

Opioid use, post-operative complications, and implant survival after unicompartmental versus total knee replacement: a population-based network study

Edward Burn, MSc,<sup>\*1</sup> James Weaver, MSc,<sup>\*2</sup> Daniel Morales, PhD,<sup>3</sup> Albert Prats-Urbe, MPH,<sup>1</sup> Antonella Delmestri, PhD,<sup>1</sup> Victoria Y Strauss, PhD,<sup>1</sup> Ying HE, PhD,<sup>1</sup> Danielle E Robinson, PhD,<sup>1</sup> Rafael Pinedo-Villanueva, PhD,<sup>1</sup> Spyros Kolovos, PhD,<sup>1</sup> Talita Duarte-Salles, PhD,<sup>4</sup> William Sproviero, PhD,<sup>5</sup> Dahai Yu, PhD,<sup>6</sup> Michel Van Speybroeck, MSc,<sup>7</sup> Ross Williams, MSc,<sup>8</sup> Luis H. John, MSc,<sup>8</sup> Nigel Hughes, MSc,<sup>7</sup> Anthony G. Sena, BA,<sup>1</sup> Ruth Costello, MSc,<sup>9</sup> Belay Birlik, MSc,<sup>9</sup> David Culliford, PhD,<sup>1,10</sup> Caroline O'Leary, MSc,<sup>11</sup> Henry Morgan, PhD,<sup>11</sup> Theresa Burkard, MSc,<sup>12,13</sup> Daniel Prieto-Alhambra, PhD,<sup>†1,14</sup> Patrick Ryan, PhD<sup>†1,15</sup>

<sup>1</sup>NDORMS, University of Oxford, Oxford, UK, <sup>2</sup>Janssen Research and Development, Titusville, NJ, USA, <sup>3</sup>University of Dundee, UK, <sup>4</sup>Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain, <sup>5</sup>Department of Psychiatry, University of Oxford, Oxford, UK, <sup>6</sup>Keele University, Keele, UK, <sup>7</sup>Janssen Pharmaceutica R&D, Beerse, Belgium, <sup>8</sup>Erasmus University Medical Center, NL, <sup>9</sup>Arthritis Research UK Centre for Epidemiology, University of Manchester, Manchester, UK, <sup>10</sup>NIHR CLAHRC Wessex, University of Southampton, UK, <sup>11</sup>IQVIA, London, UK, <sup>12</sup>University of Basel, Basel, Switzerland, <sup>13</sup>University Hospital Basel, Basel, Switzerland, <sup>14</sup>GREMPAL Research Group, Idiap Jordi Gol and CIBERFes, Universitat Autònoma de Barcelona and Instituto de Salud Carlos III, Barcelona, Spain, <sup>15</sup>Columbia University, New York, NY, USA

\*Joint first authors †Joint senior authors

Corresponding author: D Prieto-Alhambra, Botnar Research Centre, Windmill Road, OX37LD, OXFORD, UK

## Abstract

### Background

The aim of this study was to compare unicompartmental and total knee replacement (UKR and TKR), emulating the design of the Total or Partial Knee Arthroplasty Trial (TOPKAT) using routinely-collected data. The primary outcome in TOPKAT was patient-reported outcomes, with secondary outcomes including post-operative complications and implant survival.

### Methods

Five US and UK healthcare databases, part of the Observational Health Data Sciences and Informatics (OHDSI) network, were analysed. Opioid use from 91 to 365 days after surgery, as a proxy for persistent pain, was assessed. Post-operative complications (venous thromboembolism, infection, readmission, and mortality) were considered over 60 days following surgery and implant survival over five years following surgery. Propensity score matched Cox proportional hazards models were fitted for each outcome. Calibrated hazard ratios (cHRs) were generated for each database to account for observed differences in control outcomes and these were combined using meta-analysis.

### Findings

In total, 32,379 and 250,377 individuals who received UKR and TKR were matched and included in the analysis. UKR was associated with a reduced risk of post-operative opioid use (cHR from meta-analysis: 0.81 (95% CI: 0.73 to 0.90)). UKR was also associated with a reduced risk of venous thromboembolism (cHR: 0.62 (0.36 to 0.95)), but little difference was seen for infection (cHR: 0.85 (0.51 to 1.37)) and readmission (cHR: 0.79 (0.47 to 1.25)). There was insufficient evidence to conclude there was a reduction in risk of mortality. UKR was also associated with an increased risk of revision (cHR: 1.64 (1.40 to 1.94)).

### Interpretation

UKR was associated with a reduced risk of opioid use compared to TKR, which may indicate a reduced risk of persistent pain after surgery. UKR was associated with a lower risk of venous thromboembolism. UKR was also, however, associated with an increased risk of revision compared to TKR.

### Funding

This work has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking (EHDEN] grant n° [806968].

## Research in context

### Evidence before this study

Prior research has found unicompartmental and total knee replacement (UKR and TKR) to result in broadly similar patient-reported outcomes, UKR to have a lower risk of some post-operative complications, notably venous thromboembolism, infection, and mortality, but TKR to have a lower risk of revision procedures. A recent randomised controlled trial, the Total or Partial Knee Arthroplasty Trial (TOPKAT), compared UKR and TKR, with 264 patients randomised into each arm of the trial. The primary outcome for TOPKAT was post-operative patient-reported outcomes, with secondary outcomes including post-operative complications and implant survival. Consistent with previous observational studies, post-operative patient-reported outcomes were similar at 5 years and fewer complications seen for those who had UKR. However, rates of revision were seen to be similar for UKR and TKR at 5 years. Direct comparisons between the randomised evidence from TOPKAT and observational studies are, however, made difficult though due to differences in study designs.

### Added value of this study

This study emulates the TOPKAT design using routinely-collected data. Where possible, similar eligibility criteria were specified and outcomes assessed in a similar manner. Patient-reported outcomes (the primary outcome in TOPKAT) were not available, and so opioid prescriptions were used as a proxy for persistent pain following surgery. Post-operative complications and implant survival were also assessed. The findings from this study will provide further evidence to inform considerations of the relative merits of UKR and TKR.

### Implications of all the available evidence

In this study, UKR was associated with a reduced risk of post-operative opioid use between 91 to 365 days after surgery relative to TKR, and this may indicate a reduced risk of persistent pain after UKR. As seen in this study and in previous research, UKR also appears to have a lower risk of venous thromboembolism compared to TKR. However, while revision rates were similar for UKR and TKR in TOPKAT, the findings from this study support that of previous observational research showing UKR to have an increased risk of revision.

## Introduction

Knee replacement is one of the most common surgical procedures and typically leads to substantial improvements in pain, function and quality of life.<sup>1</sup> However, there is variation in how knee replacements are performed. One area of particular uncertainty is around whether to use unicompartmental or total knee replacement (UKR or TKR) for those individuals with osteoarthritis confined to a single compartment of the knee. While all the compartments of the joint are replaced in TKR, only the affected part of the joint is replaced in UKR.

With patient-reported pain and function key indications for knee and hip replacement, it follows that they should also be considered as a key measure of the effectiveness of surgery. Previous research has generally found UKR and TKR to result in broadly similar gains in patient-reported outcomes after surgery.<sup>2</sup> Both UKR and TKR are major orthopaedic procedures and so are accompanied by a risk of post-operative complications. Findings from previous research suggests that UKR, which is a quicker and less-invasive procedure relative to TKR, may have a lower risk of some post-operative complications, notably venous thromboembolism, infection, and mortality.<sup>2</sup> As well as the short-term risk of post-operative complications, patients who have had a knee and hip replacement have a long-term risk of revision surgery, in which implant components are removed, added or exchanged. Revision procedures are associated with significant morbidity for individuals, with those undergoing revision surgery generally reporting worse patient-reported outcomes before and after revision procedures compared with those undergoing primary procedures.<sup>3</sup> Observational research has consistently found UKR to have a higher risk of revision procedures compared to TKR, with the increased risk maintained over 25 years after the primary procedure.<sup>2,4</sup>

In a recently published randomised controlled trial comparing UKR and TKR, the Total or Partial Knee Arthroplasty Trial (TOPKAT), 264 patients were randomly assigned UKR with another 264 assigned TKR, with 245 and 269 going on to receive UKR and TKR, respectively. Surgeons performing the procedures were either 'equipoise' surgeons who performed both surgeries, or 'expertise' surgeons who performed only one of the procedures while another 'expertise' surgeon in the same centre performed the other. To perform a given procedure surgeons needed to have been practising it for at a year and to have performed it at least ten times in the previous year.<sup>5</sup> The trial was powered to assess the primary outcome which was self-reported pain and function, as measured by the Oxford Knee Score (OKS).<sup>5</sup> Both groups achieved substantial improvements in OKS relative to baseline scores, with the gains broadly similar across the two comparator groups. Post-operative complications and implant survival were also assessed in TOPKAT as secondary outcomes. Fewer individuals had a post-operative complication after UKR compared to TKR. In contrast to the previous observational research, UKR and TKR were also seen to have similar rates of revision after 5 years in the trial.<sup>6</sup>

The aim of this study was to emulate the TOPKAT trial design using routinely-collected data, so as to answer the same causal question. A study which uses routinely-collected data to emulate the 'target trial' should be harmonised, with similar study designs applied to allow for meaningful comparisons.<sup>7,8</sup> The primary outcome was patient-reported pain and function. As this was not possible, the effect of type of procedure (UKR or TKR) on persistent pain after surgery was considered. Secondary outcomes in the target trial included post-operative complications and implant survival, and these were also assessed in this study.

## Methods

A network cohort study was conducted across 5 observational health care databases from the US and the UK. The study period was from 1 January 2005 to 30 April 2018. The study was designed and

performed before the results of TOPKAT became available. To promote transparency and reproducibility, the full study protocol, all code lists used, and source code for the study execution are publicly available at <https://github.com/OHDSI/StudyProtocols/tree/master/UkaTkaSafetyEffectiveness>.

## Data sources

We used data from the following 5 healthcare databases: 1) IBM MarketScan® Commercial Database (CCAIE), which includes claims data from individuals in the United States (US) enrolled in employer-sponsored insurance health plans; 2) IBM MarketScan® Medicare Supplemental and Coordination of Benefits Database (MDCR), which includes claims data from older adults in the US with primary or Medicare supplemental coverage through privately insured fee-for-service, point-of-service, or capitated health plans; 3) Optum® de-identified Clinformatics® Datamart, Extended - Date of Death (Optum), which includes US patients fully insured in commercial plans or covered with administrative services only and commercial Medicare; 4) PharMetrics™ Plus (PharMetrics), an adjudicated claims database of privately insured US individuals; and, 5) The Health Improvement Network (THIN), which includes pseudonymised electronic primary care medical records from a representative sample of UK inhabitants. These 5 databases were converted to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), which enables consistent application of analyses across disparate data sources.<sup>9</sup>

## Exposure cohorts

Individuals who underwent either a UKR or TKR were identified. Study participants were required to have data captured over at least the year prior to surgery. We excluded patients using published exclusion criteria of TOPKAT,<sup>5</sup> with individuals required to be aged 40 or over at surgery, and have no prior evidence of knee arthroplasty, knee fracture, knee surgery except for diagnostic procedures, rheumatoid arthritis, inflammatory arthropathies, or septic arthritis. In addition, patients with spine, hip, or foot pathology in the year prior to surgery were also excluded. These criteria were intended to identify patients who were eligible for either type of knee replacement, and exclude patients who were not indicated for either UKR or TKR.

## Outcome definitions

Relating to patient-reported outcomes which were the primary outcome in the target trial, persistent pain after surgery was assessed using opioid use (identified by a written or dispensed opioid prescription) as a proxy, with a time-at-risk 91 days after surgery to 1 year after surgery. The 90-day washout period intended to exclude those prescriptions which could be considered as a routine consequence of undergoing surgery. Opioid use was assessed in all databases.

Post-operative complications assessed were symptomatic venous thromboembolism (identified by a diagnosis code of either deep vein thrombosis or pulmonary embolism), infection (identified by a diagnosis of an infection that could be associated with knee replacement), readmission (identified by an inpatient or emergency room visit for any cause), and all-cause mortality. Venous thromboembolism and infection were assessed in all databases, readmission in CCAIE, Optum, and MDCR, and mortality in Optum and THIN. Time-at-risk for post-operative complications was from the date of surgery to 60 days after surgery. Meanwhile, implant survival was assessed in terms of revision (identified by a relevant procedure code) with the time-at-risk from date of surgery to 5 years after surgery. Implant survival was assessed in all databases.

## Statistical methods

Propensity score matching was used to minimise confounding by observed characteristics.<sup>10</sup> A large set of patient-level baseline covariates (representing demographics, health services utilization, and prior diagnoses, medications, lab tests, and procedures) were constructed for propensity score model input. These covariates were assessed over varying time windows relative to an individual's index date, with them identified from 30 days, 365 days, 1095 days and all available days prior to the index date. Propensity scores were generated using a large-scale regularized logistic regression fitted with a Laplace prior (LASSO) and the optimal hyperparameter determined through 10-fold cross validation in order to balance baseline covariates while avoiding overfitting.<sup>11,12</sup> In the primary analyses, patients were matched on the propensity score using variable-ratio matching with a maximum ratio of UKR to TKR of 1:10. The balance of propensity score-matched cohorts was evaluated using standardized mean difference, with values of <0.1 taken to indicate negligible group differences.<sup>13</sup> Propensity score distribution plots, normalized to the preference scale, were used to evaluate empirical equipoise.<sup>14</sup>

Cox proportional hazards models, with procedure type (UKR or TKR) as the sole explanatory variable and conditioned on the matched sets, were fitted to estimate the average treatment effect among UKR patients on the outcomes listed above. Proportionality of hazards was checked visually using Kaplan-Meier plots. Cox models were also estimated for 39 pre-specified negative control conditions (detailed in Appendix Table A1) believed to have no causal relationship with type of knee replacement. To control for residual confounding, hazard ratios (HRs) for the outcomes of interest were calibrated based on the estimated residual error from negative control outcomes and synthetic positive control outcomes.<sup>15,16</sup> Empirical calibration is a process whereby the residual error of an estimator is quantified and incorporated into a calibrated version of the estimator. The calibrated HR (cHR), in this case, reflects the distribution of estimates on the negative control outcomes. For example, if the negative control estimates are on average greater than the null, an increased risk for the outcome of interest will be attenuated following calibration. The cHRs were only estimated if a sufficient number of control outcomes were observed during a given time-at-risk window. Each analysis was conducted separately in each database.

Findings across databases were combined using meta-analysis, with the inverse variance random-effects approach used.<sup>17</sup> At the request of peer review, results were meta-analysed for each of the outcomes. An  $I^2$  above 40% can, however, be taken to indicate substantial heterogeneity across databases.<sup>18</sup> Estimates for negative and positive controls were pooled before performing empirical calibration on the pooled estimates.

## Sensitivity analyses

Pre-specified sensitivity analyses were run for each of the outcomes of interest, with variations of cohort definitions, time-at-risk, and approaches to matching (Appendix Table 2).

## 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

## Results

32,379 individuals who had UKR and 250,377 who had TKR were matched using propensity scores (see Appendix Figure A1 for study flowcharts). Prior to matching, individuals undergoing UKR were younger and healthier than those undergoing TKR (Appendix Table A3). Diagnostics for propensity score matching and control outcome findings are summarised in Appendix Figure A2. After

matching, both cohorts appeared largely comparable in terms of observed characteristics (Table 1 and Appendix Figure 2). Individuals in the matched CCAE, Optum, and PharMetrics cohorts were generally younger and had fewer comorbidities compared to THIN and, in particular, MDCR. THIN covered the broadest age range of individuals. Pre-operative opioid use was well balanced for the comparator groups, with between 30% to 45% of individuals classified as an opioid user before surgery.

UKR was consistently associated with a reduced risk of opioid use after surgery relative to TKR, with cHRs for the use of opioids in the 3 to 12 months post-surgery ranging from 0.70 (0.57 to 0.90) for THIN to 0.86 (0.78 to 0.96) for Optum. The estimate from meta-analysis was 0.81 (0.73 to 0.90). The cumulative incidence of opioid use in the 3 to 12 months post-surgery was around 35% to 40% for UKR and about 5 percentage points higher for TKR in the 4 databases from the US. Opioid use was around 20% for UKR and 25% for TKR in the database from the UK (Appendix Figure A2). These findings were generally similar across sensitivity analyses. When considered up to 5 years, UKR was still associated with a reduced risk of opioid use, but the estimated effects were slightly attenuated with cHRs ranging from 0.86 (0.78 to 0.96) for CCAE to 0.90 (0.82 to 1.02) for MDCR, with no meta-analysis performed for these outcomes.

UKR was consistently associated with a lower risk of venous thromboembolism compared to TKR. The cHRs ranged between 0.47 (0.32 to 0.71) for MDCR and 0.76 (0.59 to 0.99) for CCAE, with the estimate from meta-analysis 0.62 (0.36 to 0.95), see Figure 1. Point estimates for risk of infection and readmission varied from a protective effect for UKR to no difference between the procedures, with cHRs for infection ranged from 0.73 (0.44 to 1.24) for PharMetrics to 1.04 (0.77 to 1.43) for CCAE, while cHRs for readmission ranged from 0.66 (0.46 to 0.97) for MDCR to 0.99 (0.71 to 1.48) for Optum (Figure 1). Estimates from meta-analysis were 0.85 (0.51 to 1.37) and 0.79 (0.47 to 1.25) for infection and readmission, respectively, although in both cases  $I^2$  was above 0.5. Finally, there was little evidence of an association between procedure and mortality, with a cHR of 1.26 (0.55 to 3.09) in Optum and a HR of 0.51 (0.03 to 2.51) in THIN. Findings were broadly similar across sensitivity analyses. There was stronger evidence, however, that UKR was associated with a reduced risk of readmission when considered over the year following surgery rather than 60 days in CCAE and MDCR, cHRs 0.75 (0.66 to 0.86) and 0.76 (0.64 to 0.93), respectively.

UKR was consistently associated with an increased risk of revision compared to TKR over the five years following surgery (Figure 1), with cHRs ranging from 1.48 (1.25 to 1.83) for PharMetrics to 2.16 (1.63 to 3.15) for MDCR. The estimate from meta-analysis was 1.64 (1.40 to 1.94), although  $I^2$  was 0.5. After 5 years, implant survival was generally around 97.5% to 95% for TKR and 95% to 92.5% following UKR (Appendix Figure A3). These findings were similar across the various sensitivity analyses considered.

Results for the primary analysis and each sensitivity analysis are detailed in Appendix Table A4. These can also be viewed, along with study flow charts, characteristics of study participants before and after matching, and propensity score distributions, using the interactive web-based application at <http://data.ohdsi.org/UkaTkaSafetyEffectiveness>.

## Discussion

In summary, compared to TKR, UKR was associated with a reduced risk of post-operative opioid use. This may indicate that UKR has a lower risk of post-operative persistent pain. UKR was also associated with a decreased risk of post-operative venous thromboembolism. There was insufficient



evidence to conclude there UKR led to a reduction in risk of infection, readmission, or mortality. TKR was associated with a lower risk of revision.

The primary outcome in TOPKAT was patient-reported pain and function, as measured by OKS. Outcome scores were broadly similar for the two comparator groups. The mean difference at five years was 1.04 in favour of UKR but this was not statistically significant, with a 95% confidence interval spanning -0.42 to 2.50,<sup>6</sup> and unlikely to be clinically meaningful with the minimal important difference in OKS being 5 points.<sup>19</sup> This finding is in accordance with previous research that has also generally found UKR and TKR to result in broadly similar gains in patient-reported outcomes after surgery.<sup>2</sup> In this study, however, we found UKR to have a lower risk of opioid use, with the absolute effect particularly pronounced for study participants in the US. This suggests that UKR may be associated with a lower risk of persistent pain after surgery. Although few studies have previously assessed procedure choice and opioid use, our findings are consistent with two studies that have.<sup>20,21</sup>

There were fewer post-operative complications for those who received UKR in TOPKAT, with UKR associated with a relative risk reduction of 28% (95% CI: 47% to 2%).<sup>6</sup> This finding is in line with those from previous observational studies, where UKR has been associated with a reduced risk for a range of complications relative to TKR.<sup>2,22-24</sup> In particular, a meta-analysis of previous studies of national or large multicentre databases or of joint registry data found UKR to be associated with a risk ratios of 0.39 (0.27 to 0.57) for venous thromboembolism and 0.27 (0.16 to 0.45) for mortality relative to TKR.<sup>2</sup> The results from this study confirm the risk reduction for venous thromboembolism. This risk reduction appears most pronounced for older patients, with the largest effect of procedure seen in MDCR. However, with mortality only available in two databases, there was insufficient evidence to conclude there was a reduction in risk of mortality for UKR in this study. Prior observational studies have typically accounted for differences in the observed characteristics of those undergoing the two procedures, either through propensity score matching or multivariable regression. It is notable that the additional calibration on control outcomes used in this study generally led to associations in favour of UKR being somewhat attenuated.

UKR and TKR were seen to have similar rates of revision in TOPKAT, with rates of revision around 4% at 5 years for both procedures.<sup>6</sup> This finding is in contrast to much of the body of previous observational research, which have consistently found UKR to have a higher risk of revision.<sup>2,4,22,25,26</sup> Indeed, while risk of revision over 5 years after UKR is currently around 6% in the UK, risk for TKR is approximately 2.5%.<sup>26</sup> The incidence of revision for study participants from the UK included in this study are in line with these previous findings, with revision risks seen to be slightly higher for study participants from the US. As with previous observational studies, UKR was also consistently associated with an increased risk of revision in this study. UKR can therefore be expected to have a higher risk of revision than TKR.

This analysis has been informed by data from 280,000 patients across 5 databases in 2 countries. This retrospective analysis relied though on data captured in electronic health records and administrative claims, and therefore our ability to emulate the inclusion criteria used the TOPKAT trial was limited. In particular, these data did not have radiographic information and so it was not possible to assess whether an individual's osteoarthritis was confined to one compartment of the knee. Patient-reported outcomes, the primary outcome in TOPKAT, was also not captured in the databases used and so opioid use was used as a proxy for persistent pain. This has limitations, however, as opioids may not necessarily have been taken even if dispensed. We used large-scale propensity score matching to balance the two cohorts using more than 10,000 candidate baseline characteristics. However, as with all observational studies, there remains the potential risk of



312 confounding due to unmeasured factors. We employed a large panel of negative control outcomes  
313 to mitigate the threat of systematic error. While the definitions for exposures and outcomes were  
314 clinically reviewed and relied on codes used in prior published studies,<sup>20,27–30</sup> individual cases were  
315 not validated and may be subject to misclassification. There may be measurement errors, for  
316 example with baseline characteristics, such as comorbidities, and outcomes, such as revision,  
317 potentially not recorded within the databases, in which case they would also be missed in the  
318 analysis. As patients in this study were selected on the basis of their inclusion criteria for the TOPKAT  
319 trial, the results from this study may also not necessarily generalise to those patients excluded from  
320 the trial but are eligible for both procedures. In addition, while meta-analysis was used to combine  
321 findings across databases, in a number of cases substantial heterogeneity was present and so the  
322 resulting estimates should be interpreted with caution.

323 In conclusion, with a lower risk of post-operative opioid use, UKR may be associated with a reduced  
324 risk of persistent pain compared to TKR. UKR is also associated with a lower risk of venous  
325 thromboembolism. UKR is also, however, associated with an increased risk of revision. The merit of  
326 using real-world data for assessing the effectiveness of treatments is still debated,<sup>31,32</sup> and  
327 randomised controlled trials remain the ‘gold standard’ for establishing efficacy. This study has  
328 demonstrated the value of real-world evidence for complementing the evidence produced from  
329 randomised trials.

330

331

332

## Acknowledgements

The authors would like to thank the OHDSI community for their contributions to the tools used for this analysis. THIN is a registered trademark of Cegedim SA in the United Kingdom and other countries. Reference made to the THIN database is intended to be descriptive of the data asset licensed by IQVIA. This work uses de-identified data provided by patients as a part of their routine primary care.

This activity under the European Health Data & Evidence Network (EHDEN) has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 806968. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. The sponsor of the study did not have any involvement in the writing of the manuscript or the decision to submit it for publication. The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). DPA is funded by a National Institute for Health Research Clinician Scientist award (CS-2013-13-012). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## Author contributions

All authors made substantial contributions to the conception or design of the work; DPA and PBR led the acquisition of the data; all authors were involved in the analysis and interpretation of data for the work; All authors have contributed to the drafting and revising critically the manuscript for important intellectual content; all authors have given final approval and agree to be accountable for all aspects of the work.

## Declaration of interest

All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: AS, JW, MvS, NH and PR are full-time employees of Janssen Research & Development, a pharmaceutical company of Johnson & Johnson, and shareholders in Johnson & Johnson. The Johnson & Johnson family of companies also includes DePuy Synthes, which is the maker of medical devices for joint reconstruction. DPA reports research grants from AMGEN, UCB Biopharma and Les Laboratoires Servier. APU reports grants from MRC - DTP Funding. COL is a part-time employee of IQVIA. EB, BB, DM, DR, DY, LHJ, HMS, RC, RPV, SK, TD, WS, YH, AD, RW, TB, VYS and DJC have nothing to disclose.

## Ethical approval

This study was approved by THIN's Scientific Review Committee (reference number: 18THIN100).

## References

- 1 Carr AJ, Robertsson O, Graves S, *et al.* Knee replacement. *Lancet* 2012; **379**: 1331–40.
- 2 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**: l352.
- 3 Greidanus N V., Peterson RC, Masri BA, Garbuz DS. Quality of Life Outcomes in Revision Versus Primary Total Knee Arthroplasty. *J Arthroplasty* 2011; **26**: 615–20.
- 4 Evans JT, Walker RW, Evans JP, Blom AW, Sayers A, Whitehouse MR. How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. *Lancet* 2019; **393**: 655–63.
- 5 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**. DOI:10.1186/1745-6215-14-292.
- 6 Beard DJ, Davies LJ, Cook JA, *et al.* 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. *Lancet* 2019; **6736**: 5–9.
- 7 Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016; **183**: 758–64.
- 8 Lodi S, Phillips A, Lundgren J, *et al.* Effect estimates in randomized trials and observational studies: comparing apples with apples. *Am J Epidemiol* 2019. DOI:10.1093/aje/kwz100.
- 9 Voss EA, Makadia R, Matcho A, *et al.* Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inf Assoc* 2015; **22**: 553–64.
- 10 Rubin DJ. Estimating Causal Effects from Large Data Sets Using Propensity Scores. *Ann Intern Med* 1997; **127**. DOI:10.7326/0003-4819-127-8\_Part\_2-199710151-00064.
- 11 Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. Massive parallelization of serial inference algorithms for a complex generalized linear model. *ACM Trans Model Comput Simul* 2014; **23**: 1–23.
- 12 Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J of Epidemiology* 2018; : 2005–14.
- 13 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011; **46**: 399–424.
- 14 Walker A, Patrick, Lauer, *et al.* A tool for assessing the feasibility of comparative effectiveness research. *Comp Eff Res* 2013; : 11.
- 15 Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: Why empirical calibration is needed to correct p-values. *Stat Med* 2014; **33**: 209–18.
- 16 Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *PNAS* 2018; **115**: 2571–2577.
- 17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88.
- 18 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.

- 408 19 Beard DJ, Harris K, Dawson J, *et al.* Meaningful changes for the Oxford hip and knee scores  
409 after joint replacement surgery. *J Clin Epidemiol* 2015; **68**: 73–9.
- 410 20 Bedard NA, DeMik DE, Dowdle SB, Callaghan JJ. Trends and risk factors for prolonged opioid  
411 use after unicompartmental knee arthroplasty. *Bone Joint J* 2018; **100-B**: 62–7.
- 412 21 Fuzier R, Serres I, Bourrel R, Palmaro A, Montastruc J, Lapeyre-Mestre M. Analgesic drug  
413 consumption increases after knee arthroplasty : A pharmacoepidemiological study  
414 investigating postoperative pain. *Pain* 2014; **155**: 1339–45.
- 415 22 Liddle AD, Judge A, Pandit H, Murray DW. Adverse outcomes after total and  
416 unicompartmental knee replacement in 101 330 matched patients: a study of data from the  
417 National Joint Registry for England and Wales. *Lancet* 2014; **384**: 1437–45.
- 418 23 Hunt LP, Ben-Shlomo Y, Clark EM, *et al.* 45-day mortality after 467779 knee replacements for  
419 osteoarthritis from the National Joint Registry for England and Wales: An observational study.  
420 *Lancet* 2014; **384**: 1429–36.
- 421 24 Brown NM, Sheth NP, Davis K, *et al.* Total Knee Arthroplasty Has Higher Postoperative  
422 Morbidity Than Unicompartmental Knee Arthroplasty : A Multicenter Analysis. *J Arthroplasty*  
423 2012; **27**: 86–90.
- 424 25 Price AJ, Alvand A, Troelsen A, *et al.* Knee replacement. *Lancet* 2018; **392**: 1672–82.
- 425 26 Green M, Wishart N, Young E, McCormack V, Swanson M. NJR 14th Annual Report. 2017.
- 426 27 Bawa H, Weick JW, Dirschl DR, Luu HH. Trends in Deep Vein Thrombosis Prophylaxis and  
427 Deep Vein Thrombosis Rates After Total Hip and Knee Arthroplasty. *JAAOS - J Am Acad*  
428 *Orthop Surg* 2018; **26**.  
429 [https://journals.lww.com/jaaos/Fulltext/2018/10010/Trends\\_in\\_Deep\\_Vein\\_Thrombosis\\_Pro](https://journals.lww.com/jaaos/Fulltext/2018/10010/Trends_in_Deep_Vein_Thrombosis_Prophylaxis_and.6.aspx)  
430 [phylaxis\\_and.6.aspx](https://journals.lww.com/jaaos/Fulltext/2018/10010/Trends_in_Deep_Vein_Thrombosis_Prophylaxis_and.6.aspx).
- 431 28 Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous  
432 thromboembolism using administrative and claims data. *Pharmacoepidemiol Drug Saf* 2012;  
433 **21**: 154–62.
- 434 29 Calderwood MS, Ma A, Khan YM, *et al.* Use of Medicare Diagnosis and Procedure Codes to  
435 Improve Detection of Surgical Site Infections following Hip Arthroplasty, Knee Arthroplasty,  
436 and Vascular Surgery. *Infect Control & Hosp Epidemiol* 2012; **33**: 40–9.
- 437 30 Bedard NA, DeMik DE, Dowdle SB, Owens JM, Liu SS, Callaghan JJ. Preoperative Opioid Use  
438 and Its Association With Early Revision of Total Knee Arthroplasty. *J Arthroplasty* 2018; **33**:  
439 3520–3.
- 440 31 Gerstein HC, McMurray J, Holman RR. Real-world studies no substitute for RCTs in  
441 establishing efficacy. *Lancet* 2019; **393**: 210–1.
- 442 32 Sharma M, Nazareth I, Petersen I. Observational studies of treatment effectiveness:  
443 worthwhile or worthless? *Clin Epidemiol* 2019; **11**: 35–42.

444

445

446

447

## Figures

Figure 1. Effect of procedure choice (UKR or TKR) on post-operative complications, opioid use, and revision

Numbers of propensity score matched individuals, observed events, HRs and cHRs for UKR relative to TKR. Readmission data were not available in PharMetrics and THIN. Mortality data were only available in Optum and THIN. Calibration of hazard ratios was infeasible for post-operative complications in THIN because there were too few negative control events observed during the 60-day time-at-risk. Adjusted HRs account for residual confounding identified by negative control outcomes analyses. Calibrated HRs were not estimated for 60-day outcomes in THIN due to too few control outcomes being observed. UKR: unicompartmental knee replacement; TKR: total knee replacement; HR: Hazard ratio; CI: confidence interval; IR: incidence rate; VTE: venous thromboembolism; MDCR: Medicare Supplemental Database; CCAE: Commercial Database; Optum: Optum De-Identified Clinformatics Data Mart Database; PharMetrics: PharMetrics™ Plus; THIN: The Health Improvement Network.