Bias and Unicompartmental Knee Replacement

DW Murray, A Liddle, A Judge, H Pandit

Abstract

We recently published a paper comparing the incidence of adverse outcomes after Unicompartmental and Total Knee replacement (UKR & TKR). The conclusion of this study, which was in favour of UKR, was dismissed as biased in a review in Bone & Joint 360 [1] that was subsequently reprinted in the Bone and Joint Journal [2]. Although this study is one of the least biased comparisons of UKR and TKR, this episode highlights the biases that exist both towards and against UKR. In this review we give examples of different sources of bias and how they can be avoided based on our published paper, National Registry data and a meta-analysis of published data. Our conclusion is that comparisons between UKR and TKR are open to bias. These biases are so marked in comparisons made by National Registers based on revision rate that the conclusions are misleading. For a fair comparison data from randomised studies or well matched prospective observational cohort studies, which may be based on registry data, are required and multiple outcome measures should be used. The data of this type that already exists is in favour of UKR.

Adverse Outcomes Study

The study comparing the incidence of adverse outcomes after UKR and TKR was a matched study based on data from the NJR (National Joint Registry for England and Wales) and other large datasets [3]. The incidence of all adverse outcomes except reoperation and revision were lower after UKR. To put this in perspective we stated "If 100 patients receiving TKR received UKR instead; the result would be around one less death and three more reoperations in the first four years after surgery". The Bone & Joint 360 review, with which no author was associated, dismissed this as "Bias in reporting? Quite possibly" [1]. It also suggested that we should have made the comparison at an earlier time point, missing the point of paper which was a comparison of adverse outcomes not just immediately post-operatively but also out to eight years. In fact, we chose the time point arguably least favourable to UKR. At an earlier time point, for example at one year, both the death rate and the re-operation were less for UKR than for TKR. We did not quote the comparison at eight years as the numbers of patients with this length of follow-up was low, but if we had, the message would have been starker: "If 60 patients receiving TKR received UKR instead, the result would be around one less death and three more reoperations in the first 8 years after surgery".

Perhaps the bias is not in the presentation, but in the data and the analysis. The data is, however, freely available, standard and appropriate statistical techniques were used and we followed the STROBE guidelines for reporting of observational studies [4]. As it was not possible to answer our research question using a randomised controlled trial (RCT, such as TOPKAT [5]), due to the rarity of adverse events such as mortality, we used the next best study design which is a prospective observational cohort study. The most important potential bias in this study and most comparative studies is selection bias as TKR and UKR tend to be

implanted in different patients. To avoid this bias we matched patients with propensity score analysis [6], which provides balanced baseline characteristics in both groups. As every patient that had a UKR would have had a TKR had they been treated by a non-UKR surgeon virtually all UKR patients were matched with TKR controls. In addition all patients receiving TKR with no comparable UKR controls were eliminated. There is not enough data within the NJR for adequate matching so we used data from multiple sources. Despite matching on over 20 variables the groups may still have been inadequately matched. However as discussed in our paper this is unlikely to have affected the conclusions. Furthermore another paper from the NJR based on the same data and presented by a group with no known bias in favour of UKR came to the same conclusion as we did about the short term mortality rate [7].

National Registers

Due to the way National Registers are set up, and through no fault of the individuals involved, they are inherently biased against UKR. In every National Register the failure rate of UKR is about three times higher than that of TKR, so it is recommended that UKR should not be widely used [8,9]. This conclusion is not justified because of the selection, reporting and measurement bias inherent in the Registers. The selection bias occurs because UKR tend to be implanted in younger and fitter patients than TKR and younger and fitter patients have higher revision rates. Joint registers do not collect enough data to adequately match patients: The difference in the revision rate between UKR and TKR is twice as high in NJR data as it is in a matched comparison using NJR and other data sets [3]. The reporting bias occurs because Registers focus on only one outcome measure: Revision. UKR do worse than TKR with revision as the outcome measure but with virtually all other outcomes measures UKR do better [3,10,11]. From a Registers point of view if a joint replacement is revised it is considered to have failed, whereas if it is not revised it is considered to be a success. Therefore it is considered to be a success even if the patient has worse symptoms post-op than pre-op and it is considered a failure if a patient has an excellent outcome but required minor surgery such as a wash out and bearing change. Furthermore it is considered a success if the patient dies or has a major medical complication as a result of the surgery, and in some registers a TKR is considered a success even if it has ended as a fusion or amputation.

The main bias in National Registries, which cannot be avoided, is measurement bias. There is measurement bias because the decision to revise depends on many factors [12] some of which may be influenced by the type of joint replacement. For example, as UKR is easier to revise than TKR the threshold for revision is lower [13]. Therefore a UKR with a poor outcome will be more likely to be revised than a TKR with a similarly poor outcome. There is good evidence from the New Zealand Joint Registry (NZJR) that this is one of the main reasons why UKR have a higher revision rate than UKR [13]. The NZJR, as well as collecting data about revision, collects Oxford Knee Scores (OKS) at six months after the operation and categorises the score into poor, fair, good and excellent [10,14]. This demonstrates that UKR have 30% more excellent results, whereas TKR have 60% more poor results. (Findings supported by data from other studies [10]). Therefore the higher revision rate is not because UKR have more poor results, as they actually have less. The NZJR also compares the six month OKS with the subsequent revision rate. Whatever the outcome score, the revision rate

of UKR is about five times higher than that of TKR [10,13]. Therefore factors independent of outcome score, such as the threshold for revision, increase the revision rate fivefold. The most striking difference in revision rate occurs in patients who have a worse score postoperatively than preoperatively (largely those with a postoperative OKS less than 20). These patients have a 10% chance of being revised if they have had a TKR and a 60% chance of being revised if they have had a UKR. Most surgeons would agree that it is an advantage of UKR over a TKR that it is relatively straightforward to revise if there is a problem. The consequence of it being easier to revise is that the threshold for revision is lower. As a result the revision rate is higher even though the chance of getting a poor result is lower. Therefore a comparison of UKR with TKR based purely on revision rates is misleading as it suggests that UKR have worse outcomes than TKR when the data suggest that they have better outcomes.

The NJR has recently stated that comparisons between UKR and TKR should not be based on Revision rate alone. Thus the NJR is acknowledging that it is acceptable for the revision rate for UKR to be higher than TKR provided that there are other advantages of UKR. It is therefore biased that the identification of surgeon outliers is based solely on revision rate and it is assumed the revision rate of UKR and TKR should be the same. The primary method used to identify an outlier is the funnel plot of all knee replacements based on the combined revision rate of UKR and TKR. As the revision rate of UKR is higher than TKR surgeons who do UKR are likely appear worse on the funnel plot than those that do not. Furthermore surgeons who do UKR are more likely to be outliers and outliers are more likely to be surgeons who do UKR. We are aware of a number of surgeons who are outliers for all replacements combined, yet are average for both UKR and TKR when these are analysed separately. These surgeons are being put under pressure to stop being outliers, and the only way they can rapidly do this is to stop doing UKR. There is a simple solution to this bias against UKR: When the analysis is undertaken for the combined funnel plot, adjustments should be made based on the average revision rate of UKR and TKR [15]. Following these adjustments surgeons that have average performance for both UKR and TKR separately would have average performance on the combined funnel plot and would not be outliers. The outlier system would then identify surgeons with bad results, whether they use UKR or TKR. It would encourage these surgeons to improve their results by focusing on their indications, techniques and implant design rather than encouraging surgeons to stop doing UKR.

Meta-analysis

A meta-analysis of the results of the Oxford UKR published in 2011 illustrates other biases and misunderstanding associated UKR [16]. In this meta-analysis, compared to the designer surgeons, the revision rate was four times higher in registry data and three times higher in independent series. The authors state that "the cause of this divergence can only be an issue for speculation". However it is clear from the lead author's presentations that he considers that the designer's data is not only biased but also unbelievable. The designer's data in the NJR is also substantially better than average demonstrating that the designer's publications are not fraudulent and probably not biased. We find it surprising that authors did not consider the explanation of their findings to be obvious. The results of knee replacement

depend not only on the implant but also on the indications and surgical technique. It is for this reason that there are such large discrepancies between the results of the same implant in different registries (For example the Repicci UKR was the best performing UKR in Australia [17] and the worst in Sweden [18] in 2005). Clearly the designer surgeons are likely to use the appropriate indications and surgical techniques so their results are likely to be indicative of the best results that can be achieved by a particular implant. In the NJR surgeons do on average 5 UKR per year so in this Register, and presumably in other registers, poor results would be expected. The authors [16] also state that other surgeons "cannot expect to reproduce (the) excellent results" achieved by the designers surgeon. This is incorrect. For example Keys' results, published in 2004 [19], were actually better than those of the designers, although, for some reason, these results were not included in the meta-analysis. Furthermore, since the meta-analysis, we have become aware of nine series that have been presented or published, which quote the 10 year survival of the Phase 3 Oxford UKR [20]. The survival from the designer surgeons [21] and the independent surgeons is similar. It would therefore seem that if surgeons adhere to the recommended indications and techniques then they can expect to reproduce the results of the designer surgeons. The indications for the Oxford UKR are clearly defined and are satisfied in about 50% of the cases requiring knee replacement [22]. In the nine series the surgeons tended to use the Oxford UKR in 20% to 60% of their knee replacements. Analysis of NJR data has shown that the best results are achieved with the Oxford UKR when surgeons use it in between 20% and 60% of their knee replacements [23]. NJR data shows that most surgeons who use the Oxford UKR do it in less than 10% of their knee replacements, which suggests that most surgeons are not using the recommended indications. The solution seems simple: If surgeons want to use the Oxford UKR they should adhere to the recommended indications. They would then be doing UKR on all appropriate patients and would be doing enough to have the necessary experience to get good results.

Conclusion

In conclusion comparisons between UKR and TKR are open to bias. This is particularly marked in comparisons made by National Registers based on revision rate. In these circumstances the selection, reporting and measurement bias are so great that the conclusions are misleading. As a result it is not appropriate to use these analyses to compare UKR and TKR and furthermore when identifying surgeon outliers it is not appropriate to assume the revision rate of UKR and TKR should be the same. For a fair comparison of UKR and TKR data from RCT or well matched prospective observational cohort studies is required and multiple outcome measures should be used, including adverse events, patient reported outcome measures (PROMs) and cost-effectiveness. Very little data of this type exists and the data that does exist is in favour of UKR [3,10,11,24,25]. This therefore supports the increased use of UKR. There is also good evidence to suggest that if surgeons want to do UKR they should use the recommended indications and techniques.

References

1. The swings and roundabouts of unicompartmental arthroplasty. *Bone and Joint 360 2015;4-1:17.*

2. *Bone Joint J 2015;97B-4.*

3. Liddle AD, Judge A, Pandit H, Murray DW. Adverse outcomes after total and unicompartmental knee replacement in 101,330 matched patients: a study of data from the National Joint Registry for England and Wales. *Lancet 2014;384-9952:1437-45.*

4. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, Initiative S. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med 2007;147-8:573-7.*

5. Beard D, Price A, Cook J, Fitzpatrick R, Carr A, Campbell M, Doll H, Campbell H, Arden N, Cooper C, Davies L, Murray D. Total or Partial Knee Arthroplasty Trial - TOPKAT: study protocol for a randomised controlled trial. *Trials 2013;14:292.*

6. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res 2011;46-3:399-424.*

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

8. Badawy M, Espehaug B, Indrekvam K, Havelin LI, Furnes O. Higher revision risk for unicompartmental knee arthroplasty in low-volume hospitals. *Acta Orthop 2014;85-4:342-7.*

9. Briggs T. A national review of adult elective orthopaedic services in England. Getting It Right First Time. London, 2015.

10. Liddle AD, Pandit H, Judge A, Murray DW. Patient-reported outcomes after total and unicompartmental knee arthroplasty: a study of 14,076 matched patients from the National Joint Registry for England and Wales. *Bone Joint J 2015;97-B-6:793-801.*

11. Peersman G, Jak W, Vandenlangenbergh T, Jans C, Cartier P, Fennema P. Cost-effectiveness of unicondylar versus total knee arthroplasty: a Markov model analysis. *Knee 2014;21 Suppl 1:S37-42.*

12. Wylde V, Blom AW. The failure of survivorship. *J Bone Joint Surg Br 2011;93-5:569-70.*

13. Goodfellow JW, O'Connor JJ, Murray DW. A critique of revision rate as an outcome measure: re-interpretation of knee joint registry data. *J Bone Joint Surg Br 2010;92-12:1628-31.*

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

15. Murray DW, Liddle AD, Dodd CA, Pandit H. Unicompartmental knee arthroplasty: is the glass half full or half empty? *Bone Joint J 2015;97-B-10 Suppl A:3-8.*

16. Labek G, Sekyra K, Pawelka W, Janda W, Stockl B. Outcome and reproducibility of data concerning the Oxford unicompartmental knee arthroplasty: a structured literature review including arthroplasty registry data. *Acta Orthop 2011;82-2:131-5.*

17. Australian Orthopaedic Association. National joint replacement registry annual report. Adelaide, 2005.

18. Lidgren L, Knutson K, Robertsson O. Swedish Knee Arthrolasty Register Annual Report. Lund, 2005.

19. Keys GW, Ul-Abiddin Z, Toh EM. Analysis of first forty Oxford medial unicompartmental knee replacement from a small district hospital in UK. *Knee 2004;11-5:375-7.*
20. Murray DW, Marks BE, Kontochristos L, Dodd CAF, Pandit HG. The Oxford unicompartmental knee replacement: long term results. *Chin J Joint Surg 2013;7:540-44.*
21. Pandit H, Jenkins C, Gill HS, Barker K, Dodd CA, Murray DW. Minimally invasive Oxford phase 3 unicompartmental knee replacement: results of 1000 cases. *J Bone Joint Surg Br 2011;93-2:198-204.*
22. Willis-Owen CA, Brust K, Alsop H, Miraldo M, Cobb JP. Unicondylar knee arthroplasty in the UK National Health Service: an analysis of candidacy, outcome and cost efficacy. *Knee 2009;16-6:473-8.*
23. Liddle AD, Pandit H, Judge A, Murray DW. Optimal usage of unicompartmental knee arthroplasty: a study of 41,986 cases from the National Joint Registry for England and Wales. *Bone Joint J 2015;97-B-11:1506-11.*
24. Beard D, Price A, Davies L, Cook J, Maclennan G, Campbell M, Carr A, Fitzpatrick R, Campbell H, Arden N, Doll H, Murray DW, Group TS. A multicentre randomised study comparing total or partial knee replacement - one year results of the TOPKAT Trial. *Annual meeting of the British Association for Surgery of the Knee (BASK).* Liverpool, 2016.
25. Newman J, Pydisetty RV, Ackroyd C. Unicompartmental or total knee replacement: the 15-year results of a prospective randomised controlled trial. *J Bone Joint Surg Br 2009;91-1:52-7.*