Patient reported outcomes following total and unicompartmental knee replacement: a study of 14,076 matched patients from the National Joint Registry for England and Wales

Alexander D. Liddle BSc MRCS
Arthritis Research UK Clinical Research Fellow
Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Hemant Pandit DPhil FRCS(Orth)
Orthopaedic Surgeon and Honorary Senior Lecturer
Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Andrew Judge PhD
Senior statistician and Departmental lecturer in musculoskeletal epidemiology
NIHR Musculoskeletal Biomedical Research Unit,
University of Oxford, and
MRC Lifecourse Epidemiology Unit, University of Southampton.

Prof. David W. Murray MD FRCS✉
Professor of Orthopaedic Surgery
Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Address (all authors): Botnar Research Centre
Windmill Road
Oxford OX3 7LD

Correspondence: **Prof. DW Murray**
Nuffield Dept. of Orthopaedics, Rheumatology and Musculoskeletal Sciences
Botnar Research Centre
Windmill Road
Oxford OX3 7LD

Tel: 01865 227457
Fax: 01865 227671
david.murray@ndorms.ox.ac.uk

Patient reported outcomes following total and unicompartmental knee replacement: a study of 14,076 matched patients from the National Joint Registry for England and Wales

Abstract

The use of total (TKR) or unicompartmental knee replacement (UKR) for end stage knee osteoarthritis is controversial. Whilst UKR provides faster recoveries, lower rates of morbidity and mortality, and fewer complications, the long-term revision rate for TKR is substantially lower. The effect of each intervention on patient-reported outcome remains unclear. The aim of this study was to determine whether six month patient-reported outcome measures (PROMs) are superior in patients receiving TKR or UKR, using data from a large national joint registry (NJR).

We performed a propensity-score matched cohort study comparing six month PROMs following TKR and UKR in patients enrolled in the NJR for England and Wales, and the UK national PROM collection programme. 3,519 UKR patients were matched to 10,557 TKRs. Mean six-month PROMs favoured UKR: OKS was 37.7(95% CI 37.4 to 38.0) in UKR and 36.1(35.9 to 36.3) in TKR; EQ5D index was 0.772 (0.764-0.780) in UKR and 0.751 (0.747-0.756) in TKR. UKR patients were more likely to attain excellent results (OR 1.59, 95% CI 1.47 to 1.72, $p<0.001$), and to be highly satisfied (OR 1.27, 95% CI 1.17 to 1.39, $p<0.001$), and were less likely to report complications than those receiving TKR. UKR has superior early patient-reported outcomes compared to TKR; these differences are most marked at the very best outcomes. Complications and readmission are more likely after TKR. Whilst the data presented concern short-term outcomes, they suggest that the high revision rate in UKR may not be due to poorer clinical outcomes. These factors should inform decision making in patients eligible for either procedure.

Patient reported outcomes following total and unicompartmental knee replacement: a study of 14,076 matched patients from the National Joint Registry for England and Wales

Introduction

Total knee replacement (TKR) is considered a reliable and successful treatment for end-stage osteoarthritis of the knee. National joint registries (NJRs) demonstrate revision rates below 5% at ten years with good functional improvement overall[1-3]. However, outcome varies greatly between patients, and up to 20% are dissatisfied with their result following TKR[1].

In a large proportion of patients (up to 47% by some estimates)[4], disease is restricted to either the medial or lateral tibiofemoral joints. These cases may be suitable for unicompartmental knee replacement (UKR)[5]. In UKR, only the parts of the knee affected by osteoarthritis are replaced, preserving the cruciate ligaments and the remaining, healthy joint surfaces. By retaining these structures, UKR restores the ligament-driven kinematics that exist within the native knee[6]. UKR requires a smaller incision than TKR, and is associated with shorter hospital stays and lower rates of infection, morbidity and mortality[7,8]; as a result, it is more cost-effective than TKR[4].

However, all major NJRs report that the revision rate of UKR is two to three times higher in unadjusted analyses compared to TKR[2,9,10]. As a result, surgeons are reluctant to offer UKR and significantly fewer patients receive UKR than are eligible on the basis of disease pattern[2,9]. The high revision rates in NJRs contrast with low long-term revision rates reported by high-volume centres[11,12]. Surgeon factors, patient selection and different thresholds for revision are likely to contribute to these differences[13].

The difference in revision rate may be interpreted as a manifestation of inferior functional and pain outcomes in UKR. However, it is not clear if this is the case. Fair comparison is difficult: whilst single-

centre studies report superior outcomes in UKR, patients offered UKR are generally younger and have superior function at baseline (a situation known as confounding by indication). As a result, it is not clear which intervention offers the best patient-reported outcomes. This uncertainty, coupled with the higher reported revision rate, has led some to question the continuing use of UKR[14].

The aim of this study is to compare patient-reported outcomes following TKR and UKR, using data from the NJR for England and Wales and the National Health Service (NHS) Patient Reported Outcome Measures (PROM) database, using propensity score matching methods to minimise confounding by indication.

Methods

Source of Data

The NJR for England and Wales was established in 2003 and collects information on all hip and knee replacements nationwide; it is the largest such registry in the world. Data are entered using contemporaneous surgeon-completed forms; revision operations are linked using unique patient identifiers.

PROMs have been collected since 2009 for all NHS-funded hip and knee replacements performed in England. The programme is co-ordinated by the Health and Social Care Information Centre on behalf of the Department of Health. Amongst other information, a pre- and post-operative quality of life questionnaire (the EuroQol 5 Domain, EQ5D[15]) and a procedure-specific questionnaire (the Oxford Knee Score (OKS), in the case of knee replacements[16]), are collected. Whilst failure data are collected into the long-term, PROM data are only collected pre-operatively and at six months. The choice of these intervals was a compromise between having appropriate proximity to surgery (to provide timely feedback and to avoid 'contamination' with non-operation related factors), whilst providing long enough follow-up to allow comparison between groups of patients having accounted for the

post-operative recovery period. The great majority of improvement in PROMs following joint replacement occurs in the first six months, with only minor improvements between six months and one year[17]. Long-term studies of TKR and UKR have demonstrated that PROMs remain relatively constant following this early rapid improvement, at least up to the tenth postoperative year[10,12,18].

Data linkage

552,015 knee replacements were recorded in the NJR from the beginning of data collection to 28th August 2012. Following exclusions of knee replacements other than TKR and UKR, those not performed for osteoarthritis, and 'complex primary' procedures, a total of 503,850 cases went forward to linkage. PROM records were linked to the NJR using standard patient identifiers (NHS numbers and patient demographics) by the Health and Social Care Information Centre (HSCIC) data linkage service. Those records were then linked to specific procedures on the basis of operation date (+/- one day). As in previous studies of this dataset[14], cases were excluded if either the pre- or post-operative PROM questionnaire was missing, if the questionnaires were not dated, or were performed too early (>90 days prior to surgery or <180 days following surgery) or too late (more than one year following surgery). Whilst 215,203 records were contained within the PROMs dataset, there were a large number of duplicate and incomplete records which were excluded. 89,132 complete records were available for linkage with the NJR dataset, of which 58,863 could be linked. A study flow-chart is given in Figure 1.

Outcome measures

Primary outcome measures were the six-month OKS and EQ5D. Secondary outcome measures included patient-reported satisfaction, complications, readmission and reoperations.

The OKS has twelve items relating to knee pain and function, each ranging from 0-4, and presented as an overall score on an ordinal scale between 0-48[19]. The distribution of the post-operative OKS is frequently skewed (due to the high proportion of good results following joint replacement)[19], which raises difficulties with interpretation of mean post-operative scores. To address this, the pre- and post-operative distributions of the OKS are presented graphically. Alongside the mean scores, the

proportion attaining excellent (>41 points on the OKS), good (34-41), fair (27-33) and poor (<27) results are given. These subdivisions (proposed by Kalairajah *et al*) are in widespread use[13,20]. The Minimal Clinically Important Difference (MCID) for the OKS has not been definitively quantified; it is thought to be between 3 and 5 points but may be smaller than this[19]. To distinguish the proportion of patients reporting no clinically important change in OKS in each group, a putative MCID of ≥4 points has been used.

The EQ5D comprises five questions concerning mobility, self-care, activities of daily living, pain and anxiety/depression, each with three levels (in ascending order: no problems, some problems, severe problems). These answers can be presented separately or as a weighted overall index rated from 1 (perfect health) to -0.594, the worst possible state[15,21], a visual analogue scale (VAS) of overall health is presented separately. The distribution of the overall EQ5D index again renders comparison of means problematic. Alongside means, individual scores are presented and the proportion of patients with no, some, or severe problems on each question, as well as the proportion of patients who are better, worse, the same, or non-comparable overall are compared (so called Paretian analysis)[21].

Both questionnaires are completed pre-operatively and at six months post-operatively. The pre-operative questionnaire is generally completed in person at the surgical unit. The six month questionnaire is posted to the patient and returned in a pre-paid envelope to the contractor responsible for data collection, independently of the operating unit. The second questionnaire also allows the patient to report complications and re-admissions, to report whether they are better or worse than pre-operatively and to report their overall satisfaction with the procedure.

Confounders

The NJR contains data on confounders including patient factors (age, gender, American Society of Anaesthesiologists (ASA) score, and Body Mass Index (BMI)); surgeon/unit factors (cases performed per year, unit type (public/private hospital or Independent Sector Treatment Centre (ISTC)), grade of operating surgeon); and technical factors (such as cemented or cementless fixation). The PROM

dataset contains additional confounders including pre-operative PROMs, self-reported general health and disability, whether the patient is admitted from home or care, and self-reported co-morbidities. Details of confounders are given in Table 1. In common with other NJR studies, a large proportion of the BMI data were found to be missing (20,302 of 58,863 pre-matching observations, 34.5%)[7]. Therefore, the principal analyses were performed without BMI, and sensitivity analyses were performed using those cases with complete BMI data.

Statistical methods

When comparing interventions in clinical studies, confounding variables can exert an effect on the allocation of a patient to an intervention, as well as having a direct effect on outcome. In Randomised Controlled Trials (RCTs), this effect, known as confounding by indication, is eliminated by randomly allocating patients to each intervention. As this is an observational study, this is impossible, and confounding by indication was addressed using propensity score matching.

A logistic equation was fitted to assess the effect of each of the confounders on the likelihood of a patient receiving UKR[22]. From this model, each case was assigned a propensity score, representing the probability of receiving UKR. Each UKR was matched to three TKRs by propensity score using a calliper of 0.02 standard deviations of the logit of the propensity score[23]. The matched study group comprised 3,519 UKRs and 10,557 TKRs.

Linear regression was used to compare outcomes in TKR and UKR for continuous variables; logistic regression was used for dichotomous outcomes (improvement/non-improvement, and each stratum of the Kalairajah score). Stata IC v.12.1 for Windows (Stata Corp. College station, TX) was used for all analyses.

Results

Baseline demographic data of the patients before and after matching are displayed in Table 1. Substantial differences in age, preoperative scores and gender were corrected-for in the propensity-matched model, with balance being achieved following matching.

Overall, the mean OKS increased from 19.6 points (95% confidence interval 19.6-19.7) to 35.1 (95% CI 35.0-35.2). Before matching, the unadjusted postoperative score was 34.9 points (95% CI 34.9-35.1) for TKR and 37.7 (95% CI 37.4-38.0) for UKR ($p<0.001$).

Following matching, the mean six month OKS was 36.1 (95% CI 35.9-36.3) for TKR and 37.7 (95% CI 37.6-38.0) for UKR ($p<0.001$). The pre- and post-operative distribution of OKS scores within the two groups is demonstrated in figure 2. Significantly more UKRs than TKRs achieved the best scores: an 'excellent' Kalairajah score (OKS>41) was more common in UKR than TKR (odds ratio 1.59 (95% CI 1.47-1.73), $p<0.001$). 'Good' (OKS 34-41 inclusive), 'fair' (OKS 27-33) and 'poor' (OKS <27) results were more common in TKR (table 3). 89.6% of UKRs achieved the MCID of four OKS points, compared to 86.8% of TKRs. The odds ratio for achieving the MCID was 1.31 (95% CI 1.16-1.48) in favour of UKR. Using the same MCID, 1,233 patients (8.76% of the total) made no clinically important change by six months, and 530 (3.77%) had made a clinically important deterioration. Both non-improvement and deterioration were more likely in TKR: odds ratios were 0.77 (95% CI 0.67-0.89) and 0.78 (0.63-0.96) respectively.

The mean EQ5D index was 0.751 (95% CI 0.747-0.756) for TKR and 0.772 (95% CI 0.764-0.780) for UKR ($p<0.001$). There were significantly superior results for four of the five subscales (mobility, self-care, activity level and pain) favouring UKR. The remaining subscale, anxiety/depression, demonstrated no difference between patients undergoing TKR and UKR. Again, there were more non-improvers in the TKR group. There was no significant difference between the groups in terms of EQ5D VAS. Full outcome score details are displayed in tables 2, 3 and 4.

Patient-reported complications, re-operation and satisfaction are summarised in Table 3. Bleeding, wound problems and drug reactions were significantly more common following TKR. Patients were more likely to report readmission in the six months following TKR, but reoperation rates were comparable between the groups. Patients were more likely to report excellent satisfaction levels following UKR (OR 1.27, 95% CI 1.17-1.39, p<0.001). Around 89% of patients reported that they were better after the procedure than before, regardless of whether they had undergone TKR or UKR (Table 5).

In the main analysis, BMI has been excluded due to the large amount of missing data. As a sensitivity analysis, the analyses were repeated using complete case data with and without BMI, and the whole dataset with BMI completed using multiple imputation. Introducing BMI into the model, either in complete case or imputed datasets, had a minimal effect on the estimates produced. Given the difficulties with the use of MI in datasets with large proportions of missing data, the relative balance of the groups and the very small effect of the introduction of BMI into the model, it was elected not to include BMI in the models used in the main paper.

Discussion

This study demonstrates a large improvement in PROMs six months following both TKR and UKR compared to the pre-operative state. However, UKR resulted in significantly better early outcomes, in terms of both joint-specific outcome and health-related quality of life.

Whilst the overall mean difference in scores between the two procedures was small (with the mean OKS following UKR being 1.6 points greater than that after TKR), the skewed nature of the outcome scores following two such successful procedures, together with the ceiling effect of the OKS limits the usefulness of the comparison of means. The clinical difference between the two procedures was particularly striking in terms of the attainment of the highest scores: patients undergoing UKR had

an odds ratio of 1.59 for the attainment of an “excellent” score. These findings are supported by an odds ratio of 1.27 for the attainment of excellent satisfaction as well as a significantly higher rate of patient-reported complications with UKR. The likelihood of not improving at all following surgery was significantly greater following TKR (odds ratio 1.31).

The strengths of this study include the large study population. The use of an unselected, registry-based sample leads to high external validity. Using propensity score techniques to match on multiple confounders (including preoperative functional status) overcomes the difficulties experienced by previous authors in analysing comparable patients; baseline matching obviates the need for the comparison of change scores or the adjustment for baseline variables within the final regression model.

Limitations include the short follow-up (although the reasons for selecting six months as the primary outcome in the national PROM programme are given in the methods section[14]). Whilst propensity score techniques substantially reduce the chance of confounding by indication, unmeasured confounding may still exist. Whilst multiple confounders are included in the matching process, other factors (such as more subtle differences in co-morbidity or functional status, more detailed implant factors including brand and pre-operative activity level) may represent sources of unmeasured confounding[24]. Whilst patients are matched on the basis of clinical disease severity, the radiological state of disease cannot be known without examining pre-operative radiographs, which is not feasible in this setting. Many of these limitations are inherent to an observational study design. An RCT, the total or partial knee arthroplasty trial (TOPKAT) is in progress; this should provide important evidence for the relative functional outcomes of the two procedures without these limitations[25].

The selection of patients via propensity score matching raises questions about the generalisability of the findings; however, fewer than 5% of the UKRs were excluded as unmatchable so this is unlikely to be a substantial problem. The large proportion of incomplete, duplicate or un-linkable PROM data has been noted by other authors[26], and represents a limitation, raising the possibility of selection

bias. However, the sensitivity analyses suggest the distribution of scores in the two procedures is similar before and after exclusions. Were any bias to exist, one would expect it to apply equally to both TKR and UKR and this should not change the findings of the study.

Despite both TKR and UKR being well-established, the evidence comparing their functional outcomes is limited. The only long-term RCT comparing UKR to TKR reported superior PROMs in UKR at 15 years, with 15/21 UKRs (71%) reporting 'excellent' results compared with 10/19 TKRs (53%); however, there was significant attrition in this study due to patient mortality and at the final analysis the study was not sufficiently powered to discern a significant difference in mean scores[11].

Other studies comprise cohort and case-control studies. Manzotti *et al* conducted a retrospective case-control study of 68 knees matched on age, gender and pre-operative PROMs[27], reporting a mean functional score of 83.5/100 following UKR and 78.8 following TKR (p=0.01). Noticewala *et al* reported significantly superior functional outcomes for UKR in a retrospective cohort of 198 knees (128 TKRs and 70 UKRs) at three years[28]. Similarly, Walton *et al* retrospectively compared 183 UKRs with a non-matched group of 142 TKRs, reporting a mean OKS of 37.8 in the UKR group and 35.5 in the TKR group (p=0.04)[29]. In another case-control study, Amin *et al* reported no difference in PROMs between the two procedures, but reported a higher revision rate for UKR at five years[30].

A study of the New Zealand Joint Registry compared 13,627 TKRs with 3,030 UKRs[13]. Mean OKS was 37.0 following TKR and 38.8 following UKR, with 46.4% 'excellent' results in the UKR group compared with 36.0% in the TKR group. However, no pre-operative scores were available and adjusted comparison was not attempted. A study of the Norwegian registry (972 TKRs and 372 UKRs, adjusting for baseline differences using linear regression) suggested superior PROMs for UKR, but the differences were small[31]. Baker *et al* analysed PROMs for TKR and UKR from the NJR for England and Wales, shortly after PROM data collection began[14]. They compared 23,393 TKRs with 505 UKRs, reporting unadjusted six month scores, and adjusted change scores. They demonstrated a small but statistically significant difference in post-operative OKS (34.0 (95%CI 33.9-34.2) for TKR and 35.5

(34.5-36.4) for UKR, p=0.002), but determined no significant difference in adjusted change scores between the two (mean difference 0.0 points, 95% CI -0.9-0.9, p=0.96).

This study has important implications for the future delivery of knee replacement. Both TKR and UKR have been in widespread use for over thirty years, and over 80,000 knee replacement procedures are performed in the UK alone each year[9]. In spite of this, the place of UKR and TKR in the management of end-stage osteoarthritis remains unclear. A large proportion of patients presenting for knee replacement are eligible for both TKR and UKR[4]. Whilst UKR has well-described advantages including lower associated morbidity and mortality, shorter hospital stays and greater cost-effectiveness compared to TKR[4,7,8], the use of UKR is controversial due to a high reported revision rate in reports from NJRs[2,9]. It is intuitive to believe that this difference in revision rate is due to poorer outcomes in UKR. Whilst this study only examines short term outcomes and cannot predict late deterioration in scores, the superiority of outcomes following UKR suggests that the difference in revision rates may not be due to differences in functional outcomes alone.

The higher revision rate may be partly explained by a lower threshold for revision in UKR compared to TKR[13]. UKR can be revised more easily than TKR (usually being converted to a 'primary' TKR), and there is an expectation of good results from such surgery. In patients with the worst outcomes (OKS<20), patients with UKR have a 60% chance of revision, whereas with a similarly poorly-functioning TKR the figure is around 10%[13].Other reasons for the difference in revision rate include differences in patients selected for each procedure (NJR data are unadjusted and UKR patients may be younger and more active, increasing the likelihood of future revision), and an additional mode of late failure in UKR (progression of arthritis to un-replaced parts of the knee).

Future studies of patient selection and surgical practice should clarify the risk factors for failure, and would allow surgeons and patients to make more informed choices as to the choice of knee replacement procedure. The use of large datasets from national registries, coupled with appropriate statistical techniques, should facilitate such studies. The increased reported risk of revision in UKR,

which is in part due to different threshold for revision, should be weighed against the superior functional outcomes and reduced rate of complications reported by this and previous studies when deciding whether to offer TKR or UKR to the large group of patients who are suitable for either.

Acknowledgements

We thank the patients and staff of all the hospitals in England and Wales who have contributed data to the National Joint Registry. We are grateful to the Healthcare Quality Improvement Partnership (HQIP), the NJR Steering Committee and staff at the NJR Centre for facilitating this work.

The authors have conformed to the NJR's standard protocol for data access and publication. The views expressed represent those of the authors and do not necessarily reflect those of the National Joint Registry Steering Committee or the Healthcare Quality Improvement Partnership (HQIP) who do not vouch for how the information is presented.

ADL is supported by Clinical Research Fellowships from the Arthritis Research UK (Grant number 20499) and the Royal College of Surgeons of England.

References

1. Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, Beard DJ. Knee replacement. *Lancet 2012;379-9823:1331-40.*

2. Australian Orthopaedic Association National Joint Replacement Registry Annual Report. *2013.*

3. Judge A, Arden NK, Cooper C, Kassim Javaid M, Carr AJ, Field RE, Dieppe PA. Predictors of outcomes of total knee replacement surgery. *Rheumatology 2012;51-10:1804-13.*

4. Willis-Owen CA, Brust K, Alsop H, Miraldo M, Cobb JP. Unicondylar knee arthroplasty in the UK National Health Service: an analysis of candidacy, outcome and cost efficacy. *Knee 2009;16-6:473-8.*

5. Wise BL, Niu J, Yang M, Lane NE, Harvey W, Felson DT, Hietpas J, Nevitt M, Sharma L, Torner J, Lewis CE, Zhang Y. Patterns of compartment involvement in tibiofemoral osteoarthritis in men and women and in whites and African Americans. *Arthritis Care Res 2012;64-6:847-52.*

6. Price AJ, Rees JL, Beard DJ, Gill RH, Dodd CA, Murray DW. Sagittal plane kinematics of a mobile-bearing unicompartmental knee arthroplasty at 10 years: a comparative in vivo fluoroscopic analysis. *J Arthroplasty 2004;19-5:590-7.*

7. Liddle AD, Judge A, Pandit H, Murray DW. Adverse outcomes after Total and Unicompartmental Knee Replacement: a study of 101,330 matched patients from the National Joint Registry for England and Wales. *Lancet 2014;384-9952:1437-45.*

8. Hunt LP, Ben-Shlomo Y, Clark EM, Dieppe P, Judge A, MacGregor AJ, Tobias JH, Vernon K, Blom A. 45-day mortality after 467 779 knee replacements for osteoarthritis from the National Joint Registry for England and Wales: a retrospective observational study. *Lancet 2014;384-9952:1429-36.*

9. National Joint Registry for England, Wales and Northern Ireland 10th Annual Report. *2013.*

10. The New Zealand Joint Registry Fourteen Year Report. *2013.*

11. 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;91-1:52-7.*

12. Pandit H, Jenkins C, Gill HS, Barker K, Dodd CA, Murray DW. Minimally invasive Oxford phase 3 unicompartmental knee replacement: results of 1000 cases. *J Bone Joint Surg Br 2011;93-2:198-204.*

13. Rothwell AG, Hooper GJ, Hobbs A, Frampton CM. An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand Joint Registry. *J Bone Joint Surg Br 2010;92-3:413-8.*

14. Baker PN, Petheram T, Jameson SS, Avery PJ, Reed MR, Gregg PJ, Deehan DJ. Comparison of patient-reported outcome measures following total and unicondylar knee replacement. *J Bone Joint Surg Br 2012;94-7:919-27.*

15. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy 1990;16-3:199-208.*

16. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br 1998;80-1:63-9.*

17. Browne JP, Bastaki H, Dawson J. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. *Health Qual Life Outcomes 2013;11:128.*

18. Breeman S, Campbell MK, Dakin H, Fiddian N, Fitzpatrick R, Grant A, Gray A, Johnston L, MacLennan GS, Morris RW, Murray DW, Group KATT. Five-year results of a randomised controlled trial comparing mobile and fixed bearings in total knee replacement. *Bone Joint J 2013;95-B-4:486-92.*

19. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, Dawson J. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br 2007;89-8:1010-4.*

20. Kalairajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties--a comparison between the Harris hip score and the Oxford hip score. *J Arthroplasty 2005;20-8:1037-41.*

21. Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Econ 2010;19-8:886-905.*
22. Heinze G, Juni P. An overview of the objectives of and the approaches to propensity score analyses. *Eur Heart J 2011;32-14:1704-8.*
23. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res 2011;46-3:399-424.*
24. Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ 2013;347:f6409.*
25. Beard D, Price A, Cook J, Fitzpatrick R, Carr A, Campbell M, Doll H, Campbell H, Arden N, Cooper C, Davies L, Murray D. Total or Partial Knee Arthroplasty Trial - TOPKAT: study protocol for a randomised controlled trial. *Trials 2013;14-1:292.*
26. Baker PN, Jameson SS, Deehan DJ, Gregg PJ, Porter M, Tucker K. Mid-term equivalent survival of medial and lateral unicondylar knee replacement: An analysis of data from a National Joint Registry. *J Bone Joint Surg Br 2012;94-12:1641-8.*
27. Manzotti A, Confalonieri N, Pullen C. Unicompartmental versus computer-assisted total knee replacement for medial compartment knee arthritis: a matched paired study. *Int Orthop 2007;31-3:315-9.*
28. Noticewala MS, Geller JA, Lee JH, Macaulay W. Unicompartmental knee arthroplasty relieves pain and improves function more than total knee arthroplasty. *J Arthroplasty 2012;27-S8:99-105.*
29. Walton NP, Jahromi I, Lewis PL, Dobson PJ, Angel KR, Campbell DG. Patient-perceived outcomes and return to sport and work: TKA versus mini-incision unicompartmental knee arthroplasty. *J Knee Surg 2006;19-2:112-6.*
30. Amin AK, Patton JT, Cook RE, Gaston M, Brenkel IJ. Unicompartmental or total knee arthroplasty?: Results from a matched study. *Clin Orthop Relat Res 2006;451:101-6.*
31. Lygre SH, Espehaug B, Havelin LI, Furnes O, Vollset SE. Pain and function in patients after primary unicompartmental and total knee arthroplasty. *J Bone Joint Surg Am 2010;92-18:2890-7.*

Tables

Table 1: Pre- and post-matching demographics of each group.

	Crude		Matched	
	TKR	UKR	TKR	UKR
Number of patients	55,208	3,655	10,557	3,519
	Mean (SD)			
Age (years)	69.5 (8.8)	63.9 (9.2)	64.4 (8.6)	64.4 (8.2)
Time from surgery (days)	579.8 (312.9)	620.1 (299.9)	616.1 (308.5)	621.9 (300.4)
Pre-op EQ5D overall	0.430 (0.304)	0.484 (0.289)	0.482 (0.291)	0.481 (0.289)
EQ5D anxiety score	1.38 (0.54)	1.34 (0.53)	1.34 (0.53)	1.34 (0.53)
EQ5D VAS	69.0 (19.3)	71.3 (19.1)	71.2 (18.8)	71.1 (19.2)
Pre-op OKS	19.5 (7.6)	21.9 (7.6)	21.7 (7.7)	21.8 (7.6)
	Number (%)			
Fixation (1= cemented)	50,605 (91.7)	3,201 (87.58)	9,517 (90.2)	3084 (89.5)
ASA:				
1	6094 (11.0)	859 (23.5)	2,214 (21.0)	769 (21.9)
2	41733 (75.6)	2524 (69.1)	7,664 (72.6)	2,478 (70.4)
3	7237 (13.1)	266 (7.3)	674 (6.4)	266 (7.6)
4+	144 (0.26)	6 (0.17)	5 (0.1)	6 (0.2)
Gender (male)	24,945 (45.2)	1,951 (53.4)	5,584 (53.0)	1,869 (53.1)
Unit type:				
Public hospital	41,992 (76.1)	2,737 (74.9)	7,815 (74.0)	2,636 (74.9)
Private hospital	10,573 (19.2)	728 (19.9)	2,285 (21.6)	699 (19.9)
ISTC	2,643 (4.8)	190 (5.2)	457 (4.3)	184 (5.2)
Cases per year	77.9 (53.3)	92.8 (58.8)	91.1 (60.3)	90.9 (56.8)
Operating surgeon=consultant	43024 (77.9)	3,207 (87.7)	9,238 (87.5)	3,075 (87.4)
Admitted from own home	55,018 (99.7)	3,642 (99.6)	10,518 (99.6)	3,508 (99.6)
Patient-reported disability	31,977 (57.9)	1,491 (40.8)	4,444 (42.1)	1,478 (42.0)
Patient-reported co-morbidities:				
Heart disease	5,743 (10.4)	261 (7.1)	781 (7.4)	260 (7.4)
High BP	26,392 (47.8)	1,441 (39.4)	4,197 (39.8)	1,426 (40.5)
Stoke	838 (1.5)	36 (1.0)	110 (1.0)	35 (1.0)
Circulation problems	3,900 (7.1)	125 (3.4)	381 (3.61)	125 (3.6)
Lung disease	4,152 (7.5)	248 (6.8)	728 (6.9)	242 (6.9)
Diabetes	6,647 (12.0)	322 (8.8)	984 (9.3)	321 (9.1)
Kidney disease	911 (1.7)	36 (1.0)	110 (1.0)	26 (1.0)
Nervous disease	441 (0.8)	23 (0.6)	71 (0.7)	23 (0.7)
Liver disease	263 (0.5)	11 (0.3)	41 (0.4)	11 (0.3)
Cancer	2,503 (4.5)	119 (3.3)	328 (3.2)	119 (3.4)
Depression	3,913 (7.1)	257 (7.0)	744 (7.1)	242 (6.9)

Table 2: Primary outcome measures: mean patient-reported outcomes of matched cohort. OKS – Oxford Knee Score; EQ5D – EuroQol 5 domain questionnaire; EQ5D VAS – EuroQol general health Visual Analogue Scale.

	TKR	UKR	Adjusted mean difference	Significance
6 month OKS	36.1	37.7	1.54 (1.21-1.88)	<0.001
6 month EQ5D	0.751	0.772	0.021 (0.012-0.030)	<0.001
6 month VAS	71.2	71.1	0.30 (-0.33-0.92)	0.351

Table 3: Primary outcome measures: categorical outcomes

		TKR	UKR	Total	OR (95% CI)	p
OKS subscales (Kalairajah)	Excellent (>41 points)	3,714 (35.18%)	1,604 (45.58%)	5,318	1.59 (1.47-1.73)	<0.001
	Good (34-41)	3,462 (32.78%)	994 (28.25%)	4,456	0.81 (0.74-0.88)	<0.001
	Fair (27-33)	1,635 (15.49%)	408 (11.59%)	2,043	0.72 (0.64-0.81)	<0.001
	Poor (<27)	1,746 (16.54%)	513 (14.58%)	2,259	0.87 (0.77-0.98)	0.021
OKS MCID	Improvement	9,161 (86.78%)	3,152 (89.57%)	12,313	1.31 (1.16-1.48)	<0.001
	No change	976 (9.25%)	257 (7.30%)	1,233	0.77 (0.67-0.89)	<0.001
	Deterioration	420 (3.98%)	110 (3.13%)	530	0.78 (0.63-0.96)	0.022
EQ5D change (Paretian)	Better	7,978 (75.57%)	2,748 (78.09%)	10,726	1.15 (1.05-1.26)	0.002
	Worse	1,053 (9.97%)	289 (8.21%)	1,342	0.81 (0.70-0.93)	0.002
	No change	1,088 (10.31%)	361 (10.26%)	1,449	0.99 (0.88-1.13)	0.936
	Mixed change	438 (4.15%)	121 (3.44%)	559	0.82 (0.67-1.01)	0.062

Table 4: Post-operative EQ5D subscales. Figures given are overall numbers in each group with percentages given in parentheses.

Domain	Difficulty level	TKR	UKR	p
Mobility	None	6104 (58)	2,174 (62)	<0.001
	Moderate	4,447 (42)	1344 (38)	
	Severe	6 (0.06)	1 (0.03)	
Self-care	None	9,013 (85)	3,108 (88)	<0.001
	Moderate	1,525 (14)	406 (11)	
	Severe	19 (0.2)	5 (0.1)	
Ability to undertake usual activities	None	5,296 (50)	1,949 (55)	<0.001
	Moderate	4,929 (47)	1,480 (42)	
	Severe	332 (3)	90 (3)	
Pain / discomfort	None	3,944 (37)	1,524 (43)	<0.001
	Moderate	6,094 (58)	1,819 (52)	
	Severe	519 (5)	176 (5)	
Anxiety / depression	None	8,553 (81)	2,851 (81)	0.981
	Moderate	1,795 (17)	595 (17)	
	Severe	209 (2)	73 (2)	

Table 5: Secondary outcome measures: patient-reported satisfaction, improvement, complications and readmissions.

		TKR (%)	UKR (%)	OR (95%CI)	P
Complications	Drug reaction	1,358 (13.7)	368 (11.0)	0.78 (0.69-0.88)	<0.001
	Urinary	996 (10.3)	318 (9.6)	0.93 (0.81-1.06)	0.254
	Bleeding	770 (8.0)	194 (6.0)	0.76 (0.62-0.86)	<0.001
	Wound	1,203 (12.3)	238 (7.3)	0.56 (0.48-0.65)	<0.001
	Readmitted	978 (9.3)	256 (7.3)	0.77 (0.67-0.89)	<0.001
	Reoperation	350 (3.3)	99 (2.8)	0.84 (0.67-1.06)	0.14
Satisfaction	Excellent	2,493 (23.7)	994 (28.3)	1.27 (1.17-1.39)	<0.001
	Very Good	3,905 (37.1)	1,241 (35.4)	0.93 (0.86-1.01)	0.067
	Good	2,510 (23.9)	718 (20.5)	0.82 (0.75-0.90)	<0.001
	Fair	1,220 (11.6)	424 (12.1)	1.04 (0.93-1.18)	0.429
	Poor	395 (3.8)	130 (3.7)	0.99 (0.81-1.21)	0.899
Improvement	Much better	7,627 (72.3)	2,599 (73.9)	1.08 (1.00-1.19)	0.06
	A little better	1,702 (16.1)	518 (14.7)	0.90 (0.81-1.00	0.049
	Same	496 (4.7)	187 (5.4)	1.15 (0.97-1.37)	0.108
	A little worse	436 (4.1)	118 (3.4)	0.81 (0.65-0.99)	0.041
	much worse	268 (2.5)	85 (2.4)	0.95 (0.74-1.22)	0.687

Figures

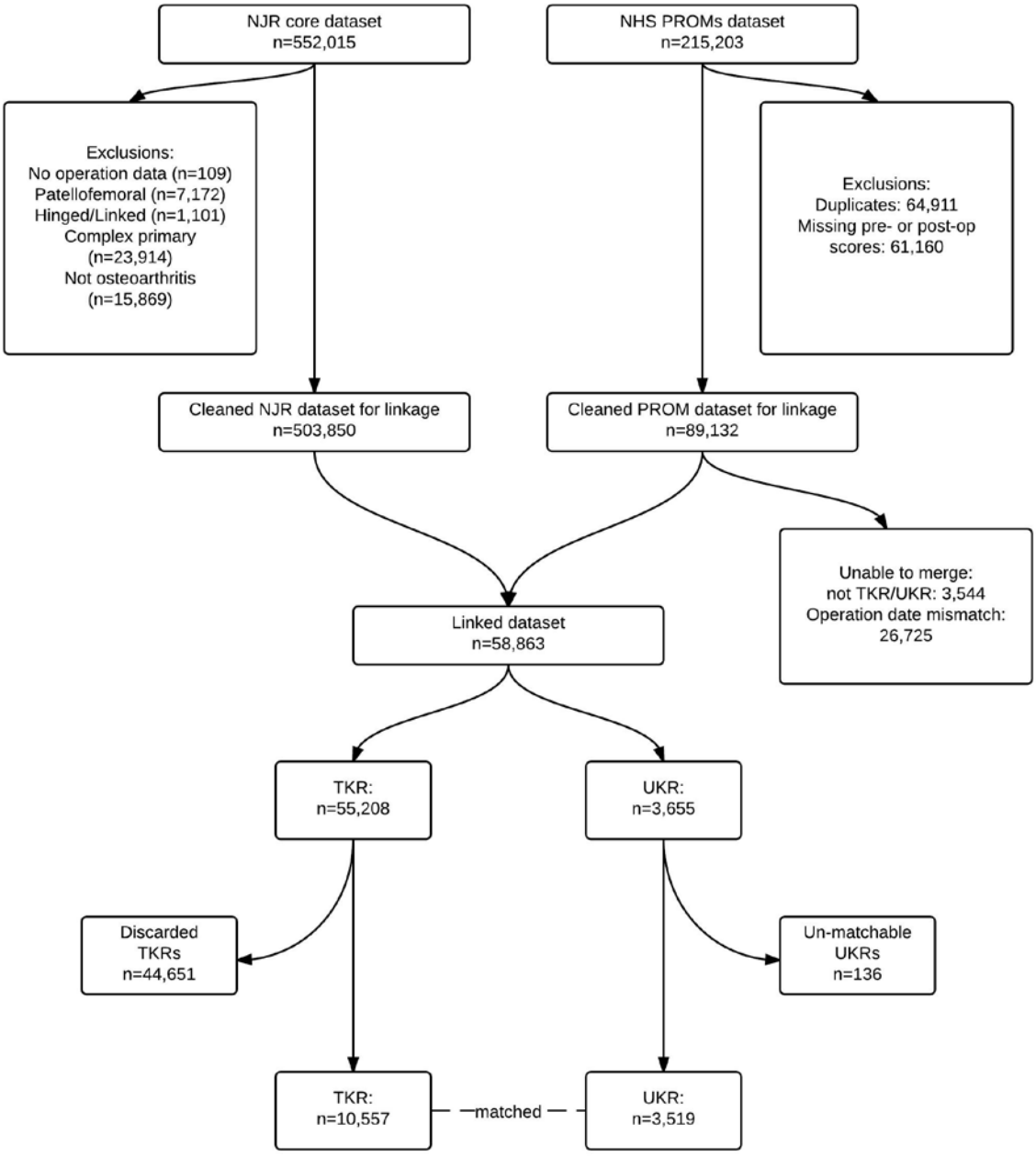

Figure 1: Data linkage and exclusion flow chart

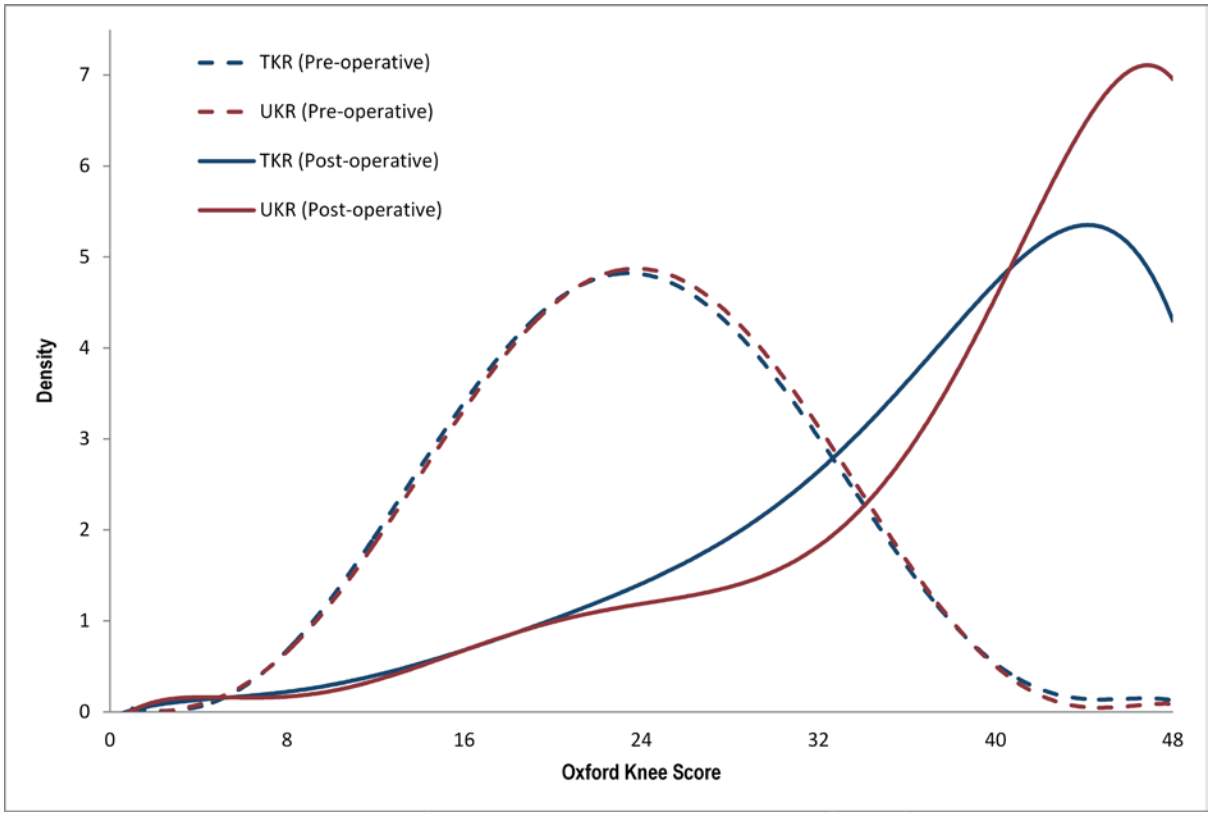

Figure 2: Kernel density plot demonstrating the pre- and post-operative OKS distributions in TKR and UKR.