



Published in final edited form as:

Knee. 2009 December ; 16(6): 494–500. doi:10.1016/j.knee.2009.04.002.

Two-Year Incidence and Predictors of Future Knee Arthroplasty in Persons with Symptomatic Knee Osteoarthritis: Preliminary Analysis of Longitudinal Data from the Osteoarthritis Initiative

Daniel L. Riddle, PT, PhD, FAPTA,

Otto D. Payton Professor in the Departments of Physical Therapy and Orthopaedic Surgery, Virginia Commonwealth University, Medical College of Virginia Campus, Richmond, VA. 23298-0224, dlriddle@vcu.edu, phone: 804-828-0234, fax: 804-828-8111

Xiangrong Kong, PhD, and

Assistant Scientist, Department of Family Population and Reproductive Health, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD 21205, kongx@vcu.edu

William A. Jiranek, M.D.

Associate Professor in the Department of Orthopaedic Surgery, Virginia Commonwealth University, Medical College of Virginia Campus, Richmond, VA. 23298-0153, wjiranek@mcvh-vcu.edu, phone: 804-828-6784

Abstract

Objective—There is little evidence to guide physicians when discussing future likelihood of knee arthroplasty with patients who have symptomatic knee osteoarthritis. Data from Osteoarthritis Initiative (OAI) was used to determine the incidence of and predictors for knee arthroplasty.

Methods—OAI data were collected on a sample of 778 persons aged 45 to 79 years with symptomatic knee osteoarthritis. An extensive set of measurements were obtained at baseline and persons were followed for 2 years to identify who underwent knee arthroplasty. Random forest analysis was used to identify optimal variables that discriminate among those who did and those who did not undergo knee arthroplasty.

Results—The two year incidence of knee arthroplasty in the cohort was 3.7% (95% CI, 2.6%, 5.3%). Because of the low number of knee arthroplasty procedures, the predictor analysis was preliminary in nature. The analysis identified several variables that could be used to assist in identifying patients at future risk for knee arthroplasty.

Conclusion—For persons at high risk of knee arthroplasty, the two year incidence of knee arthroplasty is very low. The most powerful predictors were those that accounted for disease severity and functional loss. These data could assist physicians in advising patients with knee osteoarthritis on future surgical care.

© 2009 Elsevier B.V. All rights reserved.

Correspondence to: Daniel L. Riddle.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement: No authors have financial and personal relationships with other people or organisations that could inappropriately influence this work.

Introduction

Knee arthroplasty has become the most effective surgical intervention for patients with osteoarthritic knee pain. There is, however, a lack of research that has described the incidence of knee arthroplasty or predictors of knee arthroplasty in persons with radiographically confirmed symptomatic knee osteoarthritis (OA). Knowing knee arthroplasty incidence rates in persons with symptomatic knee OA would provide insight into resource utilization for an increasingly common and costly procedure. By assessing for baseline variables that are associated with future knee arthroplasty, orthopaedists could assist patients in better planning for potential future surgical care.

Only two papers were found that prospectively assessed persons with knee arthritis to determine if predictor variables for knee arthroplasty could be identified. Hawker and colleagues studied 2128 persons 55 years and older from two regions in Ontario, Canada.¹ Subjects had either hip or knee osteoarthritis, which was defined, based on subject self-report, as difficulty with stair climbing, arising from a chair, standing and walking and at least 6 weeks of pain, swelling or stiffness in hips or knees. Subjects also had to indicate that either a hip or a knee was "troublesome" and the WOMAC summary score had to be greater than 39 out of 100. Subjects were followed for approximately 6 years to determine which subjects underwent either a hip or knee arthroplasty. Key predictors were higher baseline WOMAC summary scores, patients in their 60s versus older, self-perceived better health and willingness to consider joint arthroplasty. Willingness to consider surgery was the most powerful predictor of time to surgery (hazard ratio = 4.92 (95% CI: 3.7, 6.4)).

Momahara and colleagues determined predictors of knee arthroplasty in a cohort of 955 patients with rheumatoid arthritis and pain or tenderness in one or both knees.² The most powerful predictors found were baseline disability and pain severity and age. Both studies were conducted in countries other than the US and, with the exception of the use of lab values in one study,² relied entirely on patient self report measures to identify predictors.

The demand for knee arthroplasty is already large and will increase substantially in coming years.³ Yet, no research was found that described the incidence of knee arthroplasty or variables that predict surgical treatment choice in US patients most likely to undergo the procedure, those with symptomatic knee OA. This study describes the incidence of knee arthroplasty over a 2 year period in a cohort of 778 persons with symptomatic knee osteoarthritis from the Osteoarthritis Initiative, a large multicenter NIH funded study.⁴ In addition, predictors of knee arthroplasty over a 2 year period are reported. Given current evidence, we hypothesized that two year incidence of knee replacement would be low and that several predictors would be identified.

Materials and Methods

The Osteoarthritis Initiative

The Osteoarthritis Initiative* is a publicly and privately funded prospective longitudinal cohort study planned for a 4 year follow-up. A primary objective of the OAI study is to develop diverse cohorts of persons for the study of the natural history, risk factors, onset and progression of

*The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners."

knee tibiofemoral OA. No treatment was provided as part of the study nor were patients solicited for treatment by the investigators.

A total of three sub-cohorts: the incidence subcohort, the control subcohort, and the progression subcohort, are being followed in the OAI study. Each subcohort has racially and ethnically diverse mixes of persons between the ages of 45 and 79 years at baseline. The incidence subcohort comprises 3,285 persons at risk for knee osteoarthritis but with no symptomatic knee osteoarthritis in either knee. The control subcohort of 122 subjects had no pain and no risk factors for knee osteoarthritis at baseline. The progression subcohort, which is, specifically, the target of this study, has 1,389 persons with symptomatic knee osteoarthritis in one or both knees. When we conducted our study, a subset of 778 of the 1,389 persons in the progression subcohort had two year follow-up data available for public use. Follow-up data for the remaining 611 persons in the progression cohort were to be collected at a later time and were therefore not publically available at the time the current study was completed.

Subjects were recruited for the progression cohort via mailings to clinical populations in the four recruitment sites, advertisement in local newspapers, and presentations to churches, community and civic organizations and via a website. Of the 778 eligible persons in the progression cohort, 160 (20.6%) were currently seeing a doctor or other health professional for treatment of arthritis, indicating that the great majority were recruited from communities associated with the four recruitment sites and were not receiving care for their arthritis. Persons with symptomatic knee osteoarthritis are defined by OAI as having both of the following in at least one knee at baseline: 1) frequent knee symptoms in the past 12 months, defined as "pain, aching or stiffness in our around the knee on most days" for at least one month during the past 12 months, and 2) radiographic tibiofemoral knee OA defined as Osteoarthritis Research Society International (OARSI) atlas grades 1 to 3⁵ which are approximately equivalent to Kellgren-Lawrence (KL) grades 2 or higher as measured on a fixed flexion radiograph. A complete study design protocol can be viewed at <http://www.oai.ucsf.edu/datarelease/docs/StudyDesignProtocol.pdf>

Study Sample

Subjects were recruited from four clinical sites: 1) the University of Maryland School of Medicine in Baltimore, Maryland, 2) the Ohio State University in Columbus, Ohio, 3) the University of Pittsburgh in Pittsburgh, Pennsylvania, and 4) Memorial Hospital of Rhode Island, in Pawtucket, Rhode Island.

Exclusion criteria as applied to the progression cohort were the presence of rheumatoid arthritis, bilateral knee arthroplasty or plans to undergo bilateral knee arthroplasty in the next 3 years, bilateral OARSI stage 3 knee OA, positive pregnancy test, inability to provide a blood sample, use of ambulatory aids other than a single straight cane for more than 50% of the time, co-morbid conditions that might interfere with 4 year participation, unlikely to reside in clinic area for at least 3 years, current participation in a double-blind randomized controlled trial, unwilling to sign informed consent. In addition, because an objective of the study was to track MRI based changes, persons were excluded if they were unable to undergo 3.0Tesla MRI. Accordingly, men weighing more than 130 kgs and women weighing more than 114kg were excluded. In total, 27% (N=4796) of those screened (N=17,457) were admitted to the study. Most common reasons for exclusion were because of drop outs after the initial telephone screen (N=3,321) and admission quotas being met for certain age and gender categories (N=2,954). A total of 2,295 persons were excluded because they exceeded the prescribed weight limits for the MRI.

A total of 278 (35.7%) of the 778 persons had symptomatic knee OA on the right side, 240 (30.8%) had symptomatic left sided knee OA and 260 (33.4%) had bilateral symptomatic knee OA.

Baseline Measurements

A multitude of measurements were taken at baseline on persons admitted to the OAI study. All measurements are thoroughly described in the Design Protocol (<http://www.oai.ucsf.edu/datarelease/docs/StudyDesignProtocol.pdf>). For this study, a subset of 31 measurements was examined. The approach to selecting measures for the current study was based on the premise that measures should be easily obtained so that surgeons could potentially obtain these measures routinely on their patients. Measurements representing demographic, socioeconomic, general health, arthritis specific health, physical examination, physical performance and self-report function and pain constructs were identified. Measurements were selected for study either because they have been shown to be associated with severity of knee OA^{6,7,8} or they have been shown to be associated with knee arthroplasty.^{1,2}

Demographic and socioeconomic variables such as age, gender, race and annual income were chosen to determine if they predicted future knee arthroplasty and to provide a thorough description of the sample. Persons undergoing knee arthroplasty tend to be 60 years of age or older, female^{28,29} and Caucasian⁹ so we were interested in determining if demographic variables predicted knee arthroplasty in this cohort. Validated general health measures of comorbidity,¹⁰ depression,¹¹ smoking status, BMI and Physical Activity Scale for the Elderly (PASE) were chosen.¹² Subjects completed a variety of arthritis-related health measures and the ones chosen for this study were the following: self reports indicating whether the subject was currently seeing a doctor for arthritis, past history of hip replacement surgery, whether the subject's mother, father sister or brother had knee replacement surgery. We chose these arthritis-related health measures because they appeared to have strong face validity for an association with future knee arthroplasty. The Pain (scored 0 to 20 with higher scores indicating worse function-related pain) and Function (scored 0 to 68 with higher scores indicating worse function) subscores of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)¹³ were included as were the SF-12 Mental Component Summary (MCS) and Physical Component Summary scores.¹⁴ In addition, data from the following question were used: "Considering all ways knee pain and arthritis affect you, how are you doing today?"

A modified KL knee arthritis grade for each knee as described by Altman and colleagues also was recorded.⁵ For purposes of this study, arthritic changes on radiographs were dichotomized so that each subject was coded as either having or not having severe arthritic changes, defined as definite osteophytes and joint space greatly impaired or bone on bone (OARSI atlas grade 3). We were particularly interested in determining whether end-stage knee OA was associated with future knee arthroplasty given the high priority that orthopaedic surgeons place on OA status for surgical decisions.¹⁵ A Grade 3 OARSI score approximates a KL grade of 4. All radiographs were taken using a fixed flexion protocol^{16,17} and all radiologists were required to complete a training protocol and achieve acceptable agreement prior to data collection.

A large number of physical examination procedures were conducted and only select procedures were examined in this study. Given that knee joint effusion is a common finding in patients with end stage knee arthritis,¹⁸ the Patellar Tap test and the Bulge Sign were included.¹⁹ Goniometric measurements of the extent of knee flexion contracture and coronal plane alignment were used as were assessments of the presence of pain with active knee flexion to end-range.

Physical performance measures were also included to attempt to capture the full depth and breadth of data that could potentially predict future knee arthroplasty. For this study, measures of whether a cane was used and time to complete a 400 meter walk,²⁰ the ability to complete five chair stands,²¹ and a 20 meter walk pace were included.²² These performance-based tests were chosen to capture a variety of tasks that are commonly performed and that require varying amounts of strength, balance and endurance.

Many measures in this study were person-level measures (e.g. 400 meter walk) and therefore each subject had only 1 score. When data were obtained on both knees (e.g. Patellar tap test) the score for the involved side was used for patients with unilateral knee OA. When subjects were coded as having bilateral involvement, the most severe score was used. We used a different strategy for the WOMAC Pain and Function scores. The OAI study required subjects to complete a WOMAC for the right knee and a WOMAC for the left knee. When subjects were coded as having unilateral symptomatic knee OA (66% of all cases), we used the score for the symptomatic side. In cases of bilateral symptomatic knee OA, we used the average of the right and left WOMAC scores because these scores were originally intended to represent function and pain at the person level.¹³

Data used in the preparation of this article were obtained from the Osteoarthritis Initiative (OAI) database, which is available for public access at <http://www.oai.ucsf.edu/>. Specific datasets used were versions 0.2.1, 1.1.2 and 3.1.1.

Statistical Methods

Sample size was dictated by the number of persons in the study with symptomatic knee osteoarthritis with 12 month (754 of 778 subjects) or 24 month (729 of 778 subjects) follow-up data. The total number of persons with a knee arthroplasty in the subsequent 2 years following baseline measurement was $n = 29$. The presence of knee arthroplasty during 12 and 24 month follow-up was determined by self report and the presence of a knee implant during follow-up radiography. Given the number of predictor variables ($n = 31$) being studied and this very low incidence rate of knee arthroplasty, the method of multivariate logistic regression would not be able to yield valid coefficient estimates.²³ Therefore, an exploratory analysis of the potential predictors was conducted using a nonparametric approach.

The classification and regression tree method (CART) has been a popular nonparametric classification approach that can simultaneously model all the predictor variables.²⁴ With this dataset, however, the sample sizes of the two classes (having knee arthroplasty or not) are highly imbalanced, and CART tends to produce biased classification results favoring the class with the majority of observations. The random forest method is a more robust alternative to CART, and has the advantage of producing multiple indexes of variable importance. We therefore used random forest to analyze the dataset to identify the variables that might be predictive of the occurrence of knee arthroplasty.²⁵

Random forest is a machine learning technique that combines the bootstrap aggregating (bagging) concept and the CART method.²⁶ Instead of predicting the sample classes with only one tree classifier, random forest method builds up B tree classifiers and predicts the sample classes by following the bagging procedure to combine the predicted classes resulted from all the trees in the forest. Each tree is grown to full size using a bootstrap resample data set obtained by sampling the original data with replacement. Moreover, each split in each tree is determined from a random selection of predictor variables at each node, thus the possible correlation between the trees is reduced. A continuous variable is internally treated as a categorical variable, where each observed value of the variable is viewed as a “category”. The predicted class for each observation is taken to be the class that is assigned to the observation most frequently in the forest.

Since each tree classifier in the forest is built using the bootstrap resamples from the original data set, about 1/3 of the original samples will not be resampled and hence are independent of this tree classifier. These samples are called “out-of-bag” observations and serve as an independent test set naturally. Two variable importance indexes were explored in this analysis. One index of variable importance is the “mean decrease in accuracy”. For each tree, the prediction accuracy (classification error rate) on the out-of-bag samples is recorded. Then for a given variable, the measurements of this variable are randomly permuted among all the out-of-bag observations. This tree classifier then is run down the permuted data, and the prediction accuracy on the out-of-bag samples is recorded. The decrease of prediction accuracy after the permutation of the variable is averaged over the B trees, and is a measure of the importance of the variable in classifying the samples. Specific to this analysis, since the sample size of the class of cases (having knee arthroplasty) is much smaller than that of controls (not having knee arthroplasty), the classification error rate would be dominated by the error rate in the class of controls. Therefore, rather than using the overall prediction accuracy, the prediction accuracy in the class of cases (i.e. class (of cases)- specific error rate) was used.

The second index of variable importance we used is “mean decrease in node impurity”. In each tree classifier, from each splitting on a variable, the node impurity will decrease. The decrement of impurity caused by each variable over all the B trees can be calculated, and is another measure of variable importance in classifying the samples. The Gini index was used to measure node impurity.²⁴

In this analysis, B=1000 trees were built in the forest, and 5 (out of the total 31) predictor variables were randomly selected to determine the split at each node in each tree. In addition, among the 778 persons in the sample, the class of cases only has 29 samples. This sample size imbalance made it problematic when bootstrap resampling from the original dataset for building up each tree, as it was very likely that few of the 29 cases could be selected into the bootstrap resamples. Therefore, stratified resampling was used. That is, for each tree, the bootstrap resampling was performed separately on each class of samples. Moreover, to make the sample sizes of the two classes more balanced, for each class, 29 bootstrap resamples were sampled with replacement from the original set of samples in the class. All the analysis was performed in the R 2.6.2 environment (<http://www.r-project.org/>).

Results

Incidence and Subject Characteristics

Of the 778 persons in the study, 29 (3.7%, 95% CI = 2.6%, 5.3%) had knee replacement surgery during the 2-year follow-up. For persons with symptomatic endstage knee arthritis (modified KL scores of 4) the incidence of knee arthroplasty at 2-year follow-up was 9.7% (22/227) with a CI=6.5%, 14.2%. Compared with persons who did not have knee arthroplasty, those who did were more likely to be older, female, a high school graduate or less, and currently seeing a doctor for arthritis. Patients with knee arthroplasty also were more likely to have a BMI of greater than 30, modified KL knee OA grade of 4, knee effusion, knee flexion contracture, a slower time on the 400 meter walk, a slower pace on the 20 meter walk, a higher (worse) WOMAC Pain and Function score, a lower (worse) SF-12 PCS score and a higher (better) SF-12 MCS score (see Table 1). None of the patients who had knee arthroplasty during the two-year follow-up had knee arthroplasty prior to study admission.

Several characteristics were similarly frequent among those with and without subsequent knee arthroplasty. For example, being African American or having an annual income of less than \$50,000 per year were not associated with knee arthroplasty. Persons with no health insurance were equally likely to have knee replacement surgery as those with health insurance though the number of persons in the study with no health insurance (n= 24) was very small.

Predictors of knee arthroplasty

The two indexes of variable importance, “mean decrease in accuracy” and “mean decrease in node impurity”, are listed in Table 2 for all variables. The larger the value, the more important the measure was in predicting future knee arthroplasty. Figure 1 graphically depicts the importance indexes for all variables. The X axis represents the “mean decrease in accuracy” index and the Y axis is the “decrease in node impurity” index. If a variable is important to the classification, both indexes should be large, and its corresponding point should be in the upper-right quadrant of the graph.

From Figure 1 it can be seen that the variable “grade4klforprogressioncohort” (modified KL knee OA grade of 4) has the largest “mean decrease in accuracy” and the second largest “mean decrease in node impurity”, and is noticeably distanced from all the other variables. This knee OA severity variable is therefore the most powerful classifying variable for predicting the occurrence of knee arthroplasty in our dataset. Variables “V00HSPSS” (SF-12 PCS), “NEWWOMFUN” (WOMAC function score), “V00HSMSS” (SF-12 MCS), “V00400MTIM” (time to complete 400m walk), and “V00K00SQOL” (KOOS QOL score) are also relatively important based on the indexes of variable importance. With the exception of age and depression, demographic and general health variables seem to be least influential in predicting the occurrence of knee surgery.

To provide a more user-friendly interpretation of the key findings, three of the more important predictor variables were used to describe the proportions of patients who had knee arthroplasty (see Table 3) for the various permutations of scores. We chose the knee OA status variable (grade4klforprogressioncohort), the SF-12 Physical Component Summary score (V00HSPSS), and the 400 meter walk time (V00400MTIM) because these variables were among the most important predictors and because they represented different constructs. Cutpoints for each of the variables shown in Table 3 were chosen using the following criteria. For the knee OA modified KL grade variable, we dichotomized the variables into those with and those without end-stage knee OA. For the other two variables, which were both continuous, we divided the data into tertiles using whole number cutpoints for each variable and reported knee arthroplasty incidence rates for each variable. As seen in Table 3, when a person had a grade 4 modified KL score, an SF-12 Physical Component Summary score of <41 and a 400 meter walk time of >334 seconds, the likelihood of a knee arthroplasty was 22% (11/50). When a person had a grade 4 modified KL score, an SF-12 score of >51 and a 400 meter walk time of less than <288 seconds, the likelihood of subsequent knee arthroplasty was 0% (0/30).

Discussion

The two-year incidence of knee arthroplasty among a large cohort of persons with symptomatic knee OA was found to be very low at 3.7% (95% CI=2.6%,5.3%) Hawker and colleagues followed 2,128 Canadian persons with presumed knee or hip arthritis for 6 years and 254 persons (11.9%) underwent knee or hip arthroplasty during follow-up. Momahara and colleagues followed 955 patients from Japan with RA and knee pain and 114 patients (11.9%) had knee arthroplasty during 5 years of follow-up.

Differences in knee arthroplasty frequency among these studies and the OAI study are likely due to a variety of factors. First, the extended follow-up in the Hawker et al and Momahara et al studies could explain the differences. However, sample differences among the three studies may also explain the differences in arthroplasty incidence rates. The OAI study required all persons to have symptomatic knee OA based on well defined self-report and radiographic findings. The studies by Hawker et al and Momahara and colleagues relied on patient self-report and did not obtain radiographs to confirm the presence of OA. In addition, the study of

Hawker and colleagues reported hip and knee OA and arthroplasty results in aggregate so the incidence of knee arthroplasty in isolation could not be determined.

This appears to be the first study to provide an incidence estimate for knee arthroplasty among at-risk persons in the US. The estimate is very low but the reasons for this low prevalence are purely speculative. It could be that patients in the cohort were generally higher functioning and had less pain than most patients undergoing knee arthroplasty. It could also be that patients in the cohort received alternative non-surgical interventions that delayed the need for surgical care. Finally, it may be that some patients were appropriate candidates for knee arthroplasty but did not undergo surgical treatment. For example, only 22 of 227 (9.7%) subjects with grade 4 modified KL knee arthritis scores underwent knee arthroplasty during the 2 year follow-up. However, these remaining 205 persons had mean WOMAC pain scores of 5.5 and mean WOMAC function scores of 18.1 at baseline. Typical WOMAC scores for patients scheduled for knee arthroplasty are in the 8 to 11 range for WOMAC pain and in the 30 to 40 range for WOMAC function.^{27,28,29} As a group, persons in the OAI study with end stage OA and no subsequent arthroplasty report appreciably less pain and enhanced function as compared to patients scheduled for knee arthroplasty. A total of 50 subjects with end stage OA but no knee arthroplasty had WOMAC pain scores of 8 or higher while 32 persons had WOMAC function scale scores of 30 or higher. These data suggest that some patients may have been appropriate candidates for knee arthroplasty but without additional data we could not determine if knee arthroplasty was underutilized in the cohort. In addition, the extent of underutilization, if any, cannot be determined until consensus is reached on optimal criteria for knee arthroplasty.^{28,30,31}

The strongest predictor for knee arthroplasty, as determined by the random forest analysis was knee OA status. Other important variables were physical functional status, as measured with the PCS-12, the WOMAC Function scale or KOOS QOL. Age, 400 meter walk time and the extent of knee flexion contracture also predicted arthroplasty status. These measures were also among the most powerful predictors in the univariate logistic regression analyses summarized in Table 1. The MCS-12 score was found to be important, but the scores for patients with and without subsequent arthroplasty both exceeded the mean for the population of US adults³² suggesting that this finding is not clinically relevant. The depression scores (CESD) also were different among the two groups, where patients who subsequently had knee arthroplasty had lower depression scores than patients who did not undergo knee arthroplasty. However, mean scores for both groups were in the range of 5 to 7 and CESD scores need to be 16 or greater to indicate the presence of depression.³³ Only 1 person (3.7%) with a subsequent arthroplasty had a CESD score indicating the presence of depression, while 12.2% of the sample of persons without knee arthroplasty had depression as measured by CESD scores. The role that depression plays in the process of knee arthroplasty surgical decision making warrants more study.

Most key predictors relate to the extent of persons' knee OA and disease consequences, namely functional status and knee specific impairment. These findings are consistent with the findings of Hawker and colleagues who also found that disability severity predicted knee arthroplasty. However, Hawker found that after adjusting for disability severity, the probability of undergoing knee arthroplasty was approximately 4 times greater for persons willing to undergo the procedure at baseline. Concerns about complications and prolonged recovery appear to be important deterrents to undergoing knee arthroplasty.³⁴ The OAI dataset did not measure willingness to undergo surgery, and therefore likely did not capture a key predictor of knee arthroplasty. A lack of willingness to undergo knee arthroplasty may explain why a substantial number of appropriate surgical candidates do not undergo knee arthroplasty.

In the current study the presence of end-stage knee OA was the most powerful predictor of subsequent knee arthroplasty. On the surface, this is not surprising. Osteoarthritic disease severity is one of the more commonly mentioned factors that guide surgical decisions.¹⁵ Given that a substantial number of persons with end stage disease did not undergo arthroplasty, other variables play a substantial role in the decision to have surgery. Physical health, as measured by the PCS-12, was the most powerful predictor of knee arthroplasty among the physical function and pain self report measures examined in this study, and suggests that when it comes to undergoing knee arthroplasty, functional loss plays a more important role than pain though more research is needed.

Based on this study, we would recommend that physicians strongly consider patients' knee arthritis status and physical function including gait when advising patients about the potential for future knee arthroplasty. Other measures such as 400 meter walk time and extent of knee flexion contracture also should be considered. Use of standardized physical function measures, such as the WOMAC or the SF-12, are recommended to aid in decision making and serve as a basis for comparison to large cohort studies of patients undergoing arthroplasty.^{28,29} Patients can be advised that these variables are among the most potent predictors for knee arthroplasty in the next 2 years. These variables combined with willingness to undergo the surgery are the most powerful evidence-based predictors of future knee replacement.

This study has several strengths but also some substantial limitations. The biggest strength is the total sample size, the number of predictors and the quality of data collection. This study appears to be the largest predictive cohort study of knee arthroplasty conducted in the US. Many variables were assessed and procedures for data collection were rigidly standardized. However, this study was not population-based, like the study by Hawker et al¹ which limits our ability to make inferences to the population of patients with knee OA. The most substantial limitation was the small number of persons who subsequently underwent knee arthroplasty. This small sample of surgical patients greatly limited the development and prevented the validation of a predictive model. The resulting approach was preliminary in nature. The OAI dataset will continue to grow with additional 2 year, 3 year and 4 year follow-up data to be added. This will provide an opportunity to validate the current model with larger numbers of subjects over additional years. Even given the limited number of surgical patients, this appears to be the largest and most complete US based examination of incidence of and predictors for subsequent knee arthroplasty.

Acknowledgments

Funding Acknowledgement: "The OAI is a public-private partnership comprised of five contracts funded by the National Institutes of Health. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc." The authors of the current paper are not part of the OAI investigative team.

References

1. Hawker GA, Guan J, Croxford R, Coyte PC, Glazier RH, Harvey BJ, Wright JG, Williams JI, Badley EM. A prospective population-based study of the predictors of undergoing total joint arthroplasty. *Arthritis Rheum* 2006;54(10):3212–3220. [PubMed: 17009255]
2. Momohara S, Inoue E, Ikari K, Kawamura K, Tsukahara S, Mochizuki T, Toki H, Miyawaki M, Saito S, Hara M, Kamatani N, Yamanaka H, Tomatsu T. Risk factors for total knee arthroplasty in rheumatoid arthritis. *Mod Rheumatol* 2007;17(6):476–480. [PubMed: 18084699]
3. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am* 2007;89(4):780–785. [PubMed: 17403800]
4. The Osteoarthritis Initiative. <http://www.oai.ucsf.edu/>

5. Altman R, Hochberg M, Murphy W, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. *Osteoarthritis Cartilage* 1995;3(A):3–70. [PubMed: 8581752]
6. Dieppe P, Cushnaghan J, Tucker M, Browning S, Shepstone L. The Bristol 'OA500 study': progression and impact of the disease after 8 years. *Osteoarthritis Cartilage* 2000;8(2):63–68. [PubMed: 10772234]
7. Mazuca SA, Brandt KD, Schauwecker DS, Katz BP, Meyer JM, Lane KA, Bradley JD, Hugenberg ST, Wolfe F, Moreland LW, Heck LW, Yocum DE, Schnitzer TJ, Sharma L, Manzi S, Oddis CV. Severity of joint pain and Kellgren-Lawrence grade at baseline are better predictors of joint space narrowing than bone scintigraphy in obese women with knee osteoarthritis. *J Rheumatol* 2005 Aug; 32(8):1540–1546. [PubMed: 16078332]
8. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA* 2001 Jul 11;286(2):188–195. [PubMed: 11448282]
9. Steel N, Clark A, Lang IA, Wallace RB, Melzer D. Racial disparities in receipt of hip and knee joint replacements are not explained by need: the Health and Retirement Study 1998–2004. *J Gerontol A Biol Sci Med Sci* 2008;63(6):629–634. [PubMed: 18559639]
10. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–383. [PubMed: 3558716]
11. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measurement* 1977;1:385–401.
12. Martin KA, Rejeski WJ, Miller ME, James MK, Ettinger WH, Messier SP. Validation of the PASE in older adults with knee pain and physical disability. *Med Sci Sports Exerc* 1999;31(5):627–633. [PubMed: 10331879]
13. 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. *Rheumatol* 1988 Dec;15(12):1833–1840.
14. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220–233. [PubMed: 8628042]
15. NIH Consensus Panel. NIH Consensus Statement on total knee replacement December 8–10, 2003. *J Bone Joint Surg Am* 2004;86-A(6):1328–1335. [PubMed: 15173310]
16. Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, Yu W, Genant HK. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. *Skeletal Radiol* 2003;32(3):128–132. [PubMed: 12605275]
17. Kothari M, Guermazi A, von Ingersleben G, Miaux Y, Sieffert M, Block JE, Stevens R, Peterfy CG. Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. *Eur Radiol* 2004;14(9):1568–1573. [PubMed: 15150666]
18. Peat G, Thomas E, Duncan R, Wood L, Wilkie R, Hill J, Hay EM, Croft P. Estimating the probability of radiographic osteoarthritis in the older patient with knee pain. *Arthritis Rheum* 2007;57(5):794–802. [PubMed: 17530679]
19. Cibere J, Bellamy N, Thorne A, Esdaile JM, McGorm KJ, Chalmers A, Huang S, Peloso P, Shojania K, Singer J, Wong H, Kopec J. Reliability of the knee examination in osteoarthritis: effect of standardization. *Arthritis Rheum* 2004 Feb;50(2):458–468. [PubMed: 14872488]
20. Simonsick EM, Montgomery PS, Newman AB, Bauer DC, Harris T. Measuring fitness in healthy older adults: the Health ABC Long Distance Corridor Walk. *J Am Geriatr Soc* 2001;49:1544–1548. [PubMed: 11890597]
21. Sharma L, Song J, Felson DT, Cahue S, Chamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA* 2001;286:188–195. [PubMed: 11448282]
22. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, Buckwalter J, Dougados M, Ehrlich G, Lequesne M, Lohmander S, Murphy WA Jr, Rosario-Jansen T, Schwartz B, Trippel S. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. *Osteoarthritis Cartilage* 1996;4(4):217–243. [PubMed: 11048620]

23. Peduzzi PN, Concato J, Kemper E, Holford TR, Feinstein A. A simulation study of the number of events per variables in logistic regression analysis. *J Clin Epidemiol* 1996;99:1373–1379. [PubMed: 8970487]
24. Breiman, L.; Friedman, JH.; Olshen, RA.; Stone, CJ. *Wadsworth Statistics/Probability Series*. Belmont, CA: Wadsworth Advanced Books and Software; 1984. *Classification and Regression Trees*.
25. Breiman L. Random forests. *Machine Learning* 2001;45:5–32.
26. Breiman L. Bagging predictors. *Machine Learning* 1996;26:123–140.
27. Stratford PW, Kennedy DM, Hanna SE. Condition-specific Western Ontario McMaster Osteoarthritis Index was not superior to region-specific Lower Extremity Functional Scale at detecting change. *J Clin Epidemiol* 2004 Oct;57(10):1025–1032. [PubMed: 15528053]
28. Quintana JM, Escobar A, Arostegui I, Bilbao A, Azkarate J, Goenaga JI, Arenaza JC. Health-related quality of life and appropriateness of knee or hip joint replacement. *Arch Intern Med* 2006;166(2): 220–226. [PubMed: 16432092]
29. Lingard EA, Sledge CB, Learmonth ID. Kinemax Outcomes Group. Patient expectations regarding total knee arthroplasty: differences among the United States, United Kingdom, and Australia. *J Bone Joint Surg Am* 2006 Jun;88(6):1201–1207. [PubMed: 16757751]
30. Gossec L, Hawker G, Davis AM, Maillefert JF, Lohmander LS, Altman R, Cibere J, Conaghan PG, Hochberg MC, Jordan JM, Katz JN, March L, Mahomed N, Pavelka K, Roos EM, Suarez-Almazor ME, Zanolli G, Dougados M. OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *J Rheumatol* 2007 Jun;34(6):1432–1435. [PubMed: 17552070]
31. Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga I, Lafuente I. Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med* 2008;168(14):1576–1584. [PubMed: 18663171]
32. Ware, JE.; Kosinski, M.; Keller, SD. *SE-12: How to Score the SF-12 Physical and Mental Health Summary Scales*. Boston, MA: The Health Institute, New England Medical Center; 1995.
33. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement* 1977;1:385–401.
34. Hamel MB, Toth M, Legedza A, Rosen MP. Joint replacement surgery in elderly patients with severe osteoarthritis of the hip or knee. *Arch Intern Med* 2008;168(13):1430–1440. [PubMed: 18625924]

Table 1

Characteristics of 778 persons with symptomatic knee osteoarthritis divided by the presence or absence of knee replacement surgery up to 2 years following admission to the study

Variables	Persons without knee replacement (n=749)	Persons with knee replacement (n=29)	Odds Ratio * (95% CI) p value
Demographic			
Age in years, mean (sd)	61.6 (9.3)	66.6 (7.5)	1.07 (1.02, 1.11) .005
Female	443 (59.1)	23 (79.3)	2.65 (1.07, 6.58) .036
African American	125 (16.7)	3 (10.3)	0.58 (0.17, 1.93) .36
High school grad or less	134 (18)	10 (34.5)	2.40 (1.09, 5.28) .03
Divorced or separated	123 (16.4)	1 (3.4)	0.18 (0.03, 1.35) .10
Annual Income < \$50,000	408 (56.4)	14 (53.8)	0.90 (0.41, 1.98) .80
No healthcare coverage	23 (3.1)	1 (3.4)	0.89 (0.12, 6.85) .91
General Health			
Comorbidity score	.46 (.85)	.48 (.83)	1.03 (0.68, 1.58) .89
Depression Scale (CESD)	7.6 (7.5)	5.4 (4.8)	0.95 (0.89, 1.02) .14
BMI >30	341 (45.5)	20 (69)	2.66 (1.20, 5.92) .02
Current cigarette smoker	47 (6.3)	1 (3.4)	0.53 (0.07, 3.96) .53
Arthritis Health			
Currently seeing a doctor for arthritis	147 (19.7)	13 (44.8)	3.32 (1.56, 7.05) .002
Had hip replacement surgery	16 (2.1)	0 (0)	-----
Mother, father sister or brother had knee replacement surgery	115 (15.5)	2 (6.9)	0.40 (0.10, 1.72) .22
Considering all ways knee pain and arthritis affects you, how are you doing today?	2.7 (2.2)	3.8 (2.0)	1.22 (1.05, 1.43) .01
Physical Examination			
Kellgren Lawrence knee OA grade of 4 for most involved side	205 (27.8)	22 (75.9)	8.16 (3.43, 19.38) <.001
Knee effusion bulge sign	226 (30.4)	15 (53.6)	2.53 (1.13, 5.66) .02
Knee effusion patellar tap	81 (10.9)	9 (31.0)	3.68 (1.62, 8.35) .002
Patient initiated knee flexion pain or tenderness	332 (44.7)	18 (62.1)	2.07 (0.96, 4.43) .06
Knee flexion contracture (degrees)	1.8 (4.3)	6.1 (7.5)	1.14 (1.07, 1.21) <.001
Coronal plane alignment (≥ 5 degrees of varus or valgus)	243 (32.6)	11 (37.9)	1.26 (0.59, 2.72) .55
Physical Performance			
Time to completion of 400 meter walk test (10 second increments)	317.4 (65.6)	379.4 (125.2)	1.07 (1.03, 1.10) <.001

Variables	Persons without knee replacement (n=749)	Persons with knee replacement (n=29)	Odds Ratio* (95% CI) p value
Use cane during 400 meter walk test	12 (1.6)	1 (3.6)	2.25 (0.28, 17.94) .44
20 meter walk pace (m/sec)	1.3 (.22)	1.2 (.23)	0.12 (0.02, .54) .006
Unable to complete 5 chair stands	694 (93.2)	24 (85.7)	2.27 (0.76, 6.79) .14
Self-report Disability and Pain			
PASE	158.4 (80.1)	132.3 (78.9)	1.00 (.99, 1.0) .09
WOMAC Pain score	5.0 (3.6)	7.2 (4.0)	1.15 (1.05, 1.26) .002
WOMAC Function score	16.0 (11.9)	23.8 (12.6)	1.05 (1.02, 1.08) .001
SF-12 PCS score	45.6 (9.8)	38.2 (10.3)	0.94 (0.90, 0.97) <.001
SF-12 MCS score	53.4 (9.0)	58.1 (6.8)	1.09 (1.03, 1.16) .005
KOOS QOL score	52.1 (19.0)	43.3 (18.0)	0.98 (.96, 1.00) .016

* Odds ratios were calculated using univariate logistic regression for each independent variable listed and with knee replacement surgery used as the dependent variable.

Table 2

Values for the two importance indexes used in the Random Forest analysis to describe the impact of each variable on classification

	Mean Decrease in accuracy	Mean Decrease in Node Impurity
Demographic		
V00AGE - Age (in years)	0.016	1.743
P02SEX – Male or female	0.002	0.337
Blackrace - African American (yes or no)	0.000	0.168
Highschoolorless - High school diploma or less (yes or no)	0.002	0.553
Sepordivorce - Separated or divorced (yes or no)	0.001	0.225
V00INCOME2 - Less than \$50,000 yearly income (yes or no)	-0.004	0.277
V00HLTHCOV - No healthcare coverage (yes or no)	0.000	0.034
General Health		
V00COMORB - Comorbidity score (continuous measure)	-0.001	0.271
V00CESD - CESD depression score (continuous measure)	0.003	1.319
BMIgr30 - BMI greater than 30 (yes or no)	0.002	0.511
Currentsmoke - Current smoker (yes or no)	0.000	0.087
Arthritis Health		
P01ARTDOC - Currently receiving medical care for arthritis (yes or no)	0.005	0.500
P01HRS – Ever had hip replacement surgery (yes or no)	0.000	0.033
P01FAMKR – Anyone in family had knee arthroplasty (yes or no)	-0.001	0.159
V00KGLRS – 0 to 10 global rating of effect of knee pain/arthritis	0.005	1.143
grade4kIforprogressioncohort – Grade 4 modified Kellgren-Lawrence (yes or no)	0.060	2.215
Physical Examination		
SXKEFFBS – Bulge sign for knee effusion (yes or no)	0.005	0.574
Patellartapyes – Patellar tap test for knee effusion (yes or no)	0.004	0.464
SXKFLEXPAIN – Pain with active knee flexion (yes or no)	0.004	0.358
Newcontractureprog – Knee flexion contracture (degrees)	0.013	1.850
Varvalnew – Knee varus valgus angle (degrees)	-0.001	1.139
Physical Performance		
V00400MTIM – Time for 400 meter walk test (sec)	0.014	1.981
V00CANEUSE – Used a cane during 400 meter walk (yes or no)	0.000	0.035
V0020MPACE – Velocity (meters/sec) during 20 meter walk test	0.005	1.644
V00CS5 – Able to complete 5 repeated chair stands (yes or no)	-0.002	0.182
Self-report Disability and Pain		
V00PASE – Physical Activity Scale for the Elderly (continuous)	0.000	1.519
NEWWOMPAIN – WOMAC Pain score (continuous)	0.007	1.538
NEWWOMFUN – WOMAC Function score (continuous)	0.022	1.957

	Mean Decrease in accuracy	Mean Decrease in Node Impurity
V00HSPSS – SF-12 Physical Component Summary score (continuous)	0.019	2.300
V00HSMSS – SF-12 Mental Component Summary score (continuous)	0.014	2.095
V00KOOSQOL – KOOS Quality of Life score (continuous)	0.021	1.764

Table 3

Summary of predictions for three key predictor variables.

Modified Kellgren- Lawrence Grade	Severe knee OA (Grade 4 modified KL score) 9.7% (22/227)						Non severe knee OA (Grade 2 or 3 modified KL score) 1.3% (7/539)											
	<41 16.9% (15/89)		41-51 7.3% (6/82)		>51 1.8% (1/56)		<288 2.0% (3/152)		288 to 334 0.6% (1/167)		>334 1.4% (3/216)							
PCS score from SF-12	<288	288 to 334	>334	<288	288 to 334	>334	<288	288 to 334	>334	<288	288 to 334	>334						
400 Meter walk time*	0/12 (0%)	4/24 (16.7%)	11/50 (22%)	2/17 (11.8%)	2/34 (5.9%)	2/31 (6.5%)	0/30 (0%)	0/10 (0%)	1/10 (10.0%)	0/34 (0%)	0/41 (0%)	2/73 (2.7%)	1/47 (2.1%)	0/69 (0%)	0/50 (0%)	2/107 (1.9%)	0/72 (0%)	1/37 (2.7%)

* One subject who underwent TKA during the subsequent 2 years had a grade 2 or 3 modified KL score at baseline and did not undergo a 400 meter walk test.