A comparison of the surgical practice of potential revision outlier joint replacement surgeons with non-outliers: A case control study from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man

Chris M Penfold[1+2*]

Michael R Whitehouse[1+2]

Adrian Sayers[1]

J Mark Wilkinson[3+4]

Linda Hunt[1]

Yoav Ben-Shlomo[5+6]

Andy Judge[1+2]

Ashley W Blom[1+2]

Affiliations:

1. Musculoskeletal Research Unit, Translational Health Sciences, Bristol Medical School, 1st Floor Learning & Research Building, Southmead Hospital, Bristol, BS10 5NB, UK
2. National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol.
3. Department of Oncology and Metabolism, University of Sheffield, Sorby Wing, Northern General Hospital, Sheffield, S5 7AU, UK
4. Centre for Integrated Research into Musculoskeletal Ageing, University of Sheffield, UK
5. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
6. The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West) at University Hospitals Bristol NHS Foundation Trust, UK

* - denotes corresponding author

Corresponding author contact details:

Email: chris.penfold@bristol.ac.uk

Address: Musculoskeletal Research Unit, University of Bristol, School of Clinical Sciences, Learning and Research Building, Southmead Hospital, Bristol, BS10 5NB

Telephone: 0117 41 47872

Data sharing statement

Access to the data analysed in this study required permission from the National Joint Registry for England, Wales and Northern Ireland Research Sub-committee. http://www.njrcentre.org.uk/njrcentre/Research/Researchrequests/tabid/305/Default.aspx contains information on research data access request to the National Joint Registry.

Disclaimer

The views expressed represent those of the authors and do not necessarily reflect those of the National Joint Registry Steering Committee or Healthcare Quality Improvement Partnership, who do not vouch for how the information is presented. The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Ethics approval

Patient consent was obtained for data collection by the National Joint Registry. According to the specifications of the NHS Health Research Authority, separate informed consent and ethical approval were not required for the present study.

Contributors

CP, AB, AJ and MW designed the study. CP, AB, AS, JMW, LH, AJ, MW and YBS reviewed the published work. CP conducted the statistical analysis and wrote the report. All contributors reviewed and agreed the final version before submission. CP had full access to all the data and AB is the guarantor.

Funding

This study was funded by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. 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 and Social Care. AS was supported by a MRC fellowship MR/L01226X/1. YBS is partly funded by National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West (CLAHRC West) at University Hospitals Bristol NHS Foundation Trust.

Abstract

Background

The National Joint Registry for England, Wales, Northern Ireland and the Isle of Man (NJR) has monitored the performance of consultant surgeons performing primary total hip (THR) or knee replacements (KR) since 2007. The aims of this study were: 1) To describe the surgical practice of consultant hip and knee replacement surgeons in the National Joint Registry for England and Wales (NJR), stratified by potential outlier status for revisions. 2) To compare the practice of revision outlier and non-outlier surgeons.

Patients and Methods

We combined NJR primary THR and KR data from 2008-2017 separately with relevant anonymised NJR outlier notification records. We described the surgical practice of outliers and non-outliers by surgical workload, implant choice, and patients' clinical and demographic characteristics. We explored associations between surgeon-level factors and outlier status with conditional logistic regression models.

Results

We included 764,888 primary THRs by 3,213 surgeons and 889,954 primary KRs by 3,084 surgeons performed between 2008-2017. One hundred and eleven (3.5%) THR and 114 (3.7%) KR consultant surgeons were potential revision outliers. Surgeons who used more types of implant had increased odds of being an outlier (KR: OR/additional implant=1.35, 95%CI 1.17-1.55; THR: OR=1.12, 95%CI 1.06-1.18).

Conclusions

The use of more types of implant is associated with increased risk of being a potential revision outlier. Further research is required to understand why surgeons use many different implants and to what extent this is responsible for the effects observed here.

Keywords

Orthopaedics, joint replacement, surgeon, national joint registry, performance monitoring

Acknowledgements

We thank the patients and staff of all the hospitals who have contributed data to the National Joint Registry. We are grateful to the Healthcare Quality Improvement Partnership, the National Joint Registry Steering Committee, and staff at the National Joint Registry for facilitating this work. This study was funded by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.

Introduction

Total hip replacements (THR) and partial (unicompartmental or patellofemoral) or total knee replacements (KR) are amongst the most common elective surgical procedures performed. In total, more than 200,000 primary THRs and KRs were performed in the UK in 2017.[1,2] Elective THRs and KRs are mainly performed to relieve pain and the functional limitations of osteoarthritis (OA).[3] They are successful procedures in which most patients achieve improvements in pain and function.

In the UK, monitoring of surgical performance is undertaken in many surgical specialties, including adult cardiac surgery, oesophago-gastric cancer surgery, bowel cancer resection and hip fracture surgery,[4] as well as joint replacement surgery.[5] The National Joint Registry for England, Wales, Northern Ireland and the Isle of Man (NJR) has monitored the performance of consultant surgeons and units performing THRs and KRs since 2007. The NJR monitors two main outcomes: the rate at which surgery is performed 'to add, remove or modify one or more components or conduct a DAIR (debridement, antibiotics and implant retention) of a total joint prosthesis' (revision surgery)[6] and the rate of mortality within 90 days postoperatively. The performance of each surgeon is compared with their peers (similar comparisons are made for units) and those with mortality/revision rates outside the accepted limits are considered to be 'potential outliers'. For the purpose of this study we focussed on revision outlier consultant surgeons.

Although surgeon performance is monitored we do not know, higher rates of revision surgery aside, whether outlying surgeons differ from non-outlying surgeons. Studies of patient-level and surgeon-level factors associated with the revision risk of primary operations suggest that we might expect outlier surgeons to differ from non-outlier surgeons in respect of these factors. Outlier surgeons may perform a lower volume of operations[7,8] or use a wider range of implants[9] than non-outliers. Alternatively, they may operate on a higher proportion of patients at higher risk of revision, such as younger patients who have a higher lifetime revision risk,[10] or people with elevated body mass index (BMI).[11]

Revision outlier status is a composite surgeon-level outcome incorporating revision rate, volume of cases and patient case-mix. Previous research has focussed on outcomes at the individual-level (i.e. revision risk of individual joint replacements) and may not directly relate to this composite surgeon-level outcome. A better understanding of the surgical practice of revision outlier surgeons compared with their peers may help to inform feedback to revision outlier surgeons and improve outcomes for patients. This study has two main aims:

1. To describe the practice of revision outlier surgeons

2. To compare the practice of revision outlier and non-outlier surgeons

Patients and Methods

Data source

In this study we combined anonymised records from the NJR with anonymised records from the NJR outlier notification process. Data collection in the NJR started in 2003 and includes details of primary and revision hip and knee replacement episodes.[1] Northern Ireland and the Isle of Man were included in the NJR in 2013 and 2015 respectively, and data linkage for those periods is limited, therefore they are excluded from this analysis.

The NJR outlier notification process started in 2007 and is undertaken every six months (notifications were annual for the first three years). The method used by the NJR to identify outliers is described elsewhere.[5,12,13] Briefly, every primary operation performed up to the date of outlier monitoring by each consultant surgeon responsible for a procedure is eligible to be included. The patient time incidence rate (PTIR) is calculated, which for each surgeon is their number of primary operations revised for the first time divided by the total time their primary operations were at risk; i.e. the time until they were either revised, the patient died or the patient is alive and the implant has not yet been revised. The PTIR is used to calculate the standardised revision ratio. Revised primary operations are allocated to the consultant in charge of the original primary, regardless of who performed the revision surgery. Consultant surgeons responsible for a procedure with a standardised revision ratio above the 99.8% control limit adjusted for age, gender and indication for primary surgery are flagged as potential revision outliers (see sensitivity analyses for exception). An anonymised list of all consultant surgeons (by anonymised NJR surgeon ID) who have been identified as outliers for each outlier notification period is maintained by the NJR. We matched these records to the NJR dataset by anonymised NJR surgeon ID.

Some surgeons remained outliers over several consecutive outlier notification periods or became an outlier more than once. Since outliers may change their surgical practice after being notified of their outlier status we only included the first time each outlying surgeon became an outlier (the first 'outlier event') and excluded any consecutive periods being an outlier or subsequent outlier events.

Study samples

We defined separate study samples for THRs and KRs and included all THRs/KRs performed for any indication respectively. We excluded surgeons who had stopped performing THRs/KRs one year before the first NJR outlier notification date (October 2007). Since surgeons may become outliers several years after their last THR/KR, we classified surgeons who had not performed a THR/KR in a 12-month period as no longer performing THRs/KRs respectively. Surgeons who performed another THR/KR after this period were re-included as well as the intervening period of non-activity. We included surgeons who had stopped performing THRs/KRs in our description of how many surgeons have ever been identified as a potential revision outlier and their cumulative number of operations and revisions, but since they had performed no relevant operations in the 12 months before becoming a potential revision outlier they were excluded from further analyses.

Surgical practice

We aggregated the practice of outlier and never outlier surgeons, and the characteristics of the patients they operated on over the 12 months prior to the date of each outlier notification. We characterised surgeons' practice according to four domains (see Table 1 for full details): 1) Surgical workload, 2) Choice of implants, 3) Patients' characteristics, 4) Source of funding for THRs/KRs.

The process used by the NJR to identify KR provisional revision outliers is not broken down by type of KRs (total, unicompartmental or patellofemoral), although the feedback to surgeons from the NJR does include separate funnel plots for each KR type. In our main analyses we therefore did not distinguish between types of KR but we have included this in our sensitivity analyses (detailed below).

Statistical analysis

We analysed THRs and KRs separately. We described the overall surgical practice of 'outliers' (cases). We also described the practice of 'never outliers' (controls) but did not compare cases and controls since the study design requires that they be matched (see below), and some operations performed by never outliers could contribute to multiple outlier notification periods. Unless otherwise stated, our unit of analysis was surgeon-year, i.e. we summarised the operations performed by each surgeon over 12-months. Since there was a low number of revision outliers, we did not describe surgical behaviour in each outlier notification period to avoid potential deanonymisation.

We used a time-matched case control design (also known as 'incidence density sampling') to compare the surgical practice of outlier surgeons in the 12-months before they were identified as being an outlier with time-matched controls. All eligible controls were identified as the surgeon-years of never-outlying surgeons, which were matched with cases by the time the case was identified. Controls were only selected from surgeons who were never outliers. Outlier surgeons were included only in the period immediately prior to becoming an outlier, never outlier surgeons were eligible to be controls for multiple cases and in multiple outlier periods. We derived unadjusted and multivariable adjusted (simultaneous adjustment for all exposure variables) odds ratios for being an outlier using conditional logistic regression models with time matching by outlier notification period and robust standard errors. We assessed potential multicollinearity through the variance inflation factor (VIF) and considered a VIF ≤10 to indicate that multicollinearity was not a concern.

Sensitivity analyses

We repeated our main analysis with the following changes:

1. Since the optimal time-period over which to characterise surgical practice is unknown, we used 24 and 60-month periods to determine whether our results were dependent on the time-period.
2. Different types of KR procedures have different associated revision rates which may affect associations in our main analyses. We included further adjustment for the proportion of Unicompartmental and patellofemoral KRs performed.
3. Since we used aggregated data, low-volume surgeons contributed the same weight as high-volume surgeons, which may have biased our results. We excluded surgeons performing below the 25th percentile in terms of volume.
4. Between 2007 and 2010 the NJR outlier process used unadjusted standardised revision ratios. This may have identified surgeons with different surgical behaviour. We excluded these outlier periods and compared the results.
5. Our inclusion of age, gender, ASA grade and indication for surgery may duplicate adjustment in the NJR outlier process. We excluded these and compared the results.

Analyses were performed using Stata v15 (StataCorp).

Results

Our study sample included 764,888 primary THRs and 889,954 primary KRs performed by 3,416 consultant surgeons, of whom 3,213 performed one or more THRs and 3,084 one or more KRs. These operations were spread across a total of 33,374 surgeon-years for THRs and 33,737 surgeon-years for KRs. Two hundred and seven surgeons (6.0%) have been identified as either a THR or KR revision outlier, 18 of whom (8.7%) have been identified as both a THR and KR revision outlier (14 were simultaneous outliers). One hundred and eleven of 3,213 THR surgeons (3.5%) and 114 of 3,084 KR surgeons (3.7%)

have been identified as THR and KR revision outliers respectively. Fifteen percent (17 of 111) THR revision outliers and 21.9% (25 of 114) KR revision outliers had stopped performing primary THRs/KRs at the time of their outlier notification and were excluded from further analyses.

When they first became revision outliers, these surgeons had performed a median total of 289 (IQR 154 to 544) THRs and 338 (IQR 199 to 606) KRs as consultant in charge in the NJR and had accrued a median of 11 (IQR 8 to 21) and 15 (IQR 10 to 25) revisions for primary THRs and KRs respectively. The median time to revision for outlier surgeons was 2.2 years for THRs (IQR 0.5 to 4.4 years) and 1.9 years for KRs (IQR 1.0 to 3.4 years), compared with 2.5 years for THRs (25%-75%: 0.6 to 5.4) and 2.3 years for KRs (25%-75%: 1.2 to 4.3) for never outliers.

For our descriptive analyses and conditional logistic regression models we included only the surgeon-year for each outlying surgeon that immediately preceded their first outlier event (see Figures S1 and S2 for study sample flowcharts). Our resultant study samples were 24,684 surgeon-years for THRs (24,601 surgeon-years for never outliers, 83 surgeon-years for outliers) and 27,824 surgeon-years for KRs (27,741 surgeon-years for never outliers, 83 surgeon-years for outliers).

Description of outliers and non-outliers

A crude comparison of the surgical practice of potential outlier and non-outlier surgeons indicates differences between surgeons in these groups, many of which were consistent between THR and KR outliers. Outlying surgeons performed more operations than non-outliers in the 12 months prior to becoming an outlier (THR: 59 vs. 17; KR: 47 vs. 24, outlier and non-outlier respectively, Tables 2 and 3). Outliers used more implant combinations than non-outliers (THR: 5 vs. 3; KR: 3 vs. 2). Compared with non-outliers, a higher proportion of operations performed by outliers were on patients <55 years old (THR: 10.7% vs. 4.2%; KR: 6.7% vs. 3.6%) and privately funded (THR: 11.6% vs. 0.0%; KR: 9.8% vs. 0.0%).

We found some differences only between THR outliers and non-outliers. A higher proportion of THR outliers than non-outliers used new implants for ≥10% of their operations (59.0% vs. 31.7%), and overall a much higher proportion of THR surgeons than KR used ‘new’ implants. There was a slight difference in joint specialisation between THR outliers and non-outliers (THR: 57.2% vs. 48.6%) but not for KR outliers (KR: 58.5% vs. 54.4%). We found no difference in the patient case-mix of outliers and non-outliers according to the proportion of patients with other indications, female patients, patients with a high ASA grade, and who were obese class II/III.

Factors associated with being a revision outlier

In our multivariable adjusted regression models, use of more implants was associated with increased odds of being a revision outlier for both THRs (OR/additional implant 1.12, 95% CI 1.06 to 1.18; Table 2) and KRs (OR/additional implant 1.35, 95%CI 1.17 to 1.55; Table 3). Surgeons who conducted a higher proportion of privately funded KRs compared had increased odds of being a revision outlier (OR/additional 10% private=1.19, 95%CI 1.10 to 1.30), but this was not associated with being a THR revision outlier (OR/additional 10%=0.99, 95%CI 0.93 to 1.06). For THRs, surgeons who performed a higher proportion of THRs to other joint replacements had higher odds of being an outlier (OR/10%=1.10, 95%CI 1.02 to 1.17) and there was weak evidence that higher volume THR surgeons had higher odds of being an outlier (OR/10 THRs=1.03, 95%CI 1.00 to 1.06). In terms of patient case-mix, THR surgeons had higher odds of being an outlier if they treated a higher proportion of females (OR=1.16, 95%CI 1.05 to 1.28) and patients younger than 55 years old (OR=1.22, 95%CI 1.09 to 1.35), but lower odds of being an outlier if they performed a higher proportion of THRs for indications other than osteoarthritis (OR=0.80, 95%CI 0.66 to 0.96). VIFs were all <10 (Table S1) despite high correlation between surgeon volume and number of implants used (r_{THR} = 0.57, r_{KR} = 0.49), therefore we did not modify our regression models due to multicollinearity.

The results from our unadjusted regression models are described briefly here and in detail in Supplementary Tables S2 and S3. For both THRs and KRs, surgeons who performed more THRs/KRs, those who used more types of implant, those who treated a higher proportion of patients younger than 55 years old, and those with a higher proportion of privately funded operations had increased odds of being a revision outlier. Associations between the extent to which surgeons specialised in performing THRs/KRs and their odds of being a revision outlier were inconsistent but suggest the degree of specialisation may be associated with being an outlier. Using a higher proportion of new implants may be associated with higher odds of being a revision outlier. For THR surgeons, having a higher proportion of ASA grade III-V patients and performing a higher proportion of THRs for indications other than osteoarthritis may be associated with lower odds of being a revision outlier. Whereas treating a higher proportion of female patients may be associated with increased odds of being an outlier.

Sensitivity analyses

Changing the time frame over which surgical practice was characterised from 12 to 24 and 60 months resulted in changes to the descriptive statistics (Supplementary Tables S4 to S7). The increase in the proportion of surgeons using new implants in the sensitivity analysis is due to including more operations from the first five years of the NJR. We defined implants as ‘new’ if they were used within five years of their first recorded use in the NJR, which covers all operations performed 2003-2008.

Results of our adjusted sensitivity analyses support the main finding of our primary analysis that using more implants is associated with higher odds of being a revision outlier for both THRs (OR/additional implant =1.97, 95%CI 1.04 to 1.12; Table S4) and KRs (OR/additional implant =1.23, 95%CI 1.11 to 1.37; Table S5). The sensitivity analyses also supported the association between performing a higher proportion of privately funded KRs and being a KR revision outlier (OR=1.16, 95%CI 1.08 to 1.25). An extension of the time frame to 60 months supported our main findings (Tables S6 and S7).

Our sensitivity analysis including the proportion of unicompartmental and patellofemoral KRs (Table S8) highlights that KR revision outliers performed a higher proportion of unicompartmental KRs than non-outliers (median=4.1% vs. 0.0%) and that this was associated with an increased odds of being a KR revision outlier (OR/10 percent=1.20, 95%CI 1.12 to 1.29). Whereas the proportion of operations which were patellofemoral KRs was very low for outliers and non-outliers, and this was not associated with being a revision outlier. Further sensitivity analyses in which we excluded low volume surgeons (Tables S9 and S10), removed the first five (unadjusted) outlier periods (Tables S11 and S12) and removed covariates already present in the NJR outlier process (Tables S13 and S14) made only minor changes to our findings.

Discussion

We used a time-matched case control study to explore differences between potential revision outliers and non-outliers in England and Wales according to surgeon, procedure and patient-level factors. We found that revision outlier consultant surgeons used a larger number of different hip or knee joint replacement implants than non-outlier surgeons.

The current study has several strengths. This is the first study to use the NJR outlier notification records to explore differences in the surgical practice of potential revision outliers and non-outliers. Since becoming a potential revision outlier is a rare event, the large size of the NJR dataset and 10 years of outlier notification records were essential to enable this study. Also, we used a time-matched case-control study design, which accounted for temporal variation in surgical trends. This study also has some important limitations. We characterised surgeon behaviour immediately prior to each outlier notification period, rather than prior to each primary operation, which may have been several years earlier. Our aim was to compare the surgical practice of outlier surgeons at the time they became outliers with that of non-outliers, but this may not reflect their practice when they performed the primary operations. Therefore, we cannot infer causality from our findings. Our selection of the 12-month timeframe over which to characterise surgical behaviour was arbitrary. We varied the timeframe in our sensitivity analysis to 24 and 60 months and found differences in our results, suggesting that the timeframe over which surgical practice is characterised may need further refinement. Some NJR records have missing data, particularly BMI (~30% missing, early data collection forms did not include BMI). We aggregated these records for each surgeon across each outlier notification period excluding any missing data. This assumes that these data were missing completely at random,[14] which may not be true. This aggregation of data may have resulted in biased estimates or inaccurate standard errors. Methods to incorporate individual-level factors into hierarchical models with the outcome measured at the group-level are developing, but at present aggregation of individual-level data at the group-level with robust standard errors may be preferable.[15] We have attributed factors to surgeons which may more accurately reflect unit-level approaches to surgery, for example some units may switch to newer implants or restrict surgeons' implant selection. Accommodating the influence of unit-level decisions on surgeons' behaviour was outside the scope of this study but may be of interest for future studies. Finally, as with any observational study, there may be other confounding factors we have not included in our analysis models (residual confounding).

Our main finding, consistent between THR and KR revision outliers, was that outliers used a larger range of implants than non-outliers at the time they became an outlier. The Australian Orthopaedic Association National Joint Replacement Registry found that surgeons who use a range of implants rather than relying on a small number of implants for most of their primary operations have a higher risk of early revision.[9] A possible explanation for this is that there is a learning-curve associated with changing implants, although evidence to support this is contradictory.[16,17] If the earlier joint replacement operations performed after switching to a different implant are at increased risk of revision, then surgeons who frequently switch implants will be in the 'learning phase' for a greater proportion of their procedures and expose more patients to higher revision risks.

The finding from our descriptive and unadjusted regression models of a positive association between surgeon volume and revision outlier status implies higher volume surgeons may be 'worse' than lower volume surgeons. However, this finding may also be an expected characteristic of outlier status identified through funnel plots. Previous research reported a lower risk of revision for primary joint replacements performed by higher volume surgeons,[7,18] although recent research contests this finding.[19] The outcome we used in our study, potential revision outlier status from control limits applied to funnel plots, is a composite measure which incorporates revision rate, surgeon volume and case-mix adjustment. Seaton and Manktelow[20] estimated the probability of detecting hospitals with true poor performance using funnel plots depending on their expected number of events. They found that hospitals with a low expected number of events (either a rare outcome or low-volume hospital) have a lower probability of being identified as a true poor performer compared with hospitals performing similarly and with similar case-mix but with a higher volume of cases. Furthermore, a high-volume compared with low-volume hospital is more likely to be identified as poorly performing for a relatively minor divergence. Our positive association between surgeon volume and being a revision outlier likely reflects this volume-related characteristic of funnel plots and control limits, and low volume surgeons may be 'protected' from becoming a revision outlier as a result.

We found that consultant surgeons in charge who perform a higher proportion of privately funded operations may be at increased odds of being a revision outlier, but this was not consistent between THR and KR outliers. We used source of funding to indicate the socioeconomic mix of patients treated by surgeons. With this interpretation, our finding suggests that revision outliers may have a case-mix favouring higher socioeconomic status. Studies in countries with universal health cover found no association between patient-level socioeconomic status and risk of revision.[21,22] Alternatively, source of funding may indicate an operation being performed in a private or NHS unit, subject to misclassification since some NHS funded operations are undertaken in private units. However, NHS units treat patients with more comorbidities[23] who may be expected to have a higher risk of revision[24] than private units. Consequently, surgeons working solely in the NHS should be at higher not lower risk of being a revision outlier. The inconsistency between this finding and previous research supports further research to explore source of funding as either a patient-level or unit-level risk factor.

Risk of revision is known to differ between total, unicompartmental and patellofemoral KRs.[1] Our finding that outlier surgeons conduct a higher proportion of unicompartmental KRs is therefore expected. The association between number of KR implants used and being an outlier persisted after inclusion of type of KR in the analyses. This may be because surgeons who perform two or all of these operations will use at least two or three different implants, and some of these operations have a higher revision risk. However, the validity of the association between number of implants used and being an outlier is reinforced by being found in both THRs and KRs.

Future research could build on this study in several ways. A study to explore whether primary joint replacements performed by surgeons who use a wider range of implants have a higher revision risk would build on the main finding of this study. In addition, it would also be useful to explore whether there is an upper limit for the number of implants used by a surgeon beyond which the risk of revision is higher. This could form the basis of surgeon feedback as part of routine performance monitoring but would need to account for the use of a wider range of implants for surgeons performing different types of KRs. This could be extended to consider the optimal time frame over which surgical practice should be characterised or whether stability/instability in surgical practice is more important than 'average' practice. Exploring unit-level factors associated with revision risk would help to direct guidance on surgical practice to the most appropriate source. Finally, there are other outcomes, particularly patient reported outcomes, which may also benefit from similar research.

References

1. **National Joint Registry for England, Wales and Northern Ireland**. 14th Annual Report 2017. 2017.

2. **Scottish Arthroplasty Project**. Annual Report 2018. NHS National Services Scotland, 2018.

3. **Dreinhöfer KE, Dieppe P, Stürmer T, Gröber-Grätz D, Flören M, Günther KP, et al.** Indications for total hip replacement: Comparison of assessments of orthopaedic surgeons and referring physicians. *Ann Rheum Dis* 2006;65(10):1346–1350.

4. **Walker K, Neuburger J, Groene O, Cromwell DA, Van Der Meulen J**. Public reporting of surgeon outcomes: Low numbers of procedures lead to false complacency. *The Lancet* 2013;382(9905):1674–1677.

5. **NJR Centre**. NJR Adopted Statistical Methodology for Potential Outlier Identification. http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Stat_Meth.pdf , (date last accessed 21 September 2018).

6. **National Joint Registry for England, Wales, Northern Ireland and the Isle of Man**. NJR 16th Annual Report 2019. 2019.

7. **Ravi B, Jenkinson R, Austin PC, Croxford R, Wasserstein D, Escott B, et al.** Relation between surgeon volume and risk of complications after total hip arthroplasty: propensity score matched cohort study. *BMJ* 2014;348:g3284.

8. **Wilson SM, Marx RG, Pan TJ, Lyman S**. Meaningful thresholds for the volume-outcomes relationship in total knee arthroplasty. *J Bone Jt Surg* 2016;1683–1690.

9. **Australian Orthopaedic Association**. Hip, Knee & Shoulder Arthroplasty: Annual Report 2017. Australian Orthopaedic Association, 2017.

10. **Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, et al.** The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *The Lancet* The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY license, 2017;389(10077):1424–1430.

11. **Mouchti S, Whitehouse MR, Sayers A, Hunt LP, MacGregor A, Blom AW**. The Association of Body Mass Index with Risk of Long-Term Revision and 90-Day Mortality Following Primary Total Hip Replacement: Findings from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. *JBJS* 2018;100(24):2140–2152.

12. **Spiegelhalter DJ**. Funnel plots for comparing institutional performance. *Stat Med* 2005;24(8):1185–1202.

13. **Tucker K, Gregg P, Kay P, Porter M, Howard P, Pickford M, et al.** Monitoring the Introduction and Performance of a Joint Replacement: The United Kingdom Metal-on-Metal Alert. *J Bone Jt Surg-Am Vol* 2011;93(Suppl 3):37–42.

14. **Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al.** Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.

15. **Foster-Johnson L, Kromrey JD**. Predicting group-level outcome variables: An empirical comparison of analysis strategies. *Behav Res Methods* 2018;50(6):2461–2479.

16. **Peltola M, Malmivaara A, Paavola M**. Hip prosthesis introduction and early revision risk. *Acta Orthop* 2013;84(1):25–31.

17. **Mohaddes M, Björk M, Nemes S, Rolfson O, Jolbäck P, Kärrholm J**. No increased risk of early revision during the implementation phase of new cup designs: Analysis of 52,903 hip arthroplasties reported to the Swedish Hip Arthroplasty Register. *Acta Orthop* 2016;87(sup1):31–36.

18. **Manley M, Ong K, Lau E, Kurtz SM**. Effect of Volume on Total Hip Arthroplasty Revision Rates in the United States Medicare Population. *JBJS* 2008;90(11):2446–2451.

19. **Sayers A, Steele F, Whitehouse MR, Price A, Ben-Shlomo Y, Blom AW**. Association between surgical volume and failure of primary total hip replacement in England and Wales: findings from a prospective national joint replacement register. *BMJ Open* 2020;10(9):e033045.

20. **Seaton SE, Barker L, Lingsma HF, Steyerberg EW, Manktelow BN**. What is the probability of detecting poorly performing hospitals using funnel plots? *BMJ Qual Saf* BMJ Publishing Group Ltd, 2013;22(10):870–876.

21. **Agabiti N, Picciotto S, Cesaroni G, Bisanti L, Forastiere F, Onorati R, et al.** The influence of socioeconomic status on utilization and outcomes of elective total hip replacement: a multicity population-based longitudinal study. *Int J Qual Health Care* 2007;19(1):37–44.

22. **Dy CJ, Marx RG, Bozic KJ, Pan TJ, Padgett DE, Lyman S**. Risk factors for revision within 10 years of total knee arthroplasty. *Clin Orthop* 2014;472(4):1198–1207.

23. **Mason A, Street A, Verzulli R**. Private sector treatment centres are treating less complex patients than the NHS. *J R Soc Med* 2010;103(8):322–331.

24. **Podmore B, Hutchings A, Van Der Meulen J, Aggarwal A, Konan S**. Impact of comorbid conditions on outcomes of hip and knee replacement surgery: A systematic review and meta-analysis. *BMJ Open* 2018;8(7):1–12.

25. **Penfold CM, Blom AW, Sayers A, Wilkinson JM, Hunt L, Judge A, et al.** Understanding the uptake of new hip replacement implants in the UK: a cohort study using data from the National Joint Registry for England and Wales. *BMJ Open* [Internet] 2019 [cited 27 Nov 2019];9(11).

26. **Penfold CM, Blom AW, Wilkinson JM, Judge A, Whitehouse MR**. The Uptake of New Knee Replacement Implants in the UK: Analysis of the National Joint Registry for England and Wales. *J Arthroplasty* 2020;35(3):699-705.e3.

27. **WHO Expert Committee**. Physical Status: The Use and Interpretation of Anthropometry. Report of a WHO Expert Committee. Tech. Rep. Ser. No 854. Geneva: World Health Organization, 1995.

Table 1: A description of the variables used to describe surgeons' practice

	Variable	Definition	Details
1. Surgical workload	a. Volume of THR/KR	The total number of THR/KR performed in the time-period	Continuous variable
	b. Proportion of THRs/KRs to other joints	The proportion of all joint replacement operations performed by the consultant surgeon which were THR/KR	Median centred, continuous variable
2. Choice of implant	a. Number of implants used	The number of different combinations of femoral and acetabular components (for THR)[25] or implant component brands (for KR)[26] used	Continuous variable
	b. Proportion of new hip/knee implants	The proportion of implants used with a first recorded use in the NJR in the previous five years	1. <10% <5 yrs old 2. ≥10% <5 yrs old
3. Patient characteristics	a. BMI	The proportion of patients who were obese class II/III[27] at the time of the operation	Median centred, continuous variable
	b. ASA grade	The proportion of people who had an ASA grade III-V at the time of their operation	Median centred, continuous variable
	c. Reason for primary operation	The proportion of operations performed on people with a non-osteoarthritis indication	Median centred, continuous variable
	d. Age	The proportion of operations performed on patients <55 years old	Median centred, continuous variable
	e. Gender	The proportion of patients who were female	Median centred, continuous variable
4. Funding for operation	a. Source of funding	The proportion of operations privately funded	Median centred, continuous variable

Table 2: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for total hip replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR[2]**	**(95% CI)**	**p**
	N=2,438		n=83				
Volume of THRs (median + IQR)	17	4-44	59	38-103	1.03	1.00-1.06	0.036
Proportion of THRs to other joints	48.6%	33.3%-63.6%	57.2%	46.8%-73.2%	1.10	1.02-1.17	0.006
Number of hip implants used (median + IQR)	3	1.5-4	5	4-9	1.12	1.06-1.18	<0.001
Proportion of new hip implants							
<10% <5 yrs old (ref)	1,665	68.3%	34	41.0%	1	-	-
≥10% <5 yrs old	773	31.7%	49	59.0%	1.39	0.80-2.43	0.242
Proportion obese class II/III	9.1%	0.0%-14.3%	11.1%	7.0%-16.4%	0.96	0.86-1.07	0.455
Proportion ASA grade ≥III	14.7%	5.6%-23.3%	15.4%	7.7%-19.7%	0.96	0.88-1.04	0.286
Proportion of primary operations for other indications (median centered)[4]	7.1%	0.0%-16.1%	6.7%	2.3%-15.0%	0.80	0.66-0.96	0.017
Proportion of THRs on people <55 years old?	4.2%	0.0%-9.7%	10.7%	6.4%-14.3%	1.22	1.09-1.35	<0.001
Proportion of female patients	61.5%	56.8%-66.7%	62.2%	56.4%-70.0%	1.16	1.05-1.28	0.004
Proportion of primary operations privately funded (median centered)[4]	0.0%	0.0%-8.6%	11.6%	1.7%-23.0%	0.99	0.93-1.06	0.779

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table 3: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR[2]**	**(95% CI)**	**p**
	N=2,505		n=83				
Volume of KRs (median + IQR)	24	7-52	47	28-92	1.02	0.97-1.07	0.401
Proportion of KRs to other joints	54.4%	41.1%-75.0%	58.5%	46.2%-90.7%	1.01	0.94-1.09	0.733
Number of knee implants used (median + IQR)	2	1-3	3	2-5	1.35	1.17-1.55	<0.001
Proportion of new knee implants							
<10% <5 yrs old (ref)	2,266	90.5%	68	81.9%	1	-	-
≥10% <5 yrs old	239	9.5%	15	18.1%	1.28	0.64-2.58	0.484
Proportion obese class II/III	20.2%	12.1%-26.9%	21.7%	12.5%-30.8%	1.05	0.96-1.15	0.255
Proportion ASA grade ≥III	14.8%	7.1%-22.1%	15.6%	9.1%-21.4%	1.03	0.93-1.13	0.625
Proportion of primary operations for other indications (median centered)[4]	0.0%	0.0%-3.3%	1.9%	0.0%-4.0%	1.00	0.69-1.45	1.000
Proportion of KRs on people <55 years old?	3.6%	0.0%-7.6%	6.7%	2.9%-12.0%	0.93	0.75-1.16	0.523
Proportion of female patients	56.9%	52.1%-62.3%	56.7%	50.0%-63.0%	0.96	0.87-1.07	0.503
Proportion of primary operations privately funded (median centered)[4]	0.0%	0.0%-6.9%	9.8%	0.0%-30.0%	1.19	1.10-1.30	<0.001

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Supplementary Material

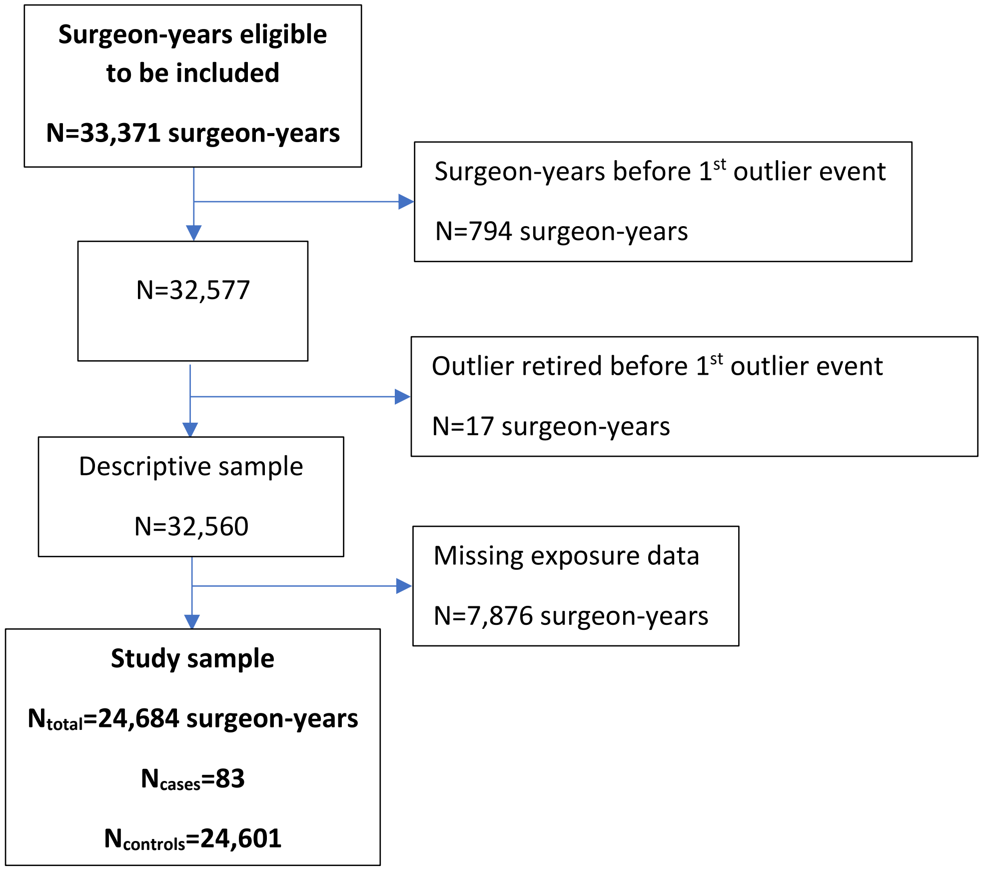

Figure S1: STROBE Flow diagram for the selection of cases and controls: Total hip replacement

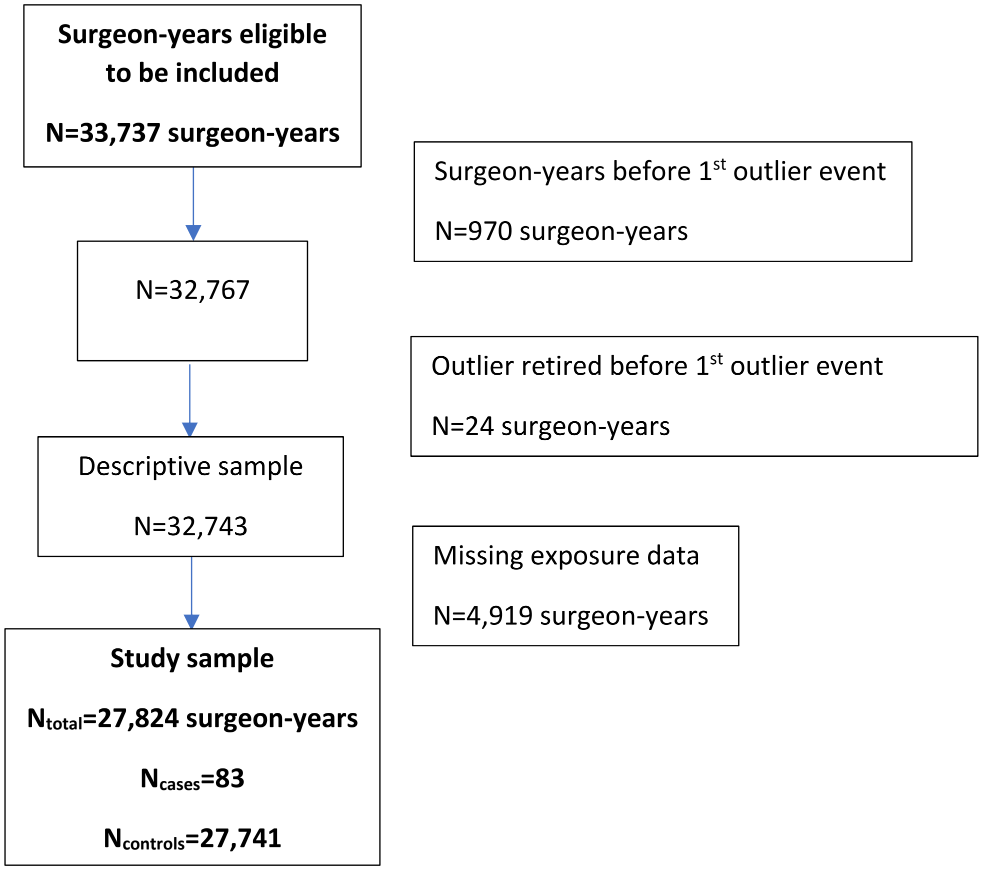

Figure S2: STROBE Flow diagram for the selection of cases and controls: Knee replacement

Table S1 Variance Inflation Factors (VIFs) from regression models of THR and KR outlier status

Variable	**THRs**	**KRs**
Volume of THRs/KRs	1.70	1.38
Proportion of THRs/KRs to other joints	1.22	1.09
Number of hip/knee implants used	1.72	1.43
Proportion of new hip/knee implants	1.15	1.04
Proportion obese class II/III	1.02	1.05
Proportion ASA grade ≥III	1.08	1.06
Proportion of primary operations for other indications	1.11	1.03
Proportion of THRs/KRs on people <55 years old?	1.11	1.09
Proportion of female patients	1.03	1.02
Proportion of primary operations privately funded	1.04	1.06

Table S2: Results from unadjusted conditional logistic regression models of being a total hip replacement revision outlier

	OR[2]	(95% CI)	p
Volume of THRs (median + IQR)	1.08	1.06-1.09	<0.001
Proportion of THRs to other joints	1.20	1.12-1.28	<0.001
Number of hip implants used (median + IQR)	1.18	1.14-1.22	<0.001
Proportion of new hip implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	2.19	1.29-3.71	0.004
Proportion obese class II/III	0.96	0.89-1.04	0.337
Proportion ASA grade ≥III	0.90	0.85-0.96	0.001
Proportion of primary operations for other indications (median centered)[4]	0.89	0.79-1.00	0.053
Proportion of THRs on people <55 years old?	0.22	1.14-1.31	<0.001
Proportion of female patients	1.07	1.01-1.14	0.024
Proportion of primary operations privately funded (median centered)[4]	1.06	1.01-1.12	0.023

1 – Odds ratios, 95% confidence intervals and p-values are from unadjusted conditional logistic regression models

2 – Odds ratios per additional 10 patients

3 – Odds ratios per additional 10 percent

Table S3: Results from unadjusted conditional logistic regression models of being a knee replacement revision outlier

	OR[2]	**(95% CI)**	**p**
Volume of KRs (median + IQR)	1.07	1.04-1.11	<0.001
Proportion of KRs to other joints	1.10	1.01-1.19	0.021
Number of knee implants used (median + IQR)	1.40	1.26-1.56	<0.001
Proportion of new knee implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	2.16	1.04-4.48	0.038
Proportion obese class II/III	1.01	0.94-1.09	0.838
Proportion ASA grade ≥III	0.95	0.85-1.06	0.343
Proportion of primary operations for other indications (median centered) [4]	1.03	0.81-1.30	0.836
Proportion of KRs on people <55 years old?	1.13	1.01-1.27	0.033
Proportion of female patients	0.97	0.90-1.05	0.530
Proportion of primary operations privately funded (median centered) [4]	1.18	1.09-1.27	<0.001

1 – Odds ratios, 95% confidence intervals and p-values are from unadjusted conditional logistic regression models

2 – Odds ratios per additional 10 patients

3 – Odds ratios per additional 10 percent

Table S4: Sensitivity analysis 1A: A description of the surgical practice over 24 months of never (controls) and ever (cases) revision outliers for total hip replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR**[2]	**(95% CI)**	**p**
	N=2,485		n=90				
Volume of THRs (median + IQR)	31	6-81	115	64-194	1.01	0.99-1.03	0.179
Proportion of THRs to other joints	48.0%	32.8%-62.6%	57.1%	47.6%-71.4%	1.14	1.02-1.26	0.021
Number of hip implants used (median + IQR)	3	2-6	8	5-11	1.09	1.04-1.13	<0.001
Proportion of new hip implants							
<10% <5 yrs old (ref)	1,529	61.5%	29	32.2%	1	-	-
≥10% <5 yrs old	956	38.5%	61	67.8%	1.66	0.99-2.78	0.055
Proportion obese class II/III	10.0%	0.0%-14.8%	11.3%	7.7%-16.7%	1.05	0.90-1.23	0.528
Proportion ASA grade ≥III	15.4%	7.4%-23.6%	14.8%	7.4%-23.4%	0.91	0.75-1.11	0.364
Proportion of primary operations for other indications (median centered) [4]	7.8%	1.7%-16.7%	7.9%	3.5%-13.6%	0.84	0.66-1.05	0.122
Proportion of THRs on people <55 years old?	5.0%	0.0%-10.0%	9.1%	6.6%-12.6%	1.08	0.87-1.34	0.475
Proportion of female patients	61.4%	57.0%-66.7%	62.1%	56.7%-65.7%	0.99	0.84-1.17	0.879
Proportion of primary operations privately funded (median centered) [4]	0.0%	0.0%-9.2%	11.7%	2.8%-25.0%	1.01	0.91-1.12	0.865

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S5: Sensitivity analysis 1B: A description of the surgical practice over 24 months of never (controls) and ever (cases) revision outliers for knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR**[2]	**(95% CI)**	**p**
	N=2,554		n=90				
Volume of KRs (median + IQR)	44	12-99	98	49-163	1.01	0.99-1.04	0.182
Proportion of KRs to other joints	53.6%	40.6%-72.7%	58.8%	45.6%-83.3%	1.03	0.95-1.12	0.517
Number of knee implants used (median + IQR)	2	1-3	4	2-6	1.23	1.11-1.37	<0.001
Proportion of new knee implants							
<10% <5 yrs old (ref)	2,237	87.6%	62	68.9%	1	-	-
≥10% <5 yrs old	317	12.4%	28	31.1%	1.41	0.79-2.52	0.241
Proportion obese class II/III	20.6%	13.1%-27.6%	21.2%	14.4%-27.8%	0.98	0.88-1.09	0.722
Proportion ASA grade ≥III	15.4%	8.3%-22.3%	16.5%	8.38%-21.1%	1.10	0.94-1.30	0.225
Proportion of primary operations for other indications (median centered) [4]	1.5%	0.0%-3.8%	2.1%	0.4%-4.3%	0.88	0.63-1.24	0.463
Proportion of KRs on people <55 years old?	4.1%	0.0%-7.7%	8.6%	4.4%-12.3%	1.09	0.97-1.22	0.152
Proportion of female patients	56.9%	52.3%-62.2%	56.4%	52.4%-60.4%	0.94	0.82-1.08	0.418
Proportion of primary operations privately funded (median centered) [4]	0.0%	0.0%-7.2%	7.9%	0.8%-26.8%	1.16	1.08-1.25	<0.001

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S6: Sensitivity analysis 1C: A description of the surgical practice over 60 months of never (controls) and ever (cases) revision outliers for total hip replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR**[2]	**(95% CI)**	**p**
	N=2524		N=99				
Volume of THRs (median + IQR)	57	12-161	224	126-403	1.00	0.99-1.01	0.606
Proportion of THRs to other joints	47.4%	32.4%-60.4%	56.3%	46.5%-71.3%	1.14	1.01-1.28	0.029
Number of hip implants used (median + IQR)	5	2-9	12	8-18	1.08	1.06-1.10	<0.001
Proportion of new hip implants							
<10% <5 yrs old (ref)	1,147	45.4%	13	13.1%	1	-	-
≥10% <5 yrs old	1,377	54.6%	86	86.9%	2.13	1.30-3.46	0.002
Proportion obese class II/III	10.4%	0.0%-15.3%	11.7%	7.7%-16.6%	1.04	0.96-1.13	0.335
Proportion ASA grade ≥III	15.7%	8.5%-23.8%	14.2%	8.5%-20.8%	1.00	0.83-1.22	0.961
Proportion of primary operations for other indications (median centered) [4]	8.3%	3.0%-16.7%	7.5%	4.0%-12.8%	0.79	0.64-0.98	0.032
Proportion of THRs on people <55 years old?	5.3%	0.0%-10.1%	9.0%	5.2%-12.0%	1.16	1.01-1.32	0.033
Proportion of female patients	61.5%	57.3%-66.7%	62.0%	57.4%-64.5%	0.87	0.77-0.98	0.024
Proportion of primary operations privately funded (median centered) [4]	0.0%	0.0%-10.3%	12.6%	2.3%-24.8%	1.00	0.94-1.06	0.960

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S7: Sensitivity analysis 1D: A description of the surgical practice over 60 months of never (controls) and ever (cases) revision outliers for knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR**[2]	**(95% CI)**	**p**
	N=2,600		n=105				
Volume of KRs (median + IQR)	84	22-206	213	109-369	1.01	1.00-1.02	0.144
Proportion of KRs to other joints	52.9%	40.7%-68.9%	56.9%	46.1%-86.1%	1.03	0.96-1.10	0.482
Number of knee implants used (median + IQR)	3	2-5	5	3-8	1.18	1.11-1.26	<0.001
Proportion of new knee implants							
<10% <5 yrs old (ref)	1,932	74.3%	53	40.5%	1	-	-
≥10% <5 yrs old	668	25.7%	52	49.5%	1.65	0.98-2.77	0.058
Proportion obese class II/III	20.8%	13.7%-27.3%	20.8%	15.7%-26.4%	1.04	0.91-1.20	0.537
Proportion ASA grade ≥III	15.4%	9.0%-22.3%	14.4%	9.3%-19.8%	1.05	0.88-1.25	0.593
Proportion of primary operations for other indications (median centered) [4]	2.0%	0.0%-4.4%	2.0%	0.9%-4.4%	0.71	0.33-1.51	0.370
Proportion of KRs on people <55 years old?	4.5%	0.8%-7.9%	7.9%	4.5%-11.7%	1.13	0.93-1.38	0.230
Proportion of female patients	56.9%	52.6%-61.8%	56.8%	53.1%-60.6%	0.93	0.77-1.14	0.489
Proportion of primary operations privately funded (median centered) [4]	0.8%	0.0%-8.0%	7.8%	0.5%-21.7%	1.10	1.00-1.20	0.043

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S8: Sensitivity analysis 2: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for knee replacements including indicators for performing partial knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	Control[1]		**Case**		**OR[2]**	**(95% CI)**	**p**
	N=2,505		n=83				
Proportion of Unicompartmental KRs	**0.0%**	**0.0%-2.8%**	**4.1%**	**0.0%-20.0%**	**1.20**	**1.12-1.29**	**<0.001**
Proportion of Patellofemoral KRs	**0.0%**	**0.0%-0.0%**	**0.0%**	**0.0%-1.5%**	**1.26**	**0.75-2.11**	**0.392**
Volume of KRs (median + IQR)	24	7-52	47	28-92	1.03	0.98-1.08	0.316
Proportion of KRs to other joints	54.4%	41.1%-75.0%	58.5%	46.2%-90.7%	1.00	0.92-1.08	0.903
Number of knee implants used (median + IQR)	2	1-3	3	2-5	1.30	1.11-1.51	0.001
Proportion of new knee implants							
<10% <5 yrs old (ref)	2,266	90.5%	68	81.9%	1	-	-
≥10% <5 yrs old	239	9.5%	15	18.1%	1.35	0.68-2.67	0.392
Proportion obese class II/III	20.2%	12.1%-26.9%	21.7%	12.5%-30.8%	1.06	0.67-1.15	0.221
Proportion ASA grade ≥III	14.8%	7.1%-22.1%	15.6%	9.1%-21.4%	1.03	0.93-1.14	0.264
Proportion of primary operations for other indications (median centered)[4]	0.0%	0.0%-3.3%	1.9%	0.0%-4.0%	1.00	0.69-1.47	0.980
Proportion of KRs on people <55 years old?	3.6%	0.0%-7.6%	6.7%	2.9%-12.0%	0.85	0.69-1.06	0.156
Proportion of female patients	56.9%	52.1%-62.3%	56.7%	50.0%-63.0%	0.99	0.89-1.09	0.779
Proportion of primary operations privately funded (median centered)[4]	0.0%	0.0%-6.9%	9.8%	0.0%-30.0%	1.16	1.06-1.27	0.001

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S9: Sensitivity analysis 3A – exclusion of low volume cases: Results from multivariable adjusted conditional logistic regression models of being a total hip replacement revision outlier

	OR[2]	**(95% CI)**	**p**
Volume of THRs (median + IQR)	1.02	0.99-1.06	0.187
Proportion of THRs to other joints	1.08	0.98-1.19	0.102
Number of hip implants used (median + IQR)	1.10	1.04-1.16	0.002
Proportion of new hip implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.30	0.75-2.27	0.348
Proportion obese class II/III	0.96	0.81-1.13	0.593
Proportion ASA grade ≥III	0.99	0.86-1.14	0.876
Proportion of primary operations for other indications (median centered) [4]	0.73	0.54-0.99	0.041
Proportion of THRs on people <55 years old?	1.41	1.14-1.75	0.002
Proportion of female patients	1.30	1.06-1.59	0.011
Proportion of primary operations privately funded (median centered) [4]	0.99	0.92-1.06	0.748

Table S10: Sensitivity analysis 3B – exclusion of low volume cases: Results from multivariable adjusted conditional logistic regression models of being a knee replacement revision outlier

	OR[2]	**(95% CI)**	**p**
Volume of KRs (median + IQR)	1.01	0.96-1.06	0.714
Proportion of KRs to other joints	0.99	0.90-1.10	0.905
Number of knee implants used (median + IQR)	1.31	1.13-1.52	<0.001
Proportion of new knee implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.30	0.67-2.52	0.446
Proportion obese class II/III	1.05	0.93-1.19	0.407
Proportion ASA grade ≥III	1.04	0.89-1.22	0.589
Proportion of primary operations for other indications (median centered) [4]	0.64	0.25-1.66	0.360
Proportion of KRs on people <55 years old?	1.13	0.78-1.63	0.532
Proportion of female patients	0.90	0.68-1.20	0.477
Proportion of primary operations privately funded (median centered) [4]	1.20	1.08-1.33	0.001

Table S11: Sensitivity analysis 4A – outlier periods 2011 onwards: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for total hip replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	OR[2]	**(95% CI)**	**p**
Volume of THRs (median + IQR)	1.01	0.95-1.06	0.839
Proportion of THRs to other joints	1.12	0.98-1.28	0.105
Number of hip implants used (median + IQR)	1.13	1.04-1.22	0.003
Proportion of new hip implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.90	1.05-3.44	0.034
Proportion obese class II/III	0.97	0.77-1.21	0.765
Proportion ASA grade ≥III	0.95	0.75-1.19	0.641
Proportion of primary operations for other indications (median centered) [4]	0.71	0.52-0.97	0.034
Proportion of THRs on people <55 years old?	1.18	0.96-1.46	0.110
Proportion of female patients	1.14	0.93-1.39	0.211
Proportion of primary operations privately funded (median centered) [4]	0.98	0.85-1.12	0.767

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S12: Sensitivity analysis 4B – outlier periods 2011 onwards: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	OR[2]	**(95% CI)**	**p**
Volume of KRs (median + IQR)	1.02	0.96-1.09	0.460
Proportion of KRs to other joints	0.99	0.90-1.08	0.821
Number of knee implants used (median + IQR)	1.36	1.12-1.64	0.001
Proportion of new knee implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.55	0.65-3.66	0.320
Proportion obese class II/III	1.04	0.90-1.19	0.629
Proportion ASA grade ≥III	1.09	1.00-1.19	0.061
Proportion of primary operations for other indications (median centered) [4]	0.94	0.57-1.53	0.794
Proportion of KRs on people <55 years old?	0.82	0.54-1.25	0.357
Proportion of female patients	1.01	0.86-1.18	0.944
Proportion of primary operations privately funded (median centered) [4]	1.10	0.97-1.24	0.125

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S13: Sensitivity analysis 5A – intermediate adjustment: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for total hip replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	OR[2]	**(95% CI)**	**p**
Proportion of THRs to other joints	1.13	1.05-1.20	<0.001
Number of hip implants used (median + IQR)	1.14	1.10-1.18	<0.001
Proportion of new hip implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.35	0.76-2.39	0.306
Proportion obese class II/III	0.96	0.87-1.07	0.501
Proportion of primary operations privately funded (median centered) [4]	1.02	0.96-1.08	0.583

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent

Table S14: Sensitivity analysis 5B – intermediate adjustment: A description of the surgical practice over 12 months of never (controls) and ever (cases) revision outliers for knee replacements, and results from multivariable adjusted conditional logistic regression models of being a revision outlier

	OR[2]	**(95% CI)**	**p**
Proportion of KRs to other joints	1.02	0.94-1.10	0.653
Number of hip implants used (median + IQR)	1.37	1.22-1.55	<0.001
Proportion of new hip implants			
<10% <5 yrs old (ref)	1	-	-
≥10% <5 yrs old	1.26	0.63-2.50	0.517
Proportion obese class II/III	1.05	0.97-1.14	0.256
Proportion of primary operations privately funded (median centered) [4]	1.18	1.08-1.29	<0.001

1 – Since controls may contribute to >1 outlier period the descriptive statistics for controls have been averaged over all eligible outlier periods

2 – Odds ratios, 95% confidence intervals and p-values are from conditional logistic regression models adjusted for all exposure variables

3 – Odds ratios per additional 10 patients

4 – Odds ratios per additional 10 percent