Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study

A-K Nilsdotter, I F Petersson, E M Roos, L S Lohmander

....

See end of article for authors' affiliations

....

Correspondence to:
Dr A-K Nilsdotter,
Department of
Orthopaedics, Lund
University Hospital, S-221
85 Lund, Sweden;
Anna.Nilsdotter@Spenshult.se

Accepted
22 January 2003
....

Objectives: To investigate prospectively long term patient relevant outcomes after unilateral total hip replacement (THR) for osteoarthritis (OA). To identify non-responders to this intervention and patient related predictors of unsatisfactory outcome.

Methods: A case-control study comparing health related quality of life of 219 patients (mean age 71) after THR with that of a matched reference group of 117 subjects without hip complaints recruited from the community. Patients and reference group answered SF-36 and WOMAC questionnaires preoperatively, at 3, 6, 12 months, and at 3.6 years (range 26–65 months) postoperatively. Supplementary questions were asked at the final follow up.

Results: 198/211 (94%) of the patients and 83/109 (76%) of the reference group participated at the final follow up. At follow up, the only difference between the two groups in the SF-36 was physical function, where patients scored worse. Patients also reported worse WOMAC function. 31% of the patients had improved by <10/100 WOMAC score points for pain and/or function at final follow up, compared with preoperatively. More pain preoperatively and higher age and postoperative low back pain predicted a worse outcome in WOMAC function.

Conclusion: 3.6 years after THR for OA, health related quality of life was similar for patients and reference group except for function, where patients had worse function. Higher age and more pain preoperatively predicted a poor outcome. Patients with hip OA with musculoskeletal comorbidities, such as low back pain and OA of the non-operated hip, have less long term functional improvement after THR.

T otal hip replacement (THR) is one of the most successful orthopaedic interventions,[1] and is the recommended treatment for severe hip osteoarthritis (OA).[2 3] However, published reports suggest considerable variability in outcomes and revision rates even within groups implanted with the same prosthesis and within the same institution. A considerable body of work has explored the factors related to implant and procedure that influence outcome after THR, sometimes in long term studies involving the analysis of some 200 000 implant procedures.[1 4 5] Although these investigations have successfully identified a number of important variables related to implant, cement, procedure, surgeon, etc, that influence an outcome usually expressed as "implant survival", patient related factors other than age or sex that influence outcome after THR have received comparatively little attention. In addition, very few studies have focused on patient relevant outcomes, as contrasted with implant survival, or are truly prospective.[6] This study is to our knowledge the first long term community based, prospective, and consecutive follow up after THR for OA with a matched reference group and where well validated patient relevant outcome instruments have been used.

The purpose of this study was to investigate long term patient relevant outcomes after unilateral total hip replacement for OA, and identify non-responders to this intervention and patient related predictors that identify them.

PATIENTS AND METHODS

Patients

Two hundred and nineteen patients (120 women) with a mean age at index surgery of 71 years (50–92) were consecutively included in the study. All patients had a primary unilateral THR performed because of primary OA between September 1995 and October 1998 at the department of orthopaedics

in Halmstad, Sweden. Primary OA was defined as idiopathic OA in contrast to secondary OA caused by metabolic, anatomical, traumatic, and inflammatory conditions.[7] Most (n=155) of the replacements were performed with both components cemented, while in 64 (29 women) the acetabular component was uncemented. The mean age of this subgroup at surgery was 62 years (50–72).

The cemented replacements were made using a cemented Lubinus acetabular component and a cemented Lubinus SP II femoral component (Link) and the hybrid replacements with an uncemented Trilogy or HGC acetabular component (Zimmer) and the cemented Lubinus SPII (Link) or Anatomic (Zimmer) femoral component.

Patients with the hybrid prosthesis were advised to partially bear weight for the first eight weeks after surgery. Surgical technique, cementing technique, rehabilitation, and follow up evaluation were otherwise identical for both groups.

Patients in need of bilateral surgery (n=14) were analysed separately at the final follow up. Patients who had recurrent dislocations of the prosthesis (n=1) during the first follow up year were excluded before analysis.

The preoperative hip radiographs were classified by two radiologists according to OARSI criteria with a radiographic atlas as a guide.[8] OA severity was graded from 0 to 3 in accordance with the degree of joint space narrowing, where 3 indicates severe OA. 71% of the patients had severe radiological OA, 28% moderate OA, and 1% mild OA.

....

Reference group

Two hundred and nineteen consecutive patients were included in the study. Eighty six of these were included from September 1997 to October 1998. During this period 258 subjects (three for each patient) were identified in the National Population Records. The subjects were matched to the patients by age, sex, and municipality. Short Form-36 (SF-36) and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questionnaires were sent to these 258 subjects with an explanatory cover letter. In the cover letter they were told not to respond if they had hip complaints. Hip complaints were defined as pain or diminished range of motion in their hips. One hundred and seventeen (45%) answered the first inquiry. Their mean age was 72 years (range 52–92, 57% women, 43% men). One hundred and forty one subjects did not answer the first inquiry. Their mean age was 72 years (range 50–90, 52% women, 48% men). No reminder letters were sent as the number was regarded as sufficient for comparison of groups. Additional letters with the same questionnaires were then sent to these subjects at the same intervals as for the patients. At the follow up after 28–65 months eight subjects had died. 83 (76%) of the remaining 109 subjects responded to the mailing. Their mean age was 71 years (range 52–86, 55% women, 45% men). These 83 subjects were used for comparison with the patients at baseline as well as at follow up.

Design of the study

Self report with the patient administered questionnaires SF-36 and WOMAC was obtained preoperatively, at 3, 6, 12 months and at a final follow up at 3.6 years (26–65 months, mean 43 months, median 40 months) after the index THR surgery.

Questionnaires

The SF-36 measures three major health attributes (functional status, wellbeing, overall health) in eight subscales. These include PF (physical function), RP (role limitations due to physical health), BP (bodily pain), GH (general health), VT (vitality), SF (social function), RE (role limitations due to emotional health), and MH (mental health).[9] The SF-36 scores are calculated on a 0–100 worst to best scale. SF-36 is translated and validated for Swedish conditions.[10] It has previously been used in follow up studies of THR.[11]

WOMAC (LK 3.0) was used as the disease-specific instrument. WOMAC is a self administered instrument validated for OA in the lower extremities and for evaluating outcome after THR.[12] It consists of 24 items grouped into three categories: pain (five questions), stiffness (two questions), and physical function (seventeen questions). It is reliable and valid for Swedish conditions.[13] To enhance the interpretation, WOMAC is transformed to a 0–100 worst to best scale.[13–15] Because this instrument was not available and validated for Swedish conditions when the study was started, it was used at baseline for the last 92 patients only. There were no differences in age and sex between these 92 patients and the 106 who were included earlier. For cross sectional analysis of outcome at 3.6 years results from 198 patients were used and for longitudinal analysis 92 patients were used.

Supplementary questionnaire

Questions about postoperative complications, preoperative and postoperative comorbidity, social circumstances, and patient satisfaction were asked at the final follow up. The patients and the reference group received the same questions except those concerning postoperative complications and patient satisfaction.

Postoperative complications

Three questions were asked about serious postoperative complications: dislocation of the prosthesis, deep infection in the hip joint, and reoperation. The self reported data were compared with data from the patients' case records. The definition of postoperative complication in this study referred to a positive answer to one of the three questions about postoperative complications.

General comorbidity

Fourteen questions were asked about intercurrent diseases preoperatively and in the present situation.[11 16] Questions were asked about the presence of 12 comorbid conditions or body areas with problems (heart, hypertension, peripheral arteries, lung, diabetes, neurological problems, cancer, ulcer, kidney disease, vision, back pain, and psychiatric disease). The questions were multiple choice (yes, no, don't know). The total number of conditions or problems reported was used as a summary variable (0, 1, 2, or more), a method shown to be valid in this kind of follow up.[11]

Musculoskeletal comorbidity

Two questions were asked about the need of walking assistance and walking distance, preoperatively and in the present situation,[17 18] two questions were asked about the need for analgesics due to pain in the operated hip joint or due to pain elsewhere. One question was asked about the experience of regional or widespread pain lasting more than three months during the past 12 months.[16] One question was asked about joint replacement in the contralateral hip or in the knees since the THR. The final questions concerned fractures in the spine, wrist, hip, or elsewhere.

Social circumstances

One question was asked about the living circumstances preoperatively and in the present situation. One question was asked about the civil status and one about the main profession and the present profession or occupation.

Patient satisfaction

One question was asked about how satisfied the patients were with the result of the operation. The question was: "Overall, how satisfied are you with the result of your hip replacement surgery". The alternative answers were: very satisfied, satisfied, dissatisfied, very dissatisfied.

Test-retest reliability

One hundred and ninety eight patients received questions about comorbidity at the final follow up. Of those, 60 patients selected at random received the same questionnaire three months later; 54 (90%) answered. In the analysis of the answers the answer "don't know" was interpreted as "no". The measurement of agreement, κ value, was 0.609–1.000.

Validity

The questions used in the questionnaire about comorbidity have in previous studies been tested for validity.[11 16–18]

Non-responders

Non-responders were identified according to three different sets of criteria. Firstly, as the patients who scored worst (the lowest quartile) in WOMAC function at follow up. Secondly, as the patients with the least absolute improvement, defined as improvement of <20/100 score units in WOMAC function between the preoperative value and that at follow up at 3.6 years. Thirdly, criteria have been suggested for the identification of non-responders and responders in a clinical trial of drug treatment for hip OA, the OARSI set of criteria.[19] These criteria are based on pain, function, and the patient's global assessment. High improvement in pain is defined as an absolute change of 30/100 score units and a relative change of 50%. High improvement in function is defined as an absolute change of 20/100 score units and a relative change of 50%. A

Table 1 Preoperative and 3.6 year postoperative mean scores and (standard deviations) for the three WOMAC subscales, for the patients, mean age 71 (50–88), and for the reference group, mean age 71 (52–86). The scale is 0–100, worst to best.

Subscale	Patients preoperative (n=92)	Patients 3.6 years postoperative (n=198)	Reference group at baseline (n=83)	Reference group 3.6 years after baseline (n=83)
WOMAC pain	45 (17.2)	82 (20.3)	88 (17.7)	87 (19.8)
WOMAC stiffness	39 (16.3)	78 (22.2)	88 (18.6)	84 (20.6)
WOMAC function	38 (14.8)	74 (21.7)	87 (18.9)	84 (21.8)

person is characterised as a responder if either of these two criteria is fulfilled. Moderate improvement is defined as (*a*) an absolute change in pain of 15/100 score units and a relative change of 25%; (*b*) an absolute change in function of 10/100 score units and a relative change of 20%; (*c*) an absolute change of 10/100 units or a relative change of 20% in patient's global assessment. A person can be labelled a responder if two of the three (*a*, *b*, *c*) criteria are fulfilled.

Statistics
Statistical analysis was done with the SPSS 10.0 package. For comparison between groups, the Mann-Whitney test was used. For comparison of preoperative and postoperative questionnaire data Wilcoxon's signed rank test was used. For comparing the frequency of comorbidities in subgroups the χ^2 test was used. Age, sex, body mass index, number of comorbid conditions, occurrence of postoperative complications, presence of type of pain, preoperative scores of SF-36 BP (bodily pain), PF (physical function), MH (mental health), and social circumstances were included in a stepwise multivariate logistic regression analysis. Odds ratios were expressed for a one year (for age) or 0–100 scale unit increase.

RESULTS
Of the 219 patients, eight died during the follow up period and 13 did not participate. Thus the results for 198 patients (94%, 106 women), with a mean age at the time of surgery of 71 years, are presented for an average follow up time of 3.6 years (range 26–65 months, median 40, mean 43) (tables 1 and 2). There was no correlation between the patients' follow up time between 26 and 65 months and the physical function score for either WOMAC or SF-36. (r_s=−0.11, p=0.122).

During the follow up period (12–65 months) 25 patients had a contralateral THR and six patients had a total knee replacement.

There was no difference in postoperative outcome between patients who received a hybrid or cemented THR as expressed by WOMAC and SF-36 when the data were adjusted for age (data not shown).

Of the 117 subjects in the reference group, eight died during the follow up period and 26 abstained from participating. Thus the results for 83 subjects (76%, 46 women) with a mean age at time of the start of the study of 71 years are presented (tables 1 and 2).

Postoperative complications
Two patients had been reoperated on after the first follow up year (one owing to recurrent hip implant dislocations and one because of a deep infection). A further three patients had recurrent dislocations after the first follow up year and one of those also sustained an infection.

Improvement in pain and function
The patients, evaluated as a group, almost doubled their absolute scores in all WOMAC subscales (p<0.0001) at the final follow up compared with the preoperative scores (table 1). They also reported improved scores at the final follow up in all SF-36 subscales (p<0.0001) except GH (general health) (p=0.2) (table 2).

Twenty three per cent of the patients reported a WOMAC function score <30 and 57% a score <40 preoperatively, while 3.3% reported a value <30 and 5.4% a value <40 at the final follow up.

Individual change in pain and function
The individual change in pain and function measured by WOMAC was calculated as the difference between preoperative and 3.6 years postoperative score (fig 1). The cut off level for a significant improvement was set at 10/100 units, previously defined as the smallest detectable clinical improvement.[20] At 3.6 years after surgery 31% of the patients had improved by ≤10 score units for pain or function compared with their value before surgery. Thus, 20/92 patients had improved by <10 score units for pain only, eight patients by <10 score units for function only, and one patient by <10 units for both pain and function. Twenty of 92 patients improved by <20/100 score units for pain only, and 15 patients by <20 units for function only, while 5 improved by <20 units

Table 2 Preoperative and 3.6 year postoperative mean scores and (standard deviations) for the SF-36 subscales, for the patients, mean age 71 (50–88), and the reference group, mean age 71 (52–86). The scale is 0–100, worst to best.

| Subscale	Patients preoperative (n=198)	Patients 3.6 years postoperative (n=198)	Reference group at baseline (n=83)	Reference group 3.6 years postoperative (n=83)
SF-36 PF	30 (19.6)	60 (25.1)	75 (20.8)	71 (26.7)
SF-36 RP	9 (21.4)	48 (44.0)	65 (42.3)	58 (44.8)
SF-36 BP	30 (16.8)	66 (26.2)	69 (27.9)	68 (27.6)
SF-36 GH	68 (19.9)	66 (22.0)	66 (24.4)	62 (24.8)
SF-36 VT	49 (20.9)	64 (24.3)	65 (25.8)	65 (26.4)
SF-36 SF	64 (26.2)	84 (22.6)	85 (22.2)	80 (27.2)
SF-36 RE	37 (42.5)	65 (42.4)	73 (34.9)	72 (42.4)
SF-36 MH	70 (21.0)	78 (20.1)	82 (19.0)	79 (19.4)
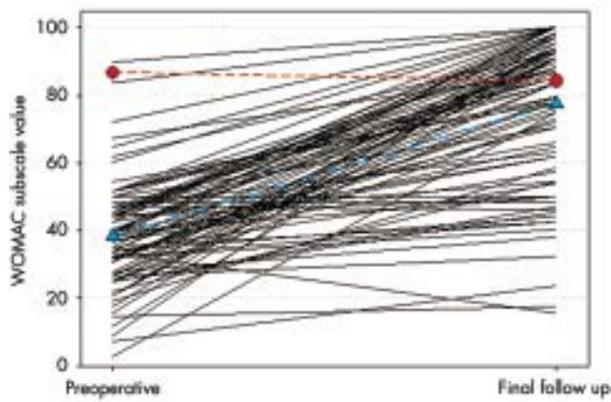
Figure 1 Preoperative (baseline) and 3.6 year postoperative THR scores of WOMAC function for OA. The scale is 0–100, worst to best. Each line represents one patient (n=92). The blue line with triangle symbols represents the average scores for the patient group. The red line with circle symbols represents average scores for WOMAC function for the reference group (n=83).

for both pain and function. Among those subjects who had improved by <20 units for pain or function at 3.6 years, 6 had also improved by <20 units at the one year follow up.

Patients who received bilateral THR
The patients who received bilateral THR during the first follow up year were examined at the 3.6 year follow up. A comparison showed no difference at this time in outcome, as measured by the SF-36, between the patients who had received bilateral THR and those with unilateral THR. However, there was a difference in the WOMAC subscale pain, with the patients who received bilateral THR scoring better (p=0.008) than the patients who received unilateral THR (table 3).

Comparisons between patients and reference group
Pain and physical function
At 3.6 years after THR the patients on average reported worse WOMAC pain (82 *v* 87, p=0.006) and worse WOMAC function (74 *v* 84, p<0.0001) than the reference group. There were no differences in the SF-36 subscales between the two groups except PF (physical function) where the patients reported worse function (60 *v* 71, p<0.0001) than the reference group (tables 1 and 2).

Frequency of comorbidities
Before THR the patients, in comparison with the reference group, reported a higher prevalence of low back pain (patients 32% *v* reference 15%, p=0.001), but a lower prevalence of pulmonary disease (patients 1% *v* reference 7%, p=0.005). No other differences between the patient and reference groups before surgery were found.

Table 4 Comparison of the prevalence of comorbid conditions between patients (n=198) and controls (n=83) at the 3.6 year follow up analysed with a χ^2 test. Results are shown as percentages

Comorbid condition	Patients	Controls	p Value
Heart	9.2	10.0	0.1
Hypertension	27.8	26.5	0.3
Peripheral arteries	12.7	13.3	0.4
Lung	0.5	7.2	0.004*
Diabetes	7.1	6.0	0.8
Neurological disease	2.0	3.6	0.7
Cancer	5.6	3.6	0.5
Peptic ulcer	1.5	2.4	0.8
Kidney disease	0	1.2	0.3
Vision problems	1.5	2.4	0.6
Low back pain	21.5	16.9	0.2
Psychiatric disease	2.6	2.4	0.3
Unilateral hip pain	20.2	4.8	0.001*
Regional pain	59.1	43.9	0.002*
Widespread pain	6.6	7.3	0.8

At the final follow up after THR there was no difference between the two groups in the prevalence of low back pain or in widespread pain. The prevalence of patients' low back pain had decreased from 32% to 21.5%. However, at this time patients had more regional pain (patients 59%, reference 44%, p=0.002) and unilateral hip pain (patients 20%, reference 5%, p=0.001). The prevalence of pulmonary disease was still lower among the patients than the reference group (0.5% and 7%, respectively, p=0.004) (table 4). At the 3.6 year follow up 33% of the patients used walking assistance, while the corresponding proportion for the reference group was 20% (p=0.03). Consistent with this, 42% of the patients and 57% of the reference group reported an unlimited walking distance, respectively (p=0.03).

Non-responders
As noted in the "Patients and methods" section, non-responders were identified in three different ways. Firstly, the lowest quartile in WOMAC function at the final follow up (25% are defined as non-responders). Secondly, an absolute improvement between baseline and 3.6 years follow up in WOMAC function of <20 score units (22% are defined as non-responders). Thirdly, by the OARSI criteria (9% are defined as non-responders) (table 5). Figure 3 illustrates the overlap between these three groups. There was no difference in age and sex between these three subgroups. We have chosen to report the results based on the first definition (lowest quartile) (see "Discussion").

Preoperative predictors for non-responders
The preoperative data for the patients who scored worst (the lowest quartile, mean score 44.8 (SD 11.4, range 15.6–61.8)) in

Table 3 Comparison between outcomes at 3.6 years after unilateral or bilateral THR. Results are shown as mean (SD)

WOMAC subscale	Unilateral THR (n=198)	Bilateral THR (n=14)
Pain	82 (20.3)*	96 (6.4)*
Stiffness	78 (22.2)	78 (24.5)
Physical function	74 (21.7)	77 (18.8)

Unilateral THR, mean age 71, bilateral THR mean age 74. The bilateral surgery was made during the first follow up year. The scale is 0–100, worst to best.
*p= 0.008.

Table 5 Stratified baseline data of WOMAC function and SF-36 PF and the percentage of patients with an improved or worse score at the final follow up

	>20 Score change	≤20 Score change
Baseline WOMAC function (n=92)		
0–29.9 (25%)	83	17
30–59.9 (66%)	78	22
60–100 (9%)	62	38
Baseline SF-36 PF (n=198)		
0–29.9 (51%)	73	27
30–59.9 (38%)	70	30
60–100 (11%)	29	71

Table 6 Univariate and multivariate logistic regression analysis of preoperative predictors and postoperative characteristics for the patients who scored worst (lowest quartile) in WOMAC function at 3.6 year follow up compared with the patients who scored better

Predictor	No	OR	95% CI	p Value	Multi-variate step (n=160)	OR	95% CI	p Value
Preoperative characteristics								
Age	189	1.09	1.04 to 1.15	0.001	1	*1.09	1.03 to 1.15	0.002
Sex	189	1.35	0.70 to 2.58	0.37				
BMI	170	1.11	1.01 to 1.22	0.03				
Comorbid conditions >2	189	1.48	0.62 to 3.51	0.38				
SF-36 BP	186	0.97	0.95 to 0.99	0.007	2	*0.97	0.94 to 0.995	0.018
SF-36 PF	183	0.98	0.96 to 0.997	0.03				
SF-36 MH	176	0.99	0.97 to 1.004	0.16				
Postoperative characteristics								
Comorbid conditions >2	189	2.06	0.93 to 4.55	0.08				
Complication	189	3.55	1.03 to 12.18	0.04				
Widespread pain	189	1.28	0.40 to 4.81	0.61				
Pain index hip	189	2.50	0.80 to 7.82	0.12				
Pain contralateral hip	189	1.95	0.75 to 5.10	0.17				
Low back pain	189	3.31	1.59 to 6.89	0.001	1	3.31	1.59 to 6.89	0.001
Living alone	189	1.50	0.77 to 2.90	0.23				

*Per one year or scale unit increase.
BP, bodily pain; PF, physical function; MH, mental health.

the WOMAC subscale physical function at 3.6 years after THR were compared in a multivariate analysis with the preoperative data for the three quartiles of patients who scored better. A higher degree of pain, as measured by the SF-36 preoperatively, and higher age predicted worse WOMAC function outcome at 3.6 years after surgery (table 6). It should be noted that the odds ratios are expressed for one year or scale unit difference. The number of preoperative comorbid conditions did not predict worse WOMAC function outcome.

In a univariate analysis, a high body mass index and worse physical function preoperatively as measured by SF-36 were predictors for worse outcome in WOMAC function at 3.6 years (table 6).

Postoperative characteristics of patients with poor function

Univariate analyses were made to compare the postoperative data at the final follow up for the patients who scored worst (lowest quartile) in the WOMAC subscale physical function with the patients who scored better. Low back pain and postoperative complications were associated with worse outcome. In a multivariate analysis, low back pain postoperatively was the only significant characteristic of the patients with a nonsuccessful result (table 6).

DISCUSSION

This is, to our knowledge, the first prospective long term follow up study of pain, function, and health related quality of life in patients with THR for OA using validated patient relevant outcome measures, including a comparison with a matched reference group. In this report we focus on the patient relevant outcome 3.6 years after unilateral THR for OA, and attempt to identify any predictors of poor outcome.

Treatment outcome can be assessed in different ways. Improvement,[21] not just the final outcome score, is an important measure for patients receiving a THR.[22] In our study the WOMAC pain and/or function score of almost one third of the patients did not improve by more than 10/100 units during the 3.6 year follow up period (figs 1 and 2). Surgeons should thus carefully consider other options before operating on patients with less preoperative pain or function scores because these patients will improve only slightly.[22]

Methodological issues

We used both a generic and a disease-specific instrument[23] because SF-36 shows a better gradient with comorbidities than the WOMAC on all three dimensions, while the WOMAC

scores provide a better gradient by the current physical condition.[11 23 24] This is the reason for using SF-36 in the comparisons between the patients and the reference group and WOMAC when trying to identify the predictors of poor outcome in hip OA. Figures 1 and 2 illustrate this difference between the WOMAC and SF-36. In fig 1, related to WOMAC, it is clear that the intervention for almost all patients has an effect on physical function (83/92), while this is more doubtful in fig 2, related to SF-36 (145/184). For the SF-36 subscale physical function (fig 2), we would expect factors other than the operated hip to influence outcome to a higher degree than for WOMAC function (fig 1).[25]

As a consequence there is a more obvious ceiling effect when using the WOMAC as outcome measurement after THR than with the SF-36 (figs 1 and 2).[26]

Several reasons motivated the identification and use in this study of a reference group matched for age, sex, and municipality. Firstly, patients compare themselves, not with other disabled people, but with how they were before the onset of symptoms.[27] Because such information is not available in a prospective study of the present kind, we chose instead to examine a matched reference group and to use these data as a surrogate. Secondly, a lack of normative data for WOMAC
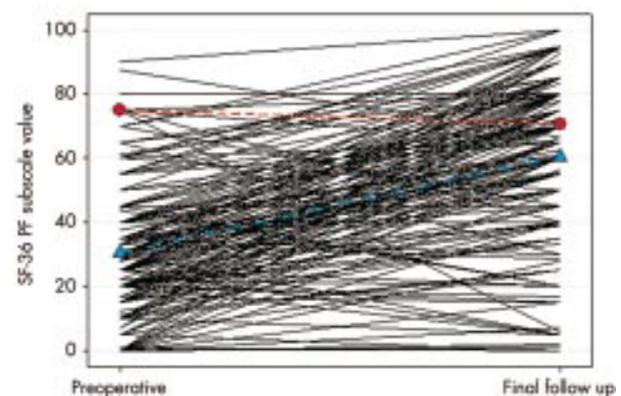

Figure 2 Preoperative (baseline) and 3.6 year postoperative THR scores of SF-36 PF (physical function) for OA. The scale is 0–100, worst to best. Each line represents one patient (n=198). The blue line with triangle symbols represents the average scores for the patient group. The red line with circle symbols represents average scores for the SF-36 PF for the reference group (n=83).

necessitated generating a matched reference group. This reference group also enabled us to gather data for SF-36, which complement the available normative Swedish SF-36 data.[10] The reference group was examined at the same times, at the same frequency, and with the same questionnaires as the THR study group. Hip problems are common in the community.[16 28 29] However, subjects with hip problems were asked not to return the questionnaire. Thus, the reference data used here should represent a "best" case.

Complications

Postoperative complications such as recurrent dislocations, deep infection, and revision surgery were as frequent as expected from Swedish national average data for THR (http://www.jru.orthop.gu.se), which suggests that the study cohort was representative in this respect. Consequently, the number of complications in relation to the number of patients was few, which makes it difficult to draw any conclusions about postoperative complications. However, patients with treated postoperative complications in this study seemed to attain a similar level of function at 3.6 years after THR as the patients without complications (data not shown). This is noteworthy, because in most follow up studies implant failure and revision are the end points for follow up.[30]

Non-responders

Non-responders were identified in three different ways. The first definition identified non-responders as the patients who scored worst (lowest quartile) in WOMAC function at the 3.6 year follow up. The second definition identified the patients with an absolute change of <20 score units on a 0–100 scale. The reason for the choice of an absolute change of 20 score units is the knowledge that in clinical trials of rehabilitation intervention and medical treatment the smallest detectable clinical improvement in WOMAC function and pain is 9–12 score units.[20 31] Knowing that the responsiveness of medical treatment in OA is about half as high (standardised response mean 0.5–0.8) as for causal surgical treatment (standardised response mean 0.8–3.1),[32–34] we decided to double the smallest detectable clinical improvement for medical treatment. This cut off point also concurs with the OARSI definition of high improvement in function.[19]

The third definition was an attempt to use an established definition of non-responders that takes into account both pain, function, and patient global assessment. The problem with using this definition in this study is that it was developed for clinical trials in pharmacological treatment of hip OA and not for surgical interventions. When these criteria were used only 8/92 (9%) patients were defined as non-responders at the 3.6 year follow up. Of those eight subjects, two reported that they were very satisfied, one satisfied, one dissatisfied with the outcome of the THR (four were missing values). Thus, even if the patient reports a bad outcome in pain and function she/he may be satisfied with the result.[35] Only 4% of all the patients reported that they were dissatisfied with the result of the THR. This is a higher degree of satisfaction after THR than that reported for total knee replacement in OA.[36] It has been stated that satisfaction is a wide concept, not necessarily relevant in outcome after THR.[37 38]

When comparing the three methods of identifying non-responders we found that the first criterion identified the greatest number of patients with a poor outcome in WOMAC function at 3.6 years. However, the overlap with the two other definitions was considerable. Because the two other methods of identifying non-responders to a great extent identified the same patients, the same predictors were identified when using these criteria (fig 3).

Determinants of poor outcome

In the present study old age predicted a poor postoperative outcome, compared with the younger patients. This finding is
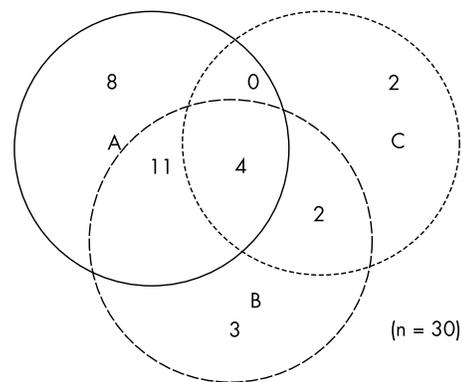

Figure 3 Non-responders to THR for OA identified in three different ways: (*a*) Defined as patients who scored worst (lowest quartile) in WOMAC function at 3.6 year follow up, (circle A, 23/92 (25%)); (*b*) defined as patients who reported an absolute improvement from baseline to 3.6 year follow up of <20 score units (circle B, 20/92 (22%)); (*c*) defined according to OARSI criteria (circle C, 8/92 (9%)). The patients included in this figure are based the longitudinal data (n=92).

consistent with a report that older people with self reported conditions restricting mobility in addition to arthritic pain in the hip or knee are at higher risk of psychological distress and physical dysfunction.[39]

We found that the number of comorbid conditions preoperatively did not predict a worse outcome postoperatively when measured by the WOMAC function. The same result was found when using the SF-36 PF (physical function) as outcome measure, even though it is expected to show a better gradient with comorbidities.

There was no difference between the patient group and the reference group in the number of comorbidities at the beginning of the study or at the follow up. This is consistent with previous observations which showed that OA is not predictive for the development of future comorbidities.[11 40]

Our results indicate that low back pain and pain in the hip not operated on is characteristic for patients who do not reach the same level of function postoperatively as the matched reference group (table 4). These findings are in concordance with a recently published study,[41] and indicate that WOMAC captures not only knee or hip pain and dysfunction but is influenced by the presence of low back pain.[42 43] The decreased prevalence of reported low back pain after THR seen in the patient cohort could in part be explained by a changed postoperative pain threshold,[44] and in part by an improved pain-free range of motion in the operated hip. All patients were offered physical therapy postoperatively and almost everyone accepted.

The general health status of people with pain in the hip or knee is comparable to that of a reference group without such pain,[39] but the health status is worse when pain in the hip or knee occurs in combination with other mobility restricting conditions—for example, pain in other joints and other musculoskeletal problems such as back pain. Pain in the hip not operated on may be a symptom of bilateral OA in these patients,[29] even in the absence of radiological change.

Limitations

A limitation of the study is the variable follow up time (26–65 months). Thus, the patients with the longest follow up time have reached an older age than the patients with a shorter follow up time. On the other hand, these patients have had a longer time for rehabilitation and recovery. However, there was no correlation between follow up time and function. The 14 patients who had been operated on the contralateral side during the first follow up year were analysed separately at the final follow up. There was no difference in outcome for

these patients, compared with the main group with unilateral THR, except that they had less pain (as measured by the WOMAC pain subscale). Pain in the non-operated hip was also more common among patients with unilateral THR than controls, suggesting the presence of bilateral OA.

CONCLUSIONS

In this prospective cohort study we have shown that patients with hip OA over 50 years of age at 3.6 years after unilateral THR report similar pain but have a lower level of physical function than a matched reference group without hip complaints. The difference in function is explained, at least in part, by the presence of musculoskeletal comorbidities such as low back pain and pain in the non-operated hip. Most patients have a significant improvement in pain and function after THR. However, at 3.6 years after surgery almost one third of the patients report only a low degree of improvement or have worsened after the THR. Some of this lack of response may be explained by old age and a high degree of pain at the time of surgery. This is supported by previous reports that have suggested surgery earlier rather than later in the course of OA.[45] A report of low back pain has a significant impact on postoperative function after THR, which is of importance when planning rehabilitation.

A major proportion of the variability in outcome after THR for OA remains unexplained. Further patient relevant outcome data collected by the use of self administered questionnaires, and a better understanding of the role of expectations and satisfaction is essential to improve our understanding of who does or does not benefit from these surgical procedures.[46] Much effort is focused on the "technical" aspects of this intervention to improve its success rate further. Perhaps we stand to gain as much or more outcome improvement by a better understanding of these "other" factors.

ACKNOWLEDGEMENTS

. . . .
Authors' affiliations
A-K Nilsdotter, I F Petersson, E M Roos, L S Lohmander, Department of Orthopaedics, Lund University Hospital, Spenshult Hospital of Rheumatic Diseases, Halmstad, Sweden

REFERENCES

1 **Herberts P**, Malchau H. Long-term registration has improved the quality of hip replacement: a review of the Swedish THR Register comparing 160,000 cases. Acta Orthop Scand 2000;71:111–21.
2 **ACR**. Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines. Arthritis Rheum 2000;43:1905–15.
3 **Pendleton A**, Arden N, Dougados M, Doherty M, Bannwarth B, Bijlsma JW, et al. EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2000;59:936–44.
4 **Birtwistle SJ**, Wilson K, Porter ML. Long-term survival analysis of total hip replacement. Ann R Coll Surg Engl. 1996;78(3 (Pt 1)):180–3.
5 **Callaghan JJ**, Albright JC, Goetz DD, Olejniczak JP, Johnston RC. Charnley total hip arthroplasty with cement. Minimum twenty-five-year follow-up. J Bone Joint Surg Am 2000;82:487–97.
6 **Britton AR**, Murray DW, Bulstrode CJ, McPherson K, Denham RA. Pain levels after total hip replacement: their use as endpoints for survival analysis. J Bone Joint Surg Br 1997;79:93–8.
7 **Sangha O**. Epidemiology of rheumatic diseases. Rheumatology (Oxford) 2000;39(suppl 2):3–12.
8 **Altman RD**, Hochberg M, Murphy WA, Wolfe FJ, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3(suppl A):3–70.
9 **Ware JE**, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). Med Care 1992;30:473–83.
10 **Sullivan M**, Karlsson J, Ware JR. The Swedish SF-36 Health Survey-I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995;41:1349–58.
11 **Bombardier C**, Melfi CA, Paul J, Green R, Hawker G, Wright J, et al. Comparison of a generic and a disease specific measure of pain and physical function after knee replacement surgery. Med Care 1995;33:131–44.
12 **Bellamy N**, Buchanan W, Goldsmith C. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes following total hip or knee arthroplasty in osteoarthritis. J Orthop Rheumatol 1988;1:95–108.
13 **Roos EM**, Klässbo M, Lohmander LS. WOMAC osteoarthritis index. Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and MacMaster Universities. Scand J Rheumatol 1999;28:210–15.
14 **Stucki G**, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, et al. [Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) arthrosis index]. Z Rheumatol 1996;55:40–9.
15 **Creamer P**, Lethbridge-Cejku M, Hochberg MC. Where does it hurt? Pain localization in osteoarthritis of the knee. Osteoarthritis Cartilage 1998;6:318–23.
16 **Bergman S**, Herrström P, Högström K, Petersson I, Svensson B, Jacobsson L. Chronic musculoskeletal pain, prevalence rates, and sociodemographic associations in a Swedish population study. J Rheumatol 2001;28:1369–77.
17 **Harris WH**. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. J Bone Joint Surg Am 1969;51:737–55.
18 **Söderman P**, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? Clin Orthop 2001:189–97.
19 **Dougados M**, Leclaire P, van der Heijde D, Bloch DA, Bellamy N, Altman RD. Response criteria for clinical trials on osteoarthritis of the knee and hip: a report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials response criteria initiative. Osteoarthritis Cartilage 2000;8:395–403.
20 **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.
21 **Felson DT**, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.
22 **MacWilliam CH**, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. Health Serv Res 1996;31:623–38.
23 **Hawker G**, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) instrument in measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22:1193–6.
24 **Jones CA**, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. J Rheumatol 2000;27:1745–52.
25 **Angst F**, Aeschlimann A, Steiner W, Stucki G. Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. Ann Rheum Dis 2001;60:834–40.
26 **Stucki G**, Liang MH, Stucki S, Katz JN, Lew RA. Application of statistical graphics to facilitate selection of health status measures for clinical practice and evaluative research. Clin Rheumatol 1999;18:101–5.
27 **Donovan J**. Patient education and the consultation: the importance of lay beliefs. Ann Rheum Dis 1991;50(suppl 3):418–21.
28 **Fear J**, Hillman M, Chamberlain MA, Tennant A. Prevalence of hip problems in the population aged 55 years and over: access to specialist care and future demand for hip arthroplasty. Br J Rheumatol 1997;36:74–6.
29 **Birrell F**, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Health impact of pain in the hip region with and without radiographic evidence of osteoarthritis: a study of new attenders to primary care. The PCR Hip Study Group. Ann Rheum Dis 2000;59:857–63.
30 **Söderman P**, Malchau H, Herberts P. Outcome after total hip arthroplasty: Part I. General health evaluation in relation to definition of failure in the Swedish National Total Hip Arthoplasty register. Acta Orthop Scand 2000;71:354–9.
31 **Angst F**, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001;45:384–91.
32 **Laupacis A**, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, et al. The effect of elective total hip replacement on health-related quality of life. J Bone Joint Surg Am 1993;75:1619–26.
33 **Nilsdotter AK**, Roos EM, Westerlund JP, Roos HP, Lohmander LS. Comparative responsiveness of measures of pain and function after total hip replacement. Arthritis Rheum 2001;45:258–62.
34 **Wright JG**, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239–46.

35 **Mancuso CA**, Salvati EA, Johanson NA, Peterson MGE, Charlson ME. Patients'expectations and satisfaction with total hip arthroplasty. J Arthroplasty 1997;12:387–96.

36 **Robertsson O**, Dunbar M, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. Acta Orthop Scand 2000;71:262–7.

37 **Söderman P**. On the validity of the results from the Swedish National Total Hip Arthroplasty register. Acta Orthop Scand Suppl 2000;71:1–33.

38 **Burton KE**, Wright V, Richards J. Patients' expectations in relation to outcome of total hip replacement surgery. Ann Rheum Dis 1979;38:471–4.

39 **Hopman-Rock M**, Odding E, Hofman A, Kraaimaat FW, Bijlsma JW. Differences in health status of older adults with pain in the hip or knee only and with additional mobility restricting conditions. J Rheumatol 1997;24:2416–23.

40 **Gabriel SE**, Crowson CS, O'Fallon WM. Comorbidity in arthritis. J Rheumatol 1999;26:2475–9.

41 **Jones CA**, Voaklander DC, Johnston DW, Suarez-Almazor ME. The effect of age on pain, function, and quality of life after total hip and knee arthroplasty. Arch Intern Med 2001;161:454–60.

42 **Wolfe F**. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology (Oxford) 1999;38:355–61.

43 **Harcourt WG**, White SH, Jones P. Specificity of the Oxford knee status questionnaire. The effect of disease of the hip or lumbar spine on patients' perception of knee disability. J Bone Joint Surg Br 2001;83:345–7.

44 **Kosek E**, Ordeberg G. Lack of pressure pain modulation by heterotopic noxious conditioning stimulation in patients with painful osteoarthritis before, but not following, surgical pain relief. Pain 2000;88:69–78.

45 **Fortin P**, Clarke A, Joseph L, Liang M, Tanzer M, Ferland D, *et al*. Outcomes of total hip and knee replacement. Arthritis Rheum 1999;42:1722–8.

46 **Lingard E**, Hashimoto H, Sledge C. Development of outcome research for total joint arthroplasty. J Orthop Sci 2000;5:175–7.

What's in the next issue

Future content

See which articles have just been accepted for publication and
preview the table of contents for the next issue a month before it is published

www.annrheumdis.com