Original Article

Arthritis Care & Research
DOI 10.1002/acr.22674

The longitudinal course of physical function in people with symptomatic knee osteoarthritis: Data from the MOST study and the OAI

Britt Elin Øiestad, PT, PhD[1], Daniel K. White, PT, ScD, MSc[2,3], Ross Booton, MMath[3], Jingbo Niu, DSc[3], Yuqing Zhang, DSc[3], Jim Torner; PhD[4], Beth Lewis, MD, PhD[5], Michael Nevitt, PhD[6], Michael LaValley, PhD[3,7], David T. Felson, MD, MPH[3,8]

[1]Department of Physical Therapy, Oslo and Akershus University College of Applied Sciences, Oslo, Norway

[2]Department of Physical Therapy, University of Delaware, Newark, DE, US

[3]Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, Boston, MA, US

[4] Department of Epidemiology, University of Iowa, Iowa, US

[5]Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, AL, US

[6]University of California, San Francisco, San Francisco, CA, US

[7]Boston University School of Public Health, Boston, US

[8]Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester, UK

Corresponding Author:

Britt Elin Øiestad

Department of Physical Therapy

Oslo and Akershus University College of Applied Sciences

Postboks 4, St. Olavs plass, 0130 Oslo, Norway

Phone: +4792803089

Britt-Elin.Oiestad@hioa.no

Manuscript Type: Original Research

Word count: 3614 (Background-Discussion)

Number of pages: 18

Number of tables: 3

Number of Figures: 3

Funding: The Research Council of Norway and NIH AR47785.The MOST study was supported by NIH grants U01-AG18820, U01-AG18832, U01-AG18947, and U01-AG19069.

Received: Mar 02, 2015; Revised: Jun 27, 2015; Accepted: Jul 21, 2015

ABSTRACT

Background: Pain and functional decline are hallmarks of knee osteoarthritis (OA). Despite this, longitudinal studies unexpectedly reveal stable or improved physical function. The aim of this study was to impute missing and pre-total knee replacement (TKR) values to describe physical function over time among people with symptomatic knee OA.

Methods: We included participants from the Multicenter Osteoarthritis Study (MOST) and the Osteoarthritis Initiative (OAI) with incident symptomatic knee OA observed during the first 30 months in MOST and 36 months in OAI. WOMAC physical function (pf), the Five Times Sit To Stand Test, and the 20-Meter Walk Test were assessed over 4,5 and 6 years, respectively. We used a multiple imputation method for missing visits, and estimated pre-TKR values close to the time of TKR using a fitted local regression smoothing curve. In mixed effect models we investigated the physical function change over time using data before and after imputation and prediction of pre-TKR values.

Results: In MOST, 225 (8%) had incident knee OA, with corresponding 577 (12.7%) in OAI. After adjusting for pre-TKR values and imputing missing values, we found that WOMAC-pf values remained stable or slightly declined over time, and the 20-Meter Walk test results changed from stable in nonimputed analyses to worsening using imputed data.

Conclusion: Data from the MOST Study and the OAI showed stable to worsening physical function over time in people with incident symptomatic knee OA after imputing missing values and adjusting pre-TKR values.

Significance and Innovations

- Longitudinal studies of knee OA unexpectedly reveal stable or improving physical function over time.
- Study participants with missing visits and total knee replacements (TKR) are often left out of analyses in longitudinal studies.
- Imputing missing values and predicting pre-TKR function up to as close as possible to the time of TKR are newer methods to analyze physical function in all persons with symptomatic knee OA.
- By imputing missing values and adding predicted pre-TKR values as close as possible to the TKR, trajectories of physical function changed from improvement to stable or worsened values over time.

BACKGROUND

Pain and functional limitations are important clinical manifestations of symptomatic knee osteoarthritis (OA) [1;2]. Since knee OA is a chronic and progressive disease, we often see in the clinics a worsening of physical function over time, and a rising number of people end up with a total knee replacement (TKR)[3]. Contrarily, longitudinal studies have found that physical function, on average, is stable, and sometimes improves in people at risk or with knee OA [4-8]. Currently, we do not know whether the average physical function values observed in longitudinal studies are true or a reflection of bias.

In this regard, there are at least two challenges that need to be addressed: First, people in such studies are often included because they are in a painful phase of their disease. The natural history of OA involves fluctuation of symptoms and these people may well have lower pain values on re-assessment in the absence of an intervention, introducing regression to the mean[9]. In addition, people entering a study in a painful phase of their disease have been shown to have more missing visits in longitudinal cohort studies[10]. The second issue with published longitudinal functional values in people with knee OA is that people lost to follow-up have been shown to be older and have poorer function than those without missing visits[11;12]. Since including subjects with complete follow-up data only may leave a study sample of people with the best physical function in the studies[11], imputation techniques to address the missing data have been introduced. However, assumptions for using multiple imputation (MI) may be hard to test and fulfill in longitudinal studies on people at risk or with knee OA. Missing data may be missing completely at random (MCAR) [13], indicating that subjects with knee OA who do not attend the study visit have reasons for not attending that are unrelated to their physical function or knee pain. However, longitudinal studies of knee OA have found that people

not attending the follow-up visits were often older, had lower educational level, longer duration of symptoms, and they had lower muscle strength[4;11]. In such cases, the missing data is suggested to be missing at random (MAR)[14] where the attendance can be predicted by subject factors other than the unobserved values for physical function and pain at the missing visit. Challenges appear if the data are missing not at random (MNAR), which means that the visit attendance does depend on the participant's unobserved physical function or pain status.

Additionally excluding those that undergo TKR as seen in previous studies[15;16], may further bias the study sample. Those undergoing TKR may have the worst physical function at baseline and decline more over time than those without TKR. Thus, excluding this group will result in describing a course of physical function among subjects who are doing relatively well over time. Including those undergoing TKR will better represent the whole population of symptomatic knee OA, but the longitudinal pre-surgical data often do not include functional status as close as possible to the time of the TKR.

Newer studies e.g. from the CHECK cohort, have addressed the issue of loss to follow-up by using MI techniques, and included pre-TKR values in the analyses[11;17], however the pre-TKR values may have been measured a relatively long time before the TKR. We know little about the longitudinal course of physical function including approaches to account for missing data and those with the poorest function that end up with TKR using data from larger longitudinal cohorts.

The aim of this study was to estimate the longitudinal course of physical function in people with incident symptomatic knee OA, accounting for missing visits and pre-TKR values. We imputed missing data and included predicted pre-TKR physical function values for those with TKR as close as possible to the time of surgery. We hypothesized that physical function would worsen significantly over time in people with symptomatic knee OA after imputing missing data and including predicted pre-TKR values of physical function (adjusted data).

MATERIAL AND METHODS

We used data from the Multicenter Osteoarthritis (MOST) Study and the Osteoarthritis Initiative (OAI). Both cohorts included community-dwelling participants at risk of developing or with established knee OA. In MOST, study participants were enrolled if they were overweight, obese, had knee symptoms, or had a history of knee injury or surgery. In OAI, additional inclusion criteria included a family history of TKR, the presence of Heberden's nodes, or a history of repetitive knee bending. People with rheumatoid arthritis, inflammatory arthritis, and end-stage disease at baseline were excluded from both studies [2].

The MOST study included 3026 people between 50-79 years from Birmingham, Alabama and Iowa City, Iowa, and it has been ongoing for 84 months, including 4 visits (baseline, 30, 60, and 84 months) [18]. OAI included 4796 people between 45-79 years at four clinical sites: Baltimore, Maryland; Pittsburgh, Pennsylvania; Pawtucket, Rhode Island, and Columbus, Ohio. Follow-up visits were conducted annually for 7 years

(www.oai.ucsf.edu) [19]. The study participants provided informed consents before first clinic visit in both cohorts.

We excluded people who had TKR at baseline (n=78 in MOST and n=63 in OAI), and people who underwent hip replacement at any time point (n=140 in MOST and n=190 in OAI), because physical function in these groups was probably also influenced by disease other than knee OA. After the exclusions, the MOST study comprised 2808 study participants, with the corresponding number of 4543 in the OAI. Of these, we included subjects with incident symptomatic knee OA. We defined subjects as having incident symptomatic knee OA if they had no symptomatic OA in both knees at baseline, but developed new disease in either knee during 30-month follow-up in the MOST Study or during 36-month follow-up period in the OAI Study.

Radiographic examination

We used data from standing posterior-anterior radiographs using a SynaFlexer frame for standardized positioning (Synarc, San Francisco, California, USA). The radiographs were read and scored by a musculoskeletal radiologist and a rheumatologist at Boston University blinded to clinical data. An adjudication panel (the two readers and DTF) resolved discrepancies between the readers. For this study, we included people with incident symptomatic knee OA defined on the basis of radiographic examination of the tibiofemoral joint (Kellgren and Lawrence grade ≥2[20]), and having knee pain in that knee on most days of the last month. We defined the subject based on their worse knee in terms of symptomatic knee OA.

Outcome measurements

The follow-up of the included subjects started from the visit when incident symptomatic knee OA was observed. Self-reported physical function was measured at all the visits in the MOST Study (baseline, 30, 60, and 84 months) and the OAI (annually from 0-84 months) using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Physical Function subscale [21]. We used the five-point Likert scale version from 0 indicating no problems to 4 indicating extreme functional problems for the subscale Physical Function (WOMAC-pf)(0-68).

Performance-based measures included the Five Times Sit To Stand Test [22] and the 20-Meter Walk Test [23]. These were measured at all visits in the MOST Study and annually for 84 months in the OAI except for the 60- and 84-month visits. For the Five Times Sit To Stand Test, the participants stood from a chair five times as quickly as they could, keeping their arms folded across their chest. The participants were told to come to a full standing position each time, and when they sat down, they were told to sit all the way down each time. We recorded total time in seconds using a stopwatch from start to finish of the test. In the 20-Meter Walk Test, the study participants were told to walk at their usual walking pace from the starting point to the end. We measured total time in seconds. For both the performance-based tests, we used the mean time of two trials.

Statistical analysis

The *crude* outcome data included participants with complete visits only, including the TKR group with original pre-TKR values. We used two methods to generate the *adjusted* outcome data: First, as many of the subjects who underwent TKR in MOST and OAI had

their last clinic visit many months before TKR, we wanted to add predicted pre-TKR outcome values (WOMAC-pf, Five Times Sit To Stand, and 20-Meter Walk) as close as possible to the time of TKR. Using data from those who underwent TKR, we constructed local regression smoothing (LOESS) curves [24] to estimate fitted curves for the outcome variables at the last visit before TKR and months from last visit to TKR (e.g in Figure 1 using MOST WOMAC-pf data). We then used the fitted curves to predict a person's WOMAC-pf value at the time of the TKR. Removing outliers did not alter these curves substantively.

We used the LOESS curves to predict the average outcome values just before the TKR and then adjusted this value for each subject by adding in the difference between the subject's measured value and the height of the LOESS curve at the time of measurement. For the WOMAC-pf in MOST this average value just before TKR was: 32.1 (Figure 1). So, for each subject the new predicted WOMAC-pf value was: predicted WOMAC-pf = 32.1 +/- the subject's deviation from the curve. For the Five Times Sit To Stand Test, the equation was: predicted Five Times Sit To Stand = 13.4 +/- deviation from the fitted curve, and for 20-Meter Walk Test: predicted 20-Meter Walk = 20.7+/- deviation from the fitted curve. In the OAI, the corresponding equations were: WOMAC-pf = 28.6 +/- deviation from the fitted curve, Five Times Sit To Stand = 13.4 +/- deviation from the fitted curve, and 20-Meter Walk = 18.4 +/- deviation from the fitted curve. We assigned this new predicted outcome values to the data as an additional pre-TKR visit.

Second, values for missing visits for the three outcome variables during the follow-up period were filled in using MI method [25] with the independent variables age, sex, and body mass index (BMI). The missing pattern for WOMAC-pf in the MOST Study

considering the visits prior to TKR or death showed that 79% had no missing visits, and 21% had some missing visits (including 1 participant (0.4%) with all visits missing). The numbers were similar for the Five Times Sit To Stand test (76% had no missing visits), and the 20-Meter Walk test (82% had no missing visits). Corresponding numbers for no missing visits in the OAI were 83%, 68%, and 79%, with none having all WOMAC-pf values missing. Among subjects who underwent TKR, 2% had some missing visits in MOST, and none had all missing visit for WOMAC-pf and 20-Meter Walk, but 4 had missing visit on Five Times Sit To Stand. In OAI, 13 (18.6% of the TKR group) had some missing visits. Only 1% of subjects undergoing TKR had all missing pre-TKR WOMAC-pf values.

In MOST, we found no statistical significant differences in age, gender, BMI or functional outcomes at the incident visit between individuals with missing WOMAC-pf data prior to TKR or death (n=48) and those with complete data. (n=177)". In OAI, those with missing functional values at any visit prior to TKR or death had poorer WOMAC-pf and performance-based functional values.

As no differences were found between those with and without missing values in MOST, it is reasonable to believe that the pattern was not MNAR. We may however, speculate that the baseline function is highly correlated with the follow-up function, and this suggests that the missing pattern in the MOST Study is at least MAR. In OAI, those with missing functional data had poorer function at any visit prior to TKR or death, but we were not able to test the assumption of MAR.

When WOMAC-pf was missing, other elements of WOMAC were missing also, preventing us from using other similar measures for imputing missing data. Therefore, we generated 5 imputed datasets including the three outcome variables in in SAS

conditioning on age, gender and BMI, and the three outcome variables. The effect estimates for the 5 data sets were pooled together across imputed datasets along with adjusted variance and taking into account the uncertainty introduced by the imputation as previously described by our group[26]. The pooled results are shown in the *adjusted* results in the Tables and Figures. Post-TKR values were set to missing.

We used mixed models to investigate the effect of time on physical function outcomes separately for crude and adjusted data in those with incident symptomatic knee OA.

RESULTS

In MOST, 225 (8%) participants had incident symptomatic knee OA during the first 30 months of follow-up. The corresponding number in OAI was 577 (12.7%) during the first 36 months of follow-up. Of these, 43 subjects (19 %) in MOST and 70 subjects (12.1 %) in OAI had TKR after their diagnosis of knee OA. Characteristics of study participants are shown in Table 1. Of those with incident symptomatic OA, 66.7% in MOST were women, and 62.1% in OAI. The mean age in MOST at 30-months was 66.5 (8.2) years, with a mean BMI of 30.9 (5.7), and 63.6 (8.6) years with a BMI of 29.7(4.8) at 12 month in the OAI. At the visit were the MOST participants had observed symptomatic knee OA, 90 (46.2%) had KL grade 2, 88 (39.1%) had KL grade 3, and 33 (14.7%) had KL grade 4. In OAI, 293 (50.8%) had KL grade 2, 208 (36%) had KL grade 3, and 76 (13.2%) had KL grade 4.

In Figure 2 and 3, we show the crude and adjusted curves for physical function over time in the MOST Study and the OAI respectively for the TKR group only, and for crude and

adjusted values for all those with incident symptomatic knee OA. The adjusted values for the TKR group show substantial worsening of knee function.

Physical function in subjects with knee OA

In the MOST Study, in crude data, we found overall improvement in WOMAC-pf over time (Table 2). In the adjusted data, no significant change over time was found. Figure 2 shows poor WOMAC-pf results over time for the pre-TKR group, and it shows slightly better function for the crude curve than for the adjusted curve for all MOST study participants. For the Five Times Sit To Stand and the 20-Meter Walk Tests, significant worsening was found over time for the crude (only Five Times Sit To Stand) and adjusted data (Table 2). After adjustment, the Five Times Sit To Stand results and the 20-Meter Walk results increased by 1.5 seconds from the first to the last visit ($p<0.003$).

In the OAI, worsening within adjusted WOMAC-pf values was seen over time (Table 3). Also, we found significantly worsening values for the crude and adjusted 20-Meter Walk data in OAI.

DISCUSSION

Among people with incident knee OA , we found physical function generally worsened over time after imputing for missing data and predicting pre-TKR physical function values of subjects as close as possible to the time of TKR. We found a significantly worse performance-based physical function over time in the MOST study. In OAI, we found a

decline in self-reported physical function after incorporating predicted pre-TKR values and data for missing visits. Overall, subjects with incident knee OA in the MOST Study and the OAI reported stable or slightly lower physical function values from the time of diagnosis to the last clinic visit, in line with others [27]. In crude results in which we did not impute missing values or pre-TKR physical function status, the trajectory of physical function was more favorable.

Other longitudinal studies from the Netherlands[5;11], including the CHECK cohort[4], and the Health and Retirement Study [6;28] showed improvement or stable physical function over time. Pisters et al.[11] reported 5-year data on a cohort of 288 knee OA patients from rehabilitation centers and hospitals. They imputed values for loss to follow up for both self-reported (WOMAC-pf), and performance based (10-Meter Walk) outcomes. They found significantly improved WOMAC-pf result, and stable 10-Meter Walk results. The improvement in WOMAC-pf was 2.7 points (9%), which the authors claimed was not clinically relevant according to other literature [29;30]. In contrast to our study, to our knowledge, none had TKR during the 5 years of follow-up and those that died were excluded. In addition, the study participants had either knee or hip OA. This may explain the different results between the studies.

Crude data from MOST and OAI showed improvement in WOMAC-pf over time, but much of the improvement may be considered to be based on regression to the mean between the first and second visit. After imputing missing values and predicting pre-TKR values, although the initial regression to the mean persisted (see figures), a later decline in function was apparent and some of this phenomenon in the MOST data disappeared (see Table 2). This was true, although to a lesser degree, in the OAI cohort

(see Table 3). We believe regression to the mean between baseline and the first follow-up was apparent both in the MOST cohort and the OAI [5;11].

Wesseling et al. [4] studied people with early knee OA in the CHECK cohort, and found stable physical function over 4-5 years. They compared the OAI and the CHECK cohort and found physical function to be poorer in the CHECK cohort. Consistent with our study using crude analyses, they coded the post-TKR values as missing and included pre-TKR values, but did not predict pre-TKR values as close as possible to the time of TKR. Inclusion of those with poorest function, even though it might be few study participants, is important to be able to generalize the results to the population with symptomatic knee OA. Our LOESS plot showed a substantial decline in physical function just before TKR, and these values should be visualized in all studies that evaluate the course of physical function in subjects with symptomatic knee OA. Few studies include or describe how they have dealt with those who go through TKR. Holla et al. [17] followed 697 people in the CHECK study with early symptomatic knee OA over 5 years. They coded post-TKR values as missing similar to our approach, and reported overall stable WOMAC-pf values over a period of 5 years, similar to Wesseling et al. using data from the same cohort. In these studies, pre-TKR values were not imputed, but the number of TKRs was relatively few compared to our cohorts.

The lack of functional decline in people with symptomatic knee OA may have different causes. While regression to the mean may play a role, it is also possible that study participants at a group level truly do not get worse over time as a result of stable disease or "response shift" [31]. A response shift of self-reported knee function indicates that the individuals learn to cope with their disability as time goes by. Many subjects with mild symptomatic OA do not have symptoms all the time, and a response shift on self-

reported questionnaires is likely to occur in this group as they know that worse periods are followed by periods with few symptoms. This may have occurred in the MOST study and OAI as their self-reported function seems better than the performance-based results. It has been shown by group-based trajectory modeling, that none of the pain trajectories exhibited substantial worsening over time [32], and this may also be true for physical function outcomes.

While we included missing data in our analysis and predicted poor pre-TKR function in those undergoing TKR, we may still have seen improvement or stability because those undergoing TKR were excluded from the analysis once they had undergone TKR, and this removed persons with the worst function. Data from both the MOST Study and the OAI showed that those who underwent TKR had a considerably worse trajectory of self-reported physical function over time than for those who did not have TKR, and many of those in the TKR group had poor function at study onset.

In this study, two large cohorts of people with incident symptomatic knee OA were followed over respectively 4.5 (MOST) and 6 (OAI) years. Imputation of missing visits and prediction of pre-TKR values closer to the time of TKR have given us the opportunity to include study participants that usually are left out of studies. MI has been considered as a superior technique compared to other methods such as last and baseline observation carried forward, for handling missing data in OA trials [13]. We were however not able to test the assumption that data were MAR, but it is reasonable to believe that the missing data in the MOST Study were MAR on the basis of no differences in baseline characteristics between those who had complete data and those who later had missing data. In addition, we were not able to test the MAR assumption in the OAI, which is a limitation with the present study.

We considered symptomatic knee OA subjects as one group; however, recent studies revealed subgroups of knee OA showing different trajectories of physical function over time [16;33]. Subgrouping subjects may in some cases lead to misclassification and the true course of physical function may be hard to detect [34]. Future studies assessing risk factors for functional decline should probably implement attempts to control for biased outcomes by using strategies such as imputation.

In conclusion, we found that imputing missing values and predicting pre-TKR function reduced some of the bias seen in the unadjusted analyses, which incorrectly suggested improvement in physical function in people with knee OA. The adjusted analyses showed either stable or worsening of physical function, which is more in line with what is observed in the clinic.

Reference List

(1) Neogi T. The epidemiology and impact of pain in osteoarthritis. Osteoarthritis Cartilage 2013; 21(9):1145-1153.

(2) White DK, Zhang Y, Niu J, Keysor JJ, Nevitt MC, Lewis CE et al. Do worsening knee radiographs mean greater chances of severe functional limitation? Arthritis Care Res (Hoboken) 2010; 62(10):1433-1439.

(3) Dowsey MM, Gunn J, Choong PF. Selecting those to refer for joint replacement: who will likely benefit and who will not? Best Pract Res Clin Rheumatol 2014; 28(1):157-171.

(4) Wesseling J, Bierma-Zeinstra SM, Kloppenburg M, Meijer R, Bijlsma JW. Worsening of pain and function over 5 years in individuals with 'early' OA is related to structural damage: data from the Osteoarthritis Initiative and CHECK (Cohort Hip & Cohort Knee) study. Ann Rheum Dis 2013.

(5) van Dijk GM, Veenhof C, Spreeuwenberg P, Coene N, Burger BJ, van SD et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. Arch Phys Med Rehabil 2010; 91(1):58-66.

(6) Dunlop DD, Semanik P, Song J, Manheim LM, Shih V, Chang RW. Risk factors for functional decline in older adults with arthritis. Arthritis Rheum 2005; 52(4):1274-1282.

(7) Mallen CD, Peat G, Thomas E, Lacey R, Croft P. Predicting poor functional outcome in community-dwelling older adults with knee pain: prognostic value of generic indicators. Ann Rheum Dis 2007; 66(11):1456-1461.

(8) Holla JF, Steultjens MP, Roorda LD, Heymans MW, Ten WS, Dekker J. Prognostic factors for the two-year course of activity limitations in early osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken) 2010; 62(10):1415-1425.

(9) Linden A. Assessing regression to the mean effects in health care initiatives. BMC Med Res Methodol 2013; 13:119.

(10) Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014; 22(12):2041-2050.

(11) Pisters MF, Veenhof C, van Dijk GM, Heymans MW, Twisk JW, Dekker J. The course of limitations in activities over 5 years in patients with knee and hip osteoarthritis with moderate functional limitations: risk factors for future functional decline. Osteoarthritis Cartilage 2012; 20(6):503-510.

(12) Soni A, Kiran A, Hart DJ, Leyland KM, Goulston L, Cooper C et al. Prevalence of reported knee pain over twelve years in a community-based cohort. Arthritis Rheum 2012; 64(4):1145-1152.

(13) Olsen IC, Kvien TK, Uhlig T. Consequences of handling missing data for treatment response in osteoarthritis: a simulation study. Osteoarthritis Cartilage 2012; 20(8):822-828.

(14) Hayati RP, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC Med Res Methodol 2015; 15:30.

(15) Peters TJ, Sanders C, Dieppe P, Donovan J. Factors associated with change in pain and disability over time: a community-based prospective observational study of hip and knee osteoarthritis. Br J Gen Pract 2005; 55(512):205-211.

(16) Sharma L, Cahue S, Song J, Hayes K, Pai YC, Dunlop D. Physical functioning over three years in knee osteoarthritis: role of psychosocial, local mechanical, and neuromuscular factors. Arthritis Rheum 2003; 48(12):3359-3370.

(17) Holla JF, van der LM, Heymans MW, Roorda LD, Bierma-Zeinstra SM, Boers M et al. Three trajectories of activity limitations in early symptomatic knee osteoarthritis: a 5-year follow-up study. Ann Rheum Dis 2014; 73(7):1369-1375.

(18) Felson DT, Niu J, Guermazi A, Roemer F, Aliabadi P, Clancy M et al. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum 2007; 56(9):2986-2992.

(19) Felson DT, Niu J, Yang T, Torner J, Lewis CE, Aliabadi P et al. Physical activity, alignment and knee osteoarthritis: data from MOST and the OAI. Osteoarthritis Cartilage 2013; 21(6):789-795.

(20) KELLGREN JH, LAWRENCE JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957; 16(4):494-502.

(21) Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15(12):1833-1840.

(22) Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med 1995; 332(9):556-561.

(23) Motyl JM, Driban JB, McAdams E, Price LL, McAlindon TE. Test-retest reliability and sensitivity of the 20-meter walk test among patients with knee osteoarthritis. BMC Musculoskelet Disord 2013; 14:166.

(24) Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots.

(25) Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991; 10(4):585-598.

(26) Niu J, Nevitt M, McCulloch C, Torner J, Lewis CE, Katz JN et al. Comparing the functional impact of knee replacements in two cohorts. BMC Musculoskelet Disord 2014; 15:145.

(27) van Leeuwen DM, Peeters GM, de Ruiter CJ, Lips P, Twisk JW, Deeg DJ et al. Effects of self-reported osteoarthritis on physical performance: a longitudinal study with a 10-year follow-up. Aging Clin Exp Res 2013; 25(5):561-569.

(28) Covinsky KE, Lindquist K, Dunlop DD, Gill TM, Yelin E. Effect of arthritis in middle age on older-age functioning. J Am Geriatr Soc 2008; 56(1):23-28.

(29) 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(9):834-840.

(30) Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005; 64(1):29-33.

(31) Razmjou H, Schwartz CE, Yee A, Finkelstein JA. Traditional assessment of health outcome following total knee arthroplasty was confounded by response shift phenomenon. J Clin Epidemiol 2009; 62(1):91-96.

(32) Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2014; 22(5):622-630.

(33) Holla JF, van der LM, Knol DL, Roorda LD, Hilberdink WK, Lems WF et al. Predictors and outcome of pain-related avoidance of activities in persons with early symptomatic knee osteoarthritis: A 5-year follow-up study. Arthritis Care Res (Hoboken) 2014.

(34) Colbert CJ, Almagor O, Chmiel JS, Song J, Dunlop D, Hayes KW et al. Excess body weight and four-year function outcomes: comparison of African Americans and whites in a prospective study of osteoarthritis. Arthritis Care Res (Hoboken) 2013; 65(1):5-14.

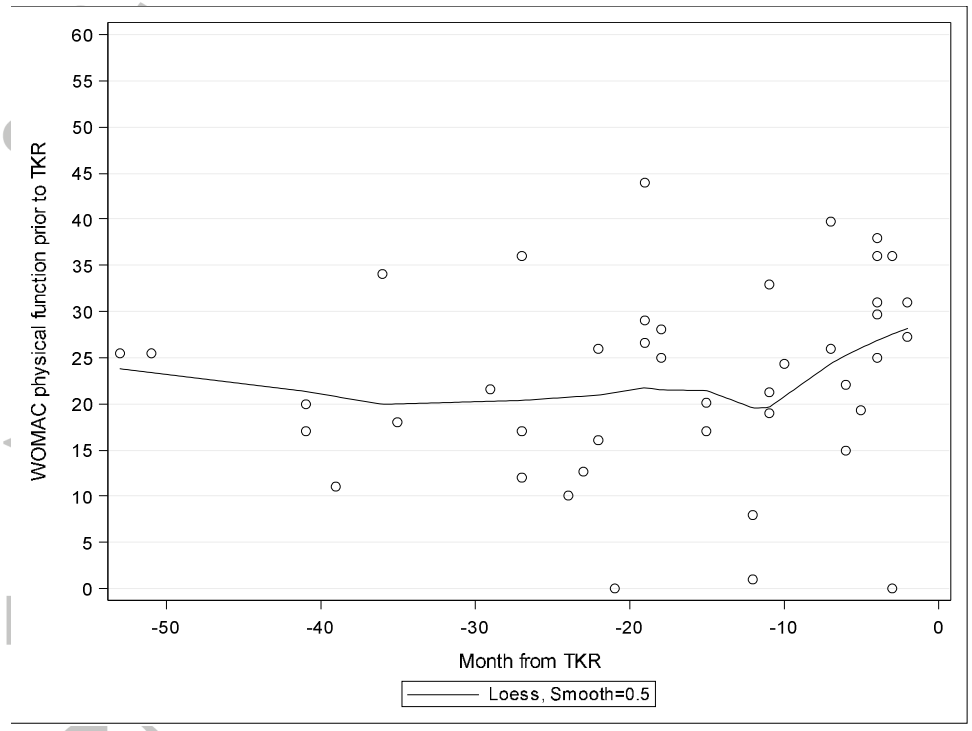

Figure 1. LOESS curve of WOMAC-pf values in the MOST Study at the last visit before total knee replacement (TKR). The time of the last visit varied from 0-60 months prior to TKR surgery as shown on the X-axis. The fitted line was used to predict pre-TKR values at the time point 0 (32.1 units + deviation from line).

Physical function in knee osteoarthritis

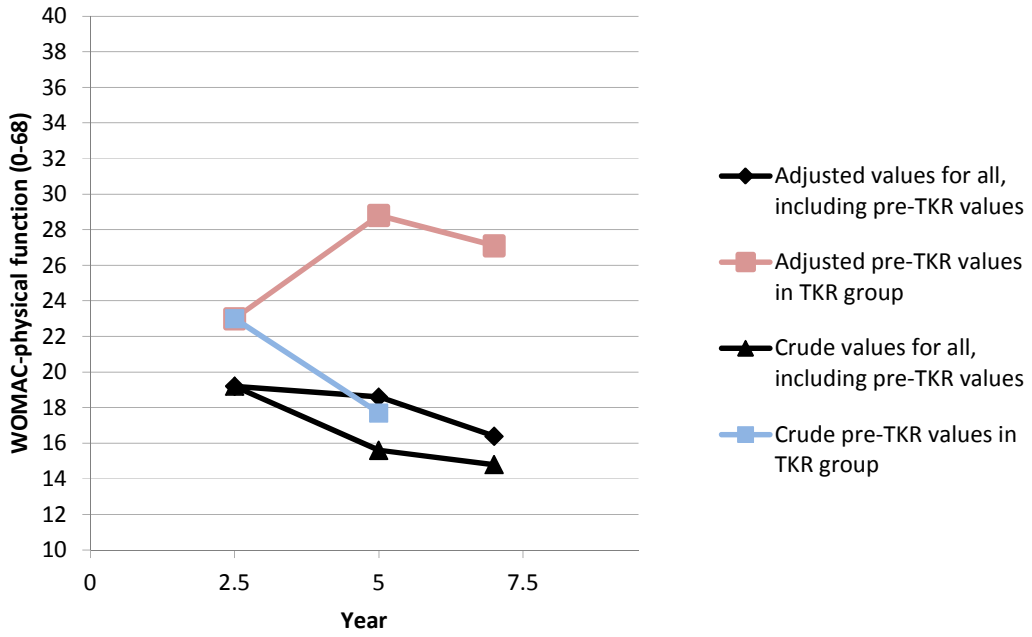

Figure 2. WOMAC physical function (0-68) over 4.5 years among MOST participants (higher WOMAC-pf values indicate worse function). For pre-total knee replacement (TKR) values, the last visit exclude the large number of subjects who had TKRs.

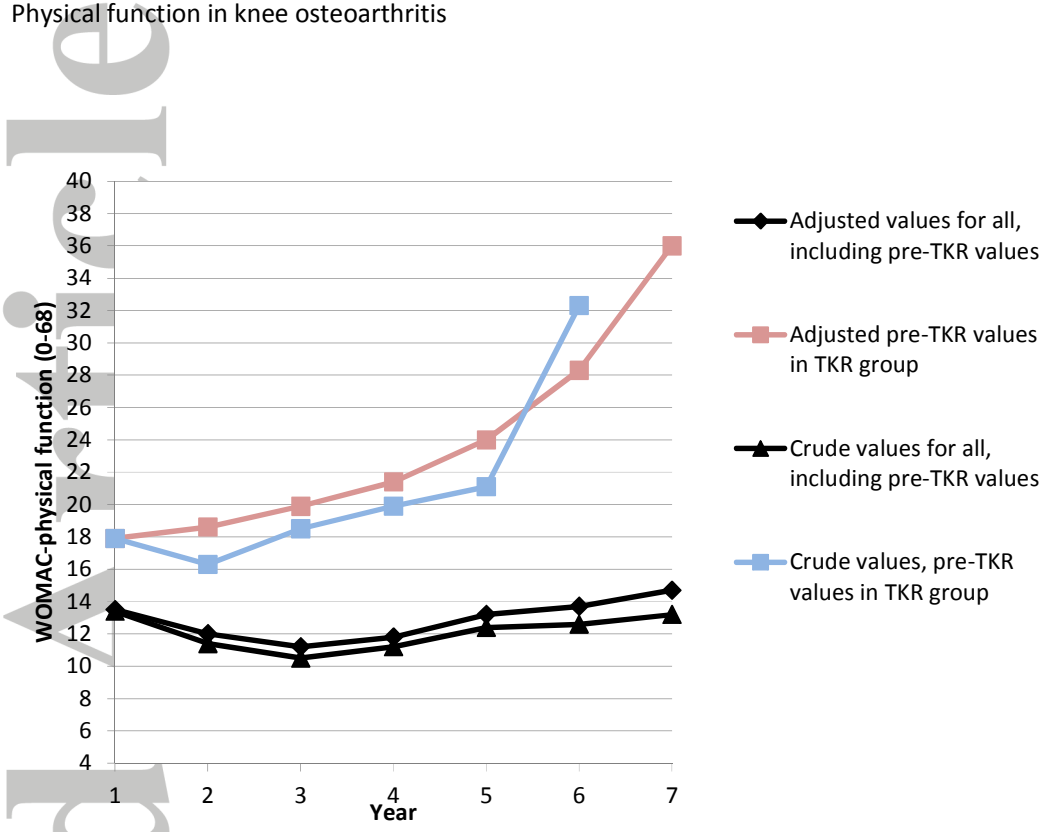

Figure 3. WOMAC physical function (0-68) over 6 years among OAI participants (higher WOMAC-pf values indicate worse function). For pre-total knee replacement (TKR) values, the last visit exclude the large number of subjects who had TKRs.

Figure 1. LOESS curve of WOMAC-pf values in the MOST Study at the last visit before total knee replacement (TKR). The time of the last visit varied from 0-60 months prior to TKR surgery as shown on the X-axis. The fitted line was used to predict pre-TKR values at the time point 0 (32.1 units + deviation from line).

Figure 2. WOMAC physical function (0-68) over 4.5 years among MOST participants (higher WOMAC-pf values indicate worse function). For pre-total knee replacement (TKR) values, the last visit exclude the large number of subjects who had TKRs.

Figure 3. WOMAC physical function (0-68) over 6 years among OAI participants (higher WOMAC-pf values indicate worse function). For pre-total knee replacement (TKR) values, the last visit exclude the large number of subjects who had TKRs.

Table 1. Characteristics of participants.

	Age	BMI	WOMAC pain
MOST (n=225)			
30-mo	66.5(8.2)	30.9(5.7)	5.9(3.4)
60-mo	69.7(8.2)	30.8(6.1)	4.6(3.4)
84-mo	71.7(8.3)	30.7(6.0)	4.4(3.5)
OAI (n=577)			
12-mo	63.6(8.6)	29.7(4.8)	4.6(3.5)
24-mo	65.2(8.6)	29.9(5.0)	3.6(3.4)
36-mo	66.1(8.6)	30.0(4.9)	3.3(3.4)
48-mo	67.1(8.6)	30.0(4.9)	3.5(3.4)
60-mo	67.9(8.6)	29.3(5.2)	3.8(3.7)
72-mo	68.4(8.4)	30.1(5.0)	3.8(3.7)
84-mo	69.2(8.5)	n/a	3.9(3.8)

Values are mean (standard deviation).

Table 2. Physical function values within crude and adjusted data in the MOST Study

Outcome		N	30 mo	N	60 mo	N	84 mo	p-value*
WOMAC-pf	Crude	223	19.2(10.5)	165	15.6(11.1)	139	14.8(10.9)	<.0001
	Adjusted	224	19.2(10.5)	222	18.6(12.8)	190	16.4(11.7)	0.0668
Five Times Sit To Stand	Crude	216	11.3(3.9)	166	12.8(3.9)	136	12.5(4.3)	<.0001
	Adjusted	218	11.3(4.0)	216	12.7(3.8)	188	12.8(5.1)	<.0001
20-Meter Walk	Crude	224	17.6(4.0)	170	17.3(3.0)	143	17.8(3.6)	0.0821
	Adjusted	225	17.6(4.0)	223	17.8(3.5)	191	18.1(4.2)	0.0026

Values are mean (standard deviation) at the actual exams in people with incident knee OA defined up to the 30-month visit. Pf, physical function; N, number of participants; mo, months; Five Times Sit To Stand and 20-Meter Walk are measured in seconds. Crude data: original values, including unadjusted pre-TKR values, but no post-TKR values. Adjusted data: original values with imputed missing values, and adjusted pre-TKR values. Higher values indicate worse function.

**P-value indicates linear trend over time from 30-84 months.

Table 3. Physical function values within crude and adjusted data in the OAI

Outcome		12 mo	24 mo	36 mo	48 mo	60 mo	72 mo	84 mo	p-value*
WOMAC-pf	N	573	525	515	496	461	347	208	
	Crude	13.4(11.0)	11.4(10.9)	10.5(10.8)	11.2(10.7)	12.4(11.7)	12.6(12.0)	13.2(12.3)	0.0816
	N	576	574	561	549	535	416	245	
	Adjusted	13.5(11.1)	12.0(11.2)	11.2(11.0)	11.8(11.1)	13.2(11.7)	13.7(12.4)	14.7(12.9)	0.0001
Five Times Sit To Stand	N	523	472	360	292	105	169	n/a	
	Crude	11.5(3.5)	11.2(3.2)	11.2(3.0)	11.3(3.8)	10.6(3.1)	11.6(4.8)	n/a	0.3636
	N	562	561	445	377	163	262	n/a	
	Adjusted	11.7(3.7)	11.4(3.8)	11.6(3.4)	11.6(4.0)	11.2(3.9)	12.1(4.9)	n/a	0.4814
20 Meter Walk	N	570	499	387	305	116	184	n/a	
	Crude	15.9(3.0)	15.8(3.5)	15.9(2.9)	16.2(2.7)	16.3(3.4)	16.7(4.3)	n/a	<.0001
	N	577	575	452	384	165	265	n/a	
	Adjusted	15.8(3.0)	15.9(3.6)	16.1(3.3)	16.4(3.0)	16.7(4.9)	17.2(4.4)	n/a	0.0002

Values are mean (standard deviation) at the actual exams from the time of incident knee osteoarthritis (OA). Pf: physical function; N, number. Crude data: original values for those with symptomatic knee OA, including unadjusted pre-TKR values, but no post-TKR values.

Adjusted data: original values with imputed missing values, and adjusted pre-TKR values *P-value indicates linear trend over time from visit 1 to visit 7, i.e., from the visit when incident symptomatic knee OA observed to 72 months after it.