Radiographic stage of osteoarthritis or sex of the patient does not predict one year outcome after total hip arthroplasty

A-K Nilsdotter, Y Aurell, A-K Siösteen, L S Lohmander, H P Roos

Abstract

Objective—To investigate whether patients with severe radiographic osteoarthritis (OA) have a different outcome at one year after total hip replacement than patients with moderate radiographic OA. To investigate sex related differences in preoperative radiographic and self reported status and in postoperative outcome.

Methods—184 patients (96 women) with a mean age at surgery of 71.4 years (50–92), with primary OA of the hip were investigated preoperatively and six and 12 months postoperatively with two self administered questionnaires, SF-36 and WOMAC. The radiographs were evaluated by two independent radiologists using an atlas. Minimal joint space narrowing, osteophytes, cysts, sclerosis, and deformity were assessed. A summary grade 0–3 was made, based on joint space, where 3 is severe OA. The reference population for SF-36 consisted of 2901 subjects matched for age and sex from the general Swedish population.

Results—162 patients fulfilled the study criteria. 113 had grade 3, 47 grade 2, and two grade 1 radiographic OA. There was no difference in preoperative or postoperative pain and physical impairment between patients with moderate and severe radiographic OA. There were no sex related differences in preoperative radiographic status, or in postoperative outcome. Neither were any differences in preoperative radiographic status of OA found in patients with previous total hip replacement of the contralateral hip, compared with those who had not been operated on before. All patients, regardless of preoperative radiographic OA stage, showed significant postoperative improvement and at one year achieved a health related quality of life similar to that of the reference group.

Conclusion—The severity of radiographic changes indicating OA often weighs heavily in the surgeon's decision to perform a total hip replacement. Yet, the findings of this study emphasise that the preoperative radiographic stage of OA has no correlation with the postoperative outcome after one year. Furthermore, this study failed to detect any sex related differences in preoperative radiographic and self reported status or in postoperative outcome of hip replacement.

(*Ann Rheum Dis* 2001;**60**:228–232)

Department of Orthopaedics, Halmstad County Hospital, Sweden
A-K Nilsdotter

Department of Radiology, Halmstad County Hospital, Sweden
Y Aurell
A-K Siösteen

Department of Orthopaedics, Lund University Hospital, Sweden
A-K Nilsdotter
L S Lohmander

Department of Orthopaedics, Helsingborg County Hospital, Sweden
H P Roos

Correspondence to:
Dr A-K Nilsdotter,
Department of
Orthopaedics, Halmstad
County Hospital, SE-30233
Halmstad, Sweden
anna.nilsdotter@adept.se

Hip arthroplasty is one of the most common orthopaedic interventions. In Sweden about 10 000 hip arthroplasties are performed annually, corresponding to an annual rate of about 400 per 100 000 inhabitants aged 50 and older.[1] The patients' experience of pain, clinical findings, functional impairment, and radiographic changes are taken into consideration when the surgical treatment decision is made.[2]

Patients and surgeons may differ in their concerns and priorities for total hip replacement.[3 4] However, patients can provide a reliable and valid judgment of health status and the benefits of treatment,[5] and there is consensus about the centrality of the patients' point of view in monitoring medical outcome.[6]

For osteoarthritis (OA), the OMERACT group has recommended four domains to be evaluated: pain, physical function, joint imaging, and patient global assessment.[7] Several questionnaires for measuring health related quality of life have been published: the Nottingham Health Profile (NHP),[8] Sickness Impact Profile (SIP)[9] and Short Form-36 (SF-36).[6] These instruments have been proved to be reliable and valid.[10 11] Since the introduction of the SF-36 in 1988 it has been the most widely used generic health instrument. It has been translated into many languages and different cultures, which makes international comparisons possible.[12]

WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) is a disease-specific measure, developed for OA in the hip and knee.[13] WOMAC has shown greater responsiveness to change over time than SF-36,[14] whereas SF-36 discriminates better between subjects based on levels of self reported general health.[15] In the measurement of outcome it is desirable to include both a generic instrument and a disease-specific instrument.[15 16]

Plain radiography is at present the most economical and easily available imaging technique for semiquantitative measurement of the morphology of OA.[17] It is vital that the techniques used to obtain and read the radiographs are standardised for comparative studies.[16] Training and standardisation can be assisted by the use of atlases.[18] Several studies of outcome after total hip replacement have been published in recent years, but none of them related the outcome to the preoperative radiographic classification.[19–22]

The purpose of this prospective study was to investigate (*a*) the relation between the preoperative radiographic status and the postopera-

tive outcome in hip arthroplasty for OA; (*b*) the relation between the preoperative radiographic status of the hip and the patient's own preoperative experience of wellbeing, pain, and physical impairment; and (*c*) if there were sex related differences in radiographic and self reported status or in postoperative outcome.

Patients and methods

Two hundred and thirty patients aged 50 and older were assigned to receive a total hip replacement because of primary OA during September 1995 to December 1997 at the department of orthopaedics in Halmstad, Sweden. One hundred and eighty four patients (96 women, 88 men) with a mean age at the time of surgery of 71.4 years (50–92) were included in the study. Forty six patients (30 women, 16 men) (mean age 72.6 years (56–88)) were not included during the early recruitment phase because of administrative errors (no relation to patient-relevant factors). There were no differences in intercurrent diseases or mental health for those not included compared with the study group.

All patients had a primary unilateral total hip replacement performed. Eight different surgeons carried out the operations, all experienced hip surgeons. The majority (n=139) of the replacements were performed with both components cemented, but in 45 (18 women, 27 men) the acetabular component was uncemented. Their mean age at surgery was 61.3 years (50–72).

RADIOGRAPHIC EVALUATION

The preoperative hip radiographs were obtained with the patient supine. Focus to film distance was 100 cm. Classification was made on plain anteroposterior views of the pelvis radiograph and the hip joint on the relevant side. Lateral views were available on the relevant side but were not evaluated. The films were independently read and classified by two radiologists (YA, A-KS) according to the OARSI criteria with a radiographic atlas as a guide.[18] Minimal joint space, and osteophytes in the acetabulum and femoral head were graded visually by the atlas (0–3). Millimetre measurements of joint space were not made. Cysts, sclerosis, and deformity were recorded as present or absent. OA was graded from 0 to 3, based on the degree of joint space narrowing only, where grade 1 is mild, grade 2 moderate, and grade 3 severe OA. Before evaluating the radiographs the radiologists had learning sessions.

In five of 162 radiographs the radiologists did not agree in evaluation of OA grade, and a consensus decision was made. There was no test on inter- and intra-observer variability because previous studies[23 24] have shown high interrater agreement when using a comparable method.

QUESTIONNAIRES

Evaluation with SF-36 was made at the hospital the day before the operation, and at six and 12 months postoperatively. The SF-36 measures three major health attributes (functional
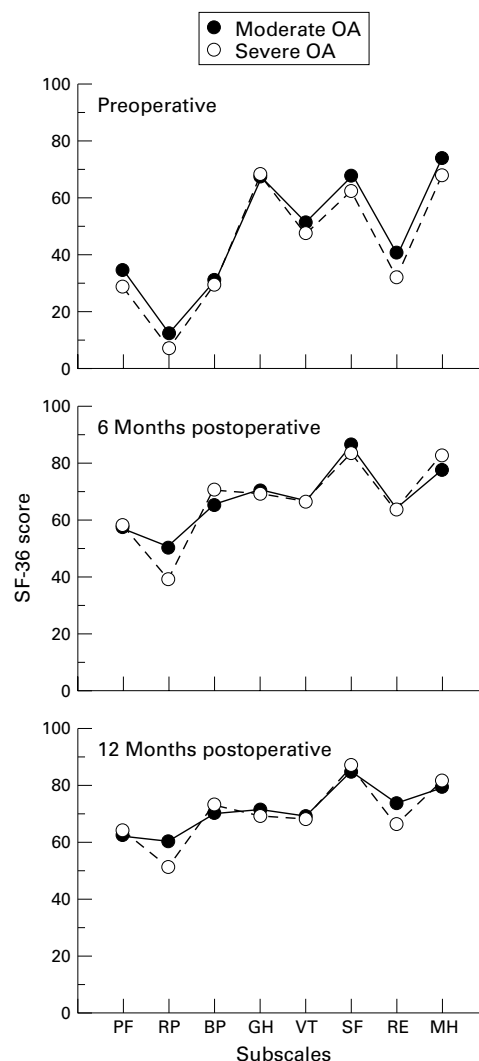

Figure 1   Mean values of SF-36 subscales preoperatively and six and 12 months postoperatively for patients with moderate (grade 2) (n=47) and severe (grade 3) (n=113) radiographic OA. The scale is 0–100, worst to best. PF = physical function; RP = role function physical; BP = bodily pain; GH = general health; VT = vitality; SF = social function; RE = role function emotional; MH = mental health.

status, wellbeing, overall health) in eight subscales. These include physical function, role limitations due to physical health, bodily pain, general health, vitality, social function, role limitations due to emotional health, and mental health.[6 25] The SF-36 scores are calculated on a 0–100 worst to best scale. Together, the eight subscales provide a health profile (fig 1). SF-36 has been translated and validated for Swedish conditions.[25] The reference population for SF-36 was collected from norms of the general Swedish population, and comprised 2901 subjects matched for age and sex.[25]

WOMAC[10] was used as the disease specific instrument. Evaluation with WOMAC was made preoperatively and at six and 12 months postoperatively. However, because this instrument was not available and validated for Swedish conditions when the study was started, it was used for the last 74 patients only. WOMAC is a self administered instrument validated for OA in the legs. It consists of 24 multiple choice

Table 1 Preoperative and postoperative mean scores and (standard deviations) of the three WOMAC subscales for radiographic moderate (grade 2) and severe (grade 3) osteoarthritis for the 55 patients investigated at all three occasions. The scale is 0–100, worst to best

WOMAC subscale	Preop Grade 2 (n=19)	Preop Grade 3 (n=36)	6 m Postop Grade 2 (n=19)	6 m Postop Grade 3 (n=36)	12 m Postop Grade 2 (n=19)	12 m Postop Grade 3 (n=36)
Pain	40.8 (12.8)	48.0 (18.1)	83.6 (16.1)	81.7 (17.4)	88.3 (15.1)	86.2 (14.7)
Stiffness	39.5 (16.3)	38.9 (16.3)	75.0 (14.0)	75.0 (17.7)	80.6 (17.3)	77.0 (20.6)
Physical function	37.0 (11.8)	40.6 (16.5)	73.4 (18.1)	73.7 (18.4)	80.4 (17.6)	80.0 (17.4)

items grouped into three categories: pain (five questions), stiffness (two questions), and physical function (17 questions). It is reliable and valid for Swedish conditions.[26] To make comparison easier with SF-36, WOMAC is transformed to a 0–100 worst to best scale.[26] The reference group for WOMAC was selected from a group matched for age, sex, and municipality and consisted of 59 subjects without hip complaints.

STATISTICS

For comparison of preoperative and postoperative questionnaire data Wilcoxon's signed rank test was used. For comparison between two subgroups the Mann-Whitney test was used. Analysis of variance (ANOVA) was used for comparison between differences in the clinical variables which could be explained by subgroup. A simple linear analysis of regression was made to assess how much of the variation in pain and physical function 12 months postoperatively could be explained by the preoperative values.

Results

Of the 184 patients with hip replacements, 22 were excluded during the study. Of these, 12 were operated on the contralateral side during the first follow up year, four refused to participate, one could not participate because of difficulties with the language, three died, and two had recurrent dislocations of the prosthesis. Thus the results for 162 patients (86 women), with a mean age at surgery of 71.2 years (50–92) are presented. Of these patients, 113 (70%) had grade 3, 47 (29%) grade 2, and 2 (1%) grade 1 radiographic OA, based on the semiquantitative assessment of minimum hip joint space. The prevalence of femoral head osteophytes was 3% for grade 0, 19% for grade 1, 20% for grade 2, and 58% for grade 3. The prevalence of osteophytes in the acetabulum was 7% for grade 1, 27% for grade 2, and 66% for grade 3.

Patients with severe preoperative radiographic OA did not differ in postoperative outcome as measured by SF-36 and WOMAC compared with patients with only moderate preoperative radiographic OA (fig 1, table 1). Furthermore, the grade of preoperative OA did not influence the severity of preoperative symptoms and impairments as measured by SF-36 and WOMAC (fig 1, tables 1 and 2). Likewise, the degree of joint space narrowing did not influence the degree of pain as measured preoperatively by SF-36 or WOMAC (p=0.373 and p=0.369, respectively). Neither did the grade of osteophytes in acetabulum or the femoral head influence the severity of preoperative pain (p=0.458 and p=0.302 for osteophytes in acetabulum and p=0.761 and p=0.465 for osteophytes in the femoral head, respectively). Consistent with these individual analyses of data, ANOVA failed to show any influence on clinical variables by the radiographic grade of OA.

Patients with a previous total hip replacement in the contralateral hip did not differ in their preoperative grade of radiographic OA or symptoms from the patients who had their first hip replacement performed (data not shown). There were no sex related differences in preoperative radiographic grade of OA or symptoms, or in postoperative outcome (data not shown).

As commented on above, the variation in pain and physical function as measured by WOMAC 12 months postoperatively could not be explained by sex of the patient or the preoperative radiographic grade of OA (ANOVA). However, there was a positive relation between

Table 2 Preoperative and 12 month postoperative mean scores and (standard deviations) of the SF-36 subscales for radiographic moderate (grade 2) and severe (grade 3) osteoarthritis. The scale is 0–100, worst to best

SF-36 subscale*	Preop Grade 2 (n=47)	Diff*	Postop Grade 2 (n=47)	Preop Grade 3 (n=113)	Diff	Postop Grade 3 (n=113)
PF	34 (19.9)	+28	62 (23.8)	29 (20.1)	+35	64 (22.9)
RP	12 (21.4)	+48	60 (41.8)	7 (20.4)	+44	51 (43.0)
BP	31 (14.0)	+39	70 (25.5)	30 (16.8)	+43	73 (24.6)
GH	67 (20.6)	+4	71 (20.3)	68 (19.1)	+1	69 (21.1)
VT	51 (20.8)	+18	69 (24.2)	47 (21.2)	+21	68 (21.1)
SF	67 (24.5)	+17	84 (27.7)	62 (26.2)	+25	87 (19.3)
RE	40 (42.2)	+33	73 (38.7)	32 (41.2)	+34	66 (40.4)
MH	73 (17.2)	+6	79 (21.3)	67 (21.0)	+14	81 (17.6)

*PF = physical function; RP = role function physical; BP = bodily pain; GH = general health; VT = vitality; SF = social function; RE = role function emotional; MH = mental health; Diff = difference between preoperative and 12 months' postoperative score.
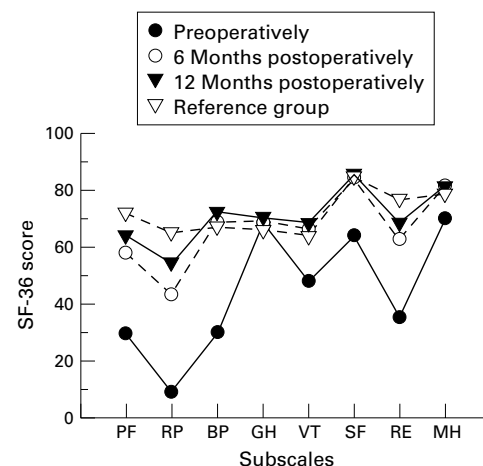

Figure 2 Pre- and postoperative mean values of SF-36 data for all patients (n=162) compared with the reference population. PF = physical function; RP = role function physical; BP = bodily pain; GH = general health; VT = vitality; SF = social function; RE = role function emotional; MH = mental health. The scale is 0–100, worst to best.

Table 3 *Preoperative and six and 12 months' postoperative mean scores and (standard deviations) of the three WOMAC subscales for all patients investigated compared with the reference group. The scale is 0–100, worst to best*

WOMAC subscale	Preop (n=74)	6 m Postop (n=74)	12 m Postop (n=74)	Reference group (n=59)
Pain	44.8 (17.1)	82.6 (16.3)	86.3 (15.5)	89.1 (18.0)
Stiffness	39.3 (15.9)	75.0 (16.4)	78.4 (19.3)	88.2 (16.3)
Physical function	39.0 (14.8)	74.1 (17.5)	78.7 (17.7)	85.1 (21.7)

the level of preoperative pain and pain at 12 months postoperatively (r^2=0.107, p=0.011) and between physical function preoperatively and at 12 months postoperatively, respectively (r^2=0.125, p<0.006) (linear regression).

Compared with the SF-36 reference population, the patients had a worse preoperative mean score for bodily pain (patients (p)=36; reference (r)=66), physical function (p=30; r=79), and for role limitations due to physical function (p=9; r=69), while the subscales general and mental health were comparable. Six months postoperatively, there was an improvement of all subscales, especially bodily pain (p from 36 to 69) and physical function (p from 30 to 58). These improvements were statistically significant at p<0.0001 for all subscales despite general health. Further improvements of all subscales appeared after 12 months. Between 6 and 12 months there were significant further improvements in physical function and role limitations due to physical function (p<0.0001). At 12 months after hip arthroplasty the health related quality of life of the patients did not differ from that of the SF-36 reference group (fig 2).

The outcome measured by WOMAC also showed a significant improvement for all three dimensions between preoperative status and six and 12 months (p<0.0001) (table 3). In contrast with the SF-36 outcome, the WOMAC outcome at 12 months postoperatively differed from the WOMAC reference group of 59 subjects (86 *v* 89 for pain (p=0.018), 78 *v* 88 for stiffness (p=0.002), and 79 *v* 85 for physical function (p=0.001)).

Discussion

The severity of the radiographic findings is an important factor in the surgeon's decision to carry out a total hip replacement.[2] There are, however, few studies of hip OA that support a relation between radiographic findings and clinical symptoms. Earlier studies of the knee[27] have shown an association of osteophytes and the presence of self reported knee pain. Previous studies showed that hip joint space narrowing was strongly associated with other radiological features and most predictive of hip pain. Furthermore, progression of hip OA could be defined by a change in joint space narrowing, and narrowing correlated with changes in clinical status.[28 29]

Here we show that the radiographic stage of OA bears no relation to clinical symptoms in patients with established hip OA.[30] We further show that the preoperative radiographic grade of OA (moderate or severe) did not influence the clinical outcome one year after total hip replacement. Because the indication for surgery at the department where this study took place required the presence of significant radiographic OA, it was not possible to draw any conclusions about the outcome of hip arthroplasty for patients with hip symptoms and only mild radiographic changes, compared with more advanced changes. A study to answer this question would require an acceptance that patients with significant symptoms of hip OA in the presence of mild or perhaps only borderline radiographic changes consistent with hip OA provide a valid indication for hip replacement surgery. Nevertheless, based on the findings of this study, we suggest that if the diagnosis of OA is unequivocal, symptoms and not the degree of radiographic change should provide the indication for surgery. Fortin and coworkers have shown that the postoperative outcome after hip replacement for OA is better when the preoperative functional status is better.[31] The results of our study, together with those of Fortin and coworkers, therefore suggest that the practice of delaying hip replacement until severe radiographic changes, pain, and functional impairment are present should be re-evaluated.

Patients from a single hospital were included in this study, so that uniform inclusion criteria for hip replacement were used. The surgical procedure was similar, except in 40 of the cases where an uncemented acetabular component was used, because this was routine in younger patients. This group of patients who were on average 10 years younger recovered more quickly and had a significantly better outcome at 12 months after hip replacement (data not shown), contrasting with the results of a previous study.[32]

In contrast with some previous studies,[32 33] we found no sex related difference in preoperative status or in postoperative outcome. This may be due to differences in healthcare system or social structure between different countries.

The present prospective, consecutive study showed a pronounced bodily pain and impaired physical function in patients with OA of the hip before total hip replacement, as measured by the SF-36 and WOMAC. Still, their general and mental health was unaffected as compared with a reference population. Total hip replacement resulted in pain relief and improvement of the physical function to a level close to that of a reference population matched for age and sex. However, the patients needed at least one year after surgery to reach the full benefits of the intervention.

1 Ingvarsson T, Hägglund G, Jónsson H, Lohmander LS. Incidence of hip replacement in Iceland 1982–1996. Acta Orthop Scand 1999;70:229–33.
2 Crawford RW, Murray DW. Total hip replacement: indications for surgery and risk factors for failure. Ann Rheum Dis 1997;56:455–7.
3 Lieberman JR, Dorey F, Shekelle P, Schumacher L, Thomas BJ, Kilgus DJ, *et al*. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. J Bone Joint Surg Am 1996;78:835–8.
4 Wright J, Young NL. The patient-specific index: asking patients what they want. J Bone Joint Surg Am 1997;79:974–83.

5 Bellamy N, Buchanan W, Goldsmith C, 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.

6 Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). Med Care 1992;30:473–83.

7 Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al*. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMER-ACT III. J Rheumatol 1997;24:799–802.

8 Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med 1981;15A:221–9.

9 Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. The Sickness Impact Profile: validation of a health status measure. Med Care 1976;14:57–67.

10 Bellamy N. Osteoarthritis clinical trials: candidate variables and clinimetric properties. J Rheumatol 1997;24:768–78.

11 Stucki G, Liang MH, Phillips C, Katz JN. The Short Form-36 is preferable to the SIP as a generic health status measure in patients undergoing elective hip arthroplasty. Arthritis Care Research 1995;8:174–81.

12 Wagner AK, Gandek B, Aaronson NK, Acquadro C, Alonso J, Apolone G, *et al*. Cross-cultural comparisons of the content of SF-36 translations across 10 countries: results from the IQOLA project. J Clin Epidemiol 1998;51:925–32.

13 Bellamy N. Instruments to assess osteoarthritis-current status and future needs. Ann Rheum Dis 1995;54:692–3.

14 Kreibich DN, Vaz M, Bourne RB, Rorabeck CH, Kim P, Hardie R, *et al*. What is the best way of assessing outcome after total knee replacement? Clinical Orthop Relat Res 1996:221–5.

15 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.

16 Dieppe PA. Recommended methodology for assessing the progression of osteoarthritis of the hip and knee joints. Osteoarthritis Cartilage 1995;3:73–7.

17 Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch D, *et al*. Design and conduct on clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Osteoarthritis Cartilage 1996;4:217–43.

18 Altman RD, Hochberg M, Murphy WA, Wolfe FJ, Lequèsne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3:3–70.

19 Gogia PP, Christensen CM, Schmidt C. Total hip replacement in patients with osteoarthritis of the hip: improvement in pain and functional status. Orthopedics 1994;17:145–50.

20 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.

21 Rissanen P, Aro S, Slätis P, Sintonen H, Paavolainen P. Health and quality of life before and after hip or knee arthroplasty. J Arthroplasty 1995;10:169–75.

22 Wiklund I, Romanus B. A comparison of quality of life before and after arthroplasty in patients who had arthrosis of the hip joint. J Bone Joint Surg Am 1991;73:765–9.

23 Hirsch R, Fernandes RJ, Pillemer SR, Hochberg MC, Lane NE, Altman RD, *et al*. Hip osteoarthritis prevalence estimates by three radiographic scoring systems. Arthritis Rheum 1998;41:361–8.

24 Bellamy N, Tesar P, Walker D, Klestov A, Muirden K, Kuhnert P, *et al*. Perceptual variation in grading hand, hip and knee radiographs: observations based on an Australian Twin Registry study of osteoarthritis. Ann Rheum Dis 1999;58:766–9.

25 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.

26 Roos E, Klässbo M, Lohmander LS. WOMAC Osteoarthritis Index-reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Scand J Rheumatol 1999;28:1–6.

27 Lethbridge-Cejku M, Scott WW Jr, Reichle R, Ettinger WH, Zonderman A, Costa P, *et al*. Association of radiographic features of osteoarthritis of the knee with knee pain: data from the Baltimore Longitudinal Study of Aging. Arthritis Care Research 1995;8:182–8.

28 Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132:514–22.

29 Dougados M, Gueguen A, Nguyen M, Berdah L, Lequèsne M, Mazieres B, *et al*. Radiological progression of hip osteoarthritis definition, risk factors and correlations with clinical status. Ann Rheum Dis 1996;55:356–62.

30 Lawrence JS, Bremner JM, Bier F. Osteo-arthrosis. Prevalence in the population and relationship between symptoms and x-ray changes. Ann Rheum Dis 1966;25:1–24.

31 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.

32 Fox KM, Hochberg M, Resnik C, Kenzora JE, Hebel R, Zimmerman SI, *et al*. Severity of radiographic findings in hip osteoarthritis associated with total hip arthroplasty. J Rheumatol 1996;23:693–7.

33 Katz JN, Wright EA, Guadagnoli E, Liang MH, Karlsson EW, Cleary PD. Differences between men and women undergoing major orthopedic surgery for degenerative arthritis. Arthritis Rheum 1994;37:687–94.