

# Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis

## A prospective cohort study of 573 patients

Claire TILBURY<sup>1</sup>, Maarten J HOLTSLAG<sup>1</sup>, Rutger L TORDOIR<sup>2</sup>, Claudia S LEICHTENBERG<sup>1</sup>, Suzan H M VERDEGAAL<sup>2</sup>, Herman M KROON<sup>3</sup>, Marta FIOCCO<sup>4</sup>, Rob G H H NELISSEN<sup>1</sup>, and Thea P M VLIET VLIELAND<sup>1</sup>

<sup>1</sup> Department of Orthopaedics, Leiden University Medical Center, Leiden; <sup>2</sup> Department of Orthopaedics, Rijnland Hospital, Leiderdorp; <sup>3</sup> Department of Radiology, Leiden University Medical Center, Leiden; <sup>4</sup> Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands.

Correspondence: c.tilbury@lumc.nl

Submitted 2014-11-16. Accepted 2015-07-26.

**Background and purpose** — There is no consensus on the impact of radiographic severity of hip and knee osteoarthritis (OA) on the clinical outcome of total hip arthroplasty (THA) and total knee arthroplasty (TKA). We assessed whether preoperative radiographic severity of OA is related to improvements in functioning, pain, and health-related quality of life (HRQoL) 1 year after THA or TKA.

**Patients and methods** — This prospective cohort study included 302 THA patients and 271 TKA patients with hip or knee OA. In the THA patients, preoperatively 26% had mild OA and 74% had severe OA; in the TKA patients, preoperatively 27% had mild OA and 73% had severe OA. Radiographic severity was determined according to the Kellgren and Lawrence (KL) classification. Clinical assessments preoperatively and 1 year postoperatively included: sociodemographic characteristics and patient-reported outcomes (PROMs): Oxford hip/knee score, hip/knee injury and osteoarthritis outcome score (HOOS/KOOS), SF36, and EQ5D. Change scores of PROMs were compared with mild OA (KL 0–2) and severe OA (KL 3–4) using a multivariate linear regression model.

**Results** — Adjusted for sex, age, preoperative scores, BMI, and Charnley score, radiographic severity of OA in THA was associated with improvement in HOOS “Activities of daily living”, “Pain”, and “Symptoms”, and SF36 physical component summary (“PCS”) scale. In TKA, we found no such associations.

**Interpretation** — The decrease in pain and improvement in function in THA patients, but not in TKA patients, was positively associated with the preoperative radiographic severity of OA.

The possible association between outcomes after THA and TKA, and preoperative radiographic severity has been addressed repeatedly. Nilsdotter et al. (2001) found that in patients undergoing THA, the preoperative radiographic stage of osteoarthritis (OA) was not related to the postoperative outcome after 1 year. Cushnaghan et al. (2009) reported that in TKA patients, improvement in physical function as measured with the Short Form (SF)-36 mean 7 years after surgery tended to be greater in patients with a higher Kellgren and Lawrence (KL) grade at baseline. Valdes et al. (2012) found that a low radiographic grade of the index joint was associated with an increased risk of postoperative pain 3 years after THA and TKA. Dowsey et al. (2012) reported that lower severity of radiographic OA was associated with poorer function and more pain after TKA. Keurentjes et al. (2013) found that both THA patients and TKA patients with severe radiographic OA had a greater improvement in the SF-36 domain “Physical functioning” than patients with mild radiographic osteoarthritis after 2–5 years of follow-up.

Although the results of most of the studies show some similarities, the designs of the studies were heterogeneous, which makes it difficult to compare them. To overcome these limitations, we determined whether the preoperative radiographic severity of OA is related to improvement in functioning, pain, and HRQoL 1 year after THA or TKA in a prospective, well-defined cohort of patients, using multivariate analysis to account for possible confounding.

## Patients and methods

### Patients and recruitment

This prospective cohort study was performed at the Department of Orthopaedics, Rijnland Hospital, the Netherlands from October 2010 through September 2012.

We aimed to include all consecutive patients undergoing a primary THA or TKA because of OA, aged 18 years or older, who were able to read and understand Dutch and were mentally and physically capable of completing questionnaires. Patients with revision of a THA or TKA, those undergoing a hemiarthroplasty, and those undergoing a THA or TKA because of tumor or rheumatoid arthritis were excluded.

1 day preoperatively, before they were admitted to hospital, all eligible patients were given oral and written information concerning the study from the treating orthopaedic surgeon. Each patient was asked to return the set of questionnaires and the informed consent form when he or she was admitted the next day, the day of surgery. For the patients who did not want to participate, only age, sex, and the type of operation (THA or TKA) were recorded.

In cases where a patient who was already included in the study underwent another joint replacement during the study period, he or she was not included for a second time.

745 patients were admitted for THA and 614 patients were admitted for TKA from October 2010 through September 2012. Primary THA for primary OA was performed in 665 patients. Primary TKA for primary OA was performed in 599 patients. These 1,264 patients met all of the selection criteria and were asked to complete a questionnaire 1 day preoperatively. Of these, 302 THA and 271 TKA patients were included in the present study (Figure 1).

### Sociodemographic and clinical characteristics

Sociodemographic characteristics (only preoperatively) included: age; sex; length (cm) and weight (kg), to calculate the body mass index; level of education (low: primary school, lower vocational education; medium: lower general secondary school, intermediate vocational education; or high: higher general secondary school, higher vocational education, university); and marital status (living alone—yes/no). In addition, it was asked whether patients had a paid job (yes/no). If not, they were asked to indicate if they were: a pensioner, a housewife/houseman, or unemployed. For comorbidity, the self-reported Charnley classification (A–C) was used. Due to an error in the preoperative knee questionnaires, we were not able to determine the Charnley classification in the TKA group.

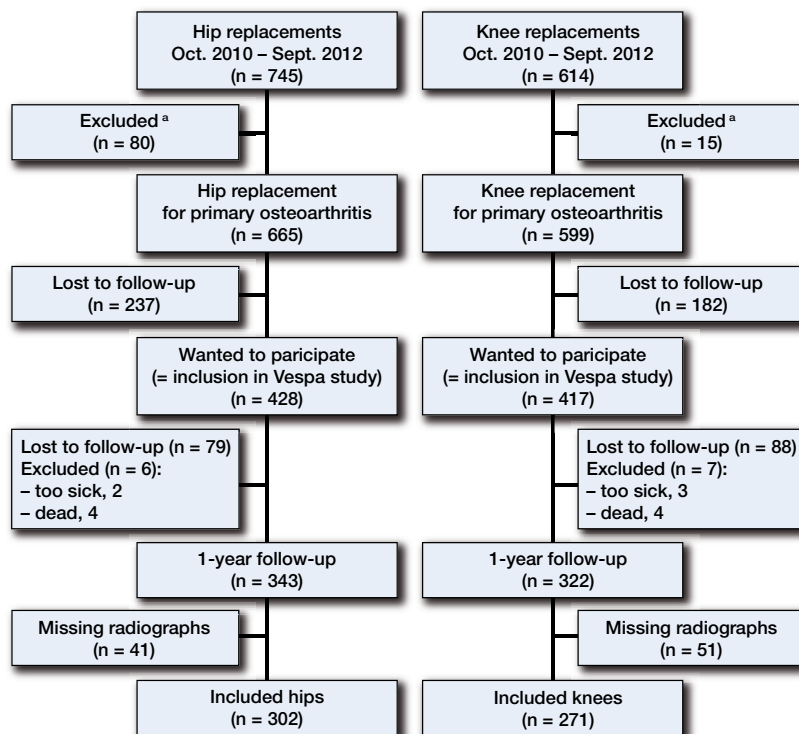


Figure 1. Flow diagram. <sup>a</sup> Reasons for exclusion were patients who did not understand Dutch or were physically or mentally unable to complete questionnaires, patients with revision surgery, undergoing a hemiarthroplasty, or undergoing THA or TKA because of a tumor or rheumatoid arthritis.

Patient-reported outcomes (PROMs) were used to describe the clinical characteristics at baseline. The hip disability and osteoarthritis outcome score (HOOS), the knee injury and osteoarthritis outcome score (KOOS), the Oxford hip score (OHS) and the Oxford knee score (OKS) were used for the preoperative and postoperative assessment of limitations (daily living, sport and recreation, function, and health-related quality of life). We used the validated Dutch versions of the HOOS, KOOS, OHS, and OKS. (Dawson et al. 1996, Dawson et al. 1998, Gosens et al. 2005, Haverkamp et al. 2005).

The Short Form-36 survey (SF-36) (Ware and Sherbourne 1992, Aaronson et al. 1998), the EuroQol 5 dimensional (EQ-5D), and the EuroQol visual analog scale (EQ-VAS) were used to assess general health-related quality of life. From the SF-36, summary component scores for physical health (PCS) and mental health (MCS) were calculated. In this study, scores from a general Dutch population were used to standardize our scores in order to apply the norm-based scoring (Fransen and Edmonds 1999).

### Preoperative radiographic severity

Preoperative supine radiographs of hips (anterior-posterior) and weight-bearing radiographs of the knees (posterior-anterior) had been taken routinely in the participating centers for preoperative templating purposes. All radiographs

were assessed by an experienced musculoskeletal radiologist (HMK), who was blinded regarding patient characteristics. The Kellgren and Lawrence (KL) grading system was used to classify the severity of OA (grade 0: indicating no OA; grade 1: doubtful OA; grade 2: minimal OA; grade 3: moderate OA; and grade 4: severe OA) (Kellgren and Lawrence 1957). 10% of the radiographs were scored twice: correlation between both readings was used to establish the intra-reader reliability (intra-class correlation, hip radiographs: 99% (95% CI: 85–93); intra-class correlation, knee radiographs: 95% (95% CI: 92–98)). The second reading was used for further statistical analyses. The KL grade in our study was classified as mild for KL 0–2 and severe for KL 3–4.

### Statistics

Comparisons between patient characteristics preoperatively that were included in the analysis and those that were excluded due to incomplete follow-up were performed with chi-squared tests (for categorical variables), or independent-samples t-tests or Mann-Whitney U-tests (for continuous variables). The choice between unpaired t-tests and Mann-Whitney U-tests was based on the distribution of the variables. For this purpose, we made normality plots of all continuous variables using the Kolmogorov-Smirnov test.

Comparisons of the preoperative characteristics and the change scores over time between the groups of patients with mild and severe OA were first done with ANOVA. Then multivariable linear regression analyses were performed, with the KL grade (mild/severe) as independent variable, for the following outcomes: the mean change in the HOOS/KOOS, OHS/OKS, SF subscales, SF summary scales, EQ score, and EQ-VAS scale. All analyses were adjusted for potential confounding factors (derived from the literature on determinants of outcome in THA and TKA and/or the above-mentioned crude analyses ( $p > 0.10$ , ANOVA)). Potential confounding factors considered were the KL grade (mild/severe); age (divided into the groups 0–65 years, 66–75 years, and  $\geq 76$  years); sex; BMI (categorized as 0–25, 26–30, and  $\geq 31$ ); and—only for the THA group—the Charnley classification (A, B, or C).

Outcome variables (dependent) were the mean change scores of the HOOS/KOOS, OHS/OKS, SF subscales, SF summary scales, EQ score, and EQ-VAS scale. Data were analyzed using the SPSS statistical package version 20.0. All statistical testing was performed with 2-tailed tests and the level of statistical significance was set at  $p \leq 0.05$  for all analyses.

### Ethics

The study protocol was reviewed and approved by the local hospital review board (Rijnland Hospital, Leiderdorp; registration number 10/07), which is affiliated to the Medical Research Ethics Committee of Leiden University Medical Center, Leiden, the Netherlands. Written informed consent to participate in the study was obtained from all patients.

## Results

### Preoperative patient characteristics

There was no statistically significant difference between the preoperative characteristics of the patients who were included in the study and those of the 41 THA and 51 TKA patients who were not included because of missing radiographs (results not shown).

### Preoperative characteristics of patients with mild or severe radiographic OA

Overall, there were no statistically significant differences between patients with mild or severe radiographic OA with respect to sex, age, BMI, preoperative HOOS/KOOS, OHS/OKS, EQ score, EQ-VAS score, or SF (all subscales), for both THA and TKA. The exceptions were a significantly higher preoperative KOOS “Sports” score and SF “MCS” score in the KL 3–4 group than in the KL 0–2 group, in both THA patients and TKA patients (Table 1, see Supplementary data).

### Crude and adjusted changes in health-related quality of life and functional outcome measurements in patients with mild or severe radiographic OA

The crude (unadjusted) mean change scores (postoperative scores minus preoperative score) in patients with mild or severe radiographic OA are shown in Tables 2a and 2b. In THA patients, the mean change scores were statistically significantly higher in patients with KL 3–4 than in patients with KL 0–2 with respect to HOOS “Symptoms”, EQ score, SF “Physical functioning” and “Bodily pain”; and SF “PCS”. In TKA patients, there were no statistically significant differences in change scores between patients with KL 0–2 and patients with 3–4 for all of the PROMs (Table 2, see Supplementary data).

Adjusting for sex, age, preoperative scores of PROMs, and BMI (and Charnley score in THA), in the THA group the severity of radiographic OA was related to 3 of 5 HOOS subscale scores (“ADL” ( $p = 0.002$ ), “Pain” ( $p = 0.004$ ), and “Symptoms” ( $p = 0.004$ )), the SF subscale score “Bodily pain” ( $p = 0.004$ ) and the SF “PCS” ( $p = 0.01$ ), but not to the HOOS “Sports” and “Quality of life” subscale scores, the EQ and EQ-VAS scales, the SF “MCS” and all the other SF subscales. In the TKA group, there was no association between radiographic severity and improvement in any of the PROMs (Table 2, see Supplementary data).

## Discussion

This prospective study in patients undergoing THA and TKA showed that changes in scores over time were greater in patients with more severe radiographic OA. The difference was statistically significant for a number of clinical outcomes in THA patients, but not in TKA patients.

Overall, our results are in line with the literature, with the majority of studies concluding that more severe radiographic OA preoperatively is associated with better outcomes in THA or TKA (Dowsey et al. 2012, Valdes et al. 2012, Keurentjes et al. 2013). Concerning THA specifically, similar to the present study, Valdes et al. (2012) reported greater improvements in pain 3 years after surgery in patients with severe radiographic OA preoperatively. Greater improvements in the SF subscale and summary scale scores were seen in patients with higher KL scores in a study by Keurentjes et al. (2013), but the differences were not statistically significant.

Regarding TKA, our study did not show any statistically significant differences between the outcomes in patients with different grades of radiographic severity, although—as in the study by Cushnaghan et al. (2009)—greater improvements were generally seen in patients with higher KL grades. In contrast, Valdes et al. (2012) and Keurentjes et al. (2013) found statistically significantly better outcomes in TKA patients with severe radiographic OA, and similar results were seen in some of the analyses in the study by Dowsey et al. (2012). Comparisons with the literature are, however, hampered by the large diversity in study designs and analyses.

It is difficult to draw conclusions about the clinical relevance of the results of our study and of previous ones. Firstly, there are several factors associated with worse outcomes after THA/TKA, such as older age, female sex, obesity, worse general health, involvement of other joints, and a lower level of education (Dieppe et al. 2009, Gossec et al. 2011). Only from large, prospective studies using a standardized set of preoperative characteristics and outcome assessments done at fixed time points can true prediction models including all potentially relevant determinants be derived, which afterwards need to be validated in multiple settings and countries. However, we can interpret the absolute change scores as observed in the different groups according to radiographic severity. A recent systematic review by Keurentjes et al. (2012) found that overall minimally clinically important differences (MICDs) in HRQoL in THA/TKA have limited precision and are not validated using external criteria. The study which is most comparable to our study is that from Clement et al. (2014). In that study, the MCID in OKS for the difference between preoperatively and 1 year postoperatively was 15.5 (95% CI: 14.7–16.4). In our study, generally patients in both the mild and severe OA groups achieved this improvement, indicating that the clinical relevance of a statistically significant difference may be limited.

A main strength of our study was the inclusion of a wide range of validated PROMs, covering all items of disease-specific outcome measures in functioning, pain, and health-related quality of life. Using all these outcome measures, both measures of pain and daily activities, we observed differences between groups according to radiographic severity. Another strength was that all radiographs were read by a single observer with extensive experience, who was blinded regard-

ing patient data. In addition, this was a prospective study with a relatively large cohort with only 20% loss to follow-up in the THA group and only 23% loss to follow-up in the TKA group.

Our study also had a number of limitations. It only included KL grading applied to the anteroposterior and posteranterior radiographs from the preoperative hip and knee.

In the study by Dowsey et al. (2012), not only KL grading but also the severity of joint space narrowing (JSN; 0–3) and osteophyte formation (0–3) using the Osteoarthritis Research Society International (OARSI) atlas, and the degree of bone attrition, were taken into account. In that study, radiographs showing advanced OA (KL 3–4) were further subdivided by including data from the individual score of JSN and bone attrition.

In addition, the patients included in the present study were a selection of all patients who underwent THA or TKA and it was carried out in 1 center in 1 country. However, the preoperative characteristics of the patients and their change scores over time are well in line with those observed in other large cohorts (Nilsson et al. 2003, Dieppe et al. 2009, Beswick et al. 2012).

In conclusion, this study shows that in patients who underwent THA, but not TKA, more severe radiographic OA preoperatively was associated with a better outcome regarding pain and function.

### Supplementary data

Tables 1 and 2 are available on the Acta Orthopaedica website, [www.actaorthopaedica.org](http://www.actaorthopaedica.org), identification number 8277.

CT, MF, and TPMVV: conception and design, analysis and interpretation of the data, drafting of the article, provision of study materials or patients, statistical expertise, and collection and assembly of data. MJH, RLT, HMK, CSL, and SHM: provision of study materials or patients, administrative, technical, or logistic support, and collection and assembly of data. HMK and RGHHN: critical revision of the article, statistical expertise.

This study was supported by the Dutch Arthritis Association (grant number LLP13).

No competing interests declared.

Aaronson N K, Muller M, Cohen P D, Essink-Bot M L, Fekkes M, Sanderman R, Sprangers M A, te V A, Verrrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 1998; 51(11): 1055–68.

Beswick A D, Wylde V, Goberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open* 2012; 2(1): e000435.

Clement N D, MacDonald D, Simpson A H. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2014; 22(8): 1933–9.

- Cushnaghan J, Bennett J, Reading I, Croft P, Byng P, Cox K, Dieppe P, Coggon D, Cooper C. Long-term outcome following total knee arthroplasty: a controlled longitudinal study. *Ann Rheum Dis* 2009; 68(5): 642-7.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br* 1996; 78(2): 185-90.
- Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br* 1998; 80(1): 63-9.
- Dieppe P, Judge A, Williams S, Ikwueke I, Guenther K P, Floeren M, Huber J, Ingvarsson T, Learmonth I, Lohmander L S, Nilsson A, Puhl W, Rowley D, Thielier R, Dreinhoefer K. Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres. *BMC Musculoskelet Disord* 2009; 10: 19.
- Dowsey M M, Dieppe P, Lohmander S, Castle D, Liew D, Choong P F. The association between radiographic severity and pre-operative function in patients undergoing primary knee replacement for osteoarthritis. *Knee* 2012; 19(6): 860-5.
- Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. *Rheumatology (Oxford)* 1999; 38(9): 807-13.
- Gosens T, Hoefnagels N H, de Vet R C, Dhert W J, van Langelaan E J, Bulstra S K, Geesink R G. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. *Acta Orthop* 2005; 76(2): 204-11.
- Gossec L, Paternotte S, Maillefert J F, Combescuré C, Conaghan P G, Davis A M, Gunther K P, Hawker G, Hochberg M, Katz J N, Kloppenburg M, Lim K, Lohmander L S, Mahomed N N, March L, Pavelka K, Punzi L, Roos E M, Sanchez-Riera L, Singh J A, Suarez-Almazor M E, Dougados M. The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. *Osteoarthritis Cartilage* 2011; 19(2): 147-54.
- Haverkamp D, Breugem S J, Sierevelt I N, Blankevoort L, van Dijk C N. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. *Acta Orthop* 2005; 76(3): 347-52.
- Kellgren J H, Lawrence J S. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957; 16(4): 494-502.
- Keurentjes J C, Van Tol F R, Fiocco M, Schoones J W, Nelissen R G. Minimal clinically important differences in health-related quality of life after total hip or knee replacement: A systematic review. *Bone Joint Res* 2012; 1(5): 71-7.
- Keurentjes J C, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert A W, Poll R G, Kroon H M, Vliet Vlieland T P, Nelissen R G. Patients with severe radiographic osteoarthritis have a better prognosis in physical functioning after hip and knee replacement: a cohort-study. *PLoS One* 2013; 8(4): e59500.
- Nilsson A K, Aurell Y, Siosteen A K, Lohmander L S, Roos H P. Radiographic stage of osteoarthritis or sex of the patient does not predict one year outcome after total hip arthroplasty. *Ann Rheum Dis* 2001; 60(3): 228-32.
- Nilsson A K, Petersson I F, Roos E M, Lohmander L S. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. *Ann Rheum Dis* 2003; 62(10): 923-30.
- Valdes A M, Doherty S A, Zhang W, Muir K R, Maciewicz R A, Doherty M. Inverse relationship between preoperative radiographic severity and post-operative pain in patients with osteoarthritis who have undergone total joint arthroplasty. *Semin Arthritis Rheum* 2012; 41(4): 568-75.
- Ware J E, Jr., Sherbourne C D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6): 473-83.