Study Title: Comparative Effectiveness of Decision Support Strategies for Joint Replacement Surgery

NCT: NCT02729831

Date: December 21, 2017

Document Title: Statistical Analysis Plan

Mangla M, Bedair H, Chang Y, Daggett S, Dwyer M, Freiberg A, Mwangi S, Talmo C, Vo H, Sepucha K. Protocol for a randomised trial evaluating the comparative effectiveness of strategies to promote shared decision making for hip and knee osteoarthritis (DECIDE-OA study) BMJ Open 2019;**9:**e024906. doi:10.1136/bmjopen-2018-024906

This section contains the analysis plan for the study and was taken from the published study protocol. Please refer to the study protocol for additional information on the design, measures, data collection and analyses:

Mangla M, Bedair H, Chang Y, Daggett S, Dwyer M, Freiberg A, Mwangi S, Talmo C, Vo H, Sepucha K. *Protocol for a randomised trial evaluating the comparative effectiveness of strategies to promote shared decision making for hip and knee osteoarthritis (DECIDE-OA study)* BMJ Open 2019;**9:**e024906. doi:10.1136/bmjopen-2018-024906

Analysis plan:

For patient-reported outcomes (decision quality, quality of life, etc.), missing data items will be handled per established protocols for the validated surveys (e.g., missing knowledge items are considered incorrect). For item-specific analysis, our primary analyses will be conducted excluding patients with missing data. The treatment received (surgical vs non-surgical) will be assessed through chart review and confirmed via patient report (T3); therefore, is not subject to missing data. Even though we cannot test the missing at random assumption, we will first compare patients with and without missing data to gain insights. As a sensitivity analysis, we will conduct several missing imputation techniques: (1) last value carried forward (LVCF), (2) single imputation with EM algorithm and (3) multiple imputation. The LVCF approach applies to follow-up missing data, which is essentially the same as assuming no change over time. Compared with single imputation, the appealing aspect of the multiple imputation approach is incorporating the variability across imputation so that the statistical uncertainty due to missing is more properly accounted for. We will compare our findings from the primary analyses with the findings from different imputation strategies to determine whether our findings are stable across different assumptions. We will also report the uncertainty associated with the treatment effect as indicated in the SE estimates from the multiple imputation analysis.

As the first step, responders and non-responders will be compared across groups to examine non-response bias. For patient-reported outcomes, missing data will be handled per established protocols for the validated surveys. We will conduct sensitivity analyses to determine the impact of missing imputation.[3] The hypotheses will be evaluated using an intention-to-treat approach. The analysis plan for the primary outcome (hypothesis 1.1) will first calculate the rate of decision quality in each group, as the percentage of patients who meet or exceed the knowledge threshold *and* receive treatment that matches their preference. A logistic regression model with the generalized estimating equations (GEE) approach will be used to compare the rates of decision quality of the DA-A and DA-B groups and account for the clustering of patients within providers.[4] Analysis will start by testing the interaction between the two intervention factors. It is plausible that an interaction between DAs and type of surgeon report exists for this analysis. Thus, the effective sample size will be limited to 117 per group when the comparisons are stratified by the type of surgeon report. The study has 89% power to detect a difference in the percentage of patients with high decision quality of 18%, from 65% in DA-B group to 83% in DA-A group.

For hypothesis 1.2, an interaction between DAs and PPR report is unlikely so there is no need to account for clustering within the same provider, thus, we will use a two-sample t-test to compare the mean knowledge score between the two groups. With approximately 560 patients from each group, we can invoke the Central Limit Theorem and use a two-sample t-test to compare mean knowledge score between the two groups, even if the knowledge score is not normally distributed. The study will have 80% power to detect a difference as small as 3.3% in total knowledge scores assuming the SD is 20%. For hypothesis 1.3, patient's treatment preference will be assessed before the surgeon visit so again, there is no need to account for clustering. A X2 test will be used to compare the percentage of patients with clear treatment preference between the two groups. Hypothesis 2.1 will use a linear regression model with the GEE approach and hypothesis 2.2 will use logistic regression with GEE approach to account for clustering of patients within surgeons for these analyses.

The heterogeneity of the treatment effect will be explored by testing the interaction between interventions and different factors on study outcomes. These factors include (1) patient characteristics (e.g., age, gender, education level, joint (hip or knee), health literacy and severity of disease), (2) provider characteristics (gender, years since graduation, surgical volume), (3) intervention compliance (whether patients reviewed the DAs) and (4) mode of DA delivery (online or paper). Linear or logistic regression models (with the GEE approach in the case of clustering within providers) will be used to test the interaction between interventions and these factors. We will also report treatment effect in each subpopulation if there are strong evidence of interactions between interventions and these factors. Some of the hypothesis testing here might be exploratory in nature. The study will have sufficient power for testing interaction for continuous outcomes (e.g., detecting meaningful ‘differences in differences’ for knowledge scores, EQ-5D scores) but not categorical outcomes (e.g., rate of high decision quality, surgical rate).

References

1. Sepucha K, Atlas SJ, Chang Y, *et al*. Patient decision aids improve decision quality and patient experience and reduce surgical rates in routine orthopaedic care: a prospective cohort study. *J Bone Joint Surg Am* 2017;99:1253–60.
2. Bland M. *Cluster randomized trials in the medical literature. Notes for talks presented at University of York*: University of York, 2005.
3. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the *EM* Algorithm. *Journal of the Royal Statistical Society: Series B* 1977;39:1–22.
4. Hardin JW, Hilbe JM. *Generalized estimating equations*. 2nd edn. Boca Raton, FL: Chapman and Hall/CRC, 2012.