EXTENDED REPORT

Greater reduction of knee than hip pain in osteoarthritis treated with naproxen, as evaluated by WOMAC and SF-36

O Svensson, M Malmenäs, L Fajutrao, E M Roos, L S Lohmander

Ann Rheum Dis 2006;**65**:781–784. doi: 10.1136/ard.2005.040519

See end of article for authors' affiliations

Correspondence to: Professor Stefan Lohmander, Lund University, Clinical Sciences Lund, Department of Orthopaedics, Lund University Hospital, SE-22185 Lund, Sweden; stefan.lohmander@med.lu.se

Accepted 26 October 2005
Published Online First 3 November 2005

Objectives: To compare the improvement of hip and knee osteoarthritis during treatment with naproxen.
Methods: Men and women aged 40 to 75 years with symptomatic osteoarthritis of the knee or hip of at least three months' duration participated in a six week placebo controlled, double blind study with naproxen 500 mg twice daily as one treatment arm. Naproxen was given to 403 patients (280 knee, 123 hip) and placebo to 108 patients (75 knee, 33 hip). WOMAC (Western Ontario and McMaster Universities osteoarthritis index) 3.1 visual analogue scale and SF-36 (36 item short form health survey) were used to assess response to treatment between baseline and week 6.
Results: There were no differences at baseline between knee and hip osteoarthritis for any of the WOMAC subscales or SF-36 domains. Improvement was between 4 and 7 mm greater for knee than for hip for all WOMAC subscales (pain, $\Delta = 4.7$ mm ($p = 0.03$); stiffness, $\Delta = 6.6$ mm ($p = 0.004$); function, $\Delta = 4.8$ mm ($p = 0.06$)). Effect size was about 0.8 for all WOMAC subscales for the knee and between 0.5 and 0.6 for the hip. Knee patients treated with naproxen improved 4.6 ($p = 0.033$) more than hip patients for SF-36 bodily pain and 10.3 ($p = 0.014$) more for SF-36 role–physical.
Conclusions: Patients with knee osteoarthritis improved more with naproxen treatment than patients with hip osteoarthritis, as monitored by WOMAC and the SF-36 domains bodily pain and role–physical. These findings warrant further investigation and strongly suggest that efficacy of treatment of osteoarthritis of knee and hip should be evaluated separately.

In clinical investigations of osteoarthritis it has been recommended that efficacy should be measured by pain, function, and global measures evaluated by the patient.[1 2] According to OARSI and OMERACT guidelines, several instruments may be considered within each domain provided they have adequate reliability, validity, and responsiveness.[2] For example, pain could be measured by a global measure of pain, the WOMAC (Western Ontario and McMaster Universities) pain subscale or as an average of the four questions of the Lequesne functional severity index that focus on pain (pain at night, pain in the standing position, pain during walking, and pain while switching from a sitting to a standing position).[1 3 4] Function in osteoarthritis trials is most commonly measured by WOMAC or the Lequesne functional severity index.[3 4] In addition, a generic instrument such as the 36 item short-form health survey (SF-36) is often used to measure health related quality of life in osteoarthritis trials.[5]

Osteoarthritis is a heterogeneous disorder.[6] Observing an effect of a treatment in one major joint does not necessarily mean that it will be equally effective in another joint. Guidelines for clinical trials recommend that patients with knee osteoarthritis should be evaluated separately from patients with hip osteoarthritis.[7] However, this increases the costs for clinical trials. It would be possible to decrease the time and cost for trials if patients with either knee or hip osteoarthritis could be combined in the same analysis. However, this would require that the response characteristics to treatment of these two joints were the same.

The present study represents an exploratory ad hoc analysis of data from a large multinational randomised clinical trial including both knee and hip osteoarthritis, comparing the gastrointestinal safety and efficacy of the COX inhibiting nitric oxide donator (CINOD) AZD3582 with naproxen.[8] The primary end point of that study was the six week incidence of endoscopic gastroduodenal ulcers. The purpose of the present study was to evaluate the relative improvement in hip and knee osteoarthritis during treatment with naproxen, as measured by the disease specific instrument WOMAC and the generic measure SF-36. For this, the effect size was separately determined for hip and knee to see whether or not the knee and hip data could be combined for trial analysis.

METHODS

Patients

We recruited men and women aged 40 to 75 years with symptomatic osteoarthritis of the knee or hip of at least three months' duration.[8] All patients had hip or knee osteoarthritis as defined by the American College of Rheumatology (ACR) global functional class I, II, or III,[9 10] and were current non-steroidal anti-inflammatory drug (NSAID) or paracetamol (acetaminophen) users.

Patients were excluded from the study if they had osteoarthritis secondary to inflammatory joint disease, a diagnosis of arthritis other than osteoarthritis, a history of gastric or duodenal bleeding within six months or gastric or duodenal ulcer within three months, NSAID hypersensitivity, a history of orthostatic hypotension, or endoscopic ulcers at baseline screening. Patients on aspirin, H_2 antagonists, antacids, misoprostol, proton pump inhibitors, or sucralfate were not eligible, and the use of these agents was prohibited throughout the study period.

Abbreviations: SF-36, 36 item short form health survey; NSAID, non-steroidal anti-inflammatory drug; OARSI, OsteoArthritis Research Society International; OMERACT, Outcome Measures in Arthritis Clinical Trials; VAS, visual analogue scale; WOMAC, Western Ontario and McMaster Universities osteoarthritis index

Study design

The protocol was a six week, double blind, randomised, parallel group, placebo controlled study with naproxen 500 mg twice daily as one treatment arm. The study was conducted from November 2001 to November 2002. It was carried out in accordance with the ethical principles in the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The institutional review board or independent ethics committee of each participating centre provided ethical approval of the study protocol. All patients gave written informed consent.

Patients were instructed to take the treatment with food at 12 hour intervals. They were considered compliant with the treatment regimen if they took at least 70% of the study drug over the course of the study. Compliance was assessed by pill count.

Patients were assessed at an initial screening visit and those who fulfilled the entry criteria discontinued NSAID treatment for a washout period of two to 10 days before the baseline visit. Patients were allowed to take paracetamol up to 4000 mg/day, provided by the investigator, for control of pain during the washout period. If the patient used paracetamol, it was requested that it was discontinued 12 hours before the baseline visit. There was no preset level of pain increase the patients had to reach during the washout to be included in the study.

Efficacy and quality of life assessments

The WOMAC osteoarthritis index, version 3.1 visual analogue scale (VAS), and the SF-36 were used to assess response to treatment at baseline and after week 6.[3 5] Patients answered the WOMAC subscales of pain, stiffness, and physical function, using a 48 hour recall period. The acute version of SF-36, with a recall period of one week, was used to evaluate physical and mental wellbeing of the patients for the domains bodily pain, physical functioning, role–physical, vitality, mental health, social functioning, role–emotional, and general health. For WOMAC the VA scales were (worst) 100–0 (best), and for SF-36 values were normalised (worst) 0–100 (best).

Efficacy was also assessed at the six week visit by the subjects' and physicians' overall rating of treatment, which were five point Likert scales (where 1 = very poor and 5 = very good) by answering the following questions:

Patient: ''How do you rate your treatment overall, taking both pain relief and everything else into consideration?''
Investigator: ''How do you rate the subject's treatment overall, taking both pain relief and everything else into consideration?''

Statistical analysis

The within-subject differences between baseline and week 6 were used to analyse changes in WOMAC and SF-36 subscale scores. An analysis of co-variance (ANCOVA) approach with adjustment for country, baseline, and location of osteoarthritis was used for the pairwise comparisons between the different osteoarthritis locations when analysing WOMAC and SF-36. Confidence intervals for relevant differences between the two locations of osteoarthritis were calculated. Statistical significance was assumed at $p<0.05$. No correction for multiple comparisons was made. Efficacy analyses were undertaken on all randomised patients who received at least one dose of study medication. The effect size was calculated by dividing the mean change from baseline with the standard deviation at baseline, without subtracting for response within the placebo group.

RESULTS

Naproxen 500 mg twice daily was given to 403 patients (280 with knee osteoarthritis and 123 with hip osteoarthritis), and placebo to 108 patients (75 knee, 33 hip). The treatment groups were well balanced at baseline for demographic characteristics (table 1).

Further, there were no differences in baseline WOMAC scores between treatment groups or between those with hip or knee osteoarthritis (table 2). The baseline scores for all SF-36 domains were similar between treatment groups and between those with knee or hip osteoarthritis, as were the proportions of subjects taking rescue medication for osteoarthritis related pain during the wash out period (data not shown). Compliance of ⩾70% of the prescribed drug intake for the knee was 96.6% for naproxen and 96.2% for placebo, and for the hip, 96.9% for naproxen and 91.9% for placebo. Rescue medication was taken by 88% of knee patients in the naproxen group and 92% in the placebo group, and by 90% of the hip patients in the naproxen group and 92% in the placebo group.

As shown in table 2, for naproxen treatment improvement at week 6 was between 4 and 7 mm greater for the knee than for the hip for all three WOMAC subscales (pain, $\Delta = 4.7$ mm ($p = 0.03$); stiffness, $\Delta = 6.6$ mm ($p = 0.004$); function, $\Delta = 4.8$ mm ($p = 0.016$)). Placebo response showed the same trend, but with non-significant statistical differences between knee and hip (pain, $\Delta = 2.9$ mm; stiffness, $\Delta = 0.1$ mm; function, $\Delta = 2.3$ mm). The effect size was about 0.8 for the knee and between 0.5 and 0.6 for the hip for the naproxen treated group.

For the SF-36 domains bodily pain and role–physical, knee patients treated with naproxen improved by 4.6 units ($p = 0.033$) and 10.3 units ($p = 0.014$) more than hip patients, respectively, with no statistically significant difference for placebo treatment between knee and hip patients.

The patients taking naproxen rated the treatment as good or very good in 75% of the cases for the knee and 66% for the hip. For placebo, the numbers were 50% for the knee and 35% for the hip. The investigators rated the naproxen treatment as good or very good for the knee in 73% of the cases, and 64%

Table 1 Demographic characteristics of study groups

	Treatment group			
	Naproxen 500 mg (n = 403)		Placebo (n = 108)	
	Knee (n = 280)	Hip (n = 123)	Knee (n = 75)	Hip (n = 33)
Sex (female)	209 (75%)	87 (71%)	62 (83%)	20(61%)
Age (years)	60.0 (8.4)	59.4 (8.92)	60.5 (8.63)	55.9 (7.96)
Primary study joint	280 (69%)	123 (31%)	75 (69%)	33 (31%)
Weight (kg)	81.1 (15.8)	75.3 (15.8)	80.3 (13.7)	77.0 (14.4)
Height (cm)	163 (10.0)	164 (9.6)	162 (8.0)	167 (9.9)
Body mass index (kg/m^2)	31 (5.3)	28 (4.6)	31 (5.4)	28 (4.5)

Values are mean (SD) or n (%).

Table 2 Change in WOMAC subscales* from baseline to week 6 and effect size for knee and hip for the naproxen and placebo treated groups

WOMAC subscale	Pain	Stiffness	Physical function
Knee (naproxen)	(n = 280)	(n = 279)	(n = 278)
BL	44.2 (20.5)	48.9 (24.1)	47.1 (21.0)
BL-6	−16.62 (20.50)	−19.39 (23.01)	−16.31 (17.89)
Effect size	0.81	0.80	0.78
Knee (placebo)	(n = 75)	(n = 75)	(n = 72)
BL	45.11 (20.3)	50.6 (25.9)	48.2 (19.8)
BL-6	−6.75 (19.12)	−9.33 (23.91)	−6.75 (18.05)
Effect size	0.33	0.36	0.34
Hip (naproxen)	(n = 123)	(n = 123)	(n = 123)
BL	47.3 (22.2)	49.3 (23.7)	49.7 (22.9)
BL-6	−12.33 (19.78)	−12.97 (21.52)	−11.66 (16.78)
Effect size	0.56	0.55	0.51
Hip (placebo)	(n = 33)	(n = 33)	(n = 33)
BL	44.7 (20.1)	48.6 (23.3)	47.5 (21.8)
BL-6	−3.24 (22.27)	−8.15 (20.88)	−4.35 (17.67)
Effect size	0.16	0.35	0.20
Difference response knee–hip (95% CI), naproxen	−4.66 (−8.83 to −0.50), p = 0.03	−6.64 (−11.16 to −2.11), p = 0.004	−4.77 (−9.82 to −0.90), p = 0.016
Difference response knee–hip (95% CI), placebo	−2.89 (−11.37 to 5.59), p = 0.50	−0.10 (−9.25 to 9.05), p = 0.98	−2.29 (−9.82 to 5.24), p = 0.55

*Values are mean (SD).
BL, baseline; CI, confidence interval; WOMAC, Western Ontario and McMaster Universities osteoarthritis index.

for the hip. For placebo, the numbers were 47% for the knee and 38% for the hip.

DISCUSSION

To our knowledge, this is the first within-trial comparison of the improvement of knee and hip osteoarthritis during treatment with naproxen. As the mean of the baseline scores were lower for the knee group than for the hip group, the greater improvement found in the knee patients during treatment with naproxen is unlikely to be caused by a regression to the mean. Is this a true difference in naproxen efficacy between hip and knee, or does the WOMAC questionnaire have a higher responsiveness for knee osteoarthritis compared with hip osteoarthritis? The fact that a higher proportion of both patients and investigators rated the overall treatment as good or very good for the knee than for the hip supports the first explanation. The subscales of SF-36 most relevant to musculoskeletal problems and most comparable to the content of WOMAC are physical functioning, role–physical, and bodily pain. Seeing the same pattern of improvement in two of these three subscales as in WOMAC implies the difference between knee and hip being a treatment effect and not a psychometric phenomenon with regard to the WOMAC instrument.

An effect size of 0.5, which was found for the hip in the present study, is considered a moderate change, while an effect size of 0.8 as found for the knee is considered large.[11–13] However, the placebo response was not subtracted from these effect size values.

The mean knee–hip differences in the WOMAC subscales pain (4.7 mm), stiffness (6.6 mm), and physical functioning (4.8 mm) were less than the proposed minimal perceptible clinical improvement (MPCI) of 10 mm for WOMAC.[14] However, in that study the baseline values were considerably higher than in the present study, at 65.0 mm VAS *v* 44.2 mm for pain (knee, naproxen; table 2), 65.9 *v* 48.9 mm for stiffness, and 63.9 *v* 47.1 mm for physical functioning, and it has been shown that higher baseline scores require larger raw changes to represent a clinically important difference.[15] Further, to be included in that study, the patients had to have at least a 15 mm increase in the pain walking score after the washout, and a washout score of ⩾40 mm.[14] Neither of this was required in the present study.

The cited studies deal with the concept of how a patient perceives a change during the course of a treatment.[14 15] This might not be the same as a perceived difference *between* treatments. Therefore, it is difficult on the basis of those studies to draw any firm conclusions as to whether the difference between hip and knee in the present study is clinically relevant or not, but using the 11-point numerical rating scale, it was concluded that a pain reduction of approximately 30% represents a clinically important difference.[15] In the present study the reduction in WOMAC knee pain for naproxen was 38% (16.6/44.2) and in WOMAC hip pain 26% (12.3/47.3). Using the suggested cut off point of 30% would imply that the reduction in pain was clinically important for the knee but not for the hip.

The results of the present study strongly influence trial power and number of patients needed per treatment arm in clinical trials. Based on the effect sizes for pain, 108 subjects with hip osteoarthritis compared with only 54 subjects with knee osteoarthritis would need to be included in a clinical trial to determine a significant difference against baseline with 80% power. The findings support the recommendation that trials concerning efficacy of treatment for osteoarthritis of the knee and hip should be stratified with respect to target joint or evaluated separately,[7] and they warrant further investigation concerning the clinical relevance for the individual patient.

ACKNOWLEDGEMENTS

The original study was supported by a grant from AstraZeneca R&D Södertälje, Sweden.[8] EMR and LSL were supported by the Swedish Research Council, the Swedish Rheumatism Association, the Kock Foundation, the King Gustaf V 80-year Anniversary Foundation, the Faculty of Medicine Lund University, and Region Skåne.

Authors' affiliations

O Svensson, M Malmenäs, L Fajutrao, AstraZeneca R&D, Södertälje, Sweden
E M Roos, L S Lohmander, Lund University, Clinical Sciences Lund, Department of Orthopaedics, Lund, Sweden

OS, MM, and LF are employees of AstraZeneca R&D, Sweden. EMR and LSL have declared no conflict of interest in relation to the subject matter of this report.

REFERENCES

1 **Dougados M**, Le Claire P, van der Heijde D, Bloch DA, Bellamy N, Altman RD. A report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials Response Criteria Initiative. *Osteoarthritis Cartilage* 2000;**8**:395–403.
2 **Pham T**, van der Heijde D, Lassere M, Altman RD, Anderson JJ, Bellamy N, *et al*. Outcome variables for osteoarthritis clinical trials: the OMERACT-OARSI set of responder criteria. *J Rheumatol* 2003;**30**:1648–54.
3 **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.
4 **Lequesne MG**, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation – value in comparison with other assessment tests. *Scand J Rheumatol* 1987;suppl 65:85–9.
5 **Ware JE**, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–83.
6 **Flores RH**, Hochberg MC. Definition and classification of osteoarthritis. In: Brandt KD, Doherty M, Lohmander LS, eds. *Osteoarthritis*, 2nd edition. Oxford: Oxford University Press, 2003:1–8.
7 **Altman R**, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, *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.
8 **Lohmander LS**, McKeith D, Svensson O, Malmenäs M, Bolin L, Kalla A, *et al*. A randomised, placebo controlled, comparative trial of the gastrointestinal safety and efficacy of AZD3582 versus naproxen in osteoarthritis. *Ann Rheum Dis* 2005;**64**:449–56.
9 **Altman R**, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, *et al*. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;**34**:505–14.
10 **Altman R**, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al*. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;**29**:1039–49.
11 **Guyatt GH**, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987;**40**:171–8.
12 **Kazis LE**, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;**27**:S178–89.
13 **Cohen J**. *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale: Lawrence Earlbaum Associates, 1988.
14 **Ehrich EW**, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. 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.
15 **Farrar JT**, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;**94**:149–58.