Author’s response to reviews

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

Authors:

Jasvinder A Singh (jasvinder.md@gmail.com)
David G Lewallen (lewallen.david@mayo.edu)

Version: 4 Date: 21 December 2012

Author’s response to reviews: see over
We thank the reviewers for insightful and detailed comments. We provide point-by-point responses to the comments. The comments have allowed us to show that the analytic results are robust and have strengthened the quality of work. Since there were numerous detailed comments, our responses are underlined.

**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.

**Response:** We regret the error. The analyses presented as main results in tables 2 and 3 in the original submission, were adjusted for the respective preoperative variable, preoperative pain for pain analyses and preoperative function for the function analyses. Thus, our findings, in fact, are adjusted for this important confounding variable.

As sensitivity analyses, we additionally adjusted each model for the other preoperative variable (preoperative pain for function outcome; preoperative function for pain outcome), and this led to results did not change (as shown below). This has been added as an appendix now.

**Results:**

**Income and outcomes after primary TKA**

"State Outcomes: In multivariable-adjusted models that adjusted for the respective preoperative variable, both lower income categories (≤$35 K and >$35-$45K) were associated with significantly lower odds of moderate severe pain, but higher odds of moderate-severe activity limitation at 2-years post primary TKA, compared to higher income (>45K) (Table 2)."

Appendix 1. Sensitivity analyses additionally adjusting multivariable analyses for preoperative activity limitation in pain outcome models and preoperative pain in activity limitation models

<table>
<thead>
<tr>
<th></th>
<th>2-year</th>
<th>5-year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderate-severe pain</td>
<td>Moderate severe functional limitation</td>
</tr>
<tr>
<td></td>
<td>Odds Ratio (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>≤$35K</td>
<td>0.62 (0.40, 0.96)</td>
<td>0.03</td>
</tr>
<tr>
<td>&gt;$35K-$45K</td>
<td>0.69 (0.50, 0.95)</td>
<td>0.02</td>
</tr>
<tr>
<td>&gt;$45K (ref)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
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.

**Response:** Income categorization is applicable for the 12-year study period, since we categorized people at 2x poverty level cut-off at the middle time-point, now clarified in the paper. For example, the poverty level cut-off in 2005 was $19,500 and at the mid-point of our study period was $17,000. Sensitivity analyses were performed changing the cut-off to 39K that revealed no change in interpretation (see below). This has been added to the manuscript, as an appendix, as well as results.

**Results:**
Sensitivity analyses that adjusted for preoperative pain in activity limitation model (and vice versa; Appendix 1), varied the income categories (Appendix 2) or examined age and BMI as continuous variables (Appendix 3), did not effect the interpretation of study results, for moderate-severe pain or moderate-severe activity limitation outcomes.

**Appendix 2. Sensitivity analyses using a different income category cut-off* in multivariable-adjusted analyses**

<table>
<thead>
<tr>
<th></th>
<th>2-year</th>
<th></th>
<th>5-year</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderate-severe pain</td>
<td>Moderate severe functional limitation</td>
<td>Moderate-severe pain</td>
<td>Moderate severe functional limitation</td>
</tr>
<tr>
<td></td>
<td>Odds Ratio (95% CI)</td>
<td>p-value</td>
<td>Odds Ratio (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>$\leq$39K*</td>
<td>0.67 (0.48, 0.93)</td>
<td>0.02</td>
<td>0.85 (0.68, 1.06)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>1.00 (0.77, 1.31)</td>
<td></td>
<td>1.00 (0.75, 1.48)</td>
<td>0.99</td>
</tr>
<tr>
<td>$&gt;$39K-$45K</td>
<td>0.64 (0.43, 0.96)</td>
<td>0.03</td>
<td>1.01 (0.78, 1.35)</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>1.00 (0.75, 1.48)</td>
<td></td>
<td>1.00 (0.75, 1.48)</td>
<td>0.76</td>
</tr>
<tr>
<td>$&gt;$45K (ref)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Income cut off was varied to 2-times the 2005 poverty cut-off of annual income of $19.5K at $39K, instead of the main analyses cut-off of $35K.
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.

**Response:** We have added a new table as suggested

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.

**Response:** We have added a new table with univariate results as suggested

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.

**Response:** All analyses for the three outcomes were specified a priori. All significant p-values in our study are 0.02 or lower, so even with the suggested correction of ~0.02, the results are still statistically significant. More importantly, the estimates are impressive with odds ratios of 0.6 and 0.7, and with all the suggested sensitivity analyses that we were asked to perform, are robust effects, since the estimates don’t change much. This supports the claim that these observations are likely “real” and not a chance finding due to the fact that we assessed three pre-specified outcomes. Given that Bonferroni correction is recommended as a conservative approach, we discuss this as a limitation and interpret our findings considering this.

“Even though we specified the three outcomes a priori, with the Bonferroni adjustment, our findings may be considered borderline significant for such a conservative approach (with a corrected p-value of 0.017 for 3 comparisons)."

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.

**Response:** We have clarified now that as planned a priori, we analyzed three outcomes and two time-points, six multivariable models were run, shown in tables 2 (2 models) and table 3 (1 model) “Multivariable-adjusted logistic regression analyses were performed separately for each of the three pre-specified outcomes at pre-specified time-points, 2- and 5-years (six models).”

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.

**Response:** The use of GEE was to account for the correlation of observations within an individual who could have undergone more than one primary TKA during the period of observation (right and left TKA).

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.

**Response:** As requested, we now provide these numbers in the results section. The response rates were 65% and 57% respectively.

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?

**Response:** We now added the detail that this was an ordinal regression analysis.

“Since improvement in knee function had 3 categories (much better, better versus same/worse [reference]), we used an ordinal logistic regression.”

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.

**Response:** We regret the confusion, we have now clarified the models were adjusted for age and BMI category, now clarified in methods and table legends. The PI is a trained epidemiologist with a master’s degree in epidemiology and training in biostatistics, who has conducted statistical analyses independently on several large datasets, in several published several peer-reviewed publications. For this study, Dr. Harmsen, a MS biostatistician assisted with these analyses, as specified in the acknowledgement section.

In addition, we conducted sensitivity analyses adjusting for age and BMI as continuous variables, with results shown in the appendix.

**Appendix 3.** Sensitivity analyses using BMI and age as continuous variables in multivariable-adjusted analyses

<table>
<thead>
<tr>
<th></th>
<th>2-year</th>
<th>5-year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderate-severe pain</td>
<td>Moderate severe functional limitation</td>
</tr>
<tr>
<td>Odds Ratio (95% CI)</td>
<td>p-value</td>
<td>Odds Ratio (95% CI)</td>
</tr>
<tr>
<td>≤$35K</td>
<td>0.59 (0.38, 0.89)</td>
<td>0.01 (0.58, 0.99)</td>
</tr>
<tr>
<td>&gt;$35K-$45K</td>
<td>0.67 (0.48, 0.92)</td>
<td>0.01 (0.78, 1.20)</td>
</tr>
<tr>
<td>&gt;$45K (ref)</td>
<td>1.00 1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
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?

**Response:** Our income data was based on census zip code data, which have been used commonly for analyses by SES. The data were for the year of survey, so if a patient underwent surgery in 2003, we used the median income for that year in his/her zip code census tract. Since we used individual year data (except 1994, 96 and 2005 approximated to the year prior), no adjustment was needed.

“The predictor of interest was the patient’s income at the time of index arthroplasty, as assessed based on patients’ zip code and the median household income for geographical area using the census data for the respective year of survey. Income was categorized into ≤$35,000, >$35,000-$45,000 and >$45,000, as previously (8) (9) (29) (30). With the exception of 1994, 1996 and 2005, data for each year were available. For these three years, we used the data from the prior years (1993, 1995 and 2004).”

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

**Response:** We agree that the direction of bias can not be certain, we have revised this statement as follows.

“Whether this biased our results towards or away from null hypothesis is not entirely clear.”

### ------------- 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?

**Response:** We have added frequencies to table 1 as requested. We have also provided the degree of missingness for each variable as suggested, as a legend.

4b. Did the table capture the variables at the operation/baseline? Or at 2-yr and 5-yr follow-up separately?

**Response:** We have specified that all variables were at baseline, except the dependent variable, outcomes of interest, that were assessed at 2- and 5-year follow-up.

“The main multivariable-adjusted analyses adjusted for several baseline variables including age category, gender, BMI category, comorbidity, ASA class, operative diagnosis, distance from the medical center and implant fixation.”
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?
Response: We have added the details regarding the response rates, and added decimal points in table 1 as suggested.

4d. “Age groups n (%)” - Why not immediately after “Mean Age”? Actually only the percentage was reported, while the frequency (n) was not.
Response: We have rearranged age variables next to each other as suggested, and added frequencies in table 1 as suggested.

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.
Response: Median household income is for is the amount which divides the income distribution in the U.S. into two equal groups, half having income above that amount, and half having income below that amount. Thus, this median is for the entire U.S. population. As expected, this is not the median for this population.

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).
Response: We have added these variables to table 1 as suggested.

Table 2
5a. Table title only states “at 2-years” but the content actually contains results at “5-years”.
Response: We regret the error and have corrected it as suggested.

5b. Please keep 2 decimal points for Odds Ratio and its CI.
Response: We have added two decimal points in table 1 as suggested.

5c. Among 8 adjusted covariates, which are also significant?
Response: We specify in table legend, which covariates are significant.

5d. Same estimated OR and 95% CI = 1.4 (1.1, 1.8), why different p-value? Just rounding error?
Table 3. Please keep 2 decimal points for 95% CI = (1.0, 3.6), p = 0.06
Response: We have added two decimal points in tables 2 and 3 as suggested.

Statistical review: Yes, and I have assessed the statistics in my report. Declaration of competing interests: I declare I have no competing interests.
Reviewer's report

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

Version: 3 Date: 31 October 2012 Reviewer: Hassan Ghomrawi

Reviewer's report:

In their revision, the authors have acknowledged many of the limitations of the paper that have not been acknowledged in the original submission. However, they have not made any efforts to address these limitations despite the availability of data. For example, they acknowledge that baseline functional status is collected but is missing on a significant number of patients. There has been no attempt to conduct any sensitivity analyses within this subset to determine if the results change, especially that with the inclusion of this variable, the only other study has found no effect of income.

Response: We regret the error in the methods, the analyses presented in the initial submission were, in fact, adjusted for pre-operative variables as has been suggested by the reviewer. In particular, each pain model was adjusted for preoperative pain outcome and each functional limitation model was adjusted for preoperative functional limitation. We have not corrected this error in the initial submission.

Additionally, we performed additional sensitivity analyses by including both preoperative pain and function in both pain and function analyses models, and present these sensitivity analyses in the results. No change in the level of significant or odds ratios were seen, confirming that these analyses are robust. These are presented as appendix, see details in the response above to a reviewer’s comment.

Another example is the income categories and how these have different purchasing power in different states. This is the main variable of interest and its impact could change with adjustment. Although the authors state this as a limitation, the authors state in the discussion that “this does not affect the observed associations” (p 9, line 5). Again the authors do not attempt to address this limitation in any type of sensitivity analysis (by restricting their analysis to the Mid-West) or for example by including the state or the census region as indicators in the analyses.

Response: We found that lower income is associated with outcomes for the entire sample. One of the main strengths of this paper was availability of a large enough to sample to have conducted analyses to detect a significant association, if one truly existed.

As suggested by the reviewer, we performed another sensitivity analyses. Since 53% of our patients lived within 100 miles of the medical center, that is representative of Mid-West (and likely indicative of similar purchasing power), we performed an analyses limited to these patients. Before performing these analyses, we realized that by reducing the sample size to <50%, our focus should be on stability of odds ratio as a measure of association and much less on p-value/confidence intervals, since by definition these would get wider/non-significant. As expected, we found that odds ratios were unchanged for the 2-year outcomes, pain and functional limitation, see below. We have now added this to the results section. This analysis also proves our belief that with smaller samples with fewer people with the outcome of interest are often underpowered, and it’s not uncommon to have Type II error, as shown by minimal changes in OR, but change in statistical significance interpretation.
| Income Range | Moderate-severe pain | | | | Moderate severe functional limitation | | | |
|-------------|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|              | Odds Ratio           | p-value         | Odds Ratio      | p-value         |                  |                  |                  |                  |                  |
|              | (95% CI)             |                 | (95% CI)        |                 |                  |                  |                  |                  |                  |
| ≤$35K        | 0.61                 | 0.02            | 0.78            | 0.07            |                  |                  |                  |                  |                  |
|              | (0.40, 0.94)         |                 | (0.60, 1.02)    |                 |                  |                  |                  |                  |                  |
| >$35K-$45K   | 0.68                 | 0.02            | 0.96            | 0.75            |                  |                  |                  |                  |                  |
|              | (0.49, 0.94)         |                 | (0.78, 1.20)    |                 |                  |                  |                  |                  |                  |
| Sensitivity Analyses limited to <100 miles (Midwest) | | | | | | | | |
| ≤$35K        | 0.67                 | 0.21            | 0.70            | 0.05            |                  |                  |                  |                  |                  |
|              | (0.36, 1.25)         |                 | (0.50, 1.00)    |                 |                  |                  |                  |                  |                  |
| >$35K-$45K   | 0.66                 | 0.05            | 0.92            | 0.53            |                  |                  |                  |                  |                  |
|              | (0.44, 1.00)         |                 | (0.70, 1.20)    |                 |                  |                  |                  |                  |                  |

The current analysis that overcomes at least the first limitation (including baseline function) is the analysis of the change variable. In this analysis, there was no effect of income on outcomes and this should be the main conclusion of the study if the authors wish to make no further adjustments to their analyses. This main finding should be acknowledged in the summary and as well as the abstract.

**Response:** We have now performed the requested analyses including preoperative variables, please see above. Since our main finding for the state variable holds true with the adjustment for baseline preoperative variables as well as three additional sensitivity analyses proposed, we consider this our main study finding. We have discussed that income was not associated with the change functional variable.