



The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial



David J Beard, Loretta J Davies, Jonathan A Cook, Graeme MacLennan, Andrew Price, Seamus Kent, Jemma Hudson, Andrew Carr, Jose Leal, Helen Campbell, Ray Fitzpatrick, Nigel Arden, David Murray, Marion K Campbell, for the TOPKAT Study Group*

Summary

Lancet 2019; 394: 746–56

Published Online

July 17, 2019

[http://dx.doi.org/10.1016/S0140-6736\(19\)31281-4](http://dx.doi.org/10.1016/S0140-6736(19)31281-4)

See [Comment](#) page 712

*TOPKAT Study Group members listed at the end of the Article

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre

(Prof D J Beard DPhil,

L J Davies MSc, J A Cook PhD,

Prof A Price FRCS,

Prof A Carr FRCS,

Prof N Arden FRCP,

Prof D Murray FRCS) and

Nuffield Department of

Population Health (S Kent PhD,

J Leal PhD, H Campbell PhD,

Prof R Fitzpatrick DPhil),

University of Oxford, Oxford,

UK; and Health Services

Research Unit, University of

Aberdeen, Aberdeen, UK

(J A Cook, Prof G MacLennan MSc,

J Hudson MSc,

Prof M K Campbell PhD)

Correspondence to:

Prof David Beard, Nuffield

Department of Orthopaedics,

Rheumatology and

Musculoskeletal Sciences, Botnar

Research Centre, University of

Oxford, Oxford OX3 7LD, UK

david.beard@ndorms.ox.ac.uk

Background Late-stage isolated medial knee osteoarthritis can be treated with total knee replacement (TKR) or partial knee replacement (PKR). There is high variation in treatment choice and little robust evidence to guide selection. The Total or Partial Knee Arthroplasty Trial (TOPKAT) therefore aims to assess the clinical effectiveness and cost-effectiveness of TKR versus PKR in patients with medial compartment osteoarthritis of the knee, and this represents an analysis of the main endpoints at 5 years.

Methods Our multicentre, pragmatic randomised controlled trial was done at 27 UK sites. We used a combined expertise-based and equipoise-based approach, in which patients with isolated osteoarthritis of the medial compartment of the knee and who satisfied general requirements for a medial PKR were randomly assigned (1:1) to receive PKR or TKR by surgeons who were either expert in and willing to perform both surgeries or by a surgeon with particular expertise in the allocated procedure. The primary endpoint was the Oxford Knee Score (OKS) 5 years after randomisation in all patients assigned to groups. Health-care costs (in UK 2017 prices) and cost-effectiveness were also assessed. This trial is registered with ISRCTN (ISRCTN03013488) and ClinicalTrials.gov (NCT01352247).

Findings Between Jan 18, 2010, and Sept 30, 2013, we assessed 962 patients for their eligibility, of whom 431 (45%) patients were excluded (121 [13%] patients did not meet the inclusion criteria and 310 [32%] patients declined to participate) and 528 (55%) patients were randomly assigned to groups. 94% of participants responded to the follow-up survey 5 years after their operation. At the 5-year follow-up, we found no difference in OKS between groups (mean difference 1.04, 95% CI -0.42 to 2.50; $p=0.159$). In our within-trial cost-effectiveness analysis, we found that PKR was more effective (0.240 additional quality-adjusted life-years, 95% CI 0.046 to 0.434) and less expensive (-£910, 95% CI -1503 to -317) than TKR during the 5 years of follow-up. This finding was a result of slightly better outcomes, lower costs of surgery, and lower follow-up health-care costs with PKR than TKR.

Interpretation Both TKR and PKR are effective, offer similar clinical outcomes, and result in a similar incidence of re-operations and complications. Based on our clinical findings, and results regarding the lower costs and better cost-effectiveness with PKR during the 5-year study period, we suggest that PKR should be considered the first choice for patients with late-stage isolated medial compartment osteoarthritis.

Funding National Institute for Health Research Health Technology Assessment Programme.

Copyright © 2019 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

Knee replacement for osteoarthritis of the knee is an effective, common procedure:^{1,2} between 2015 and 2017, 303 960 knee replacements were performed in the UK.¹ There are two main surgical options for late-stage medial compartment osteoarthritis of the knee: a partial knee replacement (PKR; otherwise known as a unicompartamental knee replacement), in which only the damaged compartment of the knee is replaced, or a total knee replacement (TKR).² There is often disagreement between surgeons regarding the best choice of implant and the best decision for these patients, who often have

identical pathology,³ which has led to variation in provision. Although fewer than 9% of knee replacements in the UK are partial,¹ it is estimated that up to 47% of patients requiring knee replacements only have unicompartamental disease and are suitable for a PKR.⁴ The best treatment option remains uncertain because there is little level 1 evidence to inform the clinical effectiveness and cost-effectiveness of either management approach.

The existing related medical literature comprises four small randomised studies,^{5–9} individual cohort studies,^{10,11} indirect comparisons,¹² and retrospective

Research in context

Evidence before this study

The evidence to guide choice in the type of arthroplasty used for late-stage isolated medial compartment knee arthritis is extremely poor. Our study began in 2009, and we searched all major databases (MEDLINE, Embase, and PubMed). We searched the medical literature for related studies between Jan 1, 1990, and Dec 31, 2008, with the search terms “total”, “partial”, “unicompartmental”, “knee replacement”, “knee arthroplasty”, “trial”, and “outcome” and no language restrictions. Before 2009, only cohort studies, indirect comparisons, and retrospective studies had been undertaken, which aimed to address specific aspects of each operation, and no large, well powered, multicentre randomised comparative controlled trials had been done. However, two small randomised trials had been done, both in the UK. The Knee Arthroplasty Trial had a component comparing partial knee replacement (PKR) with total knee replacement (TKR) but this comparison had to be terminated early because of inadequate recruitment. Another small single-centre randomised controlled trial that was published in 1998 (5 years after starting) and followed up in 2009 (after 15 years) reported good long-term follow-up, and the trial showed benefits of PKR; however, the study was underpowered for both knee score and revision frequency. Since 2009, three new randomised controlled trials have been published, but these trials were also small single-centre studies (comprising 56–72 participants per trial). To our knowledge, there are five ongoing trials. Additionally, a propensity-matched cohort study from the National Joint Registry was published in *The Lancet* in 2014, but the focus was solely on adverse outcomes (namely, revision frequency—described in terms of implant survival—and death). In this time, two systematic reviews with meta-analyses have also been published, both of which confirmed that there are

minimal randomised trial data on which to guide treatment, and there are no robust randomised trials reporting outcomes in the longer term (ie, beyond 5 years).

Added value of this study

Our addition of the first large multicentre randomised study provides the best quality evidence to date to guide knee replacement treatment globally, and these data show low concordance with well established registry and cohort data. Our results after 5 years regarding the number of re-operations and revisions do not correspond to joint registry data, in which PKR has been shown to lead to more revisions than TKR. Also, our health economic data, which had previously not been assessed to this level of detail, adds an additional dimension and further weight to our overall interpretation and guidance for practice. Our trial reflects the breadth of current practice through the inclusion of all types of implants, which is a strength, and its funding source (namely, our dissociation from industry, in a context of commonplace industry-supported research on surgical devices) helps to ensure unbiased evidence.

Implications of all the available evidence

Knee replacement is increasing in frequency, and it has an associated substantial cost implication to any health-care provider. It is also essential that patients receive the most efficacious operation for this condition. Before our study, and despite several cohort-based reports, knowledge of whether one operation type is superior, remained uncertain. Our 5-year study has indicated that both TKR and PKR are beneficial interventions but, based on our combined clinical and cost-effectiveness data and providing the operation is performed by those with adequate experience, we recommend that PKR should be offered as the treatment of choice for late-stage isolated medial compartment osteoarthritis of the knee.

studies,^{13–16} often examining very specific aspects of each procedure. To our knowledge, no one has undertaken a large, adequately powered, multicentre randomised controlled trial with long-term follow-up of patient-reported outcomes, re-operation, or revision, to directly compare PKR with TKR. A revision procedure is a subsequent further operation to replace the original implant with a new device, usually because of device failure, infection, or persistent pain. Revision is often described in terms of implant survival, and it can be done at any time from initial implantation (in some cases up to 30 or 40 years after primary arthroplasty) but early revision (before 5 years) is considered an early failure of the original surgical intervention. It is therefore an important metric in arthroplasty outcome. Re-operations are any further or subsequent operation, which can be revision surgery or other kinds of surgery. Systematic reviews^{17,18} and unadjusted data from national registries¹ show more frequent revision associated with PKR: the chance of revision with PKR is more than

double that with a TKR. When comparing matched patients receiving PKRs and TKRs 8 years after the operations, higher frequencies of revision and re-operations with PKR were also identified in an observational, registry-based study.¹⁹ However, there are no comparable data on revisions from randomised trials to-date. Although registry and cohort data can provide useful information, to provide further robust and unbiased estimates to compare the procedures, a large randomised trial was required.

The primary objective for the Total or Partial Knee Arthroplasty Trial (TOPKAT) was, therefore, to assess the clinical effectiveness and cost-effectiveness of TKR versus PKR in patients with medial compartment osteoarthritis of the knee in a formal, randomised controlled comparison. Secondary objectives included investigation of complications (including re-operations and revisions), patient satisfaction, and the cost implications of the knee replacements for patients, employers, and health-care providers. This assessment represents the main

endpoint results after 5 years of a 10-year follow-up study.

Methods

Study design and participants

The detailed study design and protocol have previously been published.²⁰ In summary, TOPKAT was a multicentre, pragmatic randomised controlled trial of patients with isolated medial compartmental osteoarthritis who were randomly assigned to receive either PKR or TKR. The trial had a combined expertise-based and equipoise-based approach. The expertise-based randomisation²¹ enabled surgeons who were not in the equipoise group to work in pairs, and each surgeon (as experts in a specific operation type) agreed to perform the operation type that they felt was appropriate (giving one surgeon in the pair who used TKR and one who used PKR); patients were then randomly assigned to a TKR or PKR surgeon in the pairs. The equipoise-based randomisation involved surgeons in the equipoise group who had sufficient experience of both operation types to perform either operation (PKR or TKR) without compromise. Minimum levels of expertise were ensured to participate in the trial. PKR-only and TKR-only surgeons were required to have had appropriate training, to have been practising their technique for at least 1 year, and to have performed their operation at least ten times in the previous year. Equipoise surgeons, who delivered both operations, were required to satisfy the criteria for both operations—ie, they had appropriate training in both operations and had performed a minimum of ten PKR and ten TKR procedures in the previous year.

We recruited patients from 27 sites across the UK (and 68 surgeons at these sites). Potential participants were identified in outpatient and pre-assessment clinics by participating surgeons. For study inclusion, participants had to have isolated osteoarthritis of the medial compartment of the knee and to satisfy general requirements for a medial PKR. Detailed inclusion and exclusion criteria are provided in the published protocol.²⁰ Patients who required simultaneous bilateral knee replacement were excluded. Participants signed a consent form during their screening visit, and they were free to withdraw from the study at any time without affecting their routine perioperative care. The study obtained approval from the UK National Research Ethics Service (Research Ethics Committee South Central Oxford C) in September, 2009 (09/H0606/88).

Randomisation and masking

Participants were randomly assigned (1:1) to receive TKR or PKR. For the randomisation procedure (in which we also used minimisation), we used a web-based randomisation service at the Centre for Healthcare Randomised Trials (Health Services Research Unit, University of Aberdeen, UK). We also used a minimisation algorithm that incorporated sex, age band (<50, 50–70, or >70 years), baseline Oxford Knee Score (OKS) band (14 or less, 15–21,

or 22 or more), and delivery unit. A delivery unit was either an equipoise surgeon or a pair of expertise surgeons with complementary expertise.²⁰ Surgeons, patients, and follow-up assessors were not masked to allocation, but the implant type was not highlighted at any stage.

Procedures

Our trial compared the TKR procedure versus the PKR procedure rather than a specific brand of TKR implant versus a specific brand of PKR implant. As such, surgeons were free to use the implant of their own choice, or that of their institution. A TKR involves all surfaces of the knee being replaced. The procedure involves excising both diseased and normal femoral condyles, the tibial plateau—and often the patella—and removing or releasing some of the ligaments. The artificial implant can be cemented in position. A PKR or unicompartmental knee replacement involves only the diseased area of the joint being replaced by artificial implants, whereas healthy compartments of the knee and ligaments are retained.

The baseline assessment included a patient self-reported questionnaire to examine the frequency and intensity of knee pain and knee function (measured by the OKS), activity, health-care resource use, and the American Knee Society Score.²² Routine pre-operative X-rays were also used.

We recorded operative details in theatre, and we routinely took postoperative X-rays. Patients attended a clinic for an American Knee Society Score assessment and a complication review 2 months, 1 year, and 5 years after their operation. All other outcomes were collected with postal questionnaires at 2 months after the operation and 1 year after randomisation and, thereafter, annually (until 5 years after randomisation).

We collected data on complications and re-operation from various sources for cross-referencing and to ensure completeness. Complication data were collected from participants in their follow-up questionnaires and clinical assessment visits. Any reported re-admissions were followed up with the patient's hospital to obtain further details. Local research teams did a final check of hospital records 5 years after the operations, to ensure that complications data were collected from all participants (ie, those who had not returned a questionnaire or attended a follow-up visit).

Outcomes

The primary outcome measure was the OKS 5 years after randomisation in all patients assigned to groups. The OKS is a patient self-reported outcome questionnaire that was specifically developed to assess function and pain after knee replacement surgery. This score is a validated and responsive measure of knee function and related pain.^{23,24}

The secondary outcome measures were assessed at 5 years after randomisation in all patients assigned to

receive an operation. These outcomes were the American Knee Society Score,²² the University of California, Los Angeles Activity Score,²⁵ the High Activity Arthroplasty Score,²⁶ the EuroQol EQ-5D-3L,²⁷ the Lund Satisfaction Score,²⁸ the frequency of complications (determining the number of patients who required re-admission and a re-operation), and three patient-reported anchor type questions regarding satisfaction, a comparison of problems with their knee before versus after the surgery, and whether they would have the operation again. The response categories are shown in the appendix (p 1). We also assessed a composite outcome for failure (a combination of re-operation and poor outcomes, as determined by an improvement in OKS of less than four points). This composite outcome assessment has not been validated and it is therefore considered an exploratory outcome.

Length of hospital stay and data on use of health-care services were collected for an health economics analysis.

Statistical analysis

A necessary sample size was calculated for the OKS and the incidence of re-operations. Our trial was powered to detect a 2.0-point difference in OKS with 80% power (or, equivalently, 90% power to detect a 3.0-point change) at a 5% (two-sided) significance level with an SD of 10.0 points. For this power, a sample size of 500 patients was required.^{29,30} The sample size was adequate to detect a change in incidence of re-operations of 7% between groups (from 5% to 12%) at 80% power. Of note, a study³¹ done and published after the TOPKAT trial was planned suggested that a 5-point difference in OKS should become the appropriate minimally important difference in OKS.

Full details of the statistical analyses are contained in the statistical analysis plan (available on request). In brief, the statistical analysis was primarily on an intention-to-treat basis and used all available participant data. The OKS scores at 5 years, which were compared with a linear regression analysis, were adjusted for minimisation factors and baseline OKS, and a cluster robust error was added for surgery delivery unit. A differential effect of expertise versus equipoise delivery of the intervention on treatment effect was explored with an interaction effect. The incidence of complications and re-operations, including revisions, and other binary variables were compared with Poisson regression, to estimate risk ratio, and they were adjusted for minimisation covariates and a cluster robust error for surgeon delivery unit. We used prespecified secondary analyses (unadjusted *t* test, a multilevel model to assess the marginal difference) and subgroup analyses (sex, OKS band, and age band, evaluated with treatment-by-subgroup interactions). To assess the effect of compliance (the operation being delivered as intended), an instrumental variable method was used to estimate the complier-average causal effect at 5 years after randomisation. We also performed a

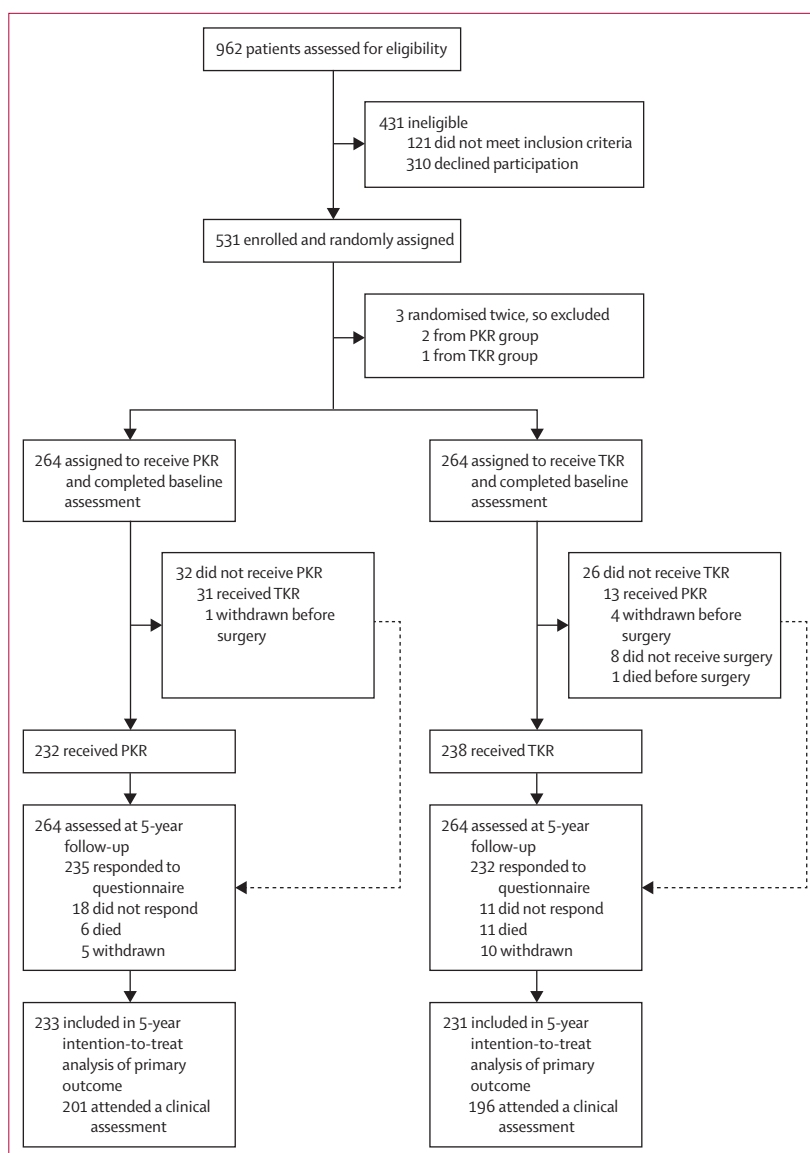


Figure 1: Trial profile

The 5-year intention-to-treat analysis includes only patients who provided results for the Oxford Knee Score assessment. PKR=partial knee replacement. TKR=total knee replacement.

post-hoc analysis of OKS that used the area under the curve (AUC), generated for each participant with the trapezoidal rule, to describe the effects of TKR and PKR on OKS across all available timepoints.

Statistical significance was judged at the two-sided 5% level and treatment effect estimates are presented with corresponding 95% CIs. Low frequency event data, such as re-operations and revisions, were also analysed on an as-treated basis, to avoid misrepresentation from incompliance to the allocated intervention (ie, crossovers).

We also did a within-trial cost-effectiveness analysis. Total costs and quality-adjusted life-years (QALYs) for all 528 participants were estimated from the date of recruitment until the earliest of death, withdrawal from

See Online for appendix

	PKR (n=264)	TKR (n=264)
Age, years	65.2 (8.8)	64.7 (8.5)
Sex		
Male	153 (58%)	153 (58%)
Female	111 (42%)	111 (42%)
Study knee		
Left	140 (53%)	141 (53%)
Right	124 (47%)	123 (47%)
Duration of osteoarthritis		
<3 years	75 (28%)	73 (28%)
3–5 years	82 (31%)	72 (27%)
6–10 years	59 (22%)	73 (28%)
>10 years	36 (14%)	30 (11%)
Missing data	12 (5%)	16 (6%)
Medical history of conditions		
Other joint problems	106 (40%)	96 (36%)
Cardiovascular	80 (30%)	86 (33%)
Diabetes	27 (10%)	26 (10%)
Gastrointestinal	17 (6%)	18 (7%)
Respiratory	19 (7%)	12 (5%)
Cancer	6 (2%)	8 (3%)
Renal or urological	8 (3%)	8 (3%)
Neurological	7 (3%)	6 (2%)
Mental health	7 (3%)	6 (2%)
Thyroid problems	3 (1%)	2 (1%)
Other*	5 (2%)	4 (2%)
Employment status		
Retired	159 (60%)	162 (61%)
Unemployed	15 (6%)	21 (8%)
Employed	82 (31%)	73 (28%)
Missing data	8 (3%)	8 (3%)
Body-mass index (n=210 vs n=221)	31.0 (4.6)	31.1 (4.8)
Extent of knee arthritis affecting mobility (Charnley ABC)		
Single	99 (38%)	119 (45%)
Both	142 (54%)	121 (46%)
Multiple arthritis or medical infirmity	6 (2%)	11 (4%)
Missing data	17 (6%)	13 (5%)
General health (n=259 vs n=260)	2.6 (0.9)	2.8 (0.9)
General health relative to 1 year ago (n=259 vs n=260)	3.3 (0.8)	3.3 (0.8)

(Table 1 continues in next column)

study, or the end of follow-up at 5 years. Both costs and QALYs were discounted at 3.5% per year. Missing data on EQ-5D score and health-care costs were imputed with multiple imputation by chained equations, assuming data was missing at random (since there was no evidence that the probability that data was missing depended on baseline participant characteristics or on lagged outcomes).³² Ten imputed datasets were created with predictive mean matching. Imputation was performed separately by treatment allocation. After imputation, the mean total health-care costs and QALYs were estimated with separate linear regression models

	PKR (n=264)	TKR (n=264)
(Continued from previous column)		
Previous treatment on study knee		
Analgesia	207 (78%)	184 (70%)
Arthroscopy	44 (17%)	47 (18%)
Arthroscopic investigative washout or debridement	44 (17%)	36 (14%)
Open or arthroscopic meniscus	33 (13%)	30 (11%)
Knee injection with steroids	19 (7%)	21 (8%)
Knee injection with viscosupp	2 (1%)	3 (1%)
Knee injection (type not stated)	4 (2%)	6 (2%)
Acupuncture	4 (2%)	5 (2%)
Chiropractor or osteopath	1 (<1%)	3 (1%)
Cartilage implantation	1 (<1%)	2 (1%)
Anterior cruciate ligament repair	1 (<1%)	0
Other†	0	3 (1%)
None	13 (5%)	26 (10%)
Problems with the other knee		
None	86 (33%)	99 (38%)
Mild	93 (35%)	74 (28%)
Moderate	63 (24%)	52 (20%)
Severe	18 (7%)	30 (11%)
Missing data	4 (2%)	9 (3%)
Contralateral knee		
TKR	24 (9%)	14 (5%)
PKR	16 (6%)	16 (6%)
Unsure	1 (<1%)	1 (<1%)
None	208 (79%)	217 (82%)
Missing	15 (6%)	16 (6%)
Oxford Knee Score	18.8 (7.0)	19.0 (7.2)
High Activity Arthroplasty Score (n=258 vs n=256)	4.8 (2.3)	4.6 (2.3)
University of California, Los Angeles Activity score (n=260 vs n=260)	3.6 (1.5)	3.7 (1.5)
American Knee Society score (objective; n=260 vs n=259)	41.0 (16.1)	42.3 (16.0)
American Knee Society score (functional; n=262 vs n=259)	59.3 (15.6)	58.7 (15.5)
EQ-5D-3L score (n=257 vs n=252)	0.428 (0.301)	0.381 (0.324)
EQ-5D visual analogue score (n=249 vs n=257)	62.8 (27.0)	60.7 (28.7)

Data are mean (SD) or n (%), unless otherwise indicated. PKR=partial knee replacement. TKR=total knee replacement. EQ-5D-3L=EuroQol (five dimensions, three levels). *Includes high body-mass index (PKR n=1; TKR n=1), antiphospholipid syndrome (TKR n=1), glaucoma (PKR n=1), cataracts (TKR n=1), anaemia (PKR n=2), appendicectomy (TKR n=1), and lipoma of intra-abdominal organs (PKR n=1). †Includes laser treatment (n=1), offloader knee brace (n=1), and aspiration (n=1).

Table 1: Baseline characteristics

controlling for treatment allocation, age band, sex, and baseline OKS band, following the main clinical analysis. For QALYs we additionally controlled for EQ-5D score at baseline. The incremental cost-effectiveness ratio was calculated by dividing the mean cost difference between PKR and TKR by the mean QALY difference. The joint

uncertainty around incremental total costs and QALYs (ie, the difference between PKR and TKR), and in the cost-effectiveness, was also investigated. All statistical analyses were done with Stata version 15, and all health economic analyses were done with R version 3.4.3.

A summary of the results, masked for implant identifier, was sent to five separate independent reviewers (experts who were familiar with orthopaedics, or trial design or interpretation, or both) plus the trial steering committee chair, to obtain an unbiased interpretation and conclusion of the trial. Any TOPKAT team members with strong personal opinions or conflicts of interest, although helpful in the setting up of the study, had little or no part in the conduct of the study or interpretation of the results. This trial is registered with ISRCTN, (ISRCTN03013488) and ClinicalTrials.gov (NCT01352247).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between Jan 18, 2010, and Sept 30, 2013, we recruited patients from 27 sites across the UK, and we included 68 surgeons from these sites. We assessed 962 patients for their eligibility, of whom 431 (45%) patients were excluded (121 [13%] patients did not meet the inclusion criteria and 310 [32%] patients declined to participate) and 528 (55%) patients were randomly assigned to groups (figure 1). Three patients were randomly assigned to groups twice in error, and the second randomisation was treated as a post-randomisation exclusion. Of the 528 patients, 264 (50%) patients were assigned to the PKR group, of whom 233 (88%) patients provided results for assessment of the primary outcome, and 264 (50%) patients were assigned to the TKR group, of whom 231 (88%) patients provided results for assessment of the primary outcome.

44 participants had a knee replacement but did not receive their allocated device (crossovers). 31 (12%) patients allocated to PKR received TKR: in six (19%) patients, the crossover occurred before surgery (the osteoarthritis pattern was not suitable for PKR [n=2]; patient decision [n=1]; inflammatory arthropathy [not suitable for PKR; n=1]; after random allocation to groups, the patient was deemed not suitable for PKR because of a contraindication [n=1]; and because of an error in communication of group allocation at the site [n=1]); and in 25 (81%) patients, an intra-operative conversion was made (the osteoarthritis pattern was not suitable for PKR [n=20]; and the anterior cruciate ligament was found to be absent or deteriorated [n=5]). 13 (5%) patients who were allocated to TKR received PKR: in eight (62%) patients, the crossover occurred before surgery (patient decision [n=5];

	Partial knee replacement		Total knee replacement		Effect size (95% CI)	p value
	Mean (SD)	n	Mean (SD)	n		
Oxford Knee Score (primary outcome)						
Baseline	18.8 (7.0)	264	19.0 (7.2)	264	1 (ref)	..
5 years	38.0 (10.1)	233	37.0 (10.6)	231	1.04 (-0.42 to 2.50)	0.159
EQ-5D-3L score						
Baseline	0.428 (0.30)	257	0.381 (0.32)	252	1 (ref)	..
5 years	0.744 (0.29)	224	0.717 (0.32)	212	0.018 (-0.033 to 0.069)	0.48
EQ-5D visual analogue scale score						
Baseline	62.8 (27.0)	249	60.7 (28.7)	257	1 (ref)	..
5 years	75.4 (16.5)	228	71.7 (19.7)	217	4.02 (1.36 to 6.67)	0.004
High Activity Arthroplasty Score						
Baseline	4.8 (2.3)	258	4.6 (2.3)	256	1 (ref)	..
5 years	7.9 (3.5)	218	7.6 (3.4)	207	0.22 (-0.24 to 0.67)	0.33
University of California, Los Angeles Activity score						
Baseline	3.6 (1.5)	260	3.7 (1.5)	260	1 (ref)	..
5 years	5.0 (1.9)	221	4.9 (2.0)	215	0.17 (-0.09 to 0.43)	0.19
American Knee Society score (objective)						
Baseline	41.0 (16.1)	260	42.3 (16.0)	259	1 (ref)	..
5 years	85.8 (16.6)	191	86.6 (16.4)	185	-0.89 (-5.18 to 3.41)	0.68
American Knee Society score (functional)						
Baseline	59.3 (15.6)	262	58.7 (15.5)	259	1 (ref)	..
5 years	82.6 (18.5)	195	81.7 (19.0)	192	0.37 (-3.81 to 4.55)	0.86

EQ-5D-3L=EuroQol (five dimensions, three levels).

Table 2: Estimated treatment effect on primary and secondary outcomes, at baseline and at 5 years

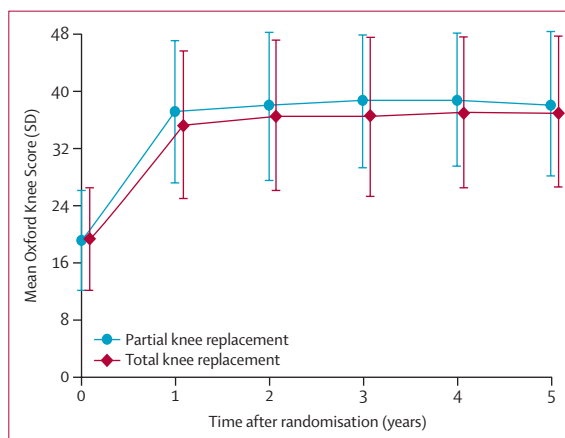


Figure 2: Oxford Knee Scores after randomisation

and error in communication of randomisation allocation at site [n=3]); and in five (38%) patients, an intra-operative decision was made (independent surgeon decision—ie, all eligibility criteria met but decision made to use PKR). The groups were well balanced on all baseline characteristics (table 1). The types of implants used in the study reflected current practice (appendix p 2).

The operation time was similar in both groups, at around 70 min (PKR mean duration 69 min [SD 18]; TKR 70 min [22]; appendix p 15). The length of the hospital stay for the surgery differed significantly between

groups: those receiving PKR stayed for a mean duration of 3.2 days (1.3), and those receiving TKR stayed for a mean duration of 4.3 days (3.6; risk ratio [RR] 0.74, 95% CI 0.63 to 0.87; $p < 0.0001$). 5 years after randomisation, 94% of patients responded to the follow-up questionnaire (88% for the primary outcome; PKR $n=233$, TKR $n=231$). Patients in both groups had superior outcomes relative to their preoperative baseline status.

OKS had substantially changed in both groups: we found a mean change in OKS of more than 18 points for both procedures.

At the 5-year follow-up, we found no difference in OKS between groups (mean difference 1.04, 95% CI -0.42 to 2.50 ; $p=0.159$; table 2; figure 2). After 5 years, the EQ-5D visual analogue scale score was better in those who received PKR than those who received TKR (4.02, 1.36 to 6.67; $p=0.004$). However, we found no differences in the other secondary outcome scores at 5 years. Patient satisfaction after 5 years did not differ between the groups (190 [82%] of 233 patients receiving PKR vs 173 [77%] of 225 patients receiving TKR; RR 1.06, 95% CI 0.99 to 1.13; $p=0.097$; table 3). However, we found differences at 5 years in the results of the two questions asking about whether the patient's knee was better to before their operation (219 [95%] of 230 PKR patients vs 200 [90%] of 222 TKR patients; 1.06, 1.01 to 1.11; $p=0.016$) and whether the patient would have the operation again (208 [91%] of 228 patients vs 183 [84%] of 217 patients; 1.08, 1.02 to 1.15; $p=0.010$). At the 5-year analysis, OKS did not differ between the surgeon groupings (ie, expertise pair vs equipose surgeons; interaction effect -1.31 , 95% CI -5.34 to 2.73 ; $p=0.518$), nor between the age band, sex, or OKS band subgroups (appendix p 14). At 5 years, there was also no effect of surgeon experience, as defined by the number of procedures of each operation type performed by the surgeons before the study (appendix p 11).

Patients receiving TKR had more complications than those receiving PKR (73 [27%] of 269 patients vs 48 [20%] of 245 patients; RR 0.72, 95% CI 0.53–0.98; $p=0.036$; appendix pp 4–6). The most common complications were unexplained pain (nine patients receiving PKR vs eight patients receiving TKR), knee stiffness (0 patients vs ten patients) and both unexplained pain and knee stiffness together (one patient vs six patients). The number of re-operations and revisions were similar between groups (table 4). Of those who had surgery (ie, as treated), 15 (6%) of 245 patients receiving PKR had a re-operation and 21 (8%) of 269 patients receiving TKR had a re-operation (0.75, 0.37–1.53; $p=0.432$). After both surgeries, ten (4%) patients had revisions, predominantly because of unexplained pain (two patients receiving PKR vs five patients receiving TKR) or bearing dislocation (three patients vs 0 patients). When we analysed the revision and complication results with intention-to-treat (ie, as allocated) data, we found similar results to those of the as-treated analyses. There were 28 failures of PKR and 38 failures of TKR, as defined by the composite outcome, which was a non-significant difference (0.74, 0.51–1.08; $p=0.118$). During the 5-year follow-up, 17 deaths were reported: six (2%) deaths in the PKR group and 11 (4%) deaths in the TKR group (appendix p 12). These deaths were predominantly due to cancer (four deaths in the PKR group vs six deaths in the TKR group) and there were two deaths in each group for unknown reasons.

	Partial knee replacement		Total knee replacement		Risk ratio (95% CI)	p value
	n (%)	Valid N	n (%)	Valid N		
Satisfied with knee	190 (82%)	233	173 (77%)	225	1.06 (0.99–1.13)	0.097
Knee better or worse, now versus before surgery	219 (95%)	230	200 (90%)	222	1.06 (1.01–1.11)	0.016
Would choose to have knee operation again	208 (91%)	228	183 (84%)	217	1.08 (1.02–1.15)	0.010

Valid N data are the number who responded to the question.

Table 3: Estimated treatment effect on self-reported anchor questions at 5 years

	Treatment received		Treatment as allocated	
	PKR (n=245)	TKR (n=269)	PKR (n=263)	TKR (n=251)
Number of participants receiving a re-operation	15 (6%)	21 (8%)	14 (5%)	22 (9%)
Number of re-operations	22 (9%)	28 (10%)	18 (7%)	32 (13%)
Number of participants with revision	10 (4%)	10 (4%)	8 (3%)	12 (5%)
Unexplained pain	2	5	1	6
Bearing dislocation	3	0	3	0
Device loosening (tibia)	2	0	1	1
Bearing dislocation and unexplained pain	1	0	1	0
Ligamentous instability	1	0	1	0
Infection	1	0	1	0
Infection and mechanical failure	0	1	0	1
Knee stiffness and unexplained pain	0	1	0	1
Ligamentous instability and malalignment	0	1	0	1
Unknown	0	2	0	2
Number of other procedures	12	18	10	20
MUA	0	12	1	11
Aspiration	3	2	3	2
Arthroscopy	3	0	3	0
Arthroscopy and debridement or washout	1	1	1	1
Debridement, exploration, or washout	2	0	0	2
Open reduction and internal fixation	1	0	1	0
Marcaine injection and MUA	2	0	1	1
Arthroscopy and MUA	0	1	0	1
Arthroscopy and biopsy	0	1	0	1
Arthroscopy and partial medial meniscectomy	0	1	0	1

Data are n or n (%). PKR=partial knee replacement. TKR=total knee replacement. MUA=manipulation under anaesthetic.

Table 4: Number of re-operations, including revisions, by treatment received and as allocated at 5 years

In our within-trial cost-effectiveness analysis, we found that PKR was more effective (0.240 additional QALYs, 95% CI 0.046 to 0.434) and less expensive than TKR (−£910, 95% CI −1503 to −317) during the 5 years of follow-up (figure 3). The probability that PKR was the most cost-effective option was more than 99.9% for all reasonable threshold values. This finding was a result of better outcome, as measured by QALYs derived from EQ-5D scores (appendix p 7) and survival, lower costs of PKR surgery, and lower follow-up health-care costs with PKR than TKR. Even assuming equal costs of the implant device, PKR was less costly and more effective than TKR. The proportion of missing data imputed to inform the cost-effectiveness analysis was 12% per period, on average (appendix p 8). Index surgery costs and cost effectiveness comparisons are given in the appendix (pp 9–10).

The planned secondary analysis of the unadjusted OKS showed similar results to the primary outcome analysis, as did a marginal estimate over the whole 5-year follow-up (mean difference 1.39, 95% CI −0.12 to 2.90; $p=0.071$). The complier-average causal effect analysis at 5 years showed no difference (0.86, −0.62 to 2.34; $p=0.254$). The AUC of the OKS across the whole study period significantly favoured PKR versus TKR. The mean AUC was 36.6 (SD 8.3) for 233 participants in the PKR group and 35.1 (9.1) for the 231 participants in the TKR group (mean difference 1.54, 95% CI 0.07 to 3.01; $p=0.040$); notably, this difference is smaller than the revised estimated minimal important difference³¹ of 5 OKS points. A kernel density plot for OKS at baseline and 5 years is shown in the appendix (p 13).

Discussion

To our knowledge, our study is the largest multicentre randomised trial to compare TKR with PKR, and it provides the most robust evidence to date on the comparative effectiveness of the two approaches for the management of medial compartmental osteoarthritis. Our trial has high generalisability, since it involves many surgeons from several centres across the UK, and the hybrid design uniquely includes surgeons both in and out of equipoise, which allowed them to perform the surgery of their stated preference and expertise.

We found that both interventions (PKR and TKR) provided medium-term benefits for patients with medial compartment osteoarthritis by producing substantial changes in outcomes (more than 18 OKS points with both surgeries). Despite no differences in the primary outcome (on cross-sectional analysis), there was consistent evidence overall to suggest a modest clinical benefit of PKR over TKR. The analysis of the OKS AUC and patient-directed questions about patient satisfaction, perceived change in their knee problem, and willingness to choose to have the operation again all support this position. The duration of hospital stay was also lower with PKR.

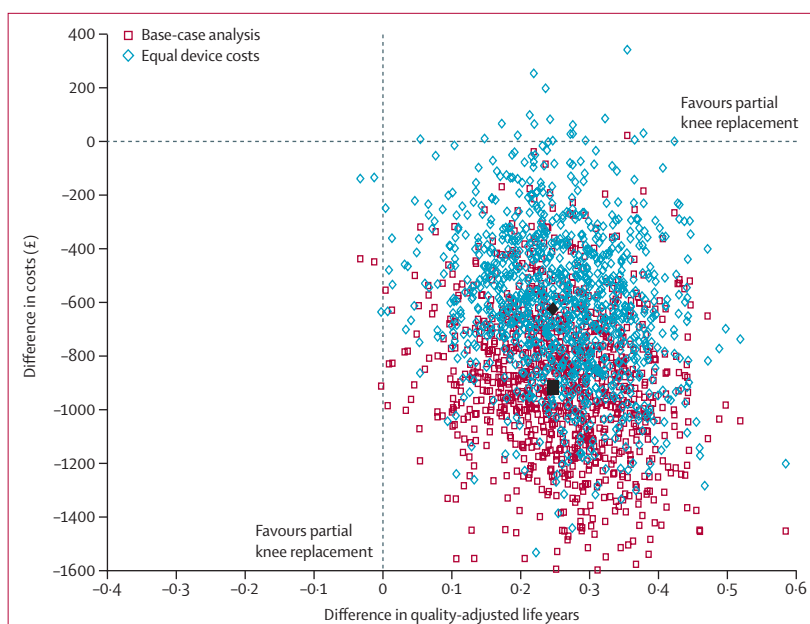


Figure 3: Cost-effectiveness of the base-case analysis versus an assumption of equal implant device costs. Large black dots are the mean values for each scenario.

Despite low power, we observed no differences between the number of re-operations and revisions between the trial groups. By contrast, nearly all global registries (which comprise non-randomised, observational data) report that PKR results in a substantially larger number of revisions than TKR (usually more than 6%).^{17,19} In TOPKAT, the frequency of complications was also observed to be increased in the TKR group. Although interpretation of these event findings in each study type is not straightforward, it is likely that non-randomised cohort data, despite matching, remain susceptible to selection bias (specific patients chosen for a particular implant). The difference might also reflect a lower threshold for revision of PKR outside of a controlled trial. However, Hawthorne-type effects and self-selection of the involved surgeons in the trial also cannot be excluded as reasoning for the lower frequency of revisions found with PKR in TOPKAT relative to other studies. It is important to emphasise that the revision and re-operation data are more susceptible to bias than some of the other variables. The sample size calculations were adequate to detect a change in the proportion of overall re-operations of 7% between groups (from 5% to 12%) with 80% power, and they might therefore be underpowered to provide concrete evidence on revisions. The anticipated differences for the power calculation were much larger than the differences in the revision frequencies found in our study. Furthermore, differences in the number of revisions between subgroups (eg, male vs female) would also remain undetected with the sample size that we used. Some of our most notable findings were in relation to the health economics data. During 5 years of follow-up in the TOPKAT trial, PKR was found

to be more effective but also less costly than TKR. Lower costs with PKR were a result of lower costs of the index surgery (even when assuming equal device costs) and lower subsequent use of health-care services, especially due to fewer outpatient visits in the first 2 years after the index surgery. These data will have most relevance for the UK health-care system, but the findings are likely to be generalisable to settings with similar health-care systems.

There are many general strengths of this study. Globally, to our knowledge, our study is the largest, most complete, and longest randomised follow-up of any comparison of TKR versus PKR to date. Our study was well powered and any type 2 errors are highly unlikely. The external validity is high, and the pragmatic design lends itself to policy support or commissioning changes. The follow-up response to the questionnaires was very high (94% for all; 88% for the primary outcome). Our study included an expertise design, which accounted for surgeons who were not in the equipoise group. Our trial is unique in that it can provide strong empirical evidence to compare with large database studies evaluating TKR and PKR, and it could highlight the limitations of relying on cohort or non-randomised systematically collected data alone. An important strength is the unbiased and independent review of our interpretation.

The weaknesses of the study include the absence of masking, some missing data for the health-care resource and EQ-5D data (12% across all time periods), and some non-compliance. It was not feasible to keep patients masked from allocation for 5 years; expectancy bias (regarding a higher benefit of PKR) cannot therefore be entirely ruled out. The missing data was equal between groups, and it was found to have no effect on results; no imputation method was required for the clinical analysis but one was used for the health economics analysis. Some non-compliance with allocation (ie, conversion from PKR to TKR from operative findings) was anticipated and clinically understandable but a complier-average causal effects analysis that we used to assess the potential effects of non-compliance on outcomes did not find any issue. The number of participants who completed follow-up was just less than the power calculation of the target of 500 participants, and this disparity is therefore unlikely to affect the results, particularly since the main analyses benefited from greater precision (not accounted for in the original sample size) because of adjustment for minimisation variables and, where relevant, also the baseline score. It is also acknowledged that the 5-year timeframe prompts some limitations. First, the within-trial cost-effectiveness analysis was not designed to accommodate for the economic implications of potential future revisions; thus the longer term implications of any potential future changes in the revision profile have not been accounted for. Second, patients in both groups are expecting the effects of their knee surgery to last longer than 5 years,

so major differences that occur later will also remain undetected. However, a 10-year follow-up has commenced and will provide a clearer picture on these data.

The magnitude of the difference observed in the primary outcome lies within the acknowledged minimally important difference in OKS at the time of study planning (a 2-point difference), although the 95% CI did not rule out the possibility that the effect of PKR could exceed 2 points. Irrespective, the observed difference is well within the more recently accepted minimally important difference in OKS (now set at 5 points).³¹ This scale of difference in OKS is unlikely to have any discernible impact on national or international practice by itself, but clinicians and patients themselves must decide whether these small magnitudes of difference, and the evidence from other variables, is sufficient to modify or support their own individual decision making. A kernel density analysis (appendix p 13) showed that there were more patients with OKS outcome scores of 40 and above in the PKR group than in the TKR group, but this finding was balanced by a greater number of TKR patients with mid-range scores. Hence, there is disagreement as to whether these data imply an important clinical signal for efficacy or a more contrived outcome to demonstrate differences. The re-operation and revision data are of note, suggesting that neither operation is associated with a high frequency of revision, a finding in contrast with previous reports that used cohort databases. The cost-effectiveness findings are undoubtedly the strongest marker, and they suggest that PKR is superior in a health-care provision context.

Both TKR and PKR remain treatment options for late-stage medial compartment arthritis of the knee. PKR should, however, be considered the first-choice procedure (based on 5-year outcomes). That said, the operation should only be done by those surgeons with sufficient experience and expertise.

Both TKR and PKR are effective, offer similar clinical outcomes, and result in similar frequencies of re-operations and complications. The number of revisions after PKR in our randomised trial were substantially lower than found in other, mainly non-randomised, cohort data. Assessments focusing on whether the patients' knees felt better after the surgery and whether they would have the surgery again were better with PKR. And, of note, PKR was more cost-effective than TKR after 5 years. Although TKR provides good outcomes, based on the combined 5-year cost-effectiveness and clinical data, PKR should likely be considered the first choice for patients with late-stage isolated medial compartment osteoarthritis.

Contributors

DJB, LJD, JAC, AP, AC, HC, RF, NA, DM, and MKC designed and developed the trial protocol. LJD was the trial manager, GM oversaw the statistical analysis, JH conducted the statistical analysis, and SK conducted the health economics analysis. GM, SK, JH, and JL had access to all the data. DJB, JAC, GM, AC, JL, and MKC interpreted the data. DJB wrote the initial draft of the manuscript and subsequent revisions

with assistance from LJD, JAC, GM, SK, JH, JL, and MKC. All authors read, provided input on, and approved the final manuscript. DJB and MKC were responsible for the decision to submit the manuscript.

TOPKAT Study Group members

Data Monitoring Committee: Karen Barker (University of Oxford), Gordon Murray, and Hamish Simpson (University of Edinburgh); Trial Steering Committee: Donna Dodwell, Simon Donell (University of East Anglia), Shawn Tavares (Royal Berkshire Hospital), Jonathan Waite (South Warwickshire NHS Foundation Trust); Trial Management Group: Nigel Arden, David Beard, Helen Campbell, Andrew Carr, Cushla Cooper, Loretta Davies, Helen Doll, Seamus Kent, Jose Leal, Francesco Fusco (until 2018), Ray Fitzpatrick, David Murray, and Andrew Price (University of Oxford), Jonathan Cook (University of Aberdeen and University of Oxford), Marion Campbell, Gordon Fernie, Alison McDonald, Anne Duncan, Mayret Castillo (until 2011; University of Aberdeen); Central Trial Office (University of Oxford): David Beard, Jonathan Cook, Cushla Cooper, Loretta Davies, Akiko Greshon, Kay Holland (until 2017), Jiyang Li (until 2018), Elena Rabiotti (until 2013), Sandra Regan (until 2012), and Victoria Stalker (until 2014); Trial Office (University of Aberdeen): Gordon Fernie, Mark Forrest, Graeme MacLennan, Alison McDonald, Gladys McPherson, Charles Boachie (until 2013), Marion Campbell, Mayret Castillo (until 2011), Diana Collins (until 2013), Janice Cruden, Anne Duncan (until 2017), Sophie Halpin (until 2015), and Beverley Smith (until 2017), and the database support team; Interpretation group: David Torgerson (University of York), Chris Maher (University of Sydney), Peter Brownson (Royal Liverpool and Broadgreen University Hospitals NHS Trust), Simon Donell (University of East Anglia), Mark Mullins (Abertawe Bro Morgannwg University Health Board), and Jane Blazeby (University of Bristol); and all principal investigators (PIs) and research teams: Ruth Jenkins, Mark Lewis (PI), and Witek Mintowt-Cyzyk (Aneurin Bevan University Health Board, Royal Gwent Hospital); David Beverland (PI), Leeann Bryce, Julie Catney, Ian Dobie, Emer Doran, and Seamus O'Brien (Belfast Health and Social Care Trust, Musgrove Park Hospital, Belfast); Fazal Ali, Heather Cripps, Amanda Whileman, Phil Williams (PI), and Julie Toms (Chesterfield Royal Hospital NHS Foundation Trust); Ellen Brown, Gillian Horner, Andrew Jennings (PI), and Glynis Rose (County Durham and Darlington NHS Foundation Trust); Frances Bamford, Wendy Goddard, Hans Marynissen (PI), Haleh Peel, and Lyndsey Richards (East Lancashire Hospitals NHS Trust, Royal Blackburn Hospital); Amanda Bell, Sunny Deo, Sarah Grayland, David Hollinghurst, Suzannah Pegler, Venkat Satish (PI), and Claire Woodruffe (Great Western Hospitals NHS Foundation Trust, Swindon); Nick London (PI), David Duffy, Caroline Bennett, and James Featherstone (Harrogate and District NHS Foundation Trust, Harrogate); Joss Cook, Kim Dearnley, Nagarajan Muthukumar (PI), Laura Onuoha, and Sarah Wilson (Hull & East Yorkshire Hospitals NHS Trust); Sandhu Banher, Eunice Emeakaroha, Jamie Horohan, Sunil Jain (PI), and Susan Thompson (Maidstone and Tunbridge Wells NHS Trust, Medway); Sarah Buckley, Aaron Ng (PI), Ajit Shetty, and Karen Simeson (Mid Yorkshire Hospitals NHS Trust); Julian Flynn, Meryl Newsom, Cheryl Padilla-Harris, and Oliver Pearce (PI) (Milton Keynes University Hospital NHS Foundation Trust); James Bidwell (PI), Alison Innes, Winifred Culley, Bill Ledingham, and Janis Stephen (NHS Grampian, Woodend Hospital, Aberdeen); Rachel Bray, Hywel Davies, Debbie Delgado, Jonathan Eldridge, Leigh Morrison, James Murray (PI), Andrew Porteous, and James Robinson (North Bristol NHS Trust); Matt Dawson (PI), Raj Dharmarajan, David Elson, Will Hage, Nicci Kelsall, and Mike Orr (North Cumbria University Hospitals NHS Trust, Carlisle); Jackie Grosvenor, SS Maheswaran (PI), Claire McCue, Hemanth Venkatesh, Michelle Wild, and Deborah Wilson (North Tees and Hartlepool NHS Foundation Trust, Stockton-On-Tees); Chris Dodd, William Jackson (PI), Pam Lovegrove, David Murray, Jennifer Piper, and Andrew Price (Oxford University Hospitals NHS Trust, Nuffield Orthopaedic Centre); Neil Bradbury, Lucy Clark, Stefanie Duncan, Genevieve Simpson, and Allister Trezies (PI) (Royal United Hospitals Bath NHS Foundation Trust, Bath); Vikram Desai (PI), Cheryl Heeley, Kramer Guy, and Rosalyn Jackson (Sherwood

Forest Hospitals NHS Foundation Trust, Kings Mill Hospital, Sutton in Ashfield); Alan Hall, Gordon Higgins (PI), Michael Hockings, David Isaac, and Pauline Mercer (South Devon Healthcare NHS Foundation Trust, Torbay); Lindsey Barber, Helen Cochrane, Janette Curtis, Julie Grindey, David Johnson (PI), and Phil Turner (Stockport NHS Foundation Trust, Stockport); David Houlihan-Burne (PI), Briony Hill, Ron Langstaff, and Mariam Nasser (The Hillingdon Hospitals NHS Trust); Mark Bowditch, Chris Martin, Steven Pryke, Bally Purewal, Chris Servant (PI), Sheeba Suresh, and Claire Tricker (The Ipswich Hospital NHS Trust, Ipswich); Robert Ashford, Manjit Attwal, Jeanette Bunga, Urjit Chatterji, Susan Cockburn, Colin Esler (PI), Steven Godsiff, Tim Green, Christina Haines, and Subash Tandon (University Hospitals of Leicester NHS Trust, Leicester); Racquel Carpio, Sarah Griffiths, Natalie Grocott, and Ian dos Remedios (PI) (University Hospitals of North Midlands NHS Trust, Stoke on Trent); David Barrett, Phil Chapman-Sheath, Caroline Grabau, Jane Moghul, William Tice (PI), and Catherine Trevithick (University Hospital Southampton NHS Foundation Trust); Rajiv Deshmukh, Mandy Howes, Kimberley Nethererton, Dipak Raj (PI), and Nikki Travis (United Lincolnshire Hospitals NHS Trust, Boston); Mohammad Maqsood, Rebecca Norton, Farzana Rashid, Alison Raynor, Mark Rowell, and Karen Warner (United Lincolnshire Hospitals NHS Trust, Lincoln).

Declaration of interests

DJB reports institutional research grant funding from the National Institute for Health Research (NIHR) and Zimmer Biomet, outside the submitted work. JAC reports grants from the NIHR Health Technology Assessment (HTA) programme and was a member of the NIHR HTA efficient trial designs board for 2 years, during the conduct of the study. GM reports grants from NIHR HTA, during the conduct of the study. AP reports consultancy fees from Zimmer Biomet, DePuy Synthes, and Smith & Nephew, and grants from NIHR and AR United Kingdom, outside the submitted work. HC reports grants from NIHR, during the conduct of the study. RF reports membership of the HTA Prioritisation Group and HTA National Stakeholder Advisory Group. NA reports grants from Merck and consultancy fees from Merck, Flexion Therapeutics, Freshfields Bruckhaus Deringer, Regeneron Pharmaceuticals, Pfizer, and Eli Lilly, outside the submitted work. DM reports grants and personal fees from Zimmer Biomet, outside the submitted work. DM also reports receiving royalties related to The Oxford Partial Knee system (by Zimmer Biomet), and from other patents relating to knee replacement. All other authors declare no competing interests.

Data sharing

The data collected for the study, including individual participant data and a data dictionary defining each field in the set, will be made available to researchers on request to the study team and with appropriate reason when accompanied by a peer-reviewed protocol, with publication and on agreement of the Trial Steering Committee. The shared data will be deidentified participant data. Data will be shared with investigator support, after approval of a proposal, with a signed data access agreement. The study protocol, statistical analysis plan, and informed consent form are available online.

Acknowledgments

The views and opinions expressed are those of the authors and do not necessarily reflect those of the Health Technology Assessment (HTA) Programme, the UK National Institute of Health Research (NIHR), National Health Service or Department of Health. The TOPKAT study is funded by the NIHR HTA Programme (number HTA 08/14/08), sponsored by the University of Oxford, and supported by Oxford Surgical Intervention Trials Unit (SITU; supported by Oxford NIHR Biomedical Research Centre) in the Royal College of Surgeons Surgical Trials Initiative. Study management was divided between the SITU (Oxford) and the Aberdeen trials centre, the Centre for Healthcare Randomised Trials. JAC held a Medical Research Council Methodology Fellowship (G1002292) for part of the study. The Health Services Research Unit is core funded by the chief scientist office of the Scottish Government Health and Social Care Directorates. We would like to thank the principal investigators and their teams at each of the TOPKAT sites.

For study documents see <https://www.journalslibrary.nihr.ac.uk/programmes/hta/081408/#>

References

- 1 National Joint Registry. 15th annual report. 2018. <http://www.njrreports.org.uk/Portals/0/PDFdownloads/NJR%2015th%20Annual%20Report%202018.pdf> (accessed Dec 27, 2018).
- 2 Price AJ, Alvand A, Troelsen A, et al. Knee replacement. *Lancet* 2018; **392**: 1672–82.
- 3 Beard DJ, Holt MD, Mullins MM, Malek S, Massa E, Price AJ. Decision making for knee replacement: variation in treatment choice for late stage medial compartment osteoarthritis. *Knee* 2012; **19**: 886–89.
- 4 Willis-Owen CA, Brust K, Alsop H, Miraldo M, Cobb JP. Unicompartmental knee arthroplasty in the UK National Health Service: an analysis of candidacy, outcome and cost efficacy. *Knee* 2009; **16**: 473–78.
- 5 Newman J, Pydisetty RV, Ackroyd C. Unicompartmental or total knee replacement: the 15-year results of a prospective randomised controlled trial. *J Bone Joint Surg Br* 2009; **91**: 52–57.
- 6 Newman JH, Ackroyd CE, Shah NA. Unicompartmental or total knee replacement? Five-year results of a prospective, randomised trial of 102 osteoarthritic knees with unicompartmental arthritis. *J Bone Joint Surg Br* 1998; **80**: 862–65.
- 7 Costa CR, Johnson AJ, Mont MA, Bonutti PM. Unicompartmental and total knee arthroplasty in the same patient. *J Knee Surg* 2011; **24**: 273–78.
- 8 Sun PF, Jia YH. Mobile bearing UKA compared to fixed bearing TKA: a randomized prospective study. *Knee* 2012; **19**: 103–06.
- 9 Kulshrestha V, Datta B, Kumar S, Mittal G. Outcome of unicompartmental knee arthroplasty vs total knee arthroplasty for early medial compartment arthritis: a randomized study. *J Arthroplasty* 2017; **32**: 1460–69.
- 10 Padgett DE, Stern SH, Insall JN. Revision total knee arthroplasty for failed unicompartmental replacement. *J Bone Joint Surg Am* 1991; **73**: 186–90.
- 11 Saldanha KA, Keys GW, Svard UC, White SH, Rao C. Revision of Oxford medial unicompartmental knee arthroplasty to total knee arthroplasty—results of a multicentre study. *Knee* 2007; **14**: 275–79.
- 12 Arirachakaran A, Choowit P, Putananon C, Muangsiri S, Kongtharvonskul J. Is unicompartmental knee arthroplasty (UKA) superior to total knee arthroplasty (TKA)? A systematic review and meta-analysis of randomized controlled trial. *Eur J Orthop Surg Traumatol* 2015; **25**: 799–806.
- 13 Johnson S, Jones P, Newman JH. The survivorship and results of total knee replacements converted from unicompartmental knee replacements. *Knee* 2007; **14**: 154–57.
- 14 Levine WN, Ozuna RM, Scott RD, Thornhill TS. Conversion of failed modern unicompartmental arthroplasty to total knee arthroplasty. *J Arthroplasty* 1996; **11**: 797–801.
- 15 Fisher DA, Watts M, Davis KE. Implant position in knee surgery: a comparison of minimally invasive, open unicompartmental, and total knee arthroplasty. *J Arthroplasty* 2003; **18** (suppl 1): 2–8.
- 16 Wilson HA, Middleton R, Abram SGF, et al. Patient relevant outcomes of unicompartmental versus total knee replacement: systematic review and meta-analysis. *BMJ* 2019; **364**: 1352.
- 17 Chawla H, van der List JP, Christ AB, Sobrero MR, Zuiderbaan HA, Pearle AD. Annual revision rates of partial versus total knee arthroplasty: a comparative meta-analysis. *Knee* 2017; **24**: 179–90.
- 18 Migliorini F, Tingart M, Niewiera M, Rath B, Eschweiler J. Unicompartmental versus total knee arthroplasty for knee osteoarthritis. *Eur J Orthop Surg Traumatol* 2018; **29**: 947–55.
- 19 Liddle AD, Judge A, Pandit H, Murray DW. Adverse outcomes after total and unicompartmental knee replacement in 101 330 matched patients: a study of data from the National Joint Registry for England and Wales. *Lancet* 2014; **384**: 1437–45.
- 20 Beard D, Price A, Cook J, et al. Total or partial knee arthroplasty trial—TOPKAT: study protocol for a randomised controlled trial. *Trials* 2013; **14**: 292.
- 21 Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ* 2005; **330**: 88.
- 22 Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res* 1989; **248**: 13–14.
- 23 Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br* 1998; **80**: 63–69.
- 24 Harris K, Dawson J, Gibbons E, et al. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. *Patient Relat Outcome Meas* 2016; **7**: 101–08.
- 25 Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. *J Arthroplasty* 1998; **13**: 890–95.
- 26 Talbot S, Hooper G, Stokes A, Zordan R. Use of a new high-activity arthroplasty score to assess function of young patients with total hip or knee arthroplasty. *J Arthroplasty* 2010; **25**: 268–73.
- 27 Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53–72.
- 28 Robertsson O, Dunbar M, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27 372 knees operated on between 1981 and 1995 in Sweden. *Acta Orthop Scand* 2000; **71**: 262–67.
- 29 Murray DW, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br* 2007; **89**: 1010–14.
- 30 Price AJ, Webb J, Topf H, et al. Rapid recovery after Oxford unicompartmental arthroplasty through a short incision. *J Arthroplasty* 2001; **16**: 970–76.
- 31 Beard DJ, Harris K, Dawson J, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol* 2015; **68**: 73–79.
- 32 Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 2014; **32**: 1157–70.