

EXTENDED REPORT

Patient Preference for Placebo, Acetaminophen (paracetamol) or Celecoxib Efficacy Studies (PACES): two randomised, double blind, placebo controlled, crossover clinical trials in patients with knee or hip osteoarthritis

T Pincus, G Koch, H Lei, B Mangal, T Sokka, R Moskowitz, F Wolfe, A Gibofsky, L Simon, S Zlotnick, J G Fort



Ann Rheum Dis 2004;**63**:931–939. doi: 10.1136/ard.2003.020313

See end of article for authors' affiliations

Correspondence to: Professor T Pincus, Division of Rheumatology and Immunology, Vanderbilt University School of Medicine, 203 Oxford House, Box 5, Nashville, TN 37232-4500, USA; t.pincus@vanderbilt.edu

Accepted 10 March 2004
Published online first
13 April 2004

Background: Acetaminophen (paracetamol) is recommended as the initial pharmacological treatment for knee or hip osteoarthritis. However, survey and clinical trial data indicate greater efficacy for non-steroidal anti-inflammatory drugs and cyclo-oxygenase-2 specific inhibitors.

Design: Two randomised, double blind, placebo controlled, crossover multicentre clinical trials, Patient Preference for Placebo, Acetaminophen or Celecoxib Efficacy Studies (PACES).

Patients: Osteoarthritis of knee or hip.

Intervention: "Wash out" of treatment; randomisation; 6 weeks of celecoxib 200 mg/day, acetaminophen 1000 mg four times a day, or placebo; second "wash out;" crossover to 6 weeks of second treatment.

Measurements: Western Ontario McMaster Osteoarthritis Index (WOMAC), visual analogue pain scale, patient preference between two treatments.

Results: Celecoxib was more efficacious than acetaminophen in both periods in both studies; WOMAC and pain scale scores differed at $p < 0.05$ in period II and both periods combined of PACES-a and in periods I and II and both periods combined in PACES-b, but not in period I of PACES-a. Acetaminophen was more efficacious than placebo, generally $p < 0.05$ in PACES-b, and > 0.05 in PACES-a. Patient preferences were 53% celecoxib v 24% acetaminophen in PACES-a ($p < 0.001$) and 50% v 32% in PACES-b ($p = 0.009$); 37% acetaminophen v 28% placebo in PACES-a ($p = 0.340$) and 48% v 24% in PACES-b ($p = 0.007$). No clinically or statistically significant differences were seen in adverse events or tolerability among the three treatment groups.

Conclusions: Greater efficacy was seen for celecoxib v acetaminophen v placebo, while adverse events and tolerability were similar. Variation in results and statistical significance in the two different trials are of interest.

Acetaminophen (paracetamol) is recommended as the initial drug treatment for patients with osteoarthritis of the knee or hip,^{1–3} based on: (a) a high level of gastrointestinal adverse events with non-specific non-steroidal anti-inflammatory drugs (NSAIDs), the primary treatment until the 1990s^{4–5}; (b) a perception that acetaminophen is safer than these traditional agents⁶; (c) clinical trials which were interpreted to indicate clinical equivalence of acetaminophen and ibuprofen⁷ or naproxen.⁸

In recent years the possible advantages of acetaminophen for most patients with osteoarthritis have been questioned.^{9–11} Two patient surveys indicated that most patients rated NSAIDs as a better treatment than acetaminophen.^{12–13} A randomised, double blind, crossover clinical trial indicated that 57% of patients reported a preference for diclofenac/misoprostol versus acetaminophen, while 21% preferred acetaminophen and 22% expressed no preference.¹⁴ Reanalysis of clinical trials which were interpreted to show similar efficacy of acetaminophen and non-steroidal anti-inflammatory drugs^{7–8} indicated greater effect sizes for ibuprofen and naproxen versus acetaminophen, similar to diclofenac/misoprostol, suggesting that absence of statistical significance resulted from low numbers of patients.⁹

A recent report indicated little efficacy of acetaminophen,¹⁵ although this study also included relatively few patients. Acetaminophen in doses of 4 g/day, as recommended for osteoarthritis by the American College of Rheumatology³ and other organisations, has been associated with a rate of gastrointestinal adverse events similar to standard NSAIDs,^{16–17} although these observations are confounded, in part, by a likelihood that acetaminophen is given to patients at highest risk for gastrointestinal events.

An important recent advance has been the development of cyclo-oxygenase-2 specific inhibitors, which have efficacy similar to that of non-specific NSAIDs, but fewer gastrointestinal adverse events.^{18–19} A recent study indicated significantly greater efficacy of rofecoxib than of acetaminophen in patients with osteoarthritis, with comparable adverse events.²⁰ In this report we present the results of two randomised, placebo controlled, two period, three treatment, crossover, double blind clinical trials, Patient Preference for

Abbreviations: MDHAQ, Multidimensional Health Assessment Questionnaire; NSAIDs, non-steroidal anti-inflammatory drugs; SF-36, Short Form-36; WOMAC, Western Ontario McMaster Osteoarthritis Index

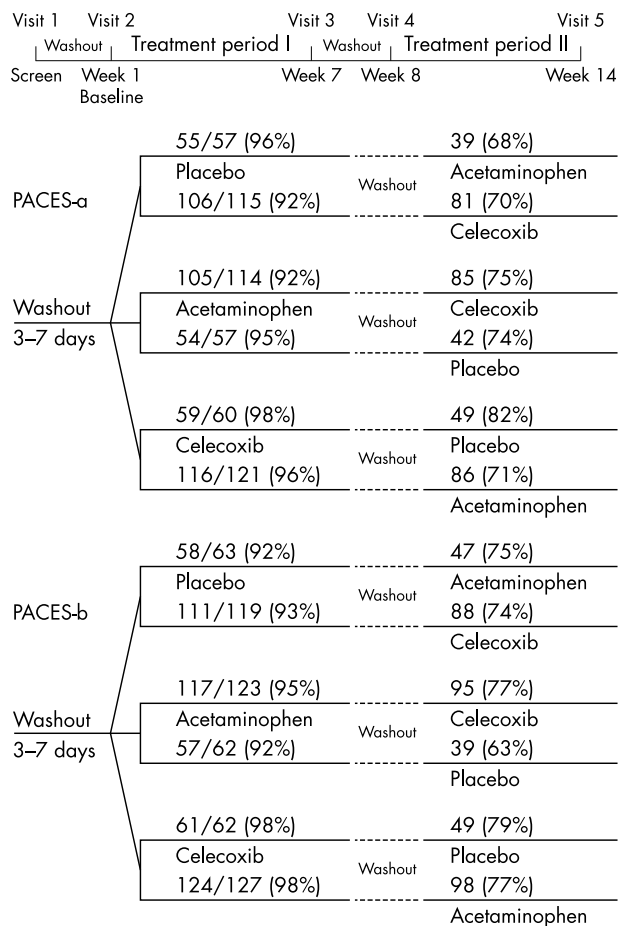


Figure 1 Summary of protocol of PACES—double blind, double dummy, clinical trials. Each patient received a drug and a placebo or two placebos in each period.

Placebo, Acetaminophen or Celecoxib Efficacy Studies (PACES), in which patients were given two of the three treatments. Results of both studies are similar, although variation of results and statistical significance in the two studies are of interest.

METHODS
Study design

PACES-a and PACES-b had an identical design—a two period, double blind, double dummy, crossover trial of celecoxib, acetaminophen, and placebo in patients with osteoarthritis of the knee or hip. Inclusion criteria were age 45 or greater, radiographic Kellgren-Lawrence grade 2–4, a score of 40–90 mm on a visual analogue pain scale, and designation by the treating physician that the patient was a candidate for long term treatment with a cyclo-oxygenase-2 specific inhibitor drug or an analgesic drug. The primary exclusion criteria were significant medical comorbidities, rheumatoid arthritis or other inflammatory arthritis, acute joint trauma, chronic pain syndrome, expected need for surgery during the course of the study, oral or parenteral corticosteroids within 2 months, or intra-articular injections of hyaluronic acid within 9 months. Women of childbearing potential were required to use contraception; pregnant or lactating women were excluded. These studies were approved by the institutional review boards of all participating study sites. All patients signed written informed consent before participation.

Patients were assigned randomly to one of six treatment sequence groups to receive a sequence of two of three treatments, each for 6 weeks (fig 1), celecoxib 200 mg/day—taken in the morning, acetaminophen 1000 mg four times a day, or placebo. Each patient received a drug and a placebo or two placebos in each period. All assessments were conducted using patient and assessor questionnaires, sent by facsimile to the Pharmacia (Pfizer) data centre in Markham, Ontario, Canada. Data were entered into a clinical database using character recognition software; each entry was verified manually by a clinical data validation specialist and double checked by a senior validation specialist. Any discrepancies were sent to the site assessors for clarification and verification.

Visit 1 was a screening visit to review inclusion and exclusion criteria and obtain informed consent. Patient volunteers were given an extensive list of NSAIDs and analgesic drugs to discontinue and not to take throughout the study, and began a 3–7 day washout period. Propoxyphene (Darvon) 65 mg up to four times per day was given as a rescue analgesic drug; codeine 60 mg or tramadol (Ultram) 100 mg, up to four times per day, were provided as alternatives to fewer than 5% of patients if propoxyphene was poorly tolerated or ineffective. Patients were instructed not to take any rescue drug within 12 hours of any visit.

Visit 2 was conducted 3–7 days after visit 1 to randomise patients and provide the 6 weeks of study drug and placebo or two placebos for period I. Visit 3 occurred at the end of the first 6 week treatment period to evaluate period I study drug. Patients who discontinued in period I before completion of 6 weeks had a visit 3 assessment at the time of discontinuation, and were invited to participate in treatment period II. Patients who continued had a second 3–7 day washout period before treatment period II. Visit 4 occurred 3–7 days after visit 3, at which time 6 weeks of the second study drug and placebo or two placebos for period II were provided. Visit 5 occurred at the end of period II, or earlier if the patient chose not to complete period II, to evaluate period II study drug. Patients were queried at visit 5 about their overall preference between the two treatments periods.

Patient and assessor questionnaires

Patients completed a standard eight page questionnaire at each visit, which included a “disease-specific” Western Ontario McMaster Osteoarthritis Index (WOMAC),^{21, 22} which contains 24 100 mm visual analogue scales to assess pain, stiffness, and function, including a general WOMAC and a WOMAC directed at the primary affected joint indicated by the patient, a “general arthritis” Multidimensional Health Assessment Questionnaire (MDHAQ),^{23–26} which includes 100 mm visual analogue scales to assess pain, global status, fatigue, and gastrointestinal distress; and a “generic” Short Form-36 (SF-36) health survey.^{27, 28}

Patients also completed a three page questionnaire unique to each visit to query general status and possible adverse events. The unique patient questionnaire for the final visit 5 included a query: “Please compare control of your arthritis during the first and second study periods,” with five response options, “much better during the first study period,” “better during the first study period,” “no difference between the first and second study periods,” “better during the second study period,” “much better during the second study period.” Responses of “much better” or “better” were merged in the analyses presented in this report.

Investigators completed unique questionnaires at each visit, which included the investigator assessment of the patient’s global status and change in global status on a 100 mm visual analogue scale, as well as a standard report

Table 1 Demographic, disease, patient questionnaire, and assessor questionnaire measures at screening by sequence group in patients randomised to receive celecoxib, acetaminophen or placebo in PACES 2 period, three treatment (six sequence) crossover clinical trials.

Variable	Celecoxib-acetaminophen (n = 121)	Acetaminophen-celecoxib (n = 114)	Celecoxib-placebo (n = 60)	Placebo-celecoxib (n = 115)	Acetaminophen-placebo (n = 57)	Placebo-acetaminophen (n = 57)	p Value
A. PACES-a							
<i>Demographic measures</i>							
Age (years)	63.4 (0.92)	63.5 (0.94)	65.5 (1.28)	62.5 (0.96)	63.7 (1.22)	62.8 (1.30)	0.583
Female (%)	66.1	64	51.7	62.6	59.6	68.4	0.445
Caucasian (%)	90.1	86.0	88.3	88.7	86.0	84.2	0.848
≤ 12 Years of education (%)	50.0	48.2	44.8	42.9	61.8	50.0	0.329
<i>Osteoarthritis measures</i>							
Duration of disease (years)	8.4 (0.74)	10.6 (0.80)	10.7 (1.29)	9.0 (0.85)	8.5 (0.97)	8.1 (1.09)	0.093
Radiographic grade (0–4)	2.77 (0.07)	2.76 (0.06)	2.67 (0.09)	2.79 (0.06)	2.75 (0.09)	2.72 (0.10)	0.916
Knee as index joint (%)	85.7	87.5	82.7	81.4	78.5	82.1	0.529
% taking NSAID for OA	68.6	72.8	66.7	67.0	73.7	75.4	0.318
WOMAC target joint (0–100)	47.0 (1.91)	51.6 (1.82)	41.2 (2.30)	45.6 (1.87)	45.0 (2.81)	45.9 (2.67)	0.031
MDHAQ pain VAS (0–100)	53.6 (1.72)	56.5 (1.95)	51.1 (2.47)	55.1 (1.78)	53.8 (2.79)	54.6 (3.16)	0.577
MDHAQ GI distress (0–100)	12.8 (1.70)	19.2 (2.29)	13.5 (2.69)	14.8 (1.81)	10.9 (2.30)	18.7 (3.57)	0.074
	(n = 127)	(n = 123)	(n = 62)	(n = 119)	(n = 62)	(n = 63)	
B. PACES-b							
<i>Demographic measures</i>							
Age (years)	63.9 (0.89)	63.4 (0.80)	62.7 (1.28)	62.8 (0.90)	64.8 (1.31)	63.4 (1.06)	0.708
Female (%)	66.1	63.4	71.0	71.4	62.9	57.1	0.410
Caucasian (%)	86.6	83.7	79.0	90.8	83.9	85.7	0.350
≤ 12 Years of education (%)	46.4	45.5	53.2	44.4	51.6	50.8	0.822
<i>Osteoarthritis measures</i>							
Duration of disease (years)	10.5 (0.95)	9.7 (0.75)	8.2 (0.87)	8.3 (0.75)	10.4 (1.32)	9.5 (1.10)	0.589
Radiographic grade (0–4)	2.81 (0.06)	2.86 (0.07)	2.84 (0.10)	2.81 (0.07)	2.94 (0.09)	2.75 (0.10)	0.782
Knee as index joint (%)	85.1	80.5	90.3	83.2	82.2	84.2	0.405
% Taking NSAID for OA	68.5	70.7	71.0	62.2	72.6	63.5	0.356
WOMAC target joint (0–100)	46.3 (1.59)	45.5 (1.84)	47.2 (2.48)	45.3 (1.87)	45.4 (2.43)	46.5 (2.43)	0.997
MDHAQ pain VAS (0–100)	55.0 (1.65)	53.3 (1.97)	53.8 (2.45)	52.3 (1.89)	53.6 (2.70)	52.6 (2.38)	0.945
MDHAQ GI distress (0–100)	14.5 (1.72)	17.5 (1.98)	17.3 (3.27)	14.4 (1.77)	18.0 (2.99)	18.7 (2.92)	0.827

Values are the mean (SE) unless stated otherwise.
VAS, visual analogue scale; GI, gastrointestinal.

used by the Food and Drug Administration for adverse events. The questionnaires are described in greater detail in the cited references and in the report of a previous study,¹⁴ in which virtually identical questionnaires were used.

Statistical methods

The three major efficacy outcomes were the total WOMAC score, 100 mm visual analogue pain scale score, and patient preference for period I treatment versus period II treatment. The sample size was planned to have at least 0.90 power at the 0.05 significance level for treatment comparisons of the total WOMAC score and the 100 mm visual analogue pain scale during period I for at least 150 patients for each of the placebo, acetaminophen, and celecoxib groups. Furthermore, the paired preference comparison would have at least 0.90 power at the 0.05 significance level with at least 100 patients for each of the two sequences for acetaminophen before/after celecoxib.

The false positive (or type I) error rate was kept at the conventional 5% level for the global tests to compare all three treatments in the study for all three major efficacy outcomes through a closed testing procedure,²⁹ with a fixed sequence of tests for which significance at all preceding steps was required to proceed to the next step. For PACES-a, the order of testing was total WOMAC for period I, followed by pain score for period I, and then patient preference response. For PACES-b the order of testing was patient preference, followed by the total WOMAC for period I, and pain score for period I. Pairwise tests between the three treatments were conducted only if the global test was found to be significant (which was the case in both studies). These methods for

multiple comparisons essentially made the total WOMAC score have the primary role in PACES-a and the patient preference have the primary role in PACES-b.

Secondary end points included patient global assessment, also measured on a 100 mm visual analogue scale, SF-36 pain scores, MDHAQ activities of daily living scale scores, investigator assessment of patient global status, and investigator assessment of patient change in global status. The last observation carried forward procedure was used to manage missing data. Sites that enrolled fewer than five patients were pooled to allow for adjustment for site in the analyses.

Analyses of total WOMAC score, pain score, and paired patient preference were conducted on three groups: (a) “intention to treat”—all patients who received a single dose of the study drug in period I; (b) “protocol adherent”—omitting patients with major protocol violations such as non-compliance; (c) “all completers”—patients with complete data for the relevant period(s). Results from all three analyses were similar. Patient preference results are presented for patients in the intention to treat population who provided preference data (requiring participation in both periods) and for the protocol adherent population. Only the intention to treat results are presented for all other end points.

For all six sequence groups, means and their corresponding standard errors were computed to describe the distribution of continuous demographic, clinical, and patient/assessor questionnaire response variables. Distributions of categorical demographic, patient questionnaire response variables, and adverse events were described for each group with frequencies and/or percentages. The extent of random imbalances in

Table 2 Baseline values and changes after 6 weeks of treatment in the WOMAC score and MDHAQ pain and gastrointestinal distress visual analogue scales at the end of each treatment period in PACES clinical trials

	Period I			Period II		
	Celecoxib (n = 181)	Acetaminophen (n = 171)	Placebo (n = 172)	Celecoxib (n = 229)	Acetaminophen (n = 178)	Placebo (n = 117)
A. PACES-a						
<i>WOMAC target joint</i>						
Baseline	48.6 (1.47)	52.8 (1.52)	50.3 (1.60)	49.0 (1.49)	47.5 (1.73)	45.1 (2.15)
Change	-10.4 (1.54)	-8.4 (1.52)	-4.8 (1.66)	-8.6 (1.15)	-4.5 (1.17)	-3.6 (1.36)
% Change*	-21.4	-15.9	-9.5	-17.6	-9.5	-8.0
<i>MDHAQ visual analogue pain scale</i>						
Baseline	62.0 (1.11)	67.5 (1.17)	64.0 (1.16)	56.8 (1.64)	54.2 (1.91)	55.1 (2.53)
Change	-19.0 (1.91)	-17.4 (1.99)	-10.5 (1.92)	-12.8 (1.56)	-5.9 (1.64)	-4.9 (2.20)
% Change*	-30.6	-25.8	-16.4	-22.5	-10.9	-8.9
<i>MDHAQ visual analogue patient gastrointestinal distress scale</i>						
Baseline	19.2 (1.71)	21.2 (1.82)	19.4 (1.75)	23.3 (1.71)	22.2 (1.87)	18.5 (2.10)
Change	-0.9 (1.62)	3.6 (1.91)	2.2 (1.86)	-0.3 (1.34)	-2.2 (1.40)	4.6 (2.50)
% Change*	-4.7	+17.0	+11.3	-1.3	-9.9	+24.9
	(n = 189)	(n = 185)	(n = 182)	(n = 242)	(n = 190)	(n = 124)
B. PACES-b						
<i>WOMAC target joint</i>						
Baseline	52.0 (1.29)	51.6 (1.27)	52.5 (1.46)	47.3 (1.47)	47.7 (1.61)	45.0 (1.83)
Change	-13.5 (1.36)	-8.4 (1.31)	-4.6 (1.49)	-10.0 (1.17)	-4.9 (1.12)	-2.4 (1.28)
% Change*	-26.0	-16.3	-8.8	-21.2	-10.3	-5.3
<i>MDHAQ visual analogue pain scale</i>						
Baseline	65.2 (1.08)	64.6 (1.24)	64.4 (1.15)	55.8 (1.68)	57.2 (1.81)	54.3 (2.21)
Change	-21.8 (1.93)	-13.8 (1.74)	-7.6 (1.99)	-13.0 (1.62)	-8.7 (1.75)	-2.0 (2.14)
% Change*	-33.4	-21.4	-11.8	-23.3	-15.2	-3.7
<i>MDHAQ visual analogue patient gastrointestinal distress scale</i>						
Baseline	22.2 (1.78)	18.7 (1.56)	21.1 (1.87)	25.4 (1.75)	23.7 (1.84)	22.0 (2.39)
Change	-1.1 (1.56)	5.0 (1.83)	2.4 (1.76)	-3.0 (1.61)	2.4 (1.65)	-0.9 (1.72)
% Change*	-5.0	+26.7	+11.4	-11.8	+10.1	-4.1

*This is the percentage of treatment mean change from treatment mean baseline, not mean percentage of individual change from baseline. Values are the mean (SE) unless otherwise stated.

comparisons of the groups at screening visit 1 was described with p values from a χ^2 /Fisher's exact test for dichotomous variables and a Kruskal-Wallis test for continuous variables.

Efficacy of the three treatments was compared for continuous variables using analysis of covariance for the change during each of periods I and II, with screening and baseline scores included as covariates, along with fixed effects for treatment and centre. Repeated measures analysis of covariance through generalised estimating equations³⁰ provided supportive analyses for both periods combined. This method had a full statistical model with covariates for treatment, period, screening (visit 1), baseline of the patient (visit 2), and baseline for the period (visit 2 for period I and visit 4 for period II), with patient as the random sampling unit. The absence of noteworthy carryover effects for the residual impact of period I on period II was confirmed through a corresponding expanded model with their additional inclusion. Patient preference was assessed using conditional logistic regression,³¹ with components for period and treatment.

The prevalence of adverse events was compared among treatment groups using conditional logistic regression. Tolerability was also assessed according to change in a gastrointestinal distress 100 mm visual analogue scale on the MDHAQ.

RESULTS

Baseline and overview

In PACES-a, 524 patients were enrolled and randomised, and 556 patients were enrolled and randomised in PACES-b (table 1). No atypical imbalances with $p < 0.05$ were seen in

values at screening between the six sequence groups or between the three period I treatment groups for age, sex, race, education level, duration of osteoarthritis, previous NSAID or analgesic use, proportion of patients who took aspirin for cardiac prophylaxis, patient identified index joint, radiographic grade of the index joint, global severity of osteoarthritis, WOMAC scores, and pain scores (table 1), other than the WOMAC pain score ($p = 0.031$) in PACES-a. This latter result is probably spurious relative to multiple comparisons and is of no consequence, as baseline scores are covariates in the analyses of covariance for the changes during periods I, II, and both periods combined. In both studies all global tests to compare the three groups were statistically significant ($p = 0.01$ in both) for the three major efficacy end points, allowing us to proceed to the test for the respective pairs of treatments. Also, all tests for carryover effects confirmed their absence.

WOMAC scores

In PACES-a, period I comparisons were addressed as the prespecified primary end point of that study, with differences between celecoxib and acetaminophen ($p = 0.180$) and between acetaminophen and placebo ($p = 0.080$) not significant, while differences between celecoxib and placebo were significant ($p = 0.002$) (table 3A). During period II and both periods combined, differences between celecoxib and acetaminophen ($p < 0.009$) and celecoxib and placebo ($p < 0.007$) were significant, but not between acetaminophen and placebo ($p = 0.080$ for both periods) (table 3A). In PACES-b, differences between celecoxib and acetaminophen, celecoxib and placebo, and acetaminophen and placebo were all significant ($p < 0.001$ to 0.03) (table 3B). Similarly,

Table 3 Estimated differences (estimate (SEM)) in WOMAC score and MDHAQ pain and gastrointestinal distress visual analogue scales between treatment groups during period I, period II, and both periods for PACES clinical trials.

	Celecoxib-acetaminophen	p Value	Celecoxib-placebo	p Value	Acetaminophen-placebo	p Value	Carryover p [§]
3A. PACES-a							
<i>WOMAC target joint</i>							
Period I*	-2.78 (2.07)	0.180	-6.47 (2.06)	0.002	-3.69 (2.10)	0.080	
Period II†	-4.03 (1.55)	0.009	-4.79 (1.77)	0.007	-0.75 (1.85)	0.683	
Both periods‡	-3.41 (1.02)	<0.001	-5.50 (1.14)	<0.001	-2.09 (1.20)	0.080	0.827
<i>MDHAQ visual analogue pain score</i>							
Period I*	-3.47 (2.66)	0.193	-9.25 (2.63)	<0.001	-5.78 (2.68)	0.031	
Period II†	-6.49 (2.17)	0.003	-7.66 (2.47)	0.002	-1.17 (2.28)	0.651	
Both periods‡	-4.88 (1.50)	0.001	-8.90 (1.73)	<0.001	-4.02 (1.73)	0.02	0.602
<i>MDHAQ patient gastrointestinal distress visual analogue scale</i>							
Period I*	-4.83 (2.30)	0.036	-2.54 (2.30)	0.269	2.29 (2.33)	0.327	
Period II†	-2.35 (1.47)	0.110	-4.97 (1.68)	0.003	-2.62 (1.75)	0.136	
Both periods‡	-1.19 (1.22)	0.333	-3.16 (1.49)	0.034	-1.97 (1.60)	0.217	0.167
B. PACES-b							
<i>WOMAC target joint</i>							
Period I*	-4.99 (1.83)	0.007	-9.09 (1.83)	<0.001	-4.10 (1.84)	0.027	
Period II†	-4.99 (1.48)	<0.001	-6.53 (1.69)	<0.001	-1.53 (1.77)	0.387	
Both periods‡	-4.62 (1.00)	<0.001	-7.70 (1.12)	<0.001	-3.08 (1.10)	0.005	0.937
<i>MDHAQ visual analogue pain score</i>							
Period I*	-7.69 (2.54)	0.003	-13.84 (2.54)	<0.001	-6.15 (2.55)	0.016	
Period II†	-4.11 (2.13)	0.054	-9.98 (2.43)	<0.001	-5.87 (2.54)	0.021	
Both periods‡	-5.87 (1.41)	<0.001	-12.24 (1.68)	<0.001	-6.38 (1.64)	<0.001	0.614
<i>MDHAQ patient gastrointestinal distress visual analogue scale</i>							
Period I*	-3.22 (2.16)	0.136	-2.54 (2.14)	0.236	0.68 (2.16)	0.755	
Period II†	-4.87 (2.00)	0.015	-0.72 (2.29)	0.754	4.15 (2.39)	0.084	
Both periods‡	-3.51 (1.35)	0.009	-1.84 (1.30)	0.156	1.68 (1.34)	0.210	0.486

*Results are from an analysis of covariance for visit 3–visit 2 with treatment, visit 1, visit 2, and centre as covariates; †results are from an analysis of covariance for visit 5–visit 4 with treatment, visit 1, visit 4, and centre as covariates; ‡results are from an analysis using GEE for visit 3–visit 2 and visit 5–visit 4 with treatment, period, visit 1, visit 2, baseline as covariates; §from an analysis using GEE with the model described in footnote ‡ expanded to include a two degrees of freedom test for carry over.

statistically significant differences were seen for period II and both periods combined. Percentage improvements from baseline in WOMAC scores averaged over all four treatment periods in PACES-a and PACES-b, were 21.6% for celecoxib (range 17.6–26.0%), 13.0% for acetaminophen (range 9.5–16.3%), and 7.9% for placebo (range 5.3–9.5%) (tables 2A and B).

Pain visual analogue scale scores

Pain VAS results were similar to the WOMAC results (tables 2 and 3). Differences between celecoxib and acetaminophen were not significant ($p = 0.193$) in PACES-a, period I, while differences between acetaminophen and placebo ($p = 0.031$) and celecoxib and placebo ($p < 0.001$) were significant (table 3A). In period II (table 2A, fig 2), differences between celecoxib and acetaminophen ($p = 0.003$) and celecoxib and placebo ($p = 0.002$) were significant, while differences between acetaminophen and placebo ($p = 0.651$) were not (table 3A). In both periods combined, differences between celecoxib and acetaminophen ($p = 0.001$), celecoxib and placebo ($p < 0.01$), and acetaminophen and placebo ($p = 0.02$) were significant (table 3A). In PACES-b (table 2B, fig 2), differences in pain scores between celecoxib and acetaminophen, celecoxib and placebo, and acetaminophen and placebo in periods I, II, and both periods combined were significant (< 0.021), other than differences between celecoxib and acetaminophen in period II ($p = 0.054$) (table 3B). Percentage improvements from baseline in pain scores averaged over all four treatment periods in PACES-a and PACES-b were 27.5% (range 22.5–33.4%) for celecoxib, 18.3% (range 10.9–25.8%) for acetaminophen, and 10.2% (range 3.7–16.4%) for placebo (tables 2A and B).

Paired patient preference

In the PACES-a “all completers” population, among 173 patients who received celecoxib and acetaminophen, 52.6% rated celecoxib as “much better” or “better”, 24.3% rated acetaminophen as “much better” or “better”, and 23.1% reported “no difference” (table 4, fig 3). Odds ratios were 2.07 for preference of celecoxib versus acetaminophen ($p < 0.001$), 2.51 for celecoxib versus placebo ($p < 0.001$), and 1.21 for acetaminophen versus placebo ($p = 0.340$). Similar results were seen for the intention to treat and protocol adherent populations (table 4).

In the PACES-b “all completers” population, odds ratios were 1.47 for patient preference of celecoxib versus acetaminophen ($p = 0.009$), 2.47 for celecoxib versus placebo ($p < 0.001$), and 1.68 for acetaminophen versus placebo ($p = 0.007$). Again, similar results were seen for both the intention to treat and protocol adherent populations (table 4).

Other efficacy end points

In both PACES-a and PACES-b, analyses of the patient global scale, MDHAQ activities of daily living scale, investigator assessment of patient global status, investigator assessment of patient change in global status, and SF-36 pain scores disclosed patterns similar to those of the primary end points (data not shown—available on request).

Adverse events

Adverse events were reported by 23–29% of patients in the three groups. No significant differences were seen between the proportion of patients reporting any gastrointestinal event, specifically diarrhoea, dyspepsia, nausea, and flatulence, as well as for upper respiratory infection and headache, or any event (table 5).

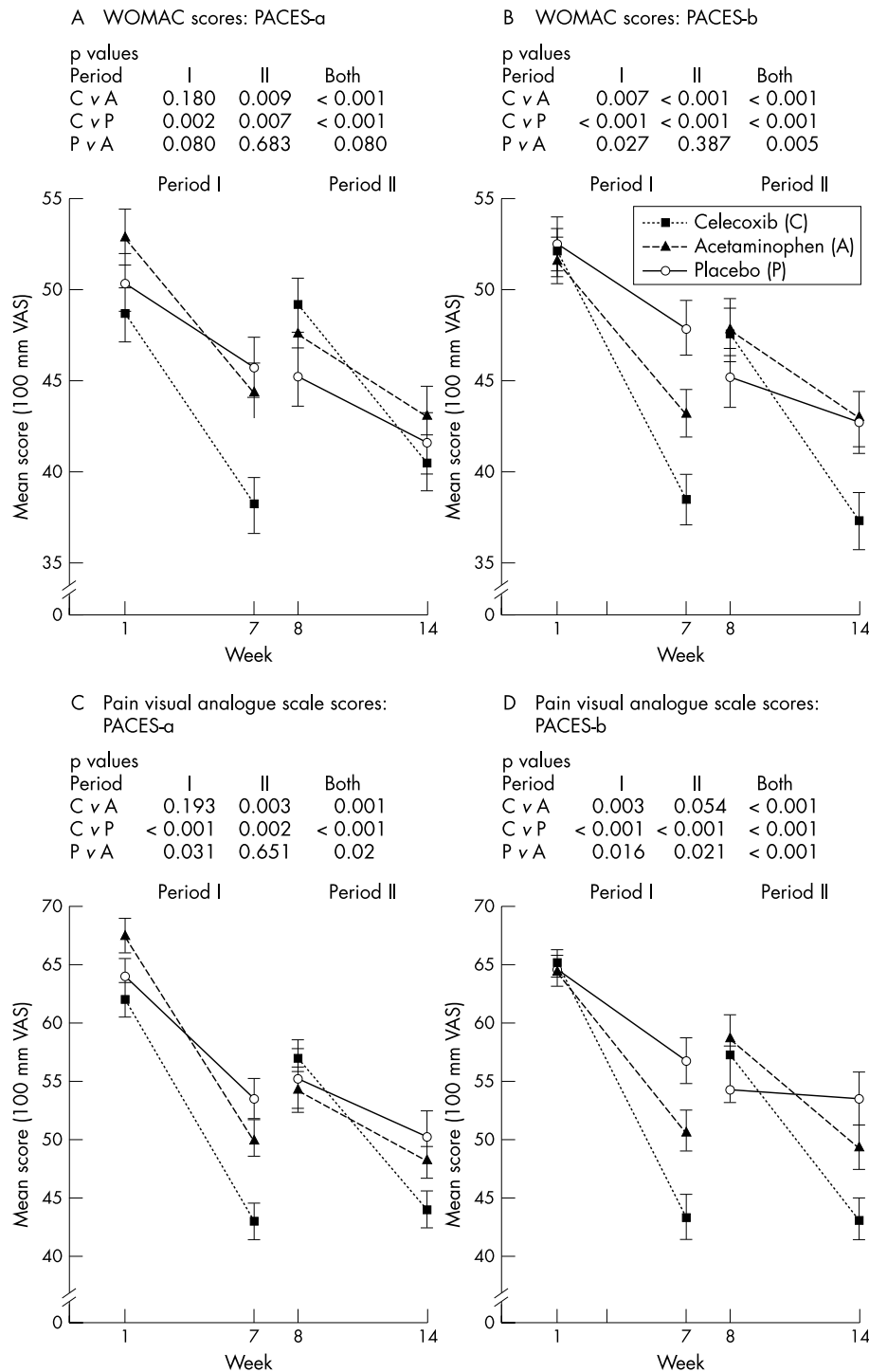


Figure 2 Change in scores for WOMAC (A and B) and pain visual analogue scale scores on the MDHAQ (C and D) of patients who received 6 weeks' treatment with celecoxib, acetaminophen, or placebo, after a 1 week washout period, in PACES-a and PACES-b.

The MDHAQ gastrointestinal distress scale indicated small changes over 6 weeks for all three treatments in both studies (tables 2A and B), suggesting no clinically significant gastrointestinal intolerance with any of the three treatments.

In PACES-a, eight adverse events were classified as serious because they required admission to hospital: one in the celecoxib group—intestinal obstruction and neuropathy; three in the acetaminophen group—one case of anxiety, one of cholelithiasis, and one of cholecystitis; and four in the

placebo group—one patient with raised liver function tests, one with urinary tract malformation and rectal disorder, one with an accidental fracture, and one with sepsis. Two of the events, the intestinal obstruction that occurred with the patient taking celecoxib and increased liver enzymes, which occurred while a patient was taking placebo, were regarded by the investigators as potentially related to the study drug. The other serious adverse events were regarded as probably not related to the study drug.

Table 4 Patient preferences in PACES clinical three treatment (six sequence) crossover clinical trials in all completers and protocol adherent patients.

	PACES-a		PACES-b	
	All completers†	Protocol adherent	All completers†	Protocol adherent
<i>Celecoxib v acetaminophen</i>				
Prefer celecoxib	91 (52.6)	66 (57.4)	95 (49.5)	84 (52.8)
Prefer acetaminophen	42 (24.3)	24 (20.9)	62 (32.3)	51 (32.1)
No preference	40 (23.1)	25 (21.7)	35 (18.3)	24 (15.1)
OR (SE)*	2.07 (1.18)	2.47 (1.22)	1.47 (1.16)	1.64 (1.18)
p Value*	<0.001	<0.001	0.009	0.002
<i>Celecoxib v placebo</i>				
Prefer celecoxib	71 (52.2)	54 (55.7)	76 (56.7)	62 (59.0)
Prefer placebo	31 (22.8)	24 (24.7)	32 (23.9)	23 (21.9)
No preference	34 (25.0)	19 (19.6)	26 (19.4)	20 (19.0)
OR (SE)*	2.51 (1.20)	2.48 (1.24)	2.47 (1.19)	2.60 (1.23)
p Value*	<0.001	<0.001	<0.001	<0.001
<i>Acetaminophen v placebo</i>				
Prefer acetaminophen	31 (37.4)	22 (36.1)	41 (47.7)	29 (44.6)
Prefer placebo	23 (27.7)	18 (29.5)	21 (24.4)	17 (26.2)
No preference	29 (34.9)	21 (34.4)	24 (27.9)	19 (29.2)
OR (SE)*	1.21 (1.22)	1.01 (1.26)	1.68 (1.21)	1.59 (1.24)
p Value*	0.340	0.977	0.007	0.033

Values are the number (%) of patients unless otherwise stated. *Results are from a conditional logistic regression model with period and treatment as fixed effects and patient as the strata variable; †all completers are patients in the intention to treat group who provided a response for patient preference.

In PACES-b, four adverse events were classified as serious, because they required admission to hospital: two in the celecoxib group—one case of cholelithiasis and one case of unstable angina; one in the acetaminophen group—chest pain, probably musculoskeletal in origin, and one in the placebo group—angina pectoris. All events were considered unrelated to the study drug by the investigators.

DISCUSSION

The data indicate a gradient of efficacy from celecoxib to acetaminophen to placebo. Although overall trends in the two studies are similar, numerical advantages in efficacy of celecoxib over acetaminophen, and acetaminophen over placebo according to WOMAC and pain scores in period I of PACES-a were not significant by the criterion of $p < 0.05$.

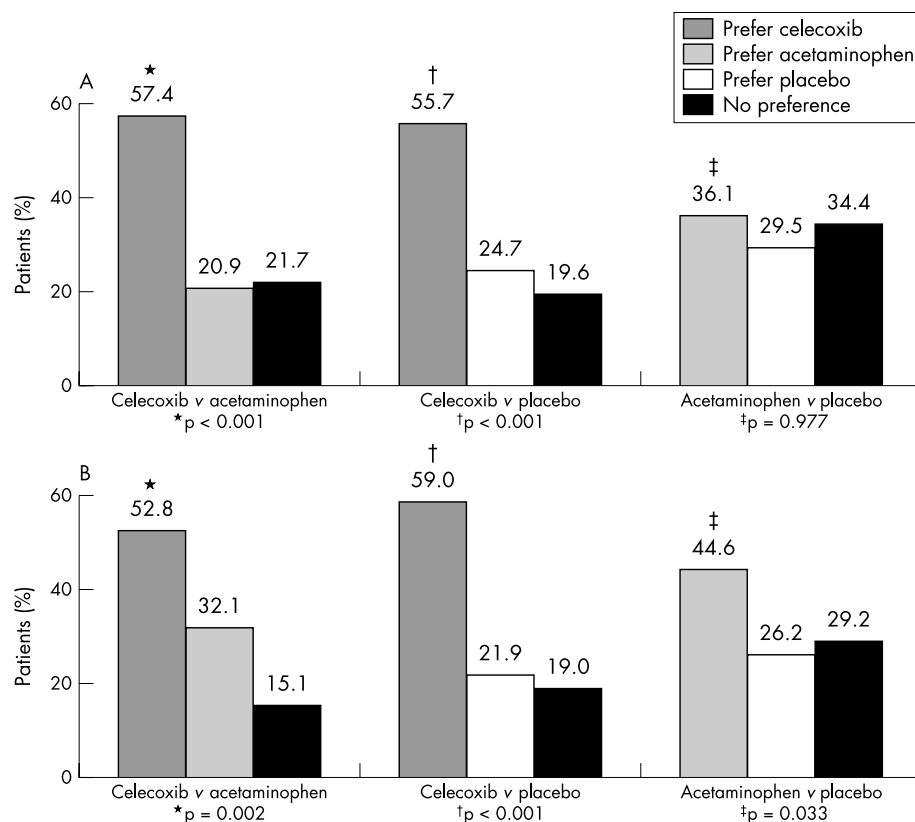


Figure 3 Patient Preference for Placebo, Acetaminophen or Celecoxib in PACES-a (A) and PACES-b (B). Each patient took two of the three treatments for 6 weeks each. At the final visit, the patient questionnaire asked: "Please compare control of your arthritis during the first and second study periods," with five response options, "much better during first study period," "better during first study period," "no difference in first and second study periods," "better during second study period," "much better during second study period." Patients who responded "much better" or "better" in either period were merged in these analyses.

Table 5 Study-emergent adverse events reported by treatment group for both periods combined* in PACES clinical trials.

	Celecoxib	Acetaminophen	Placebo	p Value†
A. PACES-a				
Patients who took drug (n)	350	300	289	
Any event	100 (28.6)	85 (28.3)	76 (26.3)	0.813
Serious events	1 (0.3)	3 (1.0)	4 (1.4)	0.999
Any gastrointestinal event	42 (12.0)	28 (9.3)	26 (9.0)	0.129
Diarrhoea	8 (2.3)	14 (4.7)	4 (1.4)	0.999
Dyspepsia	10 (2.9)	7 (2.3)	3 (1.0)	0.641
Nausea	7 (2.0)	7 (2.3)	5 (1.7)	0.872
Flatulence	8 (2.3)	4 (1.3)	1 (0.3)	0.232
Other				
Upper respiratory tract infection	16 (4.6)	17 (5.7)	9 (3.1)	0.3620
Headache	4 (1.1)	11 (3.7)	5 (1.7)	0.872
B. PACES-b				
Patients who took drug (n)	373	331	273	
Any event	103 (27.6)	87 (26.3)	63 (23.1)	0.423
Serious events	2 (0.5)	1 (0.3)	1 (0.4)	0.999
Any gastrointestinal event	35 (9.4)	30 (9.1)	17 (6.2)	0.500
Diarrhoea	6 (1.6)	11 (3.3)	4 (1.5)	0.870
Dyspepsia	6 (1.6)	6 (1.8)	2 (0.7)	0.999
Nausea	8 (2.1)	4 (1.2)	3 (1.1)	0.329
Flatulence	3 (0.8)	4 (1.2)	1 (0.4)	–
Other				
Upper respiratory tract infection	7 (1.9)	9 (2.7)	2 (0.7)	0.248
Headache	7 (1.9)	9 (2.7)	7 (2.6)	0.419

*Values are the number (%) of patients reporting at least 1 event. Percentage are based on the number of patients exposed to the drug in either period; †p values are from a conditional logistic regression model with effects for period and treatment.

Differences between celecoxib and acetaminophen were significant in period II and both periods combined in PACES-a, and periods I, II, and both periods combined in PACES-b. Patient preference data for celecoxib versus acetaminophen or placebo were significant in PACES-a and PACES-b. Patient preference for acetaminophen versus placebo was significant in PACES-b, but not in PACES-a.

The rate of adverse events was low and similar for celecoxib, acetaminophen and placebo, with few serious adverse events. All three treatments were well tolerated at comparable levels. Although the patient preference inquiry was based primarily on efficacy, the absence of gastrointestinal intolerance with celecoxib was probably incorporated by the patients into an assessment of preference.

Greater efficacy of celecoxib compared with acetaminophen was also seen using other proposed measures of improvement in osteoarthritis clinical trials. A 20% improvement criterion proposed by Case *et al*,¹⁵ was seen for WOMAC scores for three of the four periods with celecoxib, compared with none of four with acetaminophen or placebo. For the pain visual analogue scale, the 20% improvement criterion was met in four of four periods for celecoxib, two of four for acetaminophen, and none of four for placebo. A 10 mm change in WOMAC scores proposed by Ehrich *et al*¹² was seen for celecoxib in period I of PACES-a and both periods of PACES-b, and was not seen for acetaminophen or placebo in either trial.

PACES-a and PACES-b are presented as individual, rather than pooled studies, to illustrate natural variation in results and p values in two identical clinical trials. Numerical differences between celecoxib and acetaminophen were similar in PACES-a and PACES-b. The p value of 0.18 in PACES-a for differences between celecoxib and acetaminophen in period I indicates a result that could occur one in five times by chance, while the p value in PACES-b of 0.007, indicates an occurrence less than 1 in 130 times by chance. The data illustrate that a focus only on the statistical criterion of $p < 0.05$ value may not necessarily be optimal to discern differences in the efficacy of one drug versus another.

The PACES clinical trials have several limitations. Firstly, comparisons between three drugs ideally might be conducted with each patient taking all three drugs. However, pragmatic considerations suggested only two treatment periods, but a design so that three treatments could be compared. Patient attrition also presents a limitation, as in any clinical trial, although completion rates for both periods of 73% of patients in PACES-a and 74% of patients in PACES-b compare favourably with single period clinical trials in osteoarthritis.

The patient preferences for celecoxib versus placebo or acetaminophen in this study were similar to those seen for diclofenac/misoprostol compared with acetaminophen in the ACTA study,¹⁴ and trends are similar to a trial of rofecoxib, celecoxib, and acetaminophen,³³ although responses to acetaminophen were greater in that trial than in the PACES trial reported here. The rate of gastrointestinal events was considerably lower with celecoxib in PACES than with diclofenac/misoprostol in ACTA. Therefore, the results indicate greater efficacy of celecoxib versus acetaminophen for patients with osteoarthritis, with similar tolerability and safety. Although a substantially higher probability is seen that patients will respond to celecoxib compared with acetaminophen or placebo, individual variation is seen, as 1 in 3–5 patients expressed a preference for acetaminophen, and 1 in 3–5 expressed no preference in both trials.

In conclusion, we have found that the efficacy and patient preference for celecoxib is greater than that for acetaminophen, and the efficacy and patient preference for acetaminophen is greater than that for placebo. These results may have implications for an optimal pharmacological approach to the management of patients with osteoarthritis using drugs available at this time.

ACKNOWLEDGEMENTS

We recognise the contributions of Reinhard Schuller in designing, Dr Joseph Kohles in managing, and all the investigators in conducting the PACES clinical trials, and the expert assistance of Dr Barbara J Struthers in the design of the figures.

Sponsored by Pfizer Corporation.

.....

Authors' affiliations

T Pincus, T Sokka, Division of Rheumatology and Immunology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
G Koch, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
H Lei, B Mangal, S Zlotnick, J G Fort, Pfizer Corporation, USA
T Sokka, Department of Medicine, Jyväskylä Central Hospital, Jyväskylä, Finland
R Moskowitz, Division of Rheumatology, Case-Western University, Cleveland, OH 44106, USA
F Wolfe, Wichita Arthritis Center, University of Kansas School of Medicine, 1035 North Emporia, Suite 230, Wichita, KS 67214, USA
A Gibofsky, Hospital for Special Surgery – Weill Medical College of Cornell University, New York, NY, USA
L Simon, Department of Medicine, Harvard University, Boston, MA 02115, USA

REFERENCES

- Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, *et al*. Guidelines for the medical management of osteoarthritis. Part I. Osteoarthritis of the hip. *Arthritis Rheum* 1995;**38**:1535–40.
- Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, *et al*. Guidelines for medical management of osteoarthritis. Part II. Osteoarthritis of the knee. *Arthritis Rheum* 1995;**38**:1541–6.
- Altman RD, Hochberg MC, Moskowitz RW, Schnitzer TJ. Recommendations for the medical management of osteoarthritis of the hip and knee. 2000 Update. *Arthritis Rheum* 2000;**43**:1905–15.
- Griffin MR, Ray WA, Schaffner W. Nonsteroidal anti-inflammatory drug use and death from peptic ulcer in elderly persons. *Ann Intern Med* 1988;**109**:359–63.
- Fries JF, Williams CA, Bloch DA, Michel BA. Nonsteroidal anti-inflammatory drug-associated gastropathy: incidence and risk factor models. *Am J Med* 1991;**91**:213–22.
- Griffin MR, Brandt KD, Liang MH, Pincus T, Ray WA. Practical management of osteoarthritis: integration of pharmacologic and nonpharmacologic measures. *Arch Fam Med* 1995;**4**:1049–55.
- Bradley JD, Brandt KD, Katz BP, Kalasinski LA, Ryan SI. Comparison of an antiinflammatory dose of ibuprofen, an analgesic dose of ibuprofen, and acetaminophen in the treatment of patients with osteoarthritis of the knee. *N Engl J Med* 1991;**325**:87–91.
- Williams HJ, Ward JR, Egger MJ, Neuner R, Brooks RH, Clegg DO, *et al*. Comparison of naproxen and acetaminophen in a two-year study of treatment of osteoarthritis of the knee. *Arthritis Rheum* 1993;**36**:1196–206.
- Felson DT. Editorial: The verdict favors nonsteroidal antiinflammatory drugs for treatment of osteoarthritis and a plea for more evidence on other treatments. *Arthritis Rheum* 2001;**44**:1477–80.
- Hochberg MC. What a difference a year makes: reflections on the ACR recommendations for the medical management of osteoarthritis. *Curr Rheumatol Rep* 2001;**3**:473–8.
- Abramson SB. Et tu, acetaminophen? [editorial]. *Arthritis Rheum* 2002;**46**:2831–5.
- Wolfe F, Zhao S, Lane N. Preference for nonsteroidal antiinflammatory drugs over acetaminophen by rheumatic disease patients: a survey of 1,799 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Arthritis Rheum* 2000;**43**:378–85.
- Pincus T, Swearingen C, Cummins P, Callahan LF. Preference for nonsteroidal anti-inflammatory drugs versus acetaminophen and concomitant use of both types of drugs in patients with osteoarthritis. *J Rheumatol* 2000;**27**:1020–7.
- Pincus T, Koch GG, Sokka T, Lefkowitz J, Wolfe F, Jordan JM, *et al*. A randomized, double-blind, crossover clinical trial of diclofenac plus misoprostol versus acetaminophen in patients with osteoarthritis of the hip or knee. *Arthritis Rheum* 2001;**44**:1587–98.
- Case JP, Baliunas AJ, Block JA. Lack of efficacy of acetaminophen in treating symptomatic knee osteoarthritis: a randomized, double-blind, placebo-controlled comparison trial with diclofenac sodium. *Arch Intern Med* 2003;**163**:169–78.
- Rodriguez LAG, Hernández-Díaz S. The risk of upper gastrointestinal complications associated with nonsteroidal anti-inflammatory drugs, glucocorticoids, acetaminophen, and combinations of these agents. *Arthritis Res* 2001;**3**:98–101.
- Rahme E, Pettitt D, LeLorier J. Determinants and sequelae associated with utilization of acetaminophen versus traditional nonsteroidal antiinflammatory drugs in an elderly population. *Arthritis Rheum* 2002;**46**:3046–54.
- Simon LS, Weaver AL, Graham DY, Kivitz AJ, Lipsky PE, Hubbard RC, *et al*. Anti-inflammatory and upper gastrointestinal effects of celecoxib in rheumatoid arthritis. *JAMA* 1999;**282**:1921–8.
- Cannon GW, Caldwell JR, Holt P, McLean B, Seidenberg B, Bolognese J, *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.
- Geba GP, Weaver AL, Polis AB, Dixon ME, Schnitzer TJ. Efficacy of rofecoxib, celecoxib, and acetaminophen in osteoarthritis of the knee: a randomized trial. *JAMA* 2002;**287**:64–71.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;**15**:1833–40.
- Boeckstyns MEH, Backer M. Reliability and validity of the evaluation of pain in patients with total knee replacement. *Pain* 1989;**38**:29–33.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;**23**:137–45.
- Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford health assessment questionnaire. *Arthritis Rheum* 1983;**26**:1346–53.
- Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. *Ann Intern Med* 1989;**110**:259–66.
- Pincus T, Swearingen C, Wolfe F. Toward a multidimensional health assessment questionnaire (MDHAQ): assessment of advanced activities of daily living and psychological status in the patient friendly health assessment questionnaire format. *Arthritis Rheum* 1999;**42**:2220–30.
- Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–81.
- McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;**31**:247–63.
- Bauer P. Multiple testing in clinical trials. *Stat Med* 1991;**10**:871–90.
- Diggle PJ, Liang KY, Zeger SL. *Analysis of longitudinal data*. Oxford, UK: Oxford University Press, 1994.
- Stokes ME, Davis CS, Koch GG. *Categorical data analysis using the SAS system*, 2nd ed. Cary, NC: SAS Institute, Inc, 2000.
- Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000;**27**:2635–41.
- Geba GP, Weaver AL, Polis AB, Dixon ME, Schnitzer TJ. Efficacy of rofecoxib, celecoxib, and acetaminophen in osteoarthritis of the knee: a randomized trial. *JAMA* 2002;**287**:64–71.