Author's response to reviews

Title: A path analysis of the effects of the doctor-patient encounter and expectancy in an open-label randomized trial of spinal manipulation for the care of low back pain

Authors:

Mitchell Haas (mhaas@uws.edu)
Darcy Vavrek (dvavrek@uws.edu)
Moni B Neradilek (moni@mwlight.com)
Nayak Polissar (nayak@mwlight.com)

Version: 3 Date: 17 December 2013

Author's response to reviews: see over
Response to Reviewers
(MS: 4655457341051454)

We would like to thank the reviewers for their helpful comments. Please find our responses in blue below. Changes to the manuscript have also been made in blue.

REVIEWER 1

Minor Essential Revisions:
Have these data been previously analyzed and published? Should note explicitly in methods if this is a re-analysis.
This paper has not been previously published and this is not a re-analysis of the data. We certify this in the cover letter of the submission to the Journal. We have clarified this by adding “This was a preplanned secondary analysis” to the first paragraph under Statistical Analysis in the Methods section.

Discretionary Revisions:
Would it be useful to present a zero-order bivariate correlation matrix?
We have added a table of bivariate correlations (Table 2) and a subsection called “Correlation” under Results.

It would be interesting to see if we can predict the dropouts using baseline predictors. Do the authors think baseline expectancy, or initially-reported DPE would predict attrition? But, as most dropouts as noted as "personal", we may not expect much interesting to come out of those analyses -- unless participants gave personal reasons to avoid being confrontational.
Although this is an interesting idea, we agree with the reviewer that there is little chance of anything useful arising. Personal reasons for nonadherence to care generally related to changing jobs, moving, and family/unrelated health problems. There were only nine participants who truly dropped out (completely lost to follow-up), so precise effect estimates are not possible. This analysis would also be well beyond the scope of this paper.

Once we leave the safety of the baseline measurement period, I’m sure the authors know interpretation of paths becomes difficult. The authors note issues of bi-directional relationships and feedback loops. The authors seem to be biased towards seeing DPE as a construct that changes pain, though the reverse path could be argued. I tend to like doctors who give me treatments that work, and dislike doctors that give me treatments that don’t work. DPE, measured at the same time as 6-week pain, may be difficult to interpret as a causal variable. A possible limitation.
We chose the path direction based on a pre-specified interest in the effect of the DPE on outcomes in unblinded trials. However, like the reviewer, we were curious about the effect of pain on the DPE and, as a sensitivity analysis, ran the model with reverse pain-DPE paths. The path coefficients were trivial in this direction and their inclusion or exclusion in the model had little effect on other path coefficients. Therefore, we exclude this sensitivity analysis from the paper to prevent digression and undue complexity.
The authors suggest their method creates equipoise among the clinicians, but we don’t know what the equipoise would have been had the authors not trained the clinicians. It is possible that the participants would have seen no differences in clinicians’ attitudes even without the equipoise training. The concern is minor, though, because it is fairly well understood that clinicians in trials should be trained for homogeneity of the practices. A possible limitation.

The reviewer raises an interesting methodological question: is clinician training necessary to establish equipoise and if so, how much? Our experience has taught us that the training threshold required for success is likely below the comfort threshold of grant reviewers. The establishment of equipoise in our study demonstrated that extensive training was unnecessary, a fact useful to researchers. Had we failed to establish equipoise, then our minimal training would have been a study limitation.

My general understanding of path analyses is that betas give relative, but not absolute, importance of the paths. Could the authors provide R2 values in Figure 2 for each of the paths so we can more easily determine the variance explained? Of, if that is not possible, could we see R2 values inserted in the variables to see how much of the variance is predicted by the model at each step?

We have added the total $R^2$ for each dependent variable in the model to Table 3. This provides information about how much variance in an outcome was explained. As for the effect of the individual predictors on a given outcome the standardized beta coefficients are heuristically comparable to the square root of the change in $R^2$ for that variable. The standardized coefficient corresponds to a standardized change in the outcome (in multiples of the SD of the outcome) per 1 SD of the independent variable. An advantage of the standardized beta coefficient is that it contains the sign that points at the direction of the effect ($R^2$ by definition is non-negative and provides no information on the direction). Both the standardized coefficients and the $R^2$ are dimensionless and therefore allow comparison of effects across different predictor variables.

Specific Comment and Recommendations for Revisions:

Title
Discretionary Revisions:
The title implies that the authors conducted structural equation modeling but this is not the case as the authors only did a path analysis, I therefore suggest that the title be rephrased to accurately convey what has been done: could rephrase to “A path analysis of the effects of the doctor-patient encounter and expectancy in an open-label randomised trial of spinal manipulation for the care of low back pain”

Path analysis is a special case of structural equation modeling without latent variables, so the term structural equation modeling is technically correct here. For the sake of clarity, we have changed “structural equation model” to “path analysis” throughout the paper with the exception of the Statistical Analysis section under Methods where we identify the path analysis models being fitted using structural equation modeling software.

Discretionary Revisions:
Abstract
I suggest that in the abstract the authors simply state that: A path analysis was conducted to determine the effects of dose rather than saying: Structural equation modelling was used in a path analysis. This has been changes as requested. Please see the previous comment.

Introduction
The scientific background and rationale of the study was reported. However, almost the same introduction was used by the authors in their earlier article (ref 14)
The current path analysis had the same primary goal as the previous one, addressing nonspecific effects in open-label trials. We do note the differences in patient populations in the introduction, as well as technical advantages of the current study in the Discussion. The current study had five times the sample size, allowing us to include more pathways in the model and get better estimates of the path coefficients. Methodologically, we also chose to evaluate indirect and total effects in addition to the direct effects in the presentation of the results of the current study.

Method
Discretionary Revisions:
The authors could also provide evidence for the reliability and validity of the measure used to assess DPE.
To our knowledge, there has been no validation of the instrument from which these questions were taken. A brief remark has been added to the first paragraph under Limitations.

Discretionary Revisions
The authors could include the fit indices for their final model
These have been added under Statistical Analysis as parentheticals in the description of the goodness of fit tests. They were placed here rather than under Results for simplicity.

Discretionary Revisions
The key weakness of this study was to use an average of four questions to assess DPE. A greater effect of DPE may be shown if a measurement model was included rather than using an average of four items. I suggest that the authors consider including a measurement model for DPE with the four questions used to assess DPE as the indicators of a latent construct.
At the outset of our modeling, the two statisticians on this paper discussed this very option with their colleague at the University of Washington who is a structural equation modeling expert. The consensus was that measurement model for DPE was rather poor and the average approach was actually preferable. The reason we found the measurement error poor is that each latent construct was measured by only four variables. We have conducted an in-house sensitivity analysis with the measurement model for DPE and the results did not differ much from the current analysis. The DPE effects were slightly stronger in the measurement model. However, given the poorly determined measurement model for DPE we felt more comfortable with the conservative choice of the average DPE. While it is true that the latent DPE effect may be slightly underestimated with the average DPE the same is likely true for the effects of previous pain and expectations that are also not analyzed as latent. Thus, the treatment of the three different risk factors (DPE, pain and expectation) in our analysis is equal. Finally, the estimation of the measurement error model is much more computationally extensive and makes bootstrap
calculations for the coefficient confidence intervals and p-values problematic (too much computational time).

Discretionary Revisions: The authors could comment on how they evaluated the assumptions of multivariate normality and checked for multi collinearity.
We have identified the tests used under Statistical Analysis at the end of the first paragraph on Page 8.

In more detail: The assumptions of multivariate normality were checked by a visual assessment of normal quantile-quantile plots for the dependent variable residuals. No substantial departures from normality were found. As a sensitivity analysis we also estimated the model by the ADF (asymptotic distribution free) method. The results were similar to the results estimated by the maximum likelihood. Finally, the confidence intervals and p-values for the coefficients are calculated by the bootstrap which lessens a potential effect of departures from the normal distribution. Multicollinearity was assessed using the VIF (variance-inflation factor) for the linear regression that comprise the model. The largest VIF was 2.27 for the 6 week pain as a predictor 12 week expectation. These are relatively low values of VIF (values >5 or >10 have been sometimes used as a rule of thumb to define high multicollinearity).

Discretionary Revisions. The methods used to handle missing data were reported but the authors could also do a sensitivity analysis using appropriate values to test what happens if a patient drops out for example due to low DPE or high pain
This is an interesting idea but we did not carry it out as the extent of missing of values was small. Only nine participants truly dropped out (completely lost to any follow-up). The suggested analysis was extremely unlikely to change the results.

Results
Direct, indirect and total effects were well presented and a detailed interpretation of these was provided. However they is a typo error in Figure 2, PPE 12 weeks should be DPE 12 weeks.
(Minor Essential Revisions)
This has been corrected. Thanks for catching this.