Author's response to reviews

Title: Modeling early recovery of physical function following hip and knee arthroplasty

Authors:

Deborah M Kennedy (d.kennedy@utoronto.ca)
Paul W Stratford (stratfor@mcmaster.ca)
Steven E Hanna (hannas@mcmaster.ca)
Jean Wessel (wesselj@mcmaster.ca)
Jeffrey D Gollish (jeffrey.gollish@sunnybrook.ca)

Version: 2 Date: 21 August 2006

Author's response to reviews:

August 20, 2006

Sandra Le, Ph.D.
Assistant Editor
BMC-series journals
BioMed Central Ltd
Middlesex House
34-42 Cleveland Street
London W1T 4LB
UK

Tel: +44 (0)20 7631 9921
Facsimile: +44 (0)20 7631 9923
e-mail: editorial@biomedcentral.com

Dear Dr Le:

RE: Revision of Manuscript Based on Reviewer Feedback

"Modeling early recovery of physical function following hip and knee arthroplasty"

Enclosed please find our revised manuscript for consideration of publication in BMC Musculoskeletal Disorders. As recommended, a point by point list of the Reviewer's insightful comments and our response and corresponding changes in the manuscript follows.

Should there be any need for immediate contact, I wish to inform you that I will not be available from August 21 to the 28th. Thank you for all your assistance.
Responses to Reviewers

Reviewer: Nadine Foster

Major Compulsory Revisions: None

Minor Essential Revisions:

Suggestions:

Please make it clear in the abstract that the participants were having post-operative rehabilitation.

A sentence was added to the abstract and further information was added to the methods section. Due to the change in the abstract, the word count became over 350 words which resulted in some further editing to the original abstract submission.

In the methods section, add some detail to explain to the reader what types of arthroplasty were used and some information about what 'standardized inpatient treatment' consisted of and what post-operative pathways were followed.

Detail concerning the types of prostheses and further information about the treatment was added to the design section in Methods.

In Methods, page 6, I recommend adding a couple of examples of the timings of the participants'
assessments.

A sentence was added with examples of the timings of the participants' assessments.

Please explain how many patients were approached to take part in the study, which were potentially eligible, so the reader can see the difference between this figure and the 188 patients who provided informed consent. If you do not know the total number approached, state this and explain why.

The total of 188 patients represents about 87% of the patients approached during the study period. The authors did not feel confident in reporting this value as there were a few periods of interruption of recruitment and tracking such as with the outbreak of SARS and a period of unexpected absence with the research assistant.

Page 7, sentence in which 'overlapping' sample of patients is mentioned. Could it be reworded for clarity?

This sentence was reworded and is now hopefully clearer.

Page 10: error-reference Snijders should read 38

Thank you for picking up this error, it has been corrected.

Page 10 and the discussion section: please reflect on whether there are sufficient numbers in the sample for the type of analysis chosen, given that for many patients, there was only one or two assessments.

Additional information was added to the section on limitations with respect to the sample and the number of follow-up assessments.

Related to this, is there a way of estimating the precision of the results from this analysis in a similar fashion to the use of confidence intervals?

Two additional tables (4 and 5) with the parameter estimates for one of each of the performance and self-report measures have been added with 95% confidence intervals and p values.

Results and discussion: The difference between the self-report and the performance measures is quite striking. Could the authors mention the small.......Looking at the curves, one wonders if surgery was really that beneficial.

The reviewer has raised good points. A sentence has been added to indicate that further study will be required to examine the recovery curves of the 6MWT and ST to determine how much further improvement in function is obtained and to determine when patients have reached the stage of most benefit from surgery. In terms of the benefits, it is not the intent of this paper to discuss this concept--however, patients are usually pleased post THA and TKA because of the reductions in their pain levels which are not graphically presented in this paper.

Where there any complications from the surgery?

A sentence was added about complications in the results section.
Page 13: mentions positional restrictions to avoid dislocation--this type of information would be better in the methods section, with other information about the standard post-operative care of these patients.

Modifications were made to these sentences in the discussion with some of the details moved into the methods section.

Reviewer: Oliver Ethgen

Major Compulsory Revisions:

Page 10, results: What about gender distribution? In addition, we would expect more info about the population; duration of disease, time spent on waiting list etc.

Details about the gender distribution have been added to the results section. The impact of gender is described in greater detail in a recent publication 'Preoperative Function and Gender Predict Pattern of Functional Recovery Following Hip and Knee Arthroplasty' in the Journal of Arthroplasty--see reference #39. Unfortunately we did not collect information on the duration of the disease. Time spent on the waiting list was not known for all of the patients--only known for a subset who participated in a different arm (pre-operatively) of the longitudinal study.

Minor Essential Revisions:

Page 6, first paragraph This minimized the chance of dictating....This sentence is difficult to understand. Could be clearer.

The aforementioned sentence was modified and is hopefully now clearer.

Page 6, last paragraph: I am not sure I understand the loss of follow-up due to the outbreak of Severe Acute Respiratory Syndrome. Can we have more details?

A footnote about Severe Acute Respiratory Syndrome (SARS) has been added.

Page 7 & 8, measures: The authors present the different measures employed in their study. This part is a little bit disorganized. Would be better to have 1 paragraph for 1 measure.

In the first two paragraphs on the measures, the authors were attempting to provide some references regarding the psychometric properties of each of the measures in a more succinct way. Initially the authors did try presenting it as the reviewer suggested, however due to the psychometric properties of the performance measures being detailed in one study, there was a lot of redundancy when it was organized separately for each measure.

Page 10, first paragraph; [Snijders, 2002 #301] actually stands for reference [38] I think.

Thank you--all of the reviewers picked up this error. It has been corrected.

Page 11, second paragraph, line 15: Should we understand a floor effect? And not a ceiling effect? Same
We can appreciate that this will be difficult for readers as the TUG curves are going down and intuitively one would think of this as floor effect, however, improvement for the TUG measure is a lower time score and thus represents a ceiling effect. In other words, patients can not improve further even though the graphs look more floor like as the curves approach the ‘x’ axis. Some clarification was added to the text.

Reviewer 3: Jos W Twisk

Major Compulsory Revisions

After creating a model for development over time, the authors address the influence of several predictors on the development over time. I think the authors should distinguish between the effects of the predictors on the outcome (which they probably did and which is not the influence on the rate of change; it is only the influence on the average value over time) and the effects of predictors on the rate of change...... the interpretation of their results.

In the analysis, we did look at the influence of predictors of the intercept as well as the slope. The intercept is interpretable as the prediction at baseline or in this case the patients’ function at one week after surgery but is also related to the average level of the response. As a more detailed description of the predictors which does distinguish between the effects of the predictors on the outcome and rate of change is published in the recent June edition of the Journal of Arthroplasty (details in the reference list), we did not wish to duplicate results. A paragraph from this publication is included for demonstration purposes:

"Preoperative 6MWT score, gender and site of arthroplasty were all significant predictors of 6MWT distances at one week after surgery (Figure 3 and 4). Figure 3 shows that patients with higher preoperative 6MWT distance scores had higher function at one week postoperatively than patients with lower preoperative scores. A significant interaction was found with baseline preoperative 6MWT score and weeks after surgery, such that patients with higher preoperative scores are predicted to have a faster rate of improvement than those beginning with lower scores (i.e. shorter distances). The interaction of the growth parameter with site of arthroplasty indicates that the instantaneous rate of change differs for patients post THA and TKA at one week after surgery, with individuals undergoing TKA showing a slower gain in distance over time (Figure 4). Although women begin with lower 6MWT scores, their rate of improvement is similar to men."

A point is raised in the minor revision section about the use of the terms "linear" and "quadratic" component of time instead of rate of change at 1 week and the change in change rate. The authors were previously advised to simplify the use of these terms to make the interpretation clearer to clinicians. Some instances in the text have been changed to use the reviewer's recommendations.

Furthermore, table 3, in which the results of these analyses are summarized has a terrible lay-out. Please provide regression coefficients, 95% confidence intervals (and perhaps also p-values). Besides this, I suggest the authors do three steps. First, the development over time. Second, the influence of additional predictor variables and third potential interaction terms. The analysis with interaction terms should be reported separately from the other analyses, because otherwise they can't be interpreted in a straight way anymore. By the way, the interactions with time indicate that the development over time is different for different subgroups.

We can appreciate the reviewer's comment from a statistical perspective but again we were trying to provide an interpretative summary of the results, oriented to the clinical reader. We would like to retain the clinical summary with revisions to provide a clearer delineation of the predictors and interaction terms. In addition we will include sample tables (numbered 4 and 5) of the parameter estimates for one of the physical performance measures (6 MWT) and self-report measures (physical function subscale of the...
WOMAC) with the confidence intervals and p-values. Reference to these two new tables has been added to the text.

A major concern is the fact that more than 50% of the subjects only have two measurements.....At least the authors should say something about this in their discussion.

Some additional points were added to the section on limitations and the impact of the number of data points for patients.

Minor Essential Revisions

The way the authors mention the linear and quadratic component of time is not very clear....the change in change rate.

This point was addressed earlier.

In the introduction, the authors explain HLM. This part should be taken to the methods section. The way the authors describe the HLM is not very clear. I think they can do better.

The portion explaining HLM was removed from the introduction and more detail was added to the methods section.

Line 5 of the results: p< 0.60

Thank you--the error was corrected.

In the results section, the authors state that a significant interaction means that the development lines cross. This not really true....of the observed interactions.

We agree with the reviewer and were just trying to reference the interactions pertinent to the graphs in this study. A few words were added to make this clearer in the results section.

At the end of the results section, the authors state that the second degree polynomial provided a reasonable fit for the data. I wonder how the authors investigated this reasonable fit and what they mean by this statement.

Upon final model selection, various plots were investigated to look at fit in terms of the residuals (point made in Methods section) and the relationship between the actual measure scores versus the predicted scores. For example, a plot of the actual versus predicted scores for each measure was graphed to see if a straight line relationship existed.

In addition, since a quadratic is a parabola--as cautioned in the discussion section, interpretation beyond the measured points is not possible given the fact that the trajectory will trend down and clinically, one would not expect to see an average pattern of deterioration beyond the time point graphed.

Table 1: There is a mistake in the N of the LEFS score. Furthermore, in the last column the T-value and the degrees of freedom are not informative: they are only used to obtain the p-value. I suggest to leave that information out. The same holds for Table 2.
Thank you, the error in Table 1 in the N of the LEFS score has been corrected. The T-values and degrees of freedom in Table 1 and Table 2 have been removed.

When I look at the figures, I have doubts about the normality of the TUG scores....Also the stair test looks suspicious

The reviewer is correct about the normality of the TUG scores and a section about the need to transform the data using a natural logarithmic transformation prior to modeling has been added to the methods section. This information was removed from an earlier manuscript due to the focus on presenting a paper that would be pertinent to the clinical reader and not be confusing with too much methodology.

As mentioned before, more than 50% of the patients...so the development over time for the whole population is highly biased.

As noted earlier we did add further information to the section on limitations. Although the authors can not fully reject the notion that subjects with two measurements recover faster--this was not the case in all instances. The authors do know that some patients with only two measures were the result of the SARS outbreak between the end of March and June 2003. During this SARS period--no outpatients were permitted into the hospital due to the possibility of infection and therefore, all follow-ups at this time had to be cancelled.