Author’s response to reviews

Title: The agreement between chronic diseases reported by patients and derived from administrative data in patients undergoing joint arthroplasty

Authors:

Bélène Podmore (belene.podmore@lshtm.ac.uk)

Andrew Hutchings (andrew.hutchings@lshtm.ac.uk)

Sujith Konan (sujithkonan@yahoo.co.uk)

Jan van der Meulen (jan.vandermeulen@lshtm.ac.uk)

Version: 1 Date: 05 Mar 2019

Author’s response to reviews:

Dear Dr. Dirk Krüger,

Thank you very much for reviewing our manuscript. We would like to thank the reviewer for their constructive comments. We have made changes to the manuscript (BMRM-D-18-00420) which we believe substantially improve it.

Please find below a point-by-point response to the reviewer’s comments. We hope that you find our responses satisfactory.

Yours sincerely,

Bélène Podmore

Andrew Hutchings

Sujith Konan

Jan van der Meulen
Reviewer 1

Comment 1: The manuscript presents an analysis of a very large data set comparing patient-reported comorbidities with those available from health records. This is an important topic as patient-report is frequently used to assess comorbidity in the literature. Both assessment methods come with specific sources or bias and error that affect reliability of the measurement and therefore consistency of results obtained from patient-report and health records.

Response: Thank you for the comment.

Comment 2: A key point, that is also stated by the authors, is the lack of a gold-standard in this analysis. For this reason the accuracy of the patient-report cannot be assessed, but only the consistency with health records (accuracy would require comparison against the "true" comorbidity status).

Response: We agree with the reviewer that the use of the term accuracy can be problematic in the absence of a ‘gold standard’ diagnosis. We have removed the use of the term to describe our analysis and, as the reviewer suggests, replaced this with the term consistency (please see Abstract pg.2 and Conclusion section pg. 13). We refer to hospital administrative data as a reference standard (rather than a ‘gold-standard’) as recommended in the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guidelines.

Comment 3: While the authors mention the lack of a gold-standard the way the data is analysed suggests differently. By calculating sensitivity and specificity of patient ratings using health records as a criterion, it is implicitly assumed that health records provide the "true" comorbidity status, since in most of the literature sensitivity and specificity are related to a gold-standard measurement.

Response: We disagree with the reviewer that the use of sensitivity and specificity assumes a ‘true’ comorbidity status from a gold standard diagnosis. STARD reporting guidelines use the term reference standard as opposed to gold standard and are clear that sensitivity and specificity are estimated using a reference standard. Pfeiffer and Castle (Epidemiology 2005;15:595–597) also point out that gold-standards are rarer than one might think.

Comment 4: Cohen's Kappa may be the more appropriate statistic here as it does not imply which of the two measurements is more accurate. In addition, the authors should point out in the discussion that sensitivity and specificity depend on prevalence which limits the comparison of these parameters across different types of comorbidities with varying prevalences. Again,
Cohen's Kappa is more suitable for this purpose (please see for example Vach 2005: https://www.ncbi.nlm.nih.gov/pubmed/15939215).

Response: We agree with the reviewer that the Kappa statistic can be helpful and we report this statistic together with sensitivity and specificity. Vach (2005) also highlights limitations of the Kappa statistic, particularly summarizing a four-fold table into one number that provides limited information about why there may be a lack of agreement between patient-reported information and administrative data. To provide fuller information we therefore report three measures in Tables 2 and 3: sensitivity, specificity and the Kappa statistic. We have added the discussion of the limitations of the agreement measures in the Discussion section and we have added the following sentences: “The agreement measures used to compare different data sources also have known limitations. Agreement measures such as the Kappa statistic have been reported to be influenced by the prevalence of the diseases [29]. Previous studies on the validity of administrative data, have recommended the use of a minimum of four statistical measures to help mitigate these limitations [30, 31]. As a result, the prevalence, the raw frequency counts, sensitivity, specificity and the Kappa statistic were all reported.” (Pg. 32, line 8-15)

Comment 5: It may therefore be more meaningful to provide a Kappa values in Figure 2.

Response: The aim of Figure 2 is to illustrate differences by subcategories of each chronic disease group. Subcategories are defined using administrative data but patient-reported data is not at the subcategory level. Sensitivity provides an indication of which subcategories are relatively under-reported by patients (e.g. dementia). The impact of patients not reporting at the subcategory level is more problematic for specificity and Kappa because the ‘false-positive’ cell count will be inflated by including patients reporting any subcategory rather than the subcategory of interest. For this reason, we have avoided presenting measures (i.e. specificity, Kappa) that rely on the absence of a subcategory as defined using administrative data.

Comment 6: In table 3, please give confidence intervals for the Kappa values in Table 3.

Response: We have followed the reviewer’s suggestions and added 95% confidence intervals for the Kappa values and also for sensitivity and specificity in Tables 2 and 3.

Comment 7: The discussion would benefit from discussing the various sources of error or bias that affect patient-report and/or health records. This may help to better interpret the findings.

Response: We have followed the reviewer’s suggestions and clarified our discussion of the various sources of error or bias in the Discussion section and we have added the following
sentence: “This is likely to lead to an underestimation of the agreement between patient-reported chronic diseases and chronic diseases derived from administrative hospital data”. (Pg. 11, line 22-24)

Comment 8: The conclusion states "The study indicates that patients can give accurate information." It is appropriate to state that the information given by patients is mostly consistent with health records. Accuracy determined by comparing against health records again implies that health records are a gold-standard.

Response: We have followed the reviewer’s suggestion and changed the sentence to read as follows: “This study indicates that patients can give information about the presence of chronic diseases that is consistent with chronic diseases derived from hospital administrative data”.