
January 10, 2022

Methods Have Multiple Sources of Potential Bias Need to Be Addressed

Peter Schilling

Assistant Professor of Orthopaedic Surgery

Dartmouth Hitchcock Medical Center / Geisel School of Medicine at Dartmouth

Other Contributors:

Wayne Moschetti

Assistant Professor of Orthopaedic Surgery

Dartmouth Hitchcock Medical Center / Geisel School of Medicine at Dartmouth

We are concerned that the choice of analysis in this study biased the results in favor of finding a mortality benefit to technology-assisted instrumentation (when one may not actually exist).

Our concerns are three-fold:

1. Our foremost concern is that the study's model did not control for relevant hospital-level effects, thus putting the findings at risk for omitted variable bias and confounding.

Omitted variable bias occurs when a statistical model leaves out one or more relevant variables, causing the model to misattribute the effect of the missing variables to those that were included. In this study, the model does not include variables to account for hospital-level effects (hospital characteristics that are known to influence outcomes like “quality of care,” case volume, geographic setting, as well as other unobservable factors) (1). The omission of these important hospital-level effects is very likely to bias or confound the relationship between the variables of interest (e.g., technology-assisted instrumentation and mortality) because the hospital characteristics associated with quicker technology adoption are also likely to be associated with superior outcomes (e.g., perioperative mortality).

The broader literature supports the validity of our concern: First, it is well established that hospitals that are quicker to adopt technology are fundamentally different from those that lag in adoption. Supporting evidence comes from Skinner et al., a study that found that hospitals that are quicker to adopt technology tend to be major teaching hospitals with higher patient volumes located in regions with higher average income (2). Second, a plethora of studies have demonstrated significant unwanted variation in the quality

of care delivered by hospitals, including variation in joint arthroplasty outcomes.(3) (Can't find Australia specific data!) Putting these observations together, it is possible, if not likely, that a relatively small number of high-volume hospitals have been performing the majority of the technology assisted surgeries: If these high volume hospitals have superior outcomes for reasons unrelated to technology assisted instrumentation (note: well established volume-outcomes relationships [4, 5]) the study may be misattributing these “hospital-level effects” to the use of the technology, when in reality, the superior outcomes are related to a host of other hospital-level factors like procedure volume, key processes of care, etc. (6) If the above-described scenario is true, the relationship between technology assisted instrumentation and mortality would be biased toward finding a positive relationship, even if one does not truly exist.

Admittedly, there is no perfect solution for this problem in the study at hand. Ideally, one would control for hospital-level effects by adding indicator variables for hospitals within the model (in other words, one would control for hospital-level fixed effects). The problem with this approach is that mortality after joint replacement is a rare outcome, and the model may not be able to accommodate the additional complexity: Add too many variables and one ends up with other problems like overfitting. Nonetheless, modeling a rare event does not obviate the need to consider these important hospital-level effects and risk of bias. The authors should address this potential confounder at length. At a minimum, the authors should provide a rudimentary analysis of how adopting and non-adopting hospitals compare across basic characteristics such as teaching status, procedure volume, rural/urban setting, average income of the region, etc. Moreover, the authors should walk back the claim that the study demonstrates a causal relationship, and instead frame the finding as interesting and worthy of further study.

2. The study does not account for the nested-nature of its data — an omission known to inflate the number of statistically significant associations in a model (increased risk of type-I error).

Data collected in registry initiatives such as this are inherently hierarchical: Patients are said to be nested or clustered within the hospitals that treat them. This multilevel data structure has implications for the most appropriate statistical modeling. Models should account for hospital clustering because patients within a hospital do not represent truly independent observations (7). Failure to account for the multilevel nature of hierarchical data structure has been shown to dramatically increase the risk of type-I errors – in other words, researchers draw the wrong conclusions because they will more frequently obtain statistically significant results (even when the treatment effect is not present) (8). This is of particular concern when treatments and interventions are at the hospital-level, as is the case when a hospital buys technology assisted instrumentation for its surgeons to use.

We suspect that the study, as designed, is at high risk for type-I error because it does not take the non-independence of patient observations into account. Ideally, the study should use multilevel or hierarchical models, a statistical technique that accounts for the nested nature of the data (9, 10). These more advanced modeling techniques have been used across the social sciences and have been increasingly used in health services research, including orthopaedics (11, 12). Admittedly, modeling rare events (like mortality after joint replacement) with hierarchical models has its own challenges. Either way, the increased risk of type-I error in the present analysis must be addressed within the manuscript since this is yet a second reason to doubt the validity of the study's findings.

3. The study attempts to disentangle secular trends in perioperative mortality from those imparted by the technology but fails to provide sufficient information and discussion about this critical component of the analysis.

The authors conducted an observational study to evaluate the introduction of a technology on outcomes. An important limitation in using observational studies is the need to control for background changes in outcomes that occur over time (e.g., secular trends affecting outcomes that are unrelated to the intervention of interest) (13). This is critically important for this study because, as the authors are aware, there is an ongoing, worldwide temporal decline in mortality following total knee arthroplasty (14). This trend started at least 20 years ago and closely overlaps the time frame of the study. Prior studies have speculated that this long-term secular trend is related to improved patient selection and perioperative care. As such, any observational study of a joint arthroplasty intervention must account for this temporal trend, or risk misattributing the improvements in mortality to the intervention as opposed to other factors driving the temporal trend. The authors attempt to control for this secular trend by incorporating "procedure year" into the model but fail to elaborate beyond saying that mortality declined over the study period. This amount of information is insufficient to judge the adequacy of their approach. At a minimum, the authors need to report the observed time trend in mortality between 2003 and 2019 since this is the foundation for understanding how best to disentangle the impact of the technology from other drivers of perioperative mortality rates over time. The issue also deserves comment in the discussion section given just how important it is for the validity of the study's conclusions.

4. There appears to be an error in Table II.

Row 2, Column 2 indicates that for conventional instrumentation, deaths at 30 days was 0.57%; however, based on the counts provided, $32 / 20,407 = 0.16\%$ (not 0.57%). Could the authors clarify this?

Disclaimer: e-Letters represent the opinions of the individual authors and are not copy-edited or verified

by JBJS.

References

- [1] Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743–1748. doi:10.1001/jama.260.12.1743
- [2] Skinner J, Staiger D. Technology Diffusion and Productivity Growth in Health Care. The Review of Economics and Statistics. 2015;97(5):951–964. doi:10.1162/REST_a_00535
- [3] Bozic KJ, Grosso LM, Lin Z, Parzynski CS, Suter LG, Krumholz HM, Lieberman JR, Berry DJ, Bucholz R, Han L, et al. Variation in Hospital-Level Risk-Standardized Complication Rates Following Elective Primary Total Hip and Knee Arthroplasty. JBJS. 2014;96(8):640–647. doi:10.2106/JBJS.L.01639
- [4] Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Association Between Hospital and Surgeon Procedure Volume and Outcomes of Total Hip Replacement in the United States Medicare Population*. JBJS. 2001;83(11):1622–1629.
- [5] Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E. Association Between Hospital and Surgeon Procedure Volume and the Outcomes of Total Knee Replacement. JBJS. 2004;86(9):1909–1916.
- [6] Bozic KJ, Maselli J, Pekow PS, Lindenauer PK, Vail TP, Auerbach AD. The Influence of Procedure Volumes and Standardization of Care on Quality and Efficiency in Total Joint Replacement Surgery. JBJS. 2010;92(16):2643–2652. doi:10.2106/JBJS.I.01477
- [7] Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, Howard G, Peterson ED, Rathore SS, Smith SC, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation. 2006;113(3):456–462. doi:10.1161/CIRCULATIONAHA.105.170769
- [8] Austin PC, Tu JV, Alter DA. Comparing hierarchical modeling with traditional logistic regression analysis among patients hospitalized with acute myocardial infarction: should we be analyzing cardiovascular outcomes data differently? American Heart Journal. 2003;145(1):27–35. doi:10.1067/mhj.2003.23
- [9] Diez-Roux AV. Multilevel analysis in public health research. Annual Review of Public Health. 2000;21:171–192. doi:10.1146/annurev.publhealth.21.1.171
- [10] Duncan C, Jones K, Moon G. Context, composition and heterogeneity: using multilevel models in health research. Social Science & Medicine (1982). 1998;46(1):97–117. doi:10.1016/s0277-9536(97)00148-2
- [11] Schilling PL, He J, Chen S, Placzek H, Bini SA. Risk-Adjusted Cost Performance for 90-Day Total Hip Arthroplasty Episodes: Comparing US Hospitals Nationwide Before CJR. The Journal of Arthroplasty. 2020;35(12):3452–3463. doi:10.1016/j.arth.2020.06.042
- [12] Schilling PL, He J, Chen S, Placzek H, Bini S. Risk-Adjusted Cost Performance for 90-Day Total Knee Arthroplasty Episodes: Data and Methods for Comparing U.S. Hospitals Nationwide. The Journal of Bone and Joint Surgery. American Volume. 2020;102(11):971–982. doi:10.2106/JBJS.19.01017
- [13] Pape UJ, Millett C, Lee JT, Car J, Majeed A. Disentangling secular trends and policy impacts in health studies: use of interrupted time series analysis. Journal of the Royal Society of Medicine. 2013;106(4):124–129. doi:10.1258/jrsm.2012.110319
- [14] Berstock JR, Beswick AD, López-López JA, Whitehouse MR, Blom AW. Mortality After Total Knee Arthroplasty: A Systematic Review of Incidence, Temporal Trends, and Risk Factors. JBJS. 2018;100(12):1064–1070. doi:10.2106/JBJS.17.00249

Conflict of Interest: None Declared

Article Author Response

2 February 2022

Article Author(s) to Letter Writer(s)

We thank Peter Schilling and Wayne Moschetto for their letter regarding our manuscript and provide a response to their points below.

1. and 2. Regarding controlling for hospital level effects (either as a fixed or random effect), there were 275 hospitals contributing data to the analysis. The high number of hospitals and rarity of early mortality meant that including each hospital in the analysis as a random effect in a hierarchical model did not allow model convergence. Previously, when accounting for hospital level variation, we have included hospital sector (private versus public) as a fixed effect as this is where most of the between-hospital variation exists (regarding such things as volume, care pathways and take-up of new technology). Including hospital sector as a fixed effect was not significant and did not alter the estimate for the main predictor (technology assisted versus conventional instrumentation). Case volume per year at each hospital was shown to not be significant and did not alter the estimate for the main predictor. We do not have other hospital characteristics, such as measures of quality, available for analysis.

Regarding the need to “walk back” the claim that the association is causal, we only report the “association” found but we offer that the findings “may be explained by” fat emboli, which we justify.

3. Regarding the secular changes in mortality, year of procedure was used to control for this. We have published a more detailed analysis on the secular trends in mortality from the same registry, and this reference was provided as reference 6 in the manuscript (Harris IA et al. Declining early mortality after hip and knee arthroplasty. ANZ J Surg. 2020;90:119-22).

4. Regarding Table II, the writers are correct that there is a mistake in row 2, column 2. The “0.57%” should be “0.16%”.

Overall, we agree with the points raised by the writers. We do not think that hospital characteristics can explain the difference found and that we have accounted for secular trends. Unfortunately, due to the data available to us and the rarity of the outcome, not all these concerns can be answered, which is why we agree with the writers that further research from different populations needs to be performed.

Ian Harris, Michelle Lorimer, David Kirwan, Yi Peng

Email: ianharris@unsw.edu.au