Author’s response to reviews

Title: Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons

Authors:

Andrew Hutchings (andrew.hutchings@lshtm.ac.uk)
Jenny Neuburger (jenny.neuburger@lshtm.ac.uk)
Jan van der Meulen (jan.vandermeulen@lshtm.ac.uk)
Nick Black (nick.black@lshtm.ac.uk)

Version: 8 Date: 30 September 2013

Author’s response to reviews:

Thank you for inviting us to revise our paper. As one reviewer did not suggest any changes were necessary, we have responded to the other reviewer who had several useful suggestions. We have responded as follows:

(1) For foreign readers who are not well acquainted with the PROMs system of the NHS (such as the reviewer), it is difficult to understand what the problem or the starting point of this paper is. Of course, you clearly describe that problems with data linkage may lead to false estimates of recruitment rates. But when I look on the respective website:


one must get the impression that the Centre undertakes many efforts to achieve an optimal linkage of questionnaire data and hospital episodes. On the website I could read:

“The methodology that matches a HES episode to a Q1 questionnaire uses a four-stage process that looks at a combination of patient identifiable fields, provider codes, operation codes and dates. It uses a „match ranking; system in which a score is attributed to each part of the linking process, where the quality of the match is denoted by the rank, with the lowest rank (i.e. 1) being the highest quality match. The scores for each possible match are compared and the highest match is chosen. This approach has the advantages of maximising the matching rate by attempting several different matches on the same data and allowing the easy monitoring of quality and confidence of the match.”

Given this statement, it seems important that you describe the problem more precisely for the naïve, but interested, reader. It would also be helpful if you have any references that have already highlighted this linkage problem and its consequences for health policy and healthcare decisions. Otherwise, one might get the impression that you are looking into a purely academic issue.
*We agree that the Information Centre have strenuously tried to minimise recruitment bias. Despite these efforts there are clearly shortcomings (see Figure 1) and we have taken the reviewer's advice to move material from the end of the Background to the second paragraph.

(2) While the paper suggests that you demonstrate the consequences of inadequate calculation of recruitment rates for a proper assessment of PROMs, it seems you do something different. Your focus is on the first round of the PROMs survey, i.e. the baseline. But what really matters is the comparison of this first survey with the second survey (after the surgery) in order to measure change and success. If the recruitment rate is falsely calculated and may result in bias, readers of this Journal, I believe, are interested to know what role this may play for measuring change and success. Of course, an inadequate calculation of response rates in the first round may have consequences for the second round as well. However, I believe that the readers of this journal are most interested in biases in the reported outcomes, and less interested in the data linkage or recruitment procedures themselves. So, I wonder why you did not include this second round although I'm very sure you have access to these data.

Or to be more explicit: As long as you do not study the complete process of data acquisition and then make conclusions about possible bias, the paper is not very interesting for readers of a health services journal but remains more or less a technical recommendation to the Centre and the NHS about how to improve linkage and the calculation of recruitment rates. So I would strongly recommend including these additional data and their analysis.

*We would challenge that "what really matters is the comparison of this first survey with the second survey". There are two principal objectives of the use of PROMs. One is, as the reviewer points out, the assessment of patients' outcome (post-op versus pre-op) so as to compare the performance of providers.

But the second, and equally important, aim is to assess the appropriateness of patient selection as judged by the severity of their pre-operative condition. Indeed, the principal reason for government support for the Programme in England was to manage demand for these operations by identifying those patients in whom the operation may not be appropriate. For this aim, recruitment bias is of central importance and our paper demonstrates that it is a justified concern.

*As regards the first aim, and the reviewer's understandable interest in it, we can only speculate because we cannot know whether or not unmeasured confounders are evenly spread among providers. As we say in the Conclusions: "Non-recruitment will bias comparisons of providers' pre-operative case-mix and may bias comparisons of outcomes if unmeasured confounders are not evenly distributed between providers".

Whilst we share the reviewer's desire that we make explicit the impact of recruitment bias on outcome comparisons, we are not able to for this reason. His
comment has however highlighted the need for us to make the twin aims of the use of PROMs clearer at the outset. We have added in some extra sentences in the first paragraph of the Introduction and the Abstract.

(3) In the Methods section, you talk about a “multi-variable logistic regression” (page 8). However, you restrict yourself more or less to report unadjusted odds ratios (Table 4) although obviously adjustment would lead to different results as you report on page 12 (lines 3 ff). I can’t see any rationale why you don’t use adjusted odds ratios throughout the paper and consider the results for discussion. Taking into account the just-mentioned adjusted odds ratio on page 12, your conclusion that the most deprived have lower recruitment rates is obviously false. May be this is true also for the other results or the odds ratios are much lower after adjustment so that the recruitment rates do not really differ according to the four characteristics under study. So, I would strongly recommend to use adjusted ORs.

*We apologise for not making clear that the results are indeed adjusted ORs. We have sought to clarify that in the text and in the title of the new Table 3 (old Table 3 and 4 combined).

(4) I am surprised that you used “mean disease-specific PROMs” (page 8) to study whether recruitment rates are associated with the health status of the patients. As far as I learn from the website https://catalogue.ic.nhs.uk/publications/hospital/proms/proms-eng-apr-10-mar-11-final/fina-prom-eng-apr-10-mar-11-pre-post-rep1.pdf patients have to report quality of life measures when filling in the questionnaires. Why not use these patient-specific scores instead of mean scores? Again, I suppose you have access to these data without any barriers.

*All our analyses of the association of recruitment and patient characteristics are, inevitably, at a population or aggregate level not at the level of individual patients. So measures of patient characteristics have to be measures based on aggregate data. For some variables (eg sex, deprivation) the approach is categorisation. But for a continuous variable (such as PROM score) we chose to use a mean score. We could have categorised the data but that has the drawback of losing a degree of detail and analytical power. We are unclear, therefore, about the reviewer's suggestion to use patient-specific scores.

(5) The second to last paragraph of the Background explains the aim of the study. I think this should be the last paragraph. What you report in the last paragraph of the Background is a mixture of methods (first lines) and rationale for the study (more than 100% recruitment rate). This last aspect should be moved to an earlier stage in the section together with some more hints why the current procedure is not adequate.

*We agree and have done as the reviewer suggests.

(6) Your argument that the external validity is reduced (page 5, lines 6 ff,) needs, if any, 1 but not 3 references.
*We have deleted two references.

(7) I also think that 1 reference is enough for the HSCIC, and not 3 references (page 5). In contrast, I miss some references about the use of PROMs and so on. This underpins the importance of your issue.

*We have deleted two references re the HSCIC. We have included extra references in the Background to meet the reviewer’s request for more about the use of PROMs.

(8) I know that abbreviations are an easy way to communicate with colleagues. However, abbreviations make reading the text sometimes very difficult and are often unnecessary. In your case, I would recommend the following:

- use “first PROMs questionnaire” (or something like that) instead of “Q1s”
- use “hospital episodes” instead of “HIS episodes”
- use “ID” instead of “HISID”

*We have taken on board two of the reviewer’s recommendations and removed Q1 and HES episodes. We retain HESID as this is a particular form of ID (carefully described) that preserves patient confidentiality and we feel that use of the generic ‘ID’ may confuse.

(9) It is not easy or even impossible to find the categories presented on page 9 in Table 1. Only for “duplicates” and “operations cancelled” it is rather easy. Please make sure that readers will be able to find the respective categories without any problems both in the text and in Table 1.

*We agree and have clarified in the text which factors can be seen in Table 1.

(10) I think it is absolutely necessary to combine Table 1 and Table 2 into a single table. It is just one of the important messages of your paper that we must become aware that different data linkages lead to different recruitment rates. This should be visible at a glance in a single table.

*Whilst they could be merged we disagree that there is any benefit. Indeed, we feel it would make interpretation harder as it would mix up the two steps of the analysis: establishing the nature and extent of eligibility (Table 1) and the consequences of ineligibility (Table 2). We are, therefore, reluctant to merge them but will do so if the editor insists!

(11) I would recommend the same for Table 3 and Table 4. First of all, you may reduce Table 3 by half by omitting the absolute numbers and only reporting the relative numbers. It is sufficient to report the total number of participants in the heading of the table. So, I think you will have enough room to include the odds ratios (as mentioned above – the adjusted odds ratios!). Again, this would help readers to recognize a possible bias without reading columns of figures.

*We accept that this is a good suggestion and have adopted it.
I'm afraid it is counter-intuitive to report odds ratios for "non-recruitment". It will be much easier for readers to have odds ratios for "recruitment". If you change the reference group accordingly, a figure above 1.0 would mean "more recruitment" and this how most brains (talking about myself) understand such figures easier.

*While respecting the reviewer's preference, we don't agree that it is more or less intuitive to consider the odds of recruitment or of non-recruitment. Other might just as easily say its counter-intuitive to consider the odds of recruitment. We feel there is no right or wrong approach and have not reanalysed and recast this aspect of the paper.

Figures 1 – 3 (and their respective figures in the appendix) are not really enlightening. Perhaps you have an idea about how to combine this information in a better way. In any case, the x-axis is not labeled with "providers" or something like that.

*We're not sure how to respond as we felt that Figs 1-3 were quite enlightening! Again, we will defer to the judgment of the editor as to whether or not they are retained. Meanwhile we have labelled the axes.

I think the two last analyses in your paper are really unnecessary and superfluous:

-the correlation between recruitment rates and life quality scores (Figures 4 and 5 and respective figures in the appendix). Not only is the correlation coefficient very small; a closer look at the figures shows the large variability so that any conclusions would not be valid.

-the association between provider recruitment and pre-operative scores: a very theoretical analysis with only little relevance, especially in the way you present it. So it is really difficult for readers to read and understand what you did.

*We agree that the correlations are weak and non-significant. On reflection (thanks to the reviewer) we have redrafted the text to reflect this. Ironically, as a result, we think this makes it all the more important from a policy perspective to show and report the lack of any important influence of recruitment rate in this way.