global Assessment of Response to Therapy (0 to 4 point Likert scale)						
Category	Sample Size			Difference from Placebo (95% CI)		Interaction p
	Placebo	12.5 mg	25 mg	12.5 mg-Placebo	25 mg-Placebo	
All patients	283	603	593	-0.88 (-1.01, -0.74)	-1.00 (-1.14, -0.86)	
Presence of spine OA						
No	180	385	364	-0.91 (-1.08, -0.74)	-1.04 (-1.21, -0.86)	0.620
Yes	103	218	229	-0.76 (-1.00, -0.52)	-0.90 (-1.13, -0.66)	
Presence of interphalangeal OA						
No	201	382	392	-0.89 (-1.06, -0.73)	-1.06 (-1.23, -0.89)	0.304
Yes	82	221	201	-0.88 (-1.14, -0.63)	-0.93 (-1.19, -0.67)	
Presence of thumb carpal OA						
No	224	465	465	-0.83 (-0.99, -0.68)	-1.03 (-1.18, -0.87)	0.139
Yes	59	138	128	-0.98 (-1.30, -0.66)	-0.94 (-1.27, -0.62)	

inhibitors may offer another treatment option for patients who do not respond to or tolerate NSAID. Second, there exists a possibility for patients to also have variable responses to selective COX-2 inhibitors. The series of analyses presented above was undertaken in response to known information about the variable efficacy of NSAID in individual patients. Individual patients may have variable responses to different NSAID, and factors for predicting this variability have not been identified in a population with OA¹⁰⁻¹². The results of the current analyses indicate that no specific factor will predict a differential response to rofecoxib among patients with OA, but that further study of particular groups may be warranted.

REFERENCES

- Altman R, Hochberg M, Moskowitz R, et al. Recommendations for the medical management of osteoarthritis of the hip and knee (American College of Rheumatology Subcommittee on Osteoarthritis Guidelines). *Arthritis Rheum* 2000;43:1905-15.
- Dieppe P, Buckwalter JA. Osteoarthritis and related disorders — management of limb joint osteoarthritis. In: Klippel JH, Dieppe PA, Arnett FC, et al, editors. *Rheumatology*. 2nd ed. London: Mosby; 1998:9.1-9.10.
- Creamer P, Hochberg M. Osteoarthritis. *Lancet* 1997;350:503-9.
- Brandt KD. Should nonsteroidal anti-inflammatory drugs be used to treat osteoarthritis. *Controversies Clin Rheumatol* 1993;19:29-44.
- Bradley JD, Brandt KD, Katz BP, Kalasinski LA, Ryan SI. Treatment of knee osteoarthritis: relationship of clinical features of joint inflammation to the response to a nonsteroidal antiinflammatory drug or pure analgesic. *J Rheumatol* 1992;19:1950-4.
- Altman R, Asch D, Bloch DA, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. *Arthritis Rheum* 1986;29:1039-49.
- Green GS, Green ES. Nonsteroidal therapy of rheumatoid arthritis and osteoarthritis: how physicians manage treatment failures. *J Rheumatol* 1998;25:2088-93.
- Scholes D, Stergchis A, Penna PM, Normand EH, Hansten PD. Nonsteroidal antiinflammatory drug discontinuation in patients with osteoarthritis. *J Rheumatol* 1995;22:708-12.
- Walker JS, Sheather-Reid RB, Carmody JJ, Vial JH, Hay RO. Nonsteroidal antiinflammatory drugs in rheumatoid arthritis and osteoarthritis: support for the concept of “responders” and “nonresponders.” *Arthritis Rheum* 1997;40:1944-54.
- Ehrich E, Schnitzer T, Kivitz A, et al. MK-966, a highly selective COX-2 inhibitor, was effective in the treatment of osteoarthritis of the knee and hip in a 6-week placebo controlled study [abstract]. *Arthritis Rheum* 1997;40 Suppl:S85.
- Saag K, van der Heijde D, Fisher C, et al. Rofecoxib, a new cyclooxygenase 2 inhibitor, shows sustained efficacy, comparable with other nonsteroidal anti-inflammatory drugs: a 6-week and a 1-year trial in patients with osteoarthritis. *Arch Family Med* 2000;9:1124-34.

12. Day R, Morrison B, Luza A, et al. A randomized trial of the efficacy and tolerability of the COX-2 inhibitor rofecoxib vs ibuprofen in patients with osteoarthritis. *Arch Intern Med* 2000;160:1781-7.
13. Cannon GW, Caldwell JR, Holt P, et al. Rofecoxib, a specific inhibitor of cyclooxygenase 2, with clinical efficacy comparable with that of diclofenac sodium: results of a one-year, randomized, clinical trial in patients with osteoarthritis of the knee and hip. *Arthritis Rheum* 2000;43:978-87.
14. Bombardier C, Laine L, Reicin A, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. *N Eng J Med* 2000;343:1520-8.
15. Langman MJ, Jensen DM, Watson DJ, et al. Adverse upper gastrointestinal effects of rofecoxib compared with NSAIDs. *JAMA* 1999;282:1929-33.
16. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. *Lancet* 1998;351:47-52.
17. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Controlled Clin Trials* 1997;18:580-93.
18. Altman R, Brandt K, Hochberg M, et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. *Osteoarthritis Cartilage* 1996;4:217-43.
19. Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III trials in knee, hip, and hand osteoarthritis. Consensus Development at OMERACT III. *J Rheumatol* 1997;24:799-802.
20. Altman R, Brandt K, Hochberg M, et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. *Osteoarthritis Cartilage* 1996;4:217-43.
21. Bolognese J, Ehrich E, Schnitzer T, Zeng Q. Precision of composite measures of osteoarthritis efficacy compared with that of individual endpoints [abstract]. *Arthritis Rheum* 1998;41 Suppl:S361.
22. Ehrich EW, Davies GM, Watson DJ, et al. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities Osteoarthritis Index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000;27:2635-41.
23. Bullough PG. Osteoarthritis and related disorders — pathology. In: Klippel JH, Dieppe PA, Arnett FC, et al, editors. *Rheumatology*. 2nd ed. London: Mosby; 1998:8.1-8.8.
24. Cooper C. Osteoarthritis and related disorders — epidemiology. In: Klippel JH, Dieppe PA, Arnett FC, et al, editors. *Rheumatology*. 2nd ed. London: Mosby; 1998:2.1-2.8.
25. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355:1064-9.
26. Truitt KE, Sperling RS, Ettinger WH, et al. A multicenter, randomized, controlled trial to evaluate the safety profile, tolerability, and efficacy of rofecoxib in advanced elderly patients with osteoarthritis. *Aging Clin Exp Res* 2001;13:112-21.