Reviewer's report

Title: Income and Patient-Reported Outcomes (PROs) after Knee Arthroplasty

Version: 3 Date: 8 November 2012

Reviewer: George Zhang

Reviewer's report:

It is an interesting topic using a relative large data set. But there are several major methodological issues.

Pre-operative variables, which are important confounding factor in the study, should be included in the modeling. The analysis and interpretation should be first to be sound valid in a setting before it can be considered whether it can be generalized to other more general settings.

It is also questionable whether the same income categories can be appropriately used in a long term period (1993-2005; till 2010 due to 5-yr follow-up?). Some sensitivity analysis would be helpful.

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

I am not very confident to say “yes”.

Other issues were listed below.

### ------------- Major Compulsory Revisions ----------------
(The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.)

1. Method / Outcomes of Interest

The authors look at three main post-operative outcomes (2 state variables, 1 change variable). However, none of these main variables of interest provide basic frequency and distribution information, e.g.,

1a. How many patients with “moderate-severe knee pain” in each of the income categories or other patient characteristics variables. A summary table would be very essential to compare the variables listed in Table 1 under different category groups of a main outcome at a follow-up time.

1b. Before running a multivariable-adjusted logistic model, an UNADJUSTED Odd Ratio and its 95% CI are generally required to examine the impact of income on the 3 main post-operative outcomes.

2. Method / Statistical Analysis

2a. Three outcomes were evaluated at 2 time points separately. Multiple
comparisons would be an issue. If a Bonferroni correction is used, then a p-value < 0.05/3 = 0.017 or other appropriate adjustments should be used.

2b. It is unclear how many and what logistic models were actually fitted and reported. The authors seemed not to include pre-operative or baseline covariates in the logistic models, which "were performed separately for each outcome at 2- and 5-years."

However, for a main outcome, say, Improvement in knee function (3 categories: much better, better, reference), was one logistic model run at 2-year data and a different model run at 5-year data separately? If only one time point data was used, why needs to “adjust for the correlation between observations on the same subject.”? – only one outcome collected from a subject at one time point.

If both 2- and 5-year two time points data were used, then the authors never mention the effective sample size - how many patients had complete survey data both at these two time points.

2c. In additional, was the “Improvement in knee function” outcome performed separately for its 2 of 3 categories at a time point? Otherwise, for such an ordinal variable with 3 categories, a traditional logistic model is not suitable (which is for a binary variable). Was an ordinal logistic model used?

2d. The logistic modeling process is confused.

For example, Age and BMI, were they treated as continuous or categorical variables?

The authors are recommended to consult a professional statistician.

Discussion
3. “Income was based on zip code median income, and not actual personal income, which may have led to misclassification bias, which would bias our results towards null hypotheses, and make these estimates conservative.”

So the income reported in Table 1 were actually not “patient-reported” and they were all zip-code based? The zip-code income was based on which year census data? Any adjustment for inflation rate?

Why could you conclude that an inaccurate / misclassification data will lead to a conservative result? Pure speculation? The logic link is weak.

### ------------- Minor Essential Revisions --------------------
(The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.)

Table 1
4a. In general, the frequency should also be reported to give a sense how missing data were distributed.

No missing for all the variables listed in Table 1?
4b. Did the table capture the variables at the operation/baseline? Or at 2-yr and 5-yr follow-up separately?
4c. What proportion of patients had both 2- and 5-yr data? Why the mean age and SD were the same? Show one more decimal point?
4d. “Age groups n (%)” - Why not immediately after “Mean Age”
Actually only the percentage was reported, while the frequency (n) was not.
4e. 45K is the supposed “median” household cut point, however, the observed proportion is just 33% or 27%, but not close to the nominal 50% level.
4f. Three factors were not listed in Table 1, but were adjusted in Tables 2 and 3 (distance, operative diagnosis, implant fixation, all not listed in Table 1).

Table 2
5a. Table title only states “at 2-years” but the content actually contains results at “5-years”.
5b. Please keep 2 decimal points for Odds Ratio and its CI.
5c. Among 8 adjusted covariates, which are also significant?
5d. Same estimated OR and 95% CI = 1.4 (1.1, 1.8), why different p-value? Just rounding error?

Table 3
6. Please keep 2 decimal points for 95% CI = (1.0, 3.6), p = 0.06

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare I have no competing interests.