# Longitudinal Course of Physical Function in People With Symptomatic Knee Osteoarthritis: Data From the Multicenter Osteoarthritis Study and the Osteoarthritis Initiative

**BRITT ELIN ØIESTAD, PT, PhD**[1], **DANIEL K. WHITE, PT, ScD, MSc**[2], **ROSS BOOTON, MMath**[3], **JINGBO NIU, DSc**[3], **YUQING ZHANG, DSc**[3], **JIM TORNER, PhD**[4], **CORA E. LEWIS, MD, PhD**[5], **MICHAEL NEVITT, PhD**[6], **MICHAEL LaVALLEY, PhD**[7], and **DAVID T. FELSON, MD, MPH**[8]

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

[2]University of Delaware, Newark; Boston University School of Medicine, Boston, Massachusetts

[3]Boston University School of Medicine, Boston, Massachusetts

[4]University of Iowa, Iowa City

[5]University of Alabama, Birmingham

[6]University of California, San Francisco

[7]Boston University School of Medicine and School of Public Health, Boston, Massachusetts

[8]Boston University School of Medicine, Boston, Massachusetts; University of Manchester, Manchester, UK

## Abstract

**Objective**—Pain and functional decline are hallmarks of knee osteoarthritis (OA). Nevertheless, 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. Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) physical function (WOMAC-PF), the 5-times sit-to-stand test, and the 20-meter-walk test were assessed at 4 and 5 years in MOST and at 6 years in OAI. 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 calculation 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 MOST and 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.

## INTRODUCTION

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). In contrast, 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 2 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 reassessment 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 followup have been shown to be older and have poorer function than those without missing visits (11,12). Since including subjects with complete followup 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 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 (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 followup visits were often older, had a lower education level, had longer duration of symptoms, and had lower muscle strength (4,11). In such cases, the missing data are suggested to be missing at random (MAR) (14), where the attendance can be calculated 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 subjects who 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 with symptomatic knee OA, but the longitudinal pre-surgical data often do not include functional status as close as possible to the time of TKR.

Newer studies (e.g., the Cohort Hip and Cohort Knee [CHECK] study) have addressed the issue of loss to followup by using multiple imputation techniques and included pre-TKR values in the analyses (11,17), but the pre-TKR values may have been measured a relatively long time before TKR. We know little about the longitudinal course of physical function, including approaches to account for missing data and those with the poorest function who 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 calculated 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 calculated pre-TKR values of physical function (adjusted data).

## MATERIALS AND METHODS

We used data from the Multicenter Osteoarthritis Study (MOST) 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 or 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, or end-stage disease at baseline were excluded from both studies (2).

MOST included 3,026 people ages 50–79 years from Birmingham, Alabama, and Iowa City, Iowa, and it was ongoing for 84 months, including 4 visits (at baseline, and at 30, 60, and 84 months) (18). OAI included 4,796 people ages 45–79 years at 4 clinical sites: Baltimore, Maryland; Pittsburgh, Pennsylvania; Pawtucket, Rhode Island; and Columbus, Ohio. Followup visits were conducted annually for 7 years (www.oai.ucsf.edu) (19). The study participants provided informed consent before the first clinic visit in both cohorts.

We excluded people who had TKR at baseline (MOST, n = 78; OAI, n = 63) and people who underwent hip replacement at any time point (MOST, n = 140; OAI, n = 190), because physical function in these groups was probably also influenced by disease other than knee OA. After the exclusions, MOST comprised 2,808 study participants, and OAI comprised 4,543. Among these, we included subjects with incident symptomatic knee OA. We defined subjects as having this condition if they had no symptomatic OA in both knees at baseline

but developed new disease in either knee during the 30-month followup in MOST or during the 36-month followup in OAI.

### Radiographic examination

We used data from standing posterior–anterior radiographs using a SynaFlexer frame for standardized positioning (Synarc). The radiographs were read and scored by a musculoskeletal radiologist and a rheumatologist at Boston University, blinded to clinical data. An adjudication panel resolved discrepancies between the readers. For this study, we included people with incident symptomatic knee OA on the basis of radiographic examination of the tibiofemoral joint (Kellgren/Lawrence [K/L] scale grade 2) (20) and having pain in that knee on most days of the previous month. We classified the subjects based on their worst knee in terms of symptomatic knee OA.

### Outcome measurements

The followup of the included subjects started from the visit when incident symptomatic knee OA was observed. Self-reported physical function was measured at all visits in MOST (at baseline, and at 30, 60, and 84 months) and OAI (annually from 0 to 84 months), using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) physical function subscale (WOMAC-PF) (21). We used the 5-point Likert scale version (where 0 = no problems and 4 = extreme functional problems) for the physical function subscale (range 0–68).

Performance-based measures included the 5-times sit-to-stand test (22) and the 20-meter-walk test (23). These were measured at all visits in MOST and annually for 84 months in OAI, except for the 60- and 84-month visits. For the 5-times sit-to-stand test, the participants stood from a chair 5 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 performance-based tests, we used the mean time of 2 trials.

### Statistical analysis

The crude outcome data included participants with complete visits only, including the TKR group with original pre-TKR values. We used 2 methods to generate the adjusted outcome data. First, because many of the subjects who underwent TKR in MOST and OAI had their last clinic visit many months before TKR, we wanted to add calculated pre-TKR outcome values (WOMAC-PF, 5-times sit-to-stand test, and 20-meter-walk test) as close as possible to the time of TKR. Using data from those who underwent TKR, we constructed loess curves (24) to estimate fitted curves for the outcome variables at the last visit before TKR and the months from last visit to TKR (Figure 1, using MOST WOMAC-PF data). We then used the fitted curves to calculate a subject's WOMAC-PF value at the time of TKR. Removing outliers did not alter these curves substantively.

We used the loess curves to calculate the average outcome values just before 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). Thus for each subject, the new calculated WOMAC-PF value was calculated as WOMAC-PF = 32.1 +/− the subject's deviation from the curve. For the 5-times sit-to-stand test, the equation was calculated as 5-times sit-to-stand test = 13.4 +/− deviation from the fitted curve, and for the 20-meter-walk test was calculated as 20-meter-walk test = 20.7 +/− deviation from the fitted curve. In OAI, the corresponding equations were WOMAC-PF = 28.6 +/− deviation from the fitted curve, 5-times sit-to-stand test = 13.4 +/− deviation from the fitted curve, and 20-meter-walk test = 18.4 +/− deviation from the fitted curve. We assigned these new calculated outcome values to the data as an additional pre-TKR visit.

As a second way to generate adjusted outcome data, values for missing visits for the 3 outcome variables during the followup period were filled in using the multiple imputation method (25), with the independent variables age, sex, and body mass index (BMI). The missing pattern for WOMAC-PF in MOST, considering the visits prior to TKR or death, showed that 79% of subjects 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 5-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 OAI were 83%, 68%, and 79%, respectively, with no subjects missing all WOMAC-PF visits. Among subjects who underwent TKR, 2% had some missing visits in MOST, and none were missing all visits for WOMAC-PF and the 20-meter-walk test, but 4 had missing visits on the 5-times sit-to-stand test. 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, sex, 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, the pattern was probably not MNAR. We may, however, speculate that the baseline function is highly correlated with the followup function, and this correlation suggests that the missing pattern in MOST 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 data sets, including the 3 outcome variables in conditioning on age, sex, and BMI, and the 3 outcome variables using SAS software. The effect estimates for the 5 data sets were pooled together across imputed data sets, 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 for those subjects with incident symptomatic knee OA.

## RESULTS

In MOST, 225 participants (8%) had incident symptomatic knee OA during the first 30 months of followup. The corresponding number in OAI was 577 (12.7%) during the first 36 months of followup. 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 and 62.1% in OAI were women. The mean ± SD age in MOST at 30 months was 66.5 ± 8.2 years, with a mean ± SD BMI of 30.9 ± 5.7, and in OAI at 12 months, 63.6 ± 8.6 years with a BMI of 29.7 ± 4.8. At the MOST visit where we found participants with symptomatic knee OA, 90 (46.2%) had K/L grade 2, 88 (39.1%) K/L grade 3, and 33 (14.7%) K/L grade 4. In OAI, 293 (50.8%) had K/L grade 2, 208 (36%) K/L grade 3, and 76 (13.2%) K/L grade 4.

In Figures 2 and 3, we show the crude and adjusted curves for physical function over time in MOST and OAI, respectively, for the TKR group only, and for crude and adjusted values for all those subjects 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 MOST, 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 participants. For the 5-times sit-to-stand test and the 20-meter-walk test, significant worsening was found over time for the crude (the 5-times sit-to-stand test only) and adjusted data (Table 2). After adjustment, the 5-times sit-to-stand test results and the 20-meter-walk test results increased by 1.5 seconds from the first to the last visit ($P < 0.003$). In OAI, worsening within adjusted WOMAC-PF values was seen over time (Table 3). We also found significantly worsening values for the crude and adjusted 20-meter-walk test data in OAI.

## DISCUSSION

Among people with incident knee OA, we found physical function generally worsened over time after imputing for missing data and calculating 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 MOST. In OAI, we found a decline in self-reported physical function after incorporating calculated pre-TKR values and data for missing visits. Overall, subjects with incident knee OA in MOST and OAI reported stable or slightly lower physical function values from the time of diagnosis to the last clinic visit, in line with others (27). For 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. The researchers imputed values for loss to followup for both self-reported (WOMAC-PF) and performance-based (10-meter-walk) outcomes. They found significantly improved WOMAC-PF results and stable 10-meter-walk results. The improvement in WOMAC-PF was 2.7 points (9%), which the authors stated was not clinically relevant according to other literature (29,30). In contrast to our study, to our knowledge, no subjects had TKR during the 5 years of followup, and those who died were excluded. In addition, the study participants had either knee or hip OA. These differences 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 based on regression to the mean between the first and second visit. After imputing missing values and calculating pre-TKR values, although the initial regression to the mean persisted (see Figures), we found a later decline in function, and some of this phenomenon in the MOST data disappeared (Table 2). This finding was true, although to a lesser degree, in the OAI cohort (Table 3). We believe regression to the mean between baseline and the first followup was apparent both in the MOST cohort and 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 calculate pre-TKR values as close as possible to the time of TKR. Inclusion of those with poorest function, even though this inclusion 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. The researchers 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, possibly 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 response shift may have occurred in MOST and OAI, as subjects' self-reported function

seems better than the performance-based results. Group-based trajectory modeling has shown 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 exclusion removed persons with the worst function. Data from both MOST and OAI showed that those who underwent TKR had a considerably worse trajectory of self-reported physical function over time than those who did not have TKR, and many of those in the TKR group had poor function at study onset.

In this study, 2 large cohorts of people with incident symptomatic knee OA were followed over 4.5 years (MOST) and 6 years (OAI). Imputation of missing visits and calculation of pre-TKR values closer to the time of TKR have given us the opportunity to include study participants who usually are left out of studies. Multiple imputation has been considered 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 MOST 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 OAI, which is a limitation of the present study.

We considered symptomatic knee OA subjects as one group, but recent studies have 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 calculating 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 physical function, which is more in line with what is observed in the clinic.

## Acknowledgments

# REFERENCES

1. Neogi T. The epidemiology and impact of pain in osteoarthritis. Osteoarthritis Cartilage. 2013; 21:1145–53. [PubMed: 23973124]

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; 62:1433–9. [PubMed: 20506398]

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; 28:157–71. [PubMed: 24792950]

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. 2015; 74:347–53. [PubMed: 24243926]

5. Van Dijk GM, Veenhof C, Spreeuwenberg P, Coene N, Burger BJ, van Schaardenburg D, et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. Arch Phys Med Rehabil. 2010; 91:58–66. [PubMed: 20103397]

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; 52:1274–82. [PubMed: 15818691]

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; 66:1456–61. [PubMed: 17456527]

8. Holla JF, Steultjens MP, Roorda LD, Heymans MW, ten Wolde S, 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; 62:1415–25. [PubMed: 20496430]

9. Linden A. Assessing regression to the mean effects in health care initiatives. BMC Med Res Methodol. 2013; 13:119. [PubMed: 24073634]

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; 22:2041–50. [PubMed: 25305072]

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; 20:503–10. [PubMed: 22330176]

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; 64:1145–52. [PubMed: 22180258]

13. Olsen IC, Kvien TK, Uhlig T. Consequences of handling missing data for treatment response in osteoarthritis: a simulation study. Osteoarthritis Cartilage. 2012; 20:822–8. [PubMed: 22441031]

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. [PubMed: 25880850]

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; 55:205–11. [PubMed: 15808036]

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; 48:3359–70. [PubMed: 14673987]

17. Holla JF, van der Leeden M, 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; 73:1369–75. [PubMed: 23716068]

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; 56:2986–92. [PubMed: 17763427]

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; 21:789–95. [PubMed: 23523851]

20. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. Ann Rheum Dis. 1957; 16:494–502. [PubMed: 13498604]

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; 15:1833–40. [PubMed: 3068365]

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; 332:556–61. [PubMed: 7838189]

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. [PubMed: 23663561]

24. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc. 1978; 74:829–36.

25. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991; 10:585–98. [PubMed: 2057657]

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. [PubMed: 24885404]

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; 25:561–9. [PubMed: 23949969]

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; 56:23–8. [PubMed: 18184204]

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; 60:834–40. [PubMed: 11502609]

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; 64:29–33. [PubMed: 15208174]

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; 62:91–6. [PubMed: 19095168]

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; 22:622–30. [PubMed: 24662734]

33. Holla JF, van der Leeden M, 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). 2015; 67:48–57. [PubMed: 24905958]

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; 65:5–14. [PubMed: 22833527]

**Significance & Innovations**

- Longitudinal studies of knee osteoarthritis (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 calculating 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 calculated pre-TKR values as close as possible to those of TKR, trajectories of physical function changed from improvement to stable or worsened values over time.
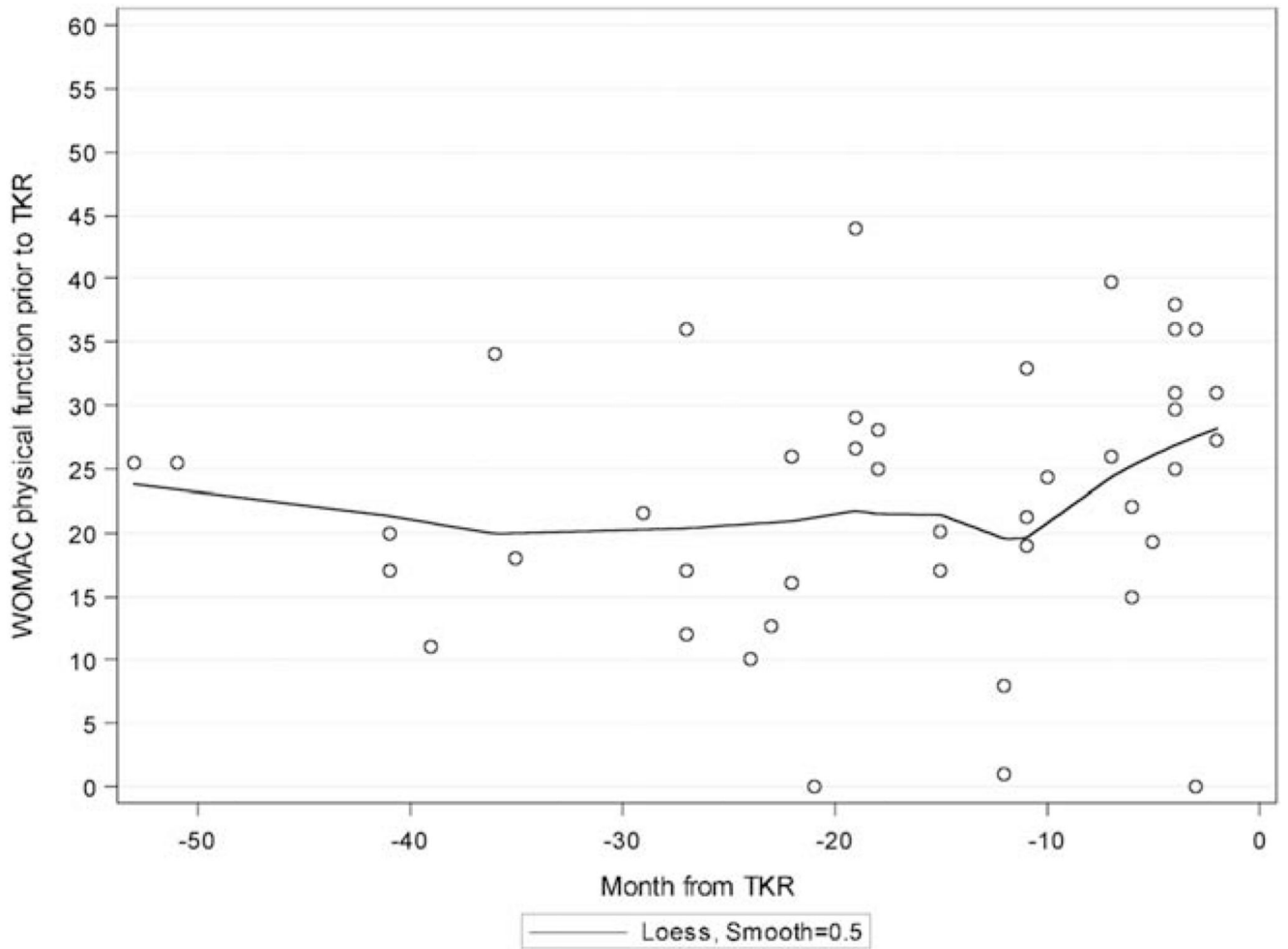
**Figure 1.**
Loess curve of Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) physical function values in the Multicenter Osteoarthritis Study at the last visit before total knee replacement (TKR). The time of the last visit varied from 0 to 60 months prior to TKR surgery, as shown on the X-axis. The fitted line was used to calculate pre-TKR values at the time point 0 (32.1 units + deviation from line).
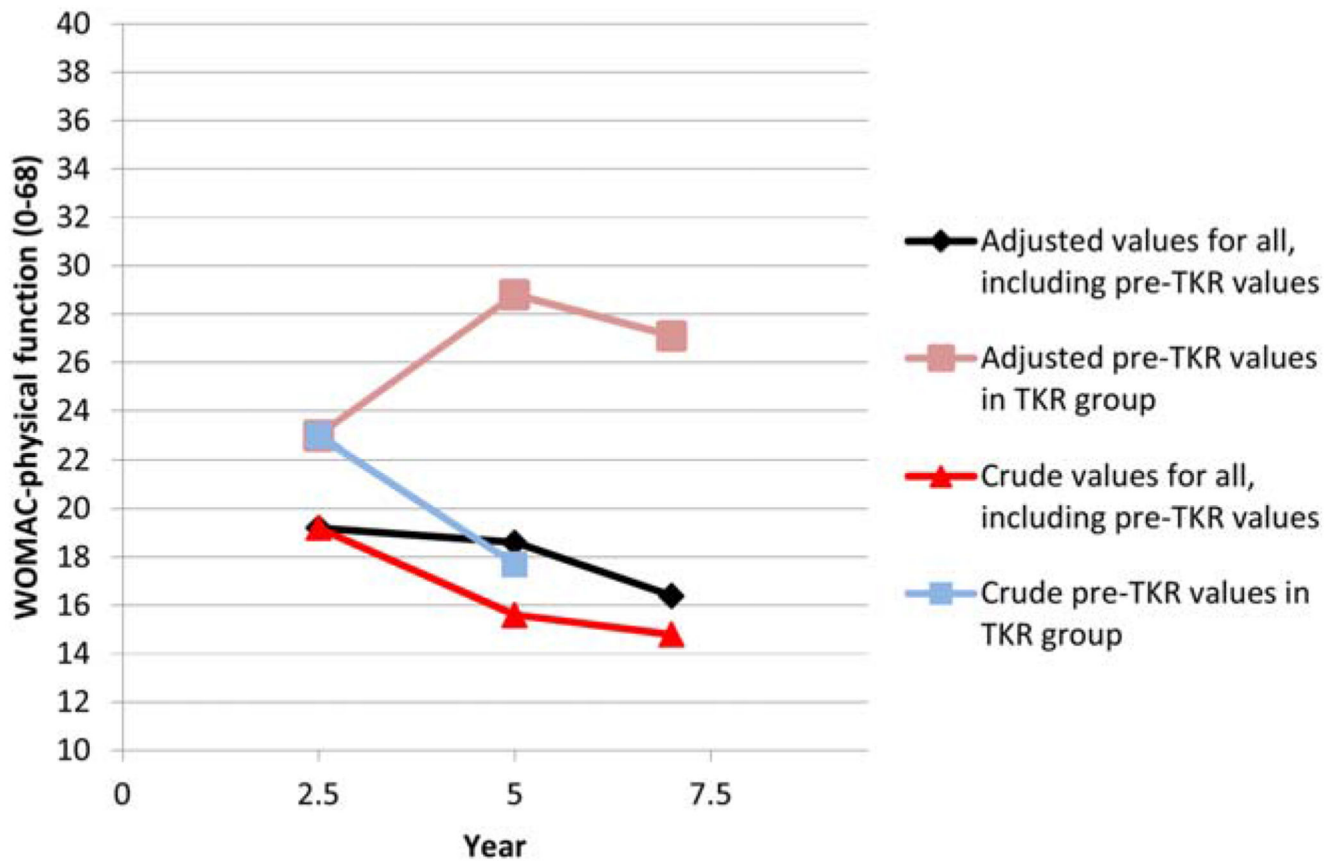
**Figure 2.**
Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) physical
function (range 0–68) over 4.5 years among Multicenter Osteoarthritis Study participants
(higher WOMAC physical function values indicate worse function). For pre–total knee
replacement (TKR) values, the last visit excludes the large number of subjects who had
TKRs.
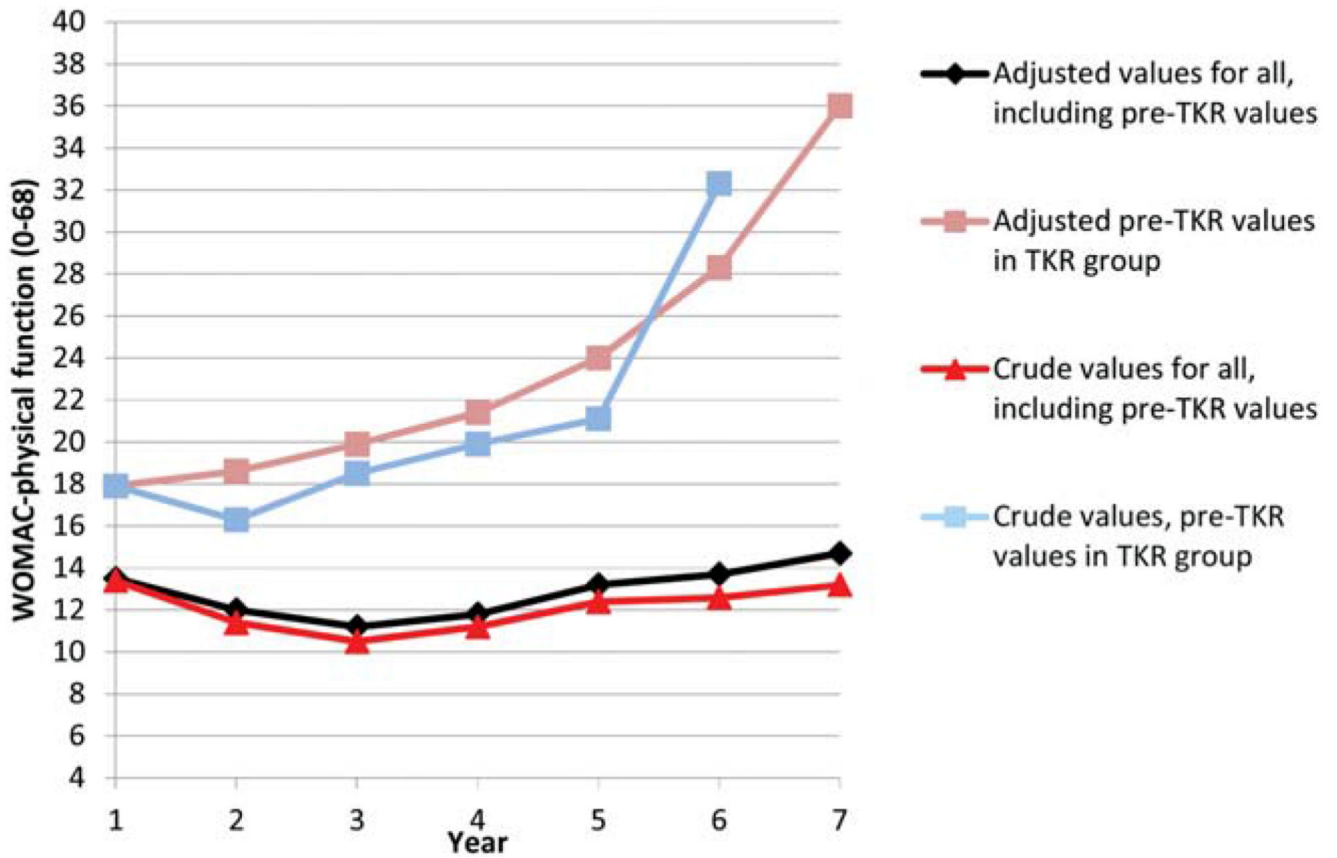
**Figure 3.**
Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) physical function (range 0–68) over 6 years among Osteoarthritis Initiative participants (higher WOMAC physical function values indicate worse function). For pre–total knee replacement (TKR) values, the last visit excludes the large number of subjects who had TKRs.

**Table 1**

**Characteristics of participants**[*]

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

[*] Values are mean ± SD. BMI = body mass index; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; MOST = Multicenter Osteoarthritis Study; OAI = Osteoarthritis Initiative; NA = not applicable.

[†] Measured at the month indicated.

**Table 2**

**Physical function values within crude and adjusted data in MOST***

| Outcome | 30 months | | 60 months | | 84 months | | $P$ [†] |
|---|---|---|---|---|---|---|---|
| | No. | Values | No. | Values | No. | Values | |
| WOMAC-PF | | | | | | | |
|   Crude | 223 | 19.2 ± 10.5 | 165 | 15.6 ± 11.1 | 139 | 14.8 ± 10.9 | < 0.0001 |
|   Adjusted | 224 | 19.2 ± 10.5 | 222 | 18.6 ± 12.8 | 190 | 16.4 ± 11.7 | 0.0668 |
| 5-times sit-to-stand test[‡] | | | | | | | |
|   Crude | 216 | 11.3 ± 3.9 | 166 | 12.8 ± 3.9 | 136 | 12.5 ± 4.3 | < 0.0001 |
|   Adjusted | 218 | 11.3 ± 4.0 | 216 | 12.7 ± 3.8 | 188 | 12.8 ± 5.1 | < 0.0001 |
| 20-meter-walk test[‡] | | | | | | | |
|   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 ± SD, unless indicated otherwise, measured at the actual examinations in people with incident knee osteoarthritis, defined up to the 30-month visit. 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. MOST = Multicenter Osteoarthritis Study; WOMAC-PF = Western Ontario and McMaster Universities Osteoarthritis Index physical function.

[†]Linear trend over time from 30 to 84 months.

[‡]Measured in seconds.

**Table 3**

**Physical function values within crude and adjusted data in OAI[*]**

| Outcome[†] | 12 | 24 | 36 | 48 | 60 | 72 | 84 | P[‡] |
|---|---|---|---|---|---|---|---|---|
| WOMAC-PF | | | | | | | | |
| No. | 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 |
| No. | 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 |
| 5-times sit-to stand test | | | | | | | | |
| No. | 523 | 472 | 360 | 292 | 105 | 169 | NA | |
| 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 | NA | 0.3636 |
| No. | 562 | 561 | 445 | 377 | 163 | 262 | NA | |
| 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 | NA | 0.4814 |
| 20-meter walk test | | | | | | | | |
| No. | 570 | 499 | 387 | 305 | 116 | 184 | NA | |
| 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 | NA | < 0.0001 |
| No. | 577 | 575 | 452 | 384 | 165 | 265 | NA | |
| 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 | NA | 0.0002 |

[*] Values are the number or the mean ± SD, unless indicated otherwise, at the actual examinations from the time of incident knee osteoarthritis (OA). 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. OAI = Osteoarthritis Initiative; WOMAC-PF = Western Ontario and McMaster Universities Osteoarthritis Index physical function; NA = not applicable.

[†] Measured from 12 to 84 months.

[‡] 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).