



OPEN ACCESS

Impact of total knee replacement practice: cost effectiveness analysis of data from the Osteoarthritis Initiative

Bart S Ferket,¹ Zachary Feldman,² Jing Zhou,¹ Edwin H Oei,³ Sita M A Bierma-Zeinstra,^{4,5} Madhu Mazumdar¹

¹Institute for Healthcare Delivery Science, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1077, New York, NY 10029, USA

²Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1077, New York, NY 10029, USA

³Department of Radiology, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA, Rotterdam, Netherlands

⁴Department of Orthopedics, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA, Rotterdam, Netherlands

⁵Department of General Practice, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA, Rotterdam, Netherlands

Correspondence to: B S Ferket bart.ferket@mountsinai.org

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2017;356:j1131 <http://dx.doi.org/10.1136/bmj.j1131>

Accepted: 13 February 2017

ABSTRACT OBJECTIVES

To evaluate the impact of total knee replacement on quality of life in people with knee osteoarthritis and to estimate associated differences in lifetime costs and quality adjusted life years (QALYs) according to use by level of symptoms.

DESIGN

Marginal structural modeling and cost effectiveness analysis based on lifetime predictions for total knee replacement and death from population based cohort data.

SETTING

Data from two studies—Osteoarthritis Initiative (OAI) and the Multicenter Osteoarthritis Study (MOST)—within the US health system.

PARTICIPANTS

4498 participants with or at high risk for knee osteoarthritis aged 45-79 from the OAI with no previous knee replacement (confirmed by baseline radiography) followed up for nine years. Validation cohort comprised 2907 patients from MOST with two year follow-up.

INTERVENTION

Scenarios ranging from current practice, defined as total knee replacement practice as performed in the OAI (with procedural rates estimated by a prediction model), to practice limited to patients with severe symptoms to no surgery.

MAIN OUTCOME MEASURES

Generic (SF-12) and osteoarthritis specific quality of life measured over 96 months, model based QALYs, costs, and incremental cost effectiveness ratios over a lifetime horizon.

RESULTS

In the OAI, total knee replacement showed improvements in quality of life with small absolute changes when averaged across levels of confounding variables: 1.70 (95% uncertainty interval 0.26 to 3.57) for SF-12 physical component summary (PCS); -10.69 (-13.39 to -8.01) for Western Ontario and McMaster Universities arthritis index (WOMAC); and 9.16 (6.35 to 12.49) for knee injury and osteoarthritis outcome score (KOOS) quality of life subscale. These improvements became larger with decreasing functional status at baseline. Provision of total knee replacement to patients with SF-12 PCS scores <35 was the optimal scenario given a cost effectiveness threshold of \$200 000/QALY, with cost savings of \$6974 (\$5789 to \$8269) and a minimal loss of 0.008 (-0.056 to 0.043) QALYs compared with current practice. These findings were reproduced among patients with knee osteoarthritis from the MOST cohort and were robust against various scenarios including increased rates of total knee replacement and mortality and inclusion of non-healthcare costs but were sensitive to increased deterioration in quality of life without surgery. In a threshold analysis, total knee replacement would become cost effective in patients with SF-12 PCS scores ≤40 if the associated hospital admission costs fell below \$14 000 given a cost effectiveness threshold of \$200 000/QALY.

CONCLUSION

Current practice of total knee replacement as performed in a recent US cohort of patients with knee osteoarthritis had minimal effects on quality of life and QALYs at the group level. If the procedure were restricted to more severely affected patients, its effectiveness would rise, with practice becoming economically more attractive than its current use.

Introduction

Osteoarthritis is a leading cause of disability worldwide,¹ resulting in pain, structural changes in the bone and joint space, and limitation of motion. Disease onset is gradual and usually begins after the age of 40. Osteoarthritis of the knee has a variable prognosis. Once present, improvement of joint structure is rare when assessed by radiography, but abatement of joint pain and disability occurs frequently.² About 12% of adults in the US are affected.³ The annual rate of total knee replacement in the US has doubled since 2000, especially in those aged 45-64.^{4,5} This disproportionate increase in this practice has been attributed to expansion of eligibility to people with less severe symptoms.⁶ The total number of procedures performed each year now exceeds 640 000, at a total annual cost of about \$10.2bn (£8.3bn, €9.6bn).⁵

WHAT IS ALREADY KNOWN ON THIS TOPIC

Rates of total knee replacement in the US have more than doubled since 2000, primarily because of expanding eligibility to patients with less symptomatic knee osteoarthritis

Up to a third of recipients of total knee replacement experience chronic pain postoperatively, and health benefits are assumed to be higher in those with poor preoperative physical functioning

WHAT THIS STUDY ADDS

Quality of life outcomes generally improve after total knee replacement, with small effects becoming larger with decreasing preoperative functional status

The practice of total knee replacement as performed in a recent US cohort of patients with knee osteoarthritis had minimal effects on QALYs at the group level and was found to be economically unattractive

Total knee replacement practice, however, could be considered cost effective if the procedure were restricted to patients with more severely affected functional status

The potential benefit of total knee replacement in patients with knee osteoarthritis should outweigh the associated costs. A recent randomized controlled trial looked at replacement compared with non-surgical management alone in 95 patients and showed large improvements in pain and physical limitations and significant increases in quality of life at 12 months.⁷ The trial population predominantly included patients with severe preoperative symptoms, as shown by low mean quality of life utility values at baseline.⁷ Many previously published uncontrolled before-after studies showed similarly large effects.^{8,9} In particular, the systematic review by Shan and colleagues described 19 studies that showed substantial improvements from baseline status, both in the intermediate and long term, for disease specific and generic health related quality of life across a broad range of domains.⁹ It is estimated, however, that up to a third of recipients of total knee replacement experience chronic pain postoperatively,^{10,11} and the health benefits of the procedure are assumed to be higher in those with poor physical functioning before surgery.¹²⁻¹⁴ This would imply that patients undergoing the procedure because of the recently expanded practice in the US might show less significant improvement in symptoms. Yet, the effectiveness of total knee replacement has been understudied in patients who are representative for the current practice population.^{10,11}

We used data from the Osteoarthritis Initiative (OAI) to estimate the effect of total knee replacement according to patients' functional status by looking at longitudinal health outcomes of patients with knee osteoarthritis with heterogeneous symptoms who underwent the procedure compared with those who did not.¹⁵ We subsequently performed a decision modeling study to evaluate the impact of the procedure on lifetime costs and quality adjusted life years (QALYs) while varying its use by level of symptoms.

Methods

Study populations

We obtained the data for our analysis from the Osteoarthritis Initiative (OAI) database, which is available for public access at <http://www.oai.ucsf.edu/>. The OAI is a multi-center cohort study of 4796 individuals with knee osteoarthritis or at risk for knee osteoarthritis who were recruited from the general population in 2005-06 across four US centers. Study participants were aged 45-79 at enrolment and were tracked with repeated follow-up evaluations for nine years. These evaluations included physical examinations, radiographs of both knees, and questionnaires on risk factors, symptoms, medical history, and quality of life. Knee osteoarthritis was defined as the patient having pain, aching, or stiffness in or around the knee on most days for at least one month during the past 12 months, and radiographically confirmed tibiofemoral osteophytes of grades 1-3 according to the Osteoarthritis Research Society International (OARSI) atlas.^{16,17} Patients eligible for the current analysis were those included in the outcomes dataset released 11 October 2015 (n=4796). To develop a decision model

for estimating lifetime outcomes, we excluded participants who had already undergone TKR at baseline, confirmed by radiography (n=62) and participants from the low osteoarthritis risk, "non-exposed" control group (n=122) who had no established risk factors, symptoms or radiographic findings of knee osteoarthritis.¹⁸ This resulted in a development sample of 4498 (table 1). OAI participants classified as having knee osteoarthritis at baseline (n=1327), as opposed to participants who were at high risk for knee osteoarthritis, were defined as the study population for estimation of the effect of total knee replacement on quality of life and use of non-surgical treatment for osteoarthritis pain, and for the base case cost effectiveness analysis. To validate the effect estimates, we repeated similar analyses with 30 months' follow-up data on 965 participants with knee osteoarthritis at baseline of the Multicenter Osteoarthritis Study (MOST).²³ To show generalizability of the base case cost effectiveness analysis, we performed a scenario analysis using the 965 MOST patients.

Modeling effect of total knee replacement on quality of life and use of non-surgical treatment

Outcomes were defined as the SF-12 physical component summary (PCS) score, the SF-12 mental component summary (MCS) score, the SF-6D utility index, the Western Ontario and McMaster Universities arthritis index (WOMAC), the quality of life subscale on the knee injury and osteoarthritis outcome score (KOOS), and self reported use of pain medication for osteoarthritis, all measured at 12, 24, 36, 48, 72, and 96 months. We evaluated the effect of total knee replacement on use of non-pharmacological treatments with measurements at 24, 48, and 96 months, as questions on these treatments were not included at the other study visits. The SF-12 instrument is a single page questionnaire measuring generic quality of life.²⁴ To estimate this, we calculated the SF-6D utility index, which can be directly derived from SF-12 by using a previously published algorithm.²⁵ The KOOS and WOMAC instruments are validated questionnaires measuring quality of life, pain, stiffness, and functionality specific for osteoarthritis.^{24,26,27} We chose to use only the KOOS quality of life subscale, which measures knee related quality of life and mental and social aspects such as awareness and lifestyle changes. These items are not well captured by the WOMAC total score, which focuses on knee symptoms and functioning. The Pearson product moment correlation coefficient for the two scores in the 1327 OAI participants with knee osteoarthritis was -0.67. Osteoarthritis pain medication included acetaminophen (paracetamol), non-steroidal anti-inflammatory drugs (NSAIDs), and cyclo-oxygenase-2 (COX-2) inhibitors. Non-pharmacological treatments included massage, chiropractic, acupuncture treatments, and other less commonly used complementary treatment options such as acupressure, chelation therapy, folk medicine, and homeopathic treatment.

To estimate the effect of total knee replacement on these longitudinally measured outcomes compared with no/delayed procedure, we used marginal

Table 1 | Characteristics of 4498 participants aged 45-79 from the Osteoarthritis Initiative (OAI). Figures are medians (interquartile range) for continuous variable and numbers (percentage) for categorical variables

Variables	High risk cohort (n=3171)	Knee osteoarthritis cohort (n=1327)	P value
Age (years)	61 (53-69)	61 (54-69)	0.77
Age group (years):			
<55	934 (29)	378 (28)	0.76
55-64	1033 (33)	432 (33)	
>64	1204 (38)	517 (39)	
Men	1302 (41)	577 (43)	0.14
Ethnicity:			
White	2626 (83)	942 (71)	<0.001
Black or African-American	466 (15)	344 (26)	
Asian	30 (1)	11 (1)	
Other	45 (1)	30 (2)	
Missing	4 (0)	0 (0)	
Annual income level (\$):			
<10 000	84 (3)	64 (5)	<0.001
10 000-<25 000	276 (9)	148 (11)	
25 000-<50 000	741(23)	337 (25)	
50 000-<100 000	1111 (35)	407 (31)	
>100 000	740 (23)	259 (20)	
Missing	219 (7)	112 (8)	
Education level:			
Less than high school graduate	84 (3)	65 (5)	<0.001
High school graduate	371 (12)	201 (15)	
Some college	719 (23)	363 (27)	
College graduate	686 (22)	256 (20)	
Some graduate school	263 (8)	106 (8)	
Graduate degree	1026 (32)	319 (24)	
Missing	22 (1)	17 (1)	
BMI	27.8 (24.6-31.1)	29.7 (26.6-33.4)	<0.001
Missing	1 (0)	3 (0)	
Modified Charlson comorbidity index:			
0	2431 (77)	914 (69)	<0.001
1	453 (14)	234 (18)	
2	172 (5)	99 (7)	
3	46 (1)	40 (3)	
≥4	38 (1)	13 (1)	
Missing	31 (1)	27 (2)	
Kellgren-Lawrence grade:			
0	1076 (34)	69 (5)	<0.001
1	579 (18)	90 (7)	
2	923 (29)	408 (31)	
3	393 (12)	481 (36)	
4	81 (3)	211 (16)	
Missing	119 (4)	68 (5)	
Self reported diagnosis of osteoarthritis in 12 months before baseline:			
Yes	472 (15)	542 (41)	<0.001
No	2568 (81)	689 (52)	
Missing	131 (4)	96 (7)	
Previous knee injury at baseline:			
Yes	1275 (40)	625 (47)	<0.001
No	1878 (59)	687 (52)	
Missing	18 (1)	15 (1)	
Previous knee surgery at baseline:			
Yes	549 (17)	450 (34)	<0.001
No	2619 (83)	875 (66)	
Missing	3 (0)	2 (0)	
SF-12* score:			
Physical component summary (PCS)	52.8 (46.4-56.1)	46.2 (38.0-52.9)	<0.001
Mental component summary (MCS)	55.3 (50.4-58.3)	55.7 (48.8-59.7)	0.11
Missing	36 (1)	21 (2)	—

(Continued)

Table 1 | Characteristics of 4498 participants aged 45-79 from the Osteoarthritis Initiative (OAI). Figures are medians (interquartile range) for continuous variable and numbers (percentage) for categorical variables

Variables	High risk cohort (n=3171)	Knee osteoarthritis cohort (n=1327)	P value
SF-6D* utility index	0.86 (0.72-0.92)	0.78 (0.66-0.86)	<0.001
Missing	36 (1)	23 (2)	—
WOMAC* total score	6.1 (1.0-17.0)	24.1 (13.0-39.0)	<0.001
Missing	14 (0)	19 (1)	—
KOOS* quality of life	75.0 (62.5-87.5)	50.0 (37.5-62.5)	<0.001
Missing	1 (0)	0 (0)	—
Use of pain medication for osteoarthritis:			
Yes	1016 (32)	620 (47)	
No	2154 (68)	706 (53)	<0.001
Missing	1 (0)	1 (0)	
Use of non-pharmacological treatment:			
Yes	284 (9)	169 (13)	
No	2883 (91)	1156 (87)	<0.001
Missing	4 (0)	2 (0)	
Missed work days in last 3 months [†] :			
Yes	64 (3)	69 (8)	
No	1928 (97)	750 (92)	<0.001
No of work days missed if ≥1	2.5 (2.0-5.2)	3.0 (2.0-5.0)	0.5

BMI=body mass index; WOMAC=Western Ontario and McMaster Universities arthritis index; KOOS=knee injury and osteoarthritis outcome score.
 *Range of scales: SF-6D, 0-1 scale (higher scores indicate better health); SF-12, 0-100 scale (higher scores indicate less severe symptoms); WOMAC, 0-100 scale (higher scores indicate more severe symptoms); KOOS quality of life, 0-100 scale (higher scores indicate less severe symptoms).
 †Measured only in those who reported being employed: n=1992 in high risk cohort and 819 in knee osteoarthritis cohort.

structural models for repeated measures defined as weighted generalized estimating equations (GEEs) with each outcome as the dependent variable.²⁸ Marginal structural models are warranted when outcome values can vary over time and can predict future treatment assignment along with other time varying confounders. For example, an increase in use of osteoarthritis pain medication could be associated with a higher likelihood of receiving total knee replacement. Each GEE included a treatment variable for the procedure, which was set to one after performance, a study visit indicator, the outcome's baseline value, and other baseline variables including age, sex, race, income, education, knee injury in medical history, knee surgery in medical history, body mass index (BMI), Charlson comorbidity score,²⁹ use of osteoarthritis pain medication, doctor's diagnosis of knee osteoarthritis, Kellgren-Lawrence radiographic grade, SF-12 scores, WOMAC total score, and KOOS quality of life. To evaluate the effectiveness of total knee replacement according to preoperative physical functioning, we included an interaction term for the procedure with baseline SF-12 PCS.³⁰⁻³¹ Within these GEEs, we included weights for treatment propensity (that is, the likelihood of having received total knee replacement) for each study visit. Weights were estimated by logistic regression models pooled for study visits with the above mentioned baseline variables, study visit, longitudinally measured BMI, Charlson comorbidity score, doctor's diagnosis of knee osteoarthritis, Kellgren-Lawrence radiographic grade, and all outcome variables. The main treatment effect estimate obtained from the marginal structural modeling should be interpreted as a "time averaged" causal effect.²⁸

Missing values were imputed 20 times with a flexible additive model including status variables and the

Nelson-Aalen estimator of the cumulative hazard for total knee replacement and death. To estimate parameter uncertainty, we re-fitted imputation, pooled logistic regression, and GEE models in 500 bootstrap datasets. We used the 2.5th and 97.5th centiles of bootstrap effect estimates for uncertainty interval limits. To validate the effect estimates from models developed with OAI patients, we performed multivariable adjusted analyses for SF-12 scores, SF-6D utility index, WOMAC, and use of osteoarthritis pain medication using 30 months' follow-up data on 965 MOST participants with knee osteoarthritis at baseline. All statistical analyses were performed with R version 3.2.0 (2015, R Foundation for Statistical Computing). For more information on these statistical analyses see appendix 1.

Modeling of lifetime outcomes and cost effectiveness of total knee replacement

We developed the KOSMOS (Knee Osteoarthritis Micro-Simulation) model to simulate the virtual life course of 1327 OAI patients by modeling the occurrence of primary total knee replacement, revision procedure, and death up to age 100. Rates of primary total knee replacement and death were modeled by cause specific multivariable Cox regression with chronological age as time scale. Revision rates were based on the literature¹⁹ using the log cumulative hazard of revision procedure as reported for different age groups, which was modeled as a linear function of log time since the primary procedure. Health related SF-6D utility scores, use of osteoarthritis pain medication, and use of non-pharmacological treatment were based on the patients' baseline and predicted 96 month values taken from the output of the GEEs with and without total knee replacement. We used linear interpolation to calculate patients' values through eight years. We estimated SF-12 PCS,

SF-12 MCS, and SF-6D scores for patients alive longer than eight years by linear extrapolation, based on the observed steady changes over time (figs A-C in appendix 3). Use of osteoarthritis pain medication and non-pharmacological treatments was assumed to remain stable after eight years (figs F and G in appendix 3), and the predicted 96 month probability of use was carried forward.

For simulated patients who survived each cycle of the model, we calculated an undiscounted QALY as the predicted SF-6D score multiplied by one year, thus resulting in a different QALY outcome for patients receiving total knee replacement in that cycle. Patients could accrue QALYs in the model until death or age 100. We calculated effect modification of procedure by SF-12 PCS scores using the latest predicted score, which was updated every eighth year until total knee replacement. We assumed that with revision procedure a beneficial effect on SF-12 PCS, use of osteoarthritis pain medication, and non-pharmacological treatment would be cancelled out by deterioration and improvement before and after the revision.^{32,33}

Costs associated with care of knee osteoarthritis and total knee replacement procedures were estimated from a US health system perspective (table 2) and were either applied for each model cycle (annual costs associated with pharmacological and non-pharmacological treatment, physician office visits, and imaging) or as a one off cost penalty (costs associated with primary and revision procedures and rehabilitation including physiotherapy). All costs were expressed in 2013 \$, with inflation rates reported in the US healthcare consumer price index.³⁴ We recalibrated the average life expectancy predicted by the KOSMOS model for the modeled patient cohort to reflect the average life expectancy as predicted by age and sex specific US 2011 life tables and validated the KOSMOS model's predictive performance in OAI and MOST data. More details on the development and validity of the KOSMOS model are provided in appendix 1.

For the base case analysis, we modeled 10 scenarios, ranging from current practice with rates as observed in the OAI, to lower rates of practice in which the procedure was performed only in individuals with lower SF-12 PCS levels (from <55-<20), to a scenario without total knee replacement. In the restricted scenarios, the underlying annual rate was kept the same as modeled for the current practice scenario, but no effects on quality of life, use of non-surgical treatment, and costs, and no procedural mortality rates were incorporated if the preceding SF-12 PCS level was greater than or equal to the threshold value.

For each scenario, we performed microsimulation of 1327 individuals for each set of 500 bootstrap equations and calculated the accrued QALYs and costs using the recommended 3% discount rate.³⁵ We calculated 95% uncertainty intervals by using the 2.5th and 97.5th centiles of 500 modeled outcomes, each averaged across 1327 individuals. We calculated incremental cost effectiveness ratios (ICERs), defined as the difference in average costs divided by the difference in average QALYs,

after ranking scenarios according to increasing costs and exclusion of dominated scenarios. Dominated scenarios were defined as programs less effective and more costly than the previous program (absolutely dominated) and programs with a larger incremental cost effectiveness ratio than the next not dominated program (extendedly dominated). We considered cost effectiveness thresholds of \$50 000, \$100 000, and \$200 000 per QALY for decision making.³⁶

Sensitivity analyses

In threshold analyses, we varied the hospital admission costs of primary and revision total knee replacement and used various percentages up to 100% for a further decline in quality of life (SF-12 PCS, SF-12 MCS, and SF-6D) in patients who were simulated to receive the procedure in the current practice scenario but were modeled to not receive the procedure in the other scenarios. For modeling the additional decline, we multiplied the main effect of follow-up time by values from 1 to 2. We used a time lag of four years after total knee replacement (equal to the difference in last follow-up measurement and median time to the procedure), so that the further decrease would reflect a long term effect of total knee replacement that we could have missed using OAI data.

We performed a scenario analysis using the 965 MOST patients after recalibration of hazard rates for total knee replacement and an analysis using EuroQol (EQ)-5d utility values as a substitute for SF-6D values by conversion of SF-12 MCS and PCS scores by a published equation.³⁷⁻³⁹ In additional analyses, we modeled an increased rate of primary total knee replacement up to 30% to investigate the impact of a further increase in procedure rates as observed in the US after 2006.⁵ We also increased the background mortality rate based on findings that patients with osteoarthritis might have an excess all cause mortality compared with the general population by multiplying mortality rates by standardized mortality ratios sampled from a lognormal distribution (mean 1.55, 95% confidence interval 1.41 to 1.70).⁴⁰ Finally, we performed an analysis exploring the potential influence of non-healthcare costs and loss of productivity not captured by a decrease in health utility,⁴¹ through inclusion of costs due to work time lost by patients and cost of informal caregiving. More details on these sensitivity analyses are provided in appendix 1.

Patient involvement

No patients were involved in developing plans for recruitment, design, or implementation of the studies, nor were they involved in developing the research questions and outcome measures. No patients were asked to advice on interpretation or writing up of results. Plans are in place for dissemination of the results of the research to the patient community. These plans include providing manuscript summaries to media sources, osteoarthritis charities, provider bodies, and patient organizations, in addition to social media announcements and institutional provision of pamphlets in health system waiting areas.

Table 2 | Input parameters for the Knee Osteoarthritis MicroSimulation (KOSMOS) model

Parameters*	Data (95% CI)	Source
Annual TKR incidence	Individualized	OAI
Annual background mortality	Individualized	OAI
Annual revision probability	Age and time dependent	NJR, 2014 ¹⁹
Hazard ratio for re-revision	1.74 (1.57 to 1.91)	Ong, 2010 ²⁰
Proportion bilateral TKR	13% (9% to 17%)	OAI
Procedural mortality with TKR at age <85	0.06% (0.04% to 0.08%)	HCUP ⁵
Procedural mortality with TKR at age ≥85	0.39% (0.20% to 0.64%)	HCUP ⁵
Annual probability of taking osteoarthritis pain medication	Individualized	OAI
Proportion of each drug type if taking osteoarthritis pain medication:		
Prescription NSAIDs	0.19	OAI
Non-prescription NSAIDs	0.55	OAI
Celecoxib	0.22	OAI
Acetaminophen	0.31	OAI
Annual probability using non-pharmacological treatment	Individualized	OAI
Proportion of each non-pharmacological treatment type if used:		
Acupuncture	0.05	OAI
Chiropractic	0.40	OAI
Massage	0.21	OAI
Other	0.34	OAI
No of annual visits for each non-pharmacological treatment type if used:		
Acupuncture	7 (3 to 13)	OAI
Chiropractic	9 (3 to 15)	OAI
Massage	7 (3 to 13)	OAI
Other	8 (3 to 14)	OAI
Prescription NSAIDs:		
Diclofenac	813	Losina, 2015 ²¹
Ibuprofen	69	Losina, 2015 ²¹
Meloxicam	467	Losina, 2015 ²¹
Nabumetone	283	Losina, 2015 ²¹
Naproxen	470	Losina, 2015 ²¹
Weighted total	460	OAI
Non-prescription NSAIDs:		
Ibuprofen	149	Losina, 2015 ²¹
Naproxen	99	Losina, 2015 ²¹
Weighted total	124	OAI
Celecoxib	3047	Losina, 2015 ²¹
Acetaminophen	71	Losina, 2015 ²¹
Non-pharmacological treatment:		
Acupuncture	725	Gore, 2012, ²² OAI
Chiropractic	494	Gore, 2012, ²² OAI
Massage	189	Gore, 2012, ²² OAI
Other	473	Gore, 2012, ²² OAI
Physician office visits	66	Losina, 2015 ²¹
Radiographic imaging	29	Losina, 2015 ²¹
Procedural costs (in 2013 \$)		
Primary TKR costs:		
Hospital costs	16 051 (15 771 to 16 331)	HCUP
Surgeon fees	1582 (1022 to 2266)	Losina, 2015 ²¹
Anesthesiologist fees	427 (276 to 612)	Losina, 2015 ²¹
Rehabilitation costs including physiotherapy	7764 (5015 to 11 121)	Losina, 2015 ²¹
Revision TKR costs:		
Hospital costs	21 014 (19 369 to 22 696)	HCUP ⁵
Surgeon fees	1812 (1170 to 2595)	Losina, 2015 ²¹
Anesthesiologist fees	494 (319 to 708)	Losina, 2015 ²¹
Rehabilitation costs including physiotherapy	7764 (5015 to 11 121)	Losina, 2015 ²¹
Quality of life utility values:		
SF-6D utility index	Individualized	OAI
One off quality of life penalty of primary TKR	−0.008 (−0.014 to −0.005)	HCUP ⁵
One off quality of life penalty of revision TKR	−0.011 (−0.044 to −0.002)	HCUP ⁵
TKR=total knee replacement.		
*β distributions used for probabilities, γ distributions for costs, Poisson distribution for counts, and lognormal distributions for one off quality of life penalties.		

Table 3 | Changes in quality of life measures and use of non-surgical treatment after total knee replacement (TKR) in four models*. Figures are effect estimates with 95% uncertainty intervals based on refitting all modeling steps in 500 bootstrap datasets given for 1327 Osteoarthritis Initiative (OAI) participants with knee osteoarthritis at baseline who were repeatedly followed up until 96 months v 965 Multicenter Osteoarthritis Study (MOST) participants with knee osteoarthritis at baseline who were followed up until 30 months

Outcome†	Model 1: unadjusted			Model 2: adjusted for baseline covariables			Model 3: adjusted for baseline and time varying covariables (OAI)			Model 4: including interaction of TKR×SF-12 PCS		
	OAI	MOST	OAI	OAI	MOST	MOST	OAI	MOST	OAI	MOST	MOST	
SF-12:												
PCS	0.76 (−0.65 to 1.76)	0.62 (−1.67 to 2.54)	1.29 (−0.09 to 2.23)	1.00 (−1.32 to 2.99)	1.70 (0.26 to 3.57)	7.12 (−0.23 to 15.77)+SF-12 PC SX−0.234 (−0.316 to 0.032)	9.95 (4.52 to 9.11)+SF-12 PC SX−0.234 (−0.482 to −0.114)					
MCS	−0.49 (−1.26 to 0.63)	2.17 (0.45 to 4.19)	−0.20 (−0.92 to 0.87)	2.29 (0.62 to 4.33)	−0.22 (−1.49 to 1.31)	−0.13 (−7.63 to 5.39)+SF-12 PC SX−0.002 (−0.122 to 0.171)	2.51 (−3.51 to 8.16)+SF-12 PC SX−0.006 (−0.148 to 0.156)					
SF-6D utility	−0.005 (−0.019 to 0.009)	0.012 (−0.008 to 0.031)	0.004 (−0.009 to 0.175)	0.015 (−0.005 to 0.036)	0.008 (−0.008 to 0.028)	0.060 (−0.003 to 0.116) + SF-12 PC SX−0.001 (−0.002 to 0.000)	0.098 (0.033 to 0.188)+SF-12 PC SX−0.002 (−0.004 to −0.001)					
WOMAC total	−9.25 (−11.00 to −6.89)	−10.33 (−13.95 to −7.41)	−9.36 (−11.08 to −7.26)	−11.02 (−14.38 to −7.94)	−10.69 (−13.39 to −8.01)	−18.35 (−26.50 to −4.12)+SF-12 PC SX0.181 (−0.102 to 0.399)	−14.89 (−30.45 to −7.39)+SF-12 PC SX0.101 (−0.092 to 0.510)					
KOOS quality of life	6.56 (3.83 to 8.43)	—	7.57 (4.83 to 9.70)	—	9.16 (6.35 to 12.49)	19.90 (6.67 to 32.91)+SF-12 PC SX−0.253 (−0.549 to 0.029)	—					
OR use of OA pain medication	1.35 (1.00 to 1.74)	1.20 (0.54 to 3.09)	1.19 (0.89 to 1.54)	1.11 (0.48 to 3.13)	0.81 (0.55 to 1.12)	1.30 (0.39 to 5.00)×SF-12 PC SX0.989 (0.957 to 1.020)	2.18 (0.14 to 50.37)×SF-12 PC SX0.983 (0.908 to 1.068)					
OR use of non-pharmacological OA treatment	0.79 (0.53 to 1.27)	—	0.80 (0.53 to 1.30)	—	0.91 (0.55 to 1.77)	2.89 (0.34 to 40.61)×SF-12 PC SX0.972 (0.915 to 1.017)	—					

PCS=physical component summary; MCS=mental component summary; WOMAC=Western Ontario and McMaster Universities arthritis index; KOOS=knee injury and osteoarthritis outcome score; OR=odds ratio; OA=osteoarthritis.

*Model 1 (unadjusted) consisted of GEEs including only TKR, visit, and baseline value of outcome as covariables; Model 2 (multivariable adjustment): GEEs extended with SF-12 PCS, age, male, African-American ethnicity, income, education, history of knee injury, history of knee surgery, BMI, Charlson comorbidity index, use of osteoarthritis pain medication, self reported diagnosis of knee osteoarthritis, Kelgren-Lawrence grade, WOMAC total, and KOOS quality of life, all measured at baseline. In MOST analyses, income and diagnosis of osteoarthritis were not included because of unavailability; Model 3: multivariable adjusted GEEs were weighted for time varying propensities of undergoing TKR. Model 3 not included for MOST because of lack of multiple periodic visits; Model 4: based on extension of model 3 for OAI and model 2 for MOST.

†Range of scales: SF-6D, 0–1 scale (higher scores indicate better health); SF-12, 0–100 scale (higher scores indicate less severe symptoms); WOMAC, 0–100 scale (higher scores indicate less severe symptoms); KOOS quality of life, 0–100 scale (higher scores indicate less severe symptoms).

Results

Study populations

The 1327 OAI participants with knee osteoarthritis at baseline had worse SF-12 PCS, SF-6D, and osteoarthritis specific quality of life scores than the 3171 at high risk for knee osteoarthritis ($P < 0.001$, table 1). There were 382 total knee replacements, of which 319 were done before the 96 month visit, and 255 were in the 1327 with knee osteoarthritis at baseline. In the MOST cohort (table A, appendix 2), 135 total knee replacements were performed in 965 participants with knee osteoarthritis at baseline, and 16 were performed in 1719 individuals at high risk of knee osteoarthritis, all before the 30 month visit.

Effect of total knee replacement on quality of life and use of non-surgical treatment

Figures A-G in appendix 3 show time trends of SF-12 PCS, SF-12 MCS, SF-6D, WOMAC, KOOS quality of life, use of osteoarthritis pain medication, and non-pharmacological treatments, specified for those who did and did not undergo total knee replacement. After adjustment for baseline and time varying confounders, the main effects of total knee replacement (that is, treatment effects averaged across confounding variable levels and follow-up time) comprised an absolute improvement of 1.70 points (95% uncertainty interval 0.26 to 3.57) on SF-12 PCS, and changes in SF-12 MCS of -0.22 (-1.49 to 1.31) and SF-6D of 0.008 (-0.008 to 0.028) point. For each unit decrease in baseline SF-12 PCS, the effect on SF-6D increased and could be calculated as $0.098 - 0.002 \times (\text{SF-12 PCS})$, suggesting that total knee replacement would become more effective if it was restricted to patients with SF-12 PCS scores < 50 . For osteoarthritis specific measures of quality of life, the procedure's main effects included improvement of the WOMAC score by 10.69 (8.01 to 13.39) and KOOS quality of life of 9.16 (6.35 to 12.49) points. Total knee replacement reduced the odds of use of medication for osteoarthritis pain, but this effect was uncertain, with an odds ratio of 0.81 (0.55 to 1.12). Use of non-pharmacological treatment did not significantly seem to change with total knee replacement (0.91 , 0.55 to 1.77). These effects were generally consistent with those obtained from multivariable adjusted analyses of MOST data, although in MOST the effect on SF-12 MCS was positive in contrast with the effect in OAI (table 3).

Lifetime outcomes and cost effectiveness of total knee replacement practice

In the current practice scenario, the lifetime likelihood of undergoing total knee replacement, as predicted for OAI, was 39.9% (95% uncertainty interval 34.5 to 45.3), and the average undiscounted life expectancy was 22.39 years (21.13 to 23.85 years). Modeled quality of life values and proportions of use of non-surgical treatments over time were generally most favorable in the current practice scenario, except for SF-12 MCS scores (figs 1-M, appendix 3). The mean discounted QALYs in the current practice scenario were 11.18 (10.66 to 11.70) and costs were $\$17\,168$ ($\$15\,307$ to $\$19\,124$).

Table 4 | Lifetime cost effectiveness outcomes for different scenarios for determining which patients are eligible for undergoing total knee replacement (TKR) with 95% uncertainty intervals based on 500 bootstrap datasets for simulations of 1327 participants from the Osteoarthritis Initiative (OAI) with knee osteoarthritis at baseline

TKR scenarios ranked according to increasing costs	Lifetime TKR likelihood (%)	Costs (\$)	QALYs	Incremental costs*	Incremental QALYs*	ICER (\$/QALY)*	% most cost effective by cost/QALY threshold (\$1000/QALY)	
							100	200
No TKR	0	7939 (7162 to 8793)	11.155 (10.634 to 11.686)	—	—	—	41.6	12.4
If SF-12 PCS <20	1.2 (0.6 to 1.9)	8181 (7390 to 9041)	11.157 (10.636 to 11.692)	242 (114 to 399)	0.003 (0.000 to 0.007)	88903	15.2	4.0
If SF-12 PCS <25	2.7 (1.7 to 3.8)	8489 (7633 to 9431)	11.160 (10.639 to 11.698)	307 (175 to 486)	0.003 (0.000 to 0.007)	108773	14.0	6.4
If SF-12 PCS <30	5.7 (4.0 to 7.2)	9159 (8298 to 10118)	11.166 (10.643 to 11.705)	671 (424 to 943)	0.005 (0.000 to 0.012)	126762	16.0	12.4
If SF-12 PCS <35	10.2 (8.1 to 12.4)	10194 (9164 to 11287)	11.172 (10.654 to 11.709)	1035 (732 to 1396)	0.006 (0.000 to 0.014)	160974	8.6	18.4
If SF-12 PCS <40	16.5 (13.4 to 19.6)	11649 (10386 to 12795)	11.179 (10.667 to 11.714)	1455 (1055 to 1895)	0.007 (−0.002 to 0.017)	217615	2.6	24.2
If SF-12 PCS <45	23.5 (19.8 to 27.1)	13193 (11734 to 14684)	11.183 (10.672 to 11.718)	1544 (1161 to 2007)	0.004 (−0.006 to 0.014)	371439	1.4	9.2
If SF-12 PCS <50	31.3 (26.8 to 36.0)	15022 (13384 to 16590)	11.184 (10.668 to 11.720)	1829 (1394 to 2305)	0.002 (−0.012 to 0.015)	1109675	0.4	4.2
If SF-12 PCS <55	37.3 (32.2 to 42.4)	16483 (14782 to 18361)	11.183 (10.665 to 11.710)	1461 (1046 to 1960)	−0.002 (−0.016 to 0.011)	Absolute dominance	0.0	1.6
Current practice	39.9 (34.5 to 45.3)	17168 (15307 to 19124)	11.180 (10.662 to 11.700)	2147 (1589 to 2776)	−0.004 (−0.027 to 0.017)	Absolute dominance	0.2	7.2

ICER=incremental cost effectiveness ratio; PCS=physical component summary.

*Calculated by comparison with preceding undominated scenario.

In the base case cost effectiveness analysis, only the ICER of total knee replacement for those with SF-12 PCS <20 fell below \$100 000/QALY. The optimal scenario given a cost effectiveness threshold of \$200 000/QALY was surgery for those with SF-12 PCS scores <35; surgery for those with SF-12 PCS <40 was borderline cost effective. Compared with current practice, restriction of surgery to those with SF-12 PCS <35 would decrease the lifetime likelihood of total knee replacement to 10.2% (95% uncertainty 8.1 to 12.4), and would save \$6974 (\$5789 to \$8269) per patient, whereas the effectiveness would be only slightly lower at −0.008 (−0.056 to 0.043) QALYs. The ICER of this strategy compared with the previous best scenario was \$160 974/QALY (table 4 and fig 1). The current practice scenario was less effective and more expensive than the more restrictive scenarios with SF-12 PCS threshold values of 40–55 and therefore dominated. None of the ICERs fell below the \$50 000/QALY threshold. The likelihood that the current practice scenario would be considered to be cost effective was low for cost effectiveness thresholds below \$200 000/QALY (fig N, appendix 3).

Sensitivity analyses

With a cost effectiveness threshold of \$200 000/QALY, restriction of surgery to those with SF-12 PCS <40 became attractive if the hospital admission costs of primary total knee replacement fell below \$14 000. If the admission costs of primary total knee replacement fell below \$8000, restriction of surgery to those with SF-12 PCS <30 would become economically attractive given a cost effectiveness threshold of \$100 000/QALY. Cost effectiveness outcomes were not sensitive to the admission costs of revision procedures.

Simulation of the MOST population with knee osteoarthritis provided much higher lifetime likelihoods of total knee replacement, but similar ICERs, although restriction of surgery to those with SF-12 PCS <40 now became the optimal scenario at a cost effectiveness threshold of \$200 000/QALY (table E in appendix 2). Use of EQ-5D utility values improved ICERs, with the ICER of restriction of surgery to those with SF-12 PCS <40 now amply falling below \$200 000/QALY (table F in appendix 2). Increasing rates of primary total knee replacement or background mortality only minimally affected incremental cost effectiveness outcomes, and restriction of surgery to those with SF-12 PCS <35 remained the optimal scenario at a cost effectiveness threshold of \$200 000/QALY (tables G and H in appendix 2).

Inclusion of costs associated with work days lost (see table D in appendix 2) and informal caregiving did not have a major impact on the cost effectiveness results: again restriction of surgery to those with SF-12 PCS <35 was the optimal scenario at a cost effectiveness threshold of \$200 000/QALY (table I in appendix 2). If patients who would receive total knee replacement in current practice, but not in the more restrictive scenarios, experienced an additional decline of 50% in quality of life over the long term, all scenarios of performing total knee replacement including current practice became

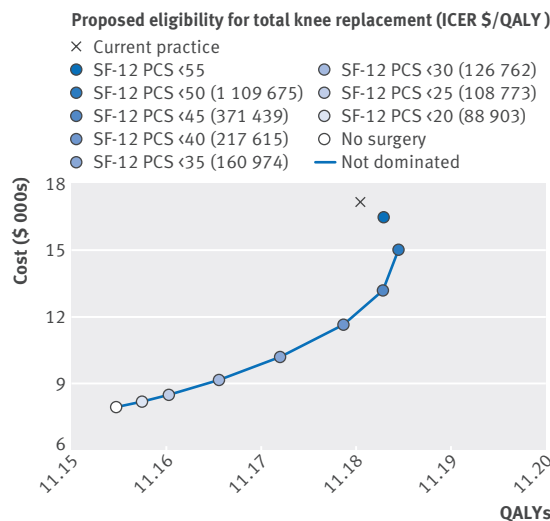


Fig 1 | Base case analysis cost effectiveness at different levels of SF-12 PCS (physical component summary). Costs (\$) and QALYs are means in OAI study population. Incremental cost effectiveness ratios (ICERs) not shown for dominated scenarios

economically attractive given a cost effectiveness threshold of \$200 000/QALY and with an additional decline of 80% given a cost effectiveness threshold of \$100 000/QALY.

Discussion

Principal findings

We evaluated the effectiveness of total knee replacement on quality of life and use of non-surgical treatment in a recent US cohort of patients with knee osteoarthritis. Compared with patients who did not undergo total knee replacement, generic quality of life scores (on SF-12 physical) and those related to osteoarthritis improved with performance of the procedure, with larger improvements generally in those with a lower SF-12 physical score at baseline. Changes in use of osteoarthritis pain medication and SF-12 mental scores were small and heterogeneous across the two cohorts. In a cost effectiveness analysis modeling the life courses of OAI patients with knee osteoarthritis with inclusion of utility values derived from the SF-12, current practice was more expensive and in some cases even less effective compared with scenarios in which total knee replacement was performed only in patients with lower physical functioning. At the group level, the economically most attractive strategy was performing the procedure in those with a SF-12 PCS <35, assuming a cost effectiveness threshold of \$200 000 per QALY. These findings were reproduced among knee osteoarthritis patients from the MOST cohort. Extension of the use of total knee replacement to those with a SF-12 physical score of ≤ 40 would become financially attractive if the hospital admission costs fell below \$14 000.

Comparison with other studies

Scenarios of total knee replacement restricted to those with lower SF-12 PCS scores provided higher QALYs in

our cost effectiveness analysis because small improvements in quality of life after the procedure became more prominent in people with lower baseline scores. A recently published randomized controlled trial showed larger effects of total knee replacement on quality of life measures than we found.⁷ The OAI patient population undergoing total knee replacement, however, had on average less severe symptoms before surgery compared with the population from that randomized controlled trial, and the mean follow-up duration was longer in the OAI. Moreover, measures of quality of life were assessed independently of care providers in the OAI, contributing to a more limited potential for reporting desirable answers on quality of life scores before and after surgery.^{42,43} Recent uncontrolled before-after studies have also shown larger effects of total knee replacement, generally including improvements in SF-12 PCS scores of 5-15 points.^{32,44-53} In these studies, changes in SF-12 MCS were heterogeneous and varied from no improvement to 5 points.^{47,49} Similar to SF-12 PCS, larger effects have been shown in before-after studies for quality of life measures specific to osteoarthritis.^{13,32,44,54,55} As with the study population in the randomized controlled trial, however, the preoperative quality of life scores for these studies were on average worse than those of our study population and the duration of follow-up was generally shorter, possibly explaining the larger effects.

In addition to our analysis, three observational studies have compared quality of life outcomes in patients who did and did not undergo total knee replacement.^{12,56,57} Two of these studies were not designed to prospectively collect data on generic and osteoarthritis related quality of life measures in people who did not and did undergo the procedure.^{12,56} In a single hospital prospective cohort study (n=174) including 30 patients who received total knee replacement, there were large improvements in SF-12 PCS (9.6 points) and total WOMAC (24.2 points) scores at 12 months after joint replacement. The effect estimation, however, included effects of hip replacements (n=21), which generally have much larger effects than knee replacements.^{47,49} Modeling studies that investigated the cost effectiveness of total knee replacement all used effects from before-after studies with larger marginal effects on quality of life, contributing to much lower estimated ICERs than we found.^{13,58,59}

Strengths and limitations of study

We estimated the effectiveness of total knee replacement versus no or delayed surgery in a large population based sample, the OAI cohort study, adjusting for baseline and time varying variables using marginal structural models. Marginal structural modeling has been shown to produce unbiased estimates of causal treatment effects.²⁸ The larger number of procedures in our analysis enabled us to evaluate effect modification by baseline symptoms. Effect estimates and modification by baseline physical functional scores obtained in OAI were generally consistent with those found in the MOST population. Nevertheless, our findings should be interpreted in light of several limita-

tions. First, the knee osteoarthritis populations of the OAI and MOST might not be representative of the current total population of patients with knee osteoarthritis in the US, limiting the generalizability of our findings. For example, younger patients might have been under-represented, contributing to fewer lower quality of life values from symptoms affecting work and other regular activities. Yet, both studies recruited participants from the general population across different regions in the US and obtained detailed information on risk factors, total knee replacements, and outcome measures independently from the hospital in which procedures were performed. Furthermore, our conclusions were unaffected by an increase in the rate of total knee replacement, resembling the most recent increase in use on a national level. Second, the study visits performed within OAI and MOST did not allow for measurement of outcomes immediately before procedures, and ignoring worsening symptoms just before could cause an underestimation of the true effect. We, however, adjusted for differences in 12 month changes in use of osteoarthritis pain medication, Kellgren-Lawrence radiographic grade, and quality of life scores assessed at the visit preceding total knee replacement. In addition, no changes in quality of life are observed at 12 months before procedures versus at one month before,⁵⁰ and an immediate worsening of symptoms as a reason for undergoing total knee replacement is not likely as patients generally defer surgery for years before finally proceeding.⁶⁰⁻⁶² Third, as our study was based on analyses of non-randomized data without an intention to treat principle, residual confounding by indication and selection bias could be possible. Fourth, the self reported outcomes available in the retrospective data, such as use of osteoarthritis pain medication and loss of productivity, could potentially have influenced outcomes because of reporting bias or non-differential overestimation and underestimation. Finally, we made several assumptions in our cost effectiveness analysis. In any modeling study, a trade-off must be made between comprehensively including consequences of each strategy and their relevance to the decision problem at hand. We developed our decision model using individual level data on quality of life, treatment use, and survival, which allowed us to incorporate correlation between the model parameters while assessing uncertainty. Unfortunately, the OAI and MOST studies did not include collection of cost data. We therefore had to estimate costs using the best available external data sources, which might not have been representative for our patient cohort.

Conclusions and policy implications

Improvements in quality of life with total knee replacement were on average smaller than previously shown. Given its limited effectiveness in individuals with less severely affected physical function, performance of total knee replacement in these patients seems to be economically unjustifiable. Considerable cost savings could be made by limiting eligibility to patients with

more symptomatic knee osteoarthritis. Only one randomized controlled trial has so far been published evaluating total knee replacement as an adjunct treatment to optimized non-surgical treatment, but it did not include results according to symptom status.⁷ Our findings emphasize the need for more research comparing total knee replacement with less expensive, more conservative interventions, particularly in patients with less severe symptoms, and research aiming to develop individualized prediction models for a better selection of patients with a predicted large net benefit from the procedure. These interventions can then be compared within cost effectiveness analyses, for which non-US data sources should be considered as well. In conclusion, the practice of total knee replacement as performed in a recent US cohort of patients with knee osteoarthritis had minimal effects on quality of life. If the procedure were restricted to patients with more severe functional status, however, its effectiveness would rise, with practice becoming economically more attractive.

We thank Dennis Fryback and Janel Hanmer for giving valuable comments on the analysis of the generic quality of life scores. The contributions of study participants of the Osteoarthritis Initiative (OAI) and Multicenter Osteoarthritis Study (MOST) are gratefully acknowledged.

Contributors: BF and ZF are joint first authors.

All authors contributed substantially to the conception and design of the work; the acquisition, analysis, and interpretation of data for the work; and the drafting the work or revising the paper critically for important intellectual content. BF and ZF searched the literature and analyzed the data. All authors had full access to all the data and take responsibility for the integrity of the data and the accuracy of the data analysis. BF developed the decision model and performed the statistical analyses and is guarantor.

Funding: The Osteoarthritis Initiative (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. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared with an OAI public use dataset and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners. Multicenter Osteoarthritis Study (MOST) receives four cooperative grants (Felson, AG18820; Torner, AG18832; Lewis, AG18947; and Nevitt, AG19069) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by MOST study investigators. This manuscript was prepared with MOST data and does not necessarily reflect the opinions or views of MOST investigators. BF is supported in part by American Heart Association grant No 16MCPRP31030016.

Competing interests: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf and declare: no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: All studies received approval from medical ethical committees, and participants gave written informed consent.

Data sharing: Patient level data are publicly available to researchers who apply to use the Osteoarthritis Initiative and Multicenter Osteoarthritis Study datasets. The protocol for the Osteoarthritis Initiative can be accessed at <https://oai.epi-ucsf.org/datarelease/docs/StudyDesignProtocol.pdf>. The protocol for the Multicenter Osteoarthritis Study can be accessed at http://most.ucsf.edu/docs/OverviewStudy_v2.0p_06.01.13.pdf. Details on the KOSMOS model and the statistical analyses allowing for replication of results are provided in appendix 1. Statistical codes and decision model are available from the corresponding author.

Transparency: The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

- 1 Cross M, Smith E, Hoy D, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1323-30. doi:10.1136/annrheumdis-2013-204763.
- 2 National Institute for Health and Care Excellence. Osteoarthritis: care and management in adults. (Clinical guideline 177) 2014. www.nice.org.uk/guidance/cg177.
- 3 Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. *J Rheumatol* 2006;33:2271-9. PMID:17013996.
- 4 Singh JA, Vessely MB, Harmsen WS, et al. A population-based study of trends in the use of total hip and total knee arthroplasty, 1969-2008. *Mayo Clin Proc* 2010;85:898-904. doi:10.4065/mcp.2010.0115.
- 5 HCUP National Inpatient Sample (NIS). Rockville, MD, 2012. <http://hcupnet.ahrq.gov/>.
- 6 Losina E, Thornhill TS, Rome BN, Wright J, Katz JN. The dramatic increase in total knee replacement utilization rates in the United States cannot be fully explained by growth in population size and the obesity epidemic. *J Bone Joint Surg Am* 2012;94:201-7. doi:10.2106/JBJSJ.01958.
- 7 Skou ST, Roos EM, Laursen MB, et al. A Randomized, Controlled Trial of Total Knee Replacement. *N Engl J Med* 2015;373:1597-606. doi:10.1056/NEJMoa1505467.
- 8 da Silva RR, Santos AAM, de Sampaio Carvalho Júnior J, Matos MA. Quality of life after total knee arthroplasty: systematic review [English Edition]. *Rev Bras Ortop* 2014;49:520-7. doi:10.1016/j.rbo.2013.10.023.
- 9 Shan L, Shan B, Suzuki A, Nouh F, Saxena A. Intermediate and long-term quality of life after total knee replacement: a systematic review and meta-analysis. *J Bone Joint Surg Am* 2015;97:156-68. doi:10.2106/JBJS.M.00372.
- 10 Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open* 2012;2:e000435. doi:10.1136/bmjopen-2011-000435.
- 11 Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? *Knee* 2007;14:417-23. doi:10.1016/j.knee.2007.06.001.
- 12 Cushnaghan J, Bennett J, Reading I, et al. Long-term outcome following total knee arthroplasty: a controlled longitudinal study. *Ann Rheum Dis* 2009;68:642-7. doi:10.1136/ard.2008.093229.
- 13 Dakin H, Gray A, Fitzpatrick R, MacLennan G, Murray D. KAT Trial Group. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ Open* 2012;2:e000332. doi:10.1136/bmjopen-2011-000332.
- 14 Ethgen O, Bruyère O, Richey F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am* 2004;86-A:963-74. doi:10.2106/00004623-200405000-00012.
- 15 Felson DT, Nevitt MC. Epidemiologic studies for osteoarthritis: new versus conventional study design approaches [vii.]. *Rheum Dis Clin North Am* 2004;30:783-97. vii. doi:10.1016/j.rdc.2004.07.005.
- 16 Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 2007;15(Suppl A):A1-56. doi:10.1016/j.joca.2006.11.009.
- 17 Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. *Osteoarthritis Cartilage* 1995;3(Suppl A):3-70. PMID:8581752.
- 18 Nevitt M, Felson D, Lester G. The Osteoarthritis Initiative: Protocol for the Cohort Study, 2006. <https://oai.epi-ucsf.org/datarelease/>
- 19 National Joint Registry. 11th Annual Report. National Joint Registry for England, Wales, and Northern Ireland, 2014. www.njrcentre.org.uk/njrcentre/Reports/PublicationsandMinutes/Annualreports/Archivedannualreports/tabid/87/Default.aspx
- 20 Ong KL, Lau E, Suggs J, Kurtz SM, Manley MT. Risk of subsequent revision after primary and revision total joint arthroplasty. *Clin Orthop Relat Res* 2010;468:3070-6. doi:10.1007/s11999-010-1399-0.
- 21 Losina E, Paltiel AD, Weinstein AM, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. *Arthritis Care Res (Hoboken)* 2015;67:203-15. doi:10.1002/acr.22412.
- 22 Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Use and costs of prescription medications and alternative treatments in patients with osteoarthritis and chronic low back pain in community-based settings. *Pain Pract* 2012;12:550-60. doi:10.1111/j.1533-2500.2012.00532.x.
- 23 Segal NA, Nevitt MC, Gross KD, et al. The Multicenter Osteoarthritis Study: opportunities for rehabilitation research. *PM R* 2013;5:647-54. doi:10.1016/j.pmrj.2013.04.014.
- 24 Ware JJ, Kosinski M, Keller SDA. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-33. doi:10.1097/00005650-199603000-00003.
- 25 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851-9. doi:10.1097/01.mlr.0000135827.18610.0d.
- 26 Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). *Arthritis Care Res (Hoboken)* 2011;63(Suppl 11):S208-28. doi:10.1002/acr.20632.
- 27 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. PMID:3068365.
- 28 Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med* 2002;21:1689-709. doi:10.1002/sim.1144.
- 29 Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? *Med Care* 1996;34:73-84. doi:10.1097/00005650-199601000-00006.
- 30 Robins JM, Hernán MA, Rotnitzky A. Effect modification by time-varying covariates. *Am J Epidemiol* 2007;166:994-1002, discussion 1003-4. doi:10.1093/aje/kwm231.
- 31 Yang L, Tsiatis A. Efficiency Study of Estimators for a Treatment Effect in a Pretest-Posttest Trial. *Am Stat* 2001;55:314-21. doi:10.1198/000313001753272466.
- 32 Greidanus NV, Peterson RC, Masri BA, Garbus DS. Quality of life outcomes in revision versus primary total knee arthroplasty. *J Arthroplasty* 2011;26:615-20. doi:10.1016/j.arth.2010.04.026.
- 33 Hartley RC, Barton-Hanson NG, Finley R, Parkinson RW. Early patient outcomes after primary and revision total knee arthroplasty. A prospective study. *J Bone Joint Surg Br* 2002;84:994-9. doi:10.1302/0301-620X.84B7.12607.
- 34 Secondary Consumer Price Index 2015. <http://www.bls.gov/cpi/tables.htm>.
- 35 Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996;276:1253-8. doi:10.1001/jama.1996.03540150055031.
- 36 Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the \$50,000-per-QALY threshold. *N Engl J Med* 2014;371:796-7. doi:10.1056/NEJMp1405158.
- 37 Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making* 2006;26:401-9. doi:10.1177/0272989X06290496.
- 38 Khanna D, Maranian P, Palta M, et al. Health-related quality of life in adults reporting arthritis: analysis from the National Health Measurement Study. *Qual Life Res* 2011;20:1131-40. doi:10.1007/s11136-011-9849-z.
- 39 Franks P, Hanmer J, Fryback DG. Relative disutilities of 47 risk factors and conditions assessed with seven preference-based health status measures in a national U.S. sample: toward consistency in cost-effectiveness analyses. *Med Care* 2006;44:478-85. doi:10.1097/01.mlr.0000207464.61661.05.
- 40 Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;342:d1165. doi:10.1136/bmj.d1165.
- 41 Weinstein MC, Siegel JE, Garber AM, et al. Productivity costs, time costs and health-related quality of life: a response to the Erasmus Group. *Health Econ* 1997;6:505-10. doi:10.1002/(SICI)1099-1050(199709)6:5<505::AID-HEC294>3.0.CO;2-I.
- 42 Bouchet C, Guillemin F, Briçon S. Nonspecific effects in longitudinal studies: impact on quality of life measures. *J Clin Epidemiol* 1996;49:15-20. doi:10.1016/0895-4356(95)00540-4.
- 43 McCarney R, Warner J, Illife S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. *BMC Med Res Methodol* 2007;7:30. doi:10.1186/1471-2288-7-30.

- 44 Alentorn-Geli E, Leal-Blanquet J, Guirro P, Hinarejos P, Pelfort X, Puig-Verdié L. Comparison of quality of life between elderly patients undergoing TKA. *Orthopedics* 2013;36:e415-9. doi:10.3928/01477447-20130327-15.
- 45 Baker P, Muthumayandi K, Gerrand C, Kleim B, Bettinson K, Deehan D. Influence of body mass index (BMI) on functional improvements at 3 years following total knee replacement: a retrospective cohort study. *PLoS One* 2013;8:e59079. doi:10.1371/journal.pone.0059079.
- 46 Cherian JJ, O'Connor MI, Robinson K, Jauregui JJ, Adleberg J, Mont MA. A Prospective, Longitudinal Study of Outcomes Following Total Knee Arthroplasty Stratified by Gender. *J Arthroplasty* 2015;30:1372-7. doi:10.1016/j.arth.2015.03.032.
- 47 Clement ND, Jenkins PJ, MacDonald D, et al. Socioeconomic status affects the Oxford knee score and short-form 12 score following total knee replacement. *Bone Joint J* 2013;95-B:52-8. doi:10.1302/0301-620X.95B1.29749.
- 48 Clement ND, MacDonald D, Burnett R, Breusch SJ. Diabetes does not influence the early outcome of total knee replacement: a prospective study assessing the Oxford knee score, short form 12, and patient satisfaction. *Knee* 2013;20:437-41. doi:10.1016/j.knee.2013.07.009.
- 49 Keurentjes JC, Blane D, Bartley M, Keurentjes JJ, Fiocco M, Nelissen RG. Socio-economic position has no effect on improvement in health-related quality of life and patient satisfaction in total hip and knee replacement: a cohort study. *PLoS One* 2013;8:e56785. doi:10.1371/journal.pone.0056785.
- 50 Mandzuk LL, McMillan DE, Bohm ER. A longitudinal study of quality of life and functional status in total hip and total knee replacement patients. *Int J Orthop Trauma Nurs* 2015;19:102-13. doi:10.1016/j.ijotn.2014.07.001.
- 51 Murray DW, MacLennan GS, Breeman S, et al. KAT group. A randomised controlled trial of the clinical effectiveness and cost-effectiveness of different knee prostheses: the Knee Arthroplasty Trial (KAT)[vii-viii]. *Health Technol Assess* 2014;18:1-235, vii-viii. doi:10.3310/hta18190.
- 52 Pivec R, Issa K, Given K, et al. A prospective, longitudinal study of patient satisfaction following total knee arthroplasty using the Short-Form 36 (SF-36) survey stratified by various demographic and comorbid factors. *J Arthroplasty* 2015;30:374-8. doi:10.1016/j.arth.2014.10.013.
- 53 Seng C, Yeo SJ, Wee JL, Subanesh S, Chong HC, Lo NN. Improved clinical outcomes after high-flexion total knee arthroplasty: a 5-year follow-up study. *J Arthroplasty* 2011;26:1025-30. doi:10.1016/j.arth.2010.09.006.
- 54 Davis AM, Perruccio AV, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthritis Cartilage* 2009;17:843-7. doi:10.1016/j.joca.2009.01.005.
- 55 Nilsson AK, Toksvig-Larsen S, Roos EMA. A 5 year prospective study of patient-relevant outcomes after total knee replacement. *Osteoarthritis Cartilage* 2009;17:601-6. doi:10.1016/j.joca.2008.11.007.
- 56 George LK, Ruiz D Jr, Sloan FA. The effects of total knee arthroplasty on physical functioning in the older population. *Arthritis Rheum* 2008;58:3166-71. doi:10.1002/art.23888.
- 57 Hamel MB, Toth M, Legedza A, Rosen MP. Joint replacement surgery in elderly patients with severe osteoarthritis of the hip or knee: decision making, postoperative recovery, and clinical outcomes. *Arch Intern Med* 2008;168:1430-40. doi:10.1001/archinte.168.13.1430.
- 58 Daigle ME, Weinstein AM, Katz JN, Losina E. The cost-effectiveness of total joint arthroplasty: a systematic review of published literature. *Best Pract Res Clin Rheumatol* 2012;26:649-58. doi:10.1016/j.berh.2012.07.013.
- 59 Losina E, Walensky RP, Kessler CL, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. *Arch Intern Med* 2009;169:1113-21, discussion 1121-2. doi:10.1001/archinternmed.2009.136.
- 60 Conner-Spady BL, Marshall DA, Hawker GA, et al. You'll know when you're ready: a qualitative study exploring how patients decide when the time is right for joint replacement surgery. *BMC Health Serv Res* 2014;14:454. doi:10.1186/1472-6963-14-454.
- 61 Sansom A, Donovan J, Sanders C, et al. Routes to total joint replacement surgery: patients' and clinicians' perceptions of need. *Arthritis Care Res (Hoboken)* 2010;62:1252-7. doi:10.1002/acr.20218.
- 62 Jacobson AF, Myerscough RP, Delambo K, et al. Patients' perspectives on total knee replacement. *Am J Nurs* 2008;108:54-63, quiz 63-4. doi:10.1097/01.NAJ.0000318000.62786.fb.

Appendix 1: Technical appendix

Appendix 2: Supplementary tables A-J

Appendix 3: Supplementary figures A-N