extent of OA in the other knee. The extent of OA in the contralateral knee should have been documented at baseline and if not balanced in the groups, the author should have used adjusted analysis to regress its effect on the outcome to bring out the true difference in the groups.

6. As brought out earlier the study also has an error of multiple testing; the allowable margin of error should have been adjusted using tools like Bonferroni correction to avoid erroneous reporting of result [7].
7. In this study, 3 different designs of implants were used, and we know that the surgical technique of using all these implants differs. In arthroplasty practice, surgeons have their preferred techniques of performing the surgery and also prefer using one design over the other. It would be good to elaborate on how the authors ensured surgical equipoise in their study. Or how did they overcome the need for surgical equipoise [9].

Vikas Kulshrestha, MS (Orth)
Department of Orthopaedics and Joint Replacement Surgery
Command Hospital Air Force
Bengaluru, India

Munish Sood, DNB (Orth)*
Department of Orthopaedics
Command Hospital
Chandimandir, India

Santhosh Kumar, DNB (Orth)
Department of Orthopaedics and Joint Replacement Surgery
Command Hospital Air Force
Bengaluru, India

*Reprint requests: Munish Sood, DNB (Orth), Department of Orthopaedics, Command Hospital (WC), Chandimandir 134109, India.

https://doi.org/10.1016/j.arth.2020.07.042

References

[1] Dowsey MM, Gould DJ, Spelman T, Pandy MG, Choong PF. A randomized controlled trial comparing a medial stabilized total knee prosthesis to a cruciate retaining and posterior stabilized design: a report of the clinical and functional outcomes following total knee replacement. J Arthroplasty 2020;35:1583–1590.e2. https://doi.org/10.1016/j.arth.2020.01.085.
[2] Cram P, Lu X, Kates SL, Singh JA, Li Y, Wolf BR. Total knee arthroplasty volume, utilization, and outcomes among Medicare beneficiaries, 1991-2010. JAMA 2012;308:1227–36. https://doi.org/10.1001/2012.jama.11153.
[3] Song SJ, Park CH, Bae DK. What to know for selecting cruciate-retaining or posterior-stabilized total knee arthroplasty. Clin Orthop Surg 2019;11:142–50. https://doi.org/10.4055/cios.2019.11.2.142-50.
[4] Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. JAMA 2019;321:1610–20. https://doi.org/10.1001/jama.2019.3087.
[5] Williams DP, Blakey CM, Hadfield SG, Murray DW, Price AJ, Field RE. Long-term trends in the Oxford knee score following total knee replacement. Bone Joint J 2013;95-B:45–51. https://doi.org/10.1302/0301-620X.95B1.28573.
[6] Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop 2016;87(Suppl 1):9–23. https://doi.org/10.1080/17453674.2016.1181816.
[7] Sedgwick P. Multiple hypothesis testing and Bonferroni's correction. BMJ 2014;349:g6284. https://doi.org/10.1136/bmj.g6284.
[8] Gunther KP, Sturmer T, Sauerland S, Zeissig I, Sun Y, Kessler S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998;57:717–23. https://doi.org/10.1136/ard.57.12.717.
[9] Kulshrestha V, Sood M, Kanade S, Kumar S, Datta B, Mittal G. Early outcomes of medial pivot total knee arthroplasty compared to posterior-stabilized design: a randomized controlled trial. Clin Orthop Surg 2020;12:178–86. https://doi.org/10.4055/cios19141.

Response to Letter to the Editor on "A Randomized Controlled Trial Comparing a Medial Stabilized Total Knee Prosthesis to a Cruciate Retaining and Posterior Stabilized Design: A Report of the Clinical and Functional Outcomes Following Total Knee Replacement"

In Reply:

Authors: We thank Sood et al for their interest in our paper and recognize that as one of only a few properly constructed randomized controlled trials (RCTs) that look at prosthetic design and its relationship to outcomes that it would and should draw appropriate attention. Much of the detail which Sood et al seek is already embedded within the body of our original manuscript. However, it may be opportune to guide Sood et al's better understanding as raised by their questions. Moreover, we had preceded the conduct of this trial by registering the protocol with the relevant registry of clinical trials and also published the protocol as part of accepted research transparency and this information is also captured in our original manuscript.

1. Multi-arm clinical trials are performed for finding treatment for a disease or condition with no proven effective treatment as yet, wherein we want to expedite the process of clinical research and decrease the need for multiple trials. Use of multi-arm trial is to decrease the number of subjects required and to arrive at a clinical solution at the earliest. Total knee replacement using posterior stabilized and cruciate retaining implants is a well-established surgical treatment of osteoarthritis (OA) of the knee [1]. The efficacy of both these prosthetic designs is scientifically proven and is equivalent [2]. The new design of implant used (medial pivot design) should have been compared to anyone of the design in one-to-one comparison using a conventional 2-armed trial which would have improved comparability and avoided the problems of multiple testing of hypothesis and inducing errors of inter-comparability [3].

Authors: Our preference for a multi-arm trial was not to minimize the required sample size needed or the time required to implement a possible clinical solution (although these can of course be additional benefits of such a design), but rather to enable a comparison of the primary medial stabilized design (for which supporting evidence is inconclusive as detailed in the "Introduction" section of our paper) against 2 alternate designs (posterior stabilized and cruciate retaining) for which the evidence is less equivocal. The 3-way comparison also better reflects the clinical reality where a surgeon may choose between the 3 competing designs, rather than a pair of them.

We disagree that our choice of design introduced comparability issues into the analysis. The comparison of baseline characteristics by design group (Manuscript Table 1) clearly demonstrates that the

DOI of original article: https://doi.org/10.1016/j.arth.2020.07.040.

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2020.07.042.

randomization ensured good balance of baseline factors across groups. All variables were associated with acceptable balance (ie, $P > .05$) with the exception of age, body mass index, and comorbidity score. These 3 factors were then used as adjusting covariates in the final analysis to ensure that the observed differences in outcome between the 3 designs were independent of these factors. This strategy to ensure comparability was established a priori and conducted as per the study protocol. This is described in the "Statistical Analysis" sub-section of the "Methods" section.

With regards to adjusting for multiple testing, the study was designed as a 3-group, *parallel* RCT (as stated in the "Trial Design" sub-section of the "Methods" section). Being a parallel design, the primary analysis compares outcomes in patients randomized to the medial stabilized design against either the fixed-bearing cruciate retaining design or, separately (or in parallel) the second posterior stabilized design comparator. Thus, correction for multiple testing was not indicated. Please refer our response to item 6 below for further detail.

2. Patient-reported outcomes after total knee arthroplasty (Oxford Knee Score) show optimum recovery only by 1 year after surgery which peaks at 2 years [4]. If we look at the literature and recommendations of the working group [5], it is advised to report on early outcomes of total knee arthroplasty only at 1 year of follow-up. Till 1 year, the time has a major effect on recovery, and it would confound with any outcomes assessed to look at the effect of implant design in comparison to others.

Authors: We question the choice of references used by Sood et al to support their statements.

Williams DP, Blakey CM, Hadfield SG, Murray DW, Price AJ, Field RE. Long-term trends in the Oxford knee score following total knee replacement. Bone Joint J 2013;95-B:45–51.

As the title suggests, the primary purpose of the study by Williams et al was to report on the long-term trends in the Oxford Knee Score (OKS) over a 10-year period. Data for this study were derived from a registry that only records annual OKS, therefore no 6-month OKS data were included in their analyses. Although variations in individual item responses were noted, there was literally no difference in mean OKS scores at 1-year (34.3), 2-year (34.4), and 3-year (34.1) time points. In addition, Williams et al acknowledge that one of the largest reported series relating to the OKS (n = 20,885) report a mean 6-month OKS of 33.6, closely aligning with their 1-year score.

Rolfson O, Bohm E, Franklin P, et al. Patient-reported outcome measures in arthroplasty registries Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop 2016;87(Suppl. 1):9–23. https://doi.org/10.1080/17453674.2016.1181816.

As the title suggests, the primary purpose of the study by Rolfson et al was to develop recommendations for reporting of outcome measures in arthroplasty registries as opposed to clinical trial reporting, whereby data collection in registries is generally limited to a single post-operative time point.

With regard to our study, we highlight a systematic review of the optimal timing to assess the Oxford Hip and Knee Scores after total hip and knee joint replacement [6] which concluded that the published evidence for knee replacement surgery presents a less consistent message than for hip replacement. "While it is clear that most of the improvement in the OKS occurs in the first six months after knee replacement surgery there is inconsistent evidence (from eight out of 13 patient groups) of a minimally important difference occurring between the benefits measured at six and at 12 months."

Notwithstanding the inconsistencies reported in the literature, we took a conservative approach by including a secondary analysis of all study outcomes at 12 month as detailed in Table 2 of the manuscript and elsewhere in the "Results" section. Our side-by-side comparison of 6-month and 12-month OKS indicates that there was no improvement in scores between the 6-month and 12-month time points in any of the 3-study arm (Manuscript Table 2 and Supplementary Table 2). We disagree that time may confound outcome comparisons at either 6 or 12 months, as all subjects across all 3 design groups were assessed as per protocol and there was no significant attrition or loss-to-follow up between the groups.

3. The author has not stated the precise hypothesis for the trial and comparisons being planned—being a 3-armed trial whether all 3 groups are being compared together or one to the other in tandem. Three-arm trial hypotheses should be defined precisely and comparison predefined. Planned primary, secondary, and exploratory comparisons should be stated at the beginning [3].

Authors: The null hypothesis is clearly stated in the final sentence of the fourth paragraph of the "Introduction" section and the intended comparisons described in detail in both the "Introduction" and the "Statistical Analysis" sub-section of the "Methods."

4. As the study was an equivalence trial it needed to use 2-sided alpha error for their sample size calculations which would result in need of larger sample size. With an allocation of 1:1:1, in a 3-armed trial there to maintain power it would need a larger sample size. In addition, a multi-arm trial which is testing 3 groups so at least 2 outcomes of equal importance being tested while comparing the 3 groups the allowable error rate needs to be kept at <0.025 following rule of Bonferroni correction [7]. Therefore, a multi-arm equivalence trial allowing 2-sided alpha error of $P < .025$ to have at least a power of 90% as planned by the author in their protocol would require significantly more than 30 patients in each arm knowing that there would be dropouts. It is requested that the authors may share their method of sample size calculation.

Authors: The study was a superiority trial, not an equivalence trial. As stated in the first paragraph of the "Statistical Analysis" sub-section, the aim of the primary analysis (and the basis for the sample size calculation) was to detect a minimally clinically important difference in the primary OKS outcome, not to establish equivalence.

Regarding the observation around Bonferroni adjustment for multiple comparisons, please refer to our response to item 6 below.

Our sample size calculation, including the method used (baseline-adjusted analysis of covariance) and input estimates for minimum effect size (minimally clinically important difference) and outcome variance (between-patient standard deviation), is described in detail in the first paragraph of the "Statistical Analysis" sub-section.

5. OA of the knee has bilateral affliction in most cases [8]. Although randomization ensures comparability, however, we need to document at baseline and control for all clinically important predictors of outcome. In this study, the outcome assessment tools were patient performance (time up and go test and 6-meter walk test) and patient-reported outcome measures like Knee Society Score which would definitely be affected by the

extent of OA in the other knee. The extent of OA in the contralateral knee should have been documented at baseline and if not balanced in the groups, the author should have used adjusted analysis to regress its effect on the outcome to bring out the true difference in the groups.

Authors: As Manuscript Table 1 demonstrates, the randomization was successful in balancing the vast majority of documented clinically important predictors. It is thus reasonable to assume in the context of an appropriately powered RCT that any confounding of outcome attributable to the extent of osteoarthritis (OA) in the contralateral knee would be equally balanced between the 3 design groups and this in fact was the case. Our institution holds an in-depth arthroplasty registry (SMART Registry) whereby patient, clinical, and patient-reported outcome measures are recorded in all patients who undergo elective hip and knee replacement [9]. Contralateral knee OA severity is routinely recorded. In the present study, the presence of OA in the contralateral knee (Kellgren Lawrence $\geq$ 2) by group was 19 (73%) in those who received a posterior stabilized design, 21 (78%) for those who received a cruciate retaining design, and 25 (86%) in the medial pivot ($P = .474$); therefore, contralateral knee OA would not have been included in our adjusted model.

6. As brought out earlier the study also has an error of multiple testing; the allowable margin of error should have been adjusted using tools like Bonferroni correction to avoid erroneous reporting of result [7].

Authors: Per the reference provided by Sood et al, Bonferroni correction should be applied to *multiple hypothesis testing* when there are *multiple pre-stated primary outcomes*. In that regard we proposed a single hypothesis with one primary outcome, assessed using a single measure which was self-reported pain and function using the OKS. As we hypothesized a priori that both primary endpoints (pain and function) would be positive, success is declared only if both primary endpoints are statistically significant in favor of the experimental treatment, therefore no multiplicity issue exists. As such, we did not lower the type 1 error threshold to adjust for co-primary endpoints. This is in keeping with Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials recommendations for analyzing multiple endpoints in trials of pain treatment [10]. All other endpoints were listed as secondary outcomes in the manuscript and the implications of assessing multiple secondary outcome measures were noted in the discussion.

7. In this study, 3 different designs of implants were used, and we know that the surgical technique of using all these implants differs. In arthroplasty practice, surgeons have their preferred techniques of performing the surgery and also prefer using one design over the other. It would be good to elaborate on how the authors ensured surgical equipoise in their study. Or how did they overcome the need for surgical equipoise [11].

Authors: Allocations to treatment were based on randomization, not surgeon preference. Furthermore, there was no difference in the proportion of the 3 designs being used by any of the 4 surgeons in the study (refer Manuscript Table 1). Thus, any potential bias attributable to surgeon experience or preference was negated by the randomization itself. Prior to commencement of the study, all study surgeons underwent cadaver training with the instrumentation system and also conducted a minimum number of procedures to familiarize themselves with the workflow. This has been highlighted in our “Methodology” section.

Acknowledgments

As declared in our original manuscript, financial support for the conduct of the study was provided by Medacta International to the study institution. The funder had no role in the study’s design, conduct, or reporting.

Michelle M. Dowsey, BHealSci, MEpi, PhD
Department of Surgery
The University of Melbourne
St Vincent’s Hospital
Melbourne, Victoria, Australia

Department of Orthopaedics
St Vincent’s Hospital
Melbourne, Victoria, Australia

Tim D. Spelman, GradCert [Stats], PhD
Department of Surgery
The University of Melbourne
St Vincent’s Hospital
Melbourne, Victoria, Australia

Peter F. Choong, MD, FRACS, FAOrthA*
Department of Surgery
The University of Melbourne
St Vincent’s Hospital
Melbourne, Victoria, Australia

Department of Orthopaedics
St Vincent’s Hospital
Melbourne, Victoria, Australia

*Reprint requests: Peter F. Choong, MD, FRACS, FAOrthA, Department of Surgery, St. Vincent’s Hospital, Level 2, Clinical Sciences Building, 29 Regent Street, Fitzroy, 3065 Victoria, Australia.

https://doi.org/10.1016/j.arth.2020.07.042

References

[1] Cram P, Lu X, Kates SL, Singh JA, Li Y, Wolf BR. Total knee arthroplasty volume, utilization, and outcomes among Medicare beneficiaries, 1991-2010. JAMA 2012;308:1227–36.
[2] Song SJ, Park CH, Bae DK. What to know for selecting cruciate-retaining or posterior-stabilized total knee arthroplasty. Clin Orthop Surg 2019;11:142–50.
[3] Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the consort 2010 statement. JAMA 2019;321:1610–20.
[4] Williams DP, Blakey CM, Hadfield SG, Murray DW, Price AJ, Field RE. Long-term trends in the Oxford knee score following total knee replacement. Bone Joint J 2013;95-B:45–51.
[5] Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries report of the patient-reported outcome measures Working Group of the International Society of arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop 2016;87(Suppl 1):9–23.
[6] Browne JP, Bastaki H, Dawson J. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. Health Qual Life Outcomes 2013;11:128.
[7] Sedgwick P. Multiple hypothesis testing and Bonferroni's correction. BMJ 2014;349:g6284.
[8] Gunther KP, Sturmer T, Sauerland S, Zeissig I, Sun Y, Kessler S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998;57:717–23.

[9] Dowsey MM, Smith AJ, Choong PFM. Latent Class Growth Analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients. Osteoarthritis Cartil 2015;23:2141–9.
[10] Turk DC, Dworkin RH, McDermott MP, Bellamy N, Burke LB, Chandler JM, et al. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials. Pain 2008;139:485–93.
[11] Kulshrestha V, Sood M, Kanade S, Kumar S, Datta B, Mittal G. Early outcomes of medial pivot total knee arthroplasty compared to posterior-stabilized design: a randomized controlled trial. Clin Orthop Surg 2020;12:178–86.

Letter to the Editor on "Midterm Results of a Contemporary, Porous-Coated Acetabular System in Patients Undergoing Primary Total Hip Replacement for Degenerative Hip Disease: A Prospective, Multicenter Study"

Dear Editors,

We commend the scientific work by Wilson et al [1] titled "*Midterm Results of a Contemporary, Porous-Coated Acetabular System in Patients Undergoing Primary Total Hip Replacement for Degenerative Hip Disease: A Prospective, Multicenter Study.*" The authors have truly added value to the literature in the field of hip arthroplasty through their evaluation of the safety of acetabular prostheses.

In the study, the acetabular prostheses used comprised of both ceramic and polyethylene liners. The authors attempted to provide a comparison to pre-existing literature by performing a rudimentary review of studies reporting revision rates of hip arthroplasties performed with insertions of ceramic- or polyethylene-lined acetabular cups. However, in their analysis, the authors used a pooled simple average to calculate the average revision rate of acetabular prostheses (table 6). In addition, revision rates were evaluated as an amalgam of both ceramic and polyethylene liner groups, with the authors citing no significant differences between the revision rates for implants made from both materials. We believe this to not be the best method of representing the average revision rates of the included studies.

Evidence from randomized controlled trials (RCTs) has found marginal differences in the rate of complications and revisions be-

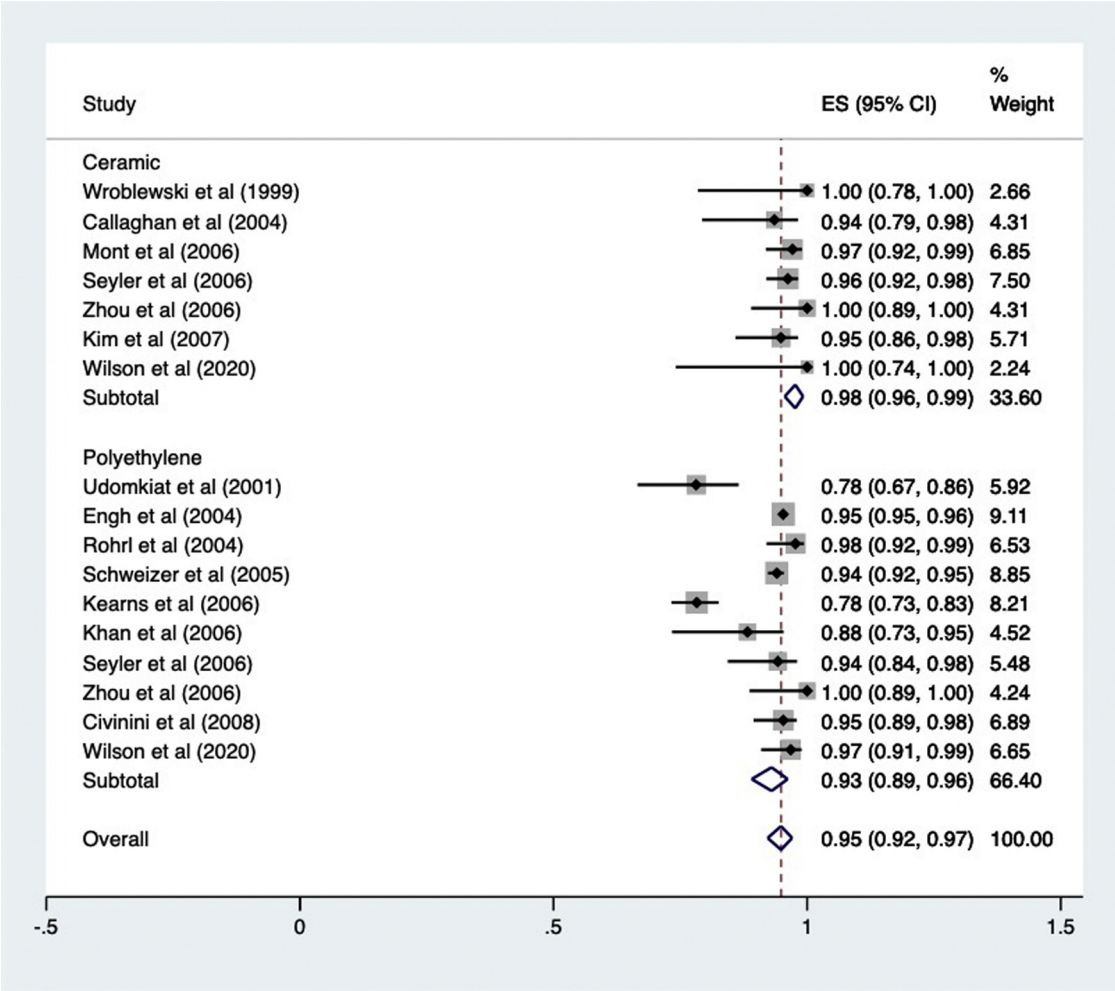

Fig. 1. Meta analysis of survival rates in ceramic and polyethylene implants.

DOI of original article: https://doi.org/10.1016/j.arth.2020.02.068.

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2020.07.079.