Reviewer's report

Title: Comparing measures of comorbidity and functional status for risk adjustment to evaluate colorectal cancer surgery: a retrospective data-linkage study

Version: 3
Date: 3 December 2014
Reviewer: Niels Peek

Reviewer's report:

The manuscript of Dr Dobbins and co-workers addresses an interesting problem: the use of indices of co-morbidity and functional status, derived from routinely collected, administrative data, in risk adjustment. While the study, in my opinion, is of potential interest to the readership of BMC Medical Informatics and Decision Making, there are serious methodological issues that should be addressed before we can decide on publication.

1. The number of missing values for the ECOG and ASA scores is substantial causing a considerable risk of selection bias, especially since the ECOG scores are clearly not missing at random. Multiple imputation should be used to address this issue: a complete case analysis alone is not reliable.

2. Co-morbidities are likely underreported in administrative data, causing the Charlson index to be systematically underestimated in this study; perhaps the same holds for the ECOG and ASA scores. Part of the underreporting can often be repaired by considering not only coded diagnoses but also looking at interventions and medication. For instance, cardiovascular disorders can often be established by looking at medication. Please consider extending your analysis with this opportunity, or explain why you did not use it. At the very least, include comprehensive lists of codes that we used to derive the scores (linking codes to the constituent parts of the scores), and give a detailed, descriptive account of your data in terms of the constituting elements of all three scores.

3. As I understand from the manuscript, all three scores were coded in categories (e.g. "0", "1", and "2 or more" for Charlson) in the regression models, which is motivated by sparsity of higher scores. Please clarify whether indeed this approach was used. I don't think it is good idea, because high scores are often very informative, all of which is lost through truncation. It is not necessary either: logistic regression can perfectly cope with covariates with skewed distributions. So, I advise to code these score numerically. If the assumption of linearity is considered problematic, a secondary analysis could use categorized scores. At the very least, include descriptive statistics of the raw (numerical) scores in the manuscript.

4. Please give a more elaborate description of the peer groups (l. 108), such that this part of the analysis is more informative for readers outside NSW (or inside
NSW but not familiar with the peer group definitions).

5. Why adjust for peer groups in the analysis (l. 143)? Probably these groups represent, to some extent, unmeasured confounding factors at patient level. Correcting for these groups is not useful in this study, because it will dilute the influence of co-morbidity and functional status.

6. In addition to the rank correlations (l. 169) it would be valuable to see crosstables of the three scores. (Probably good to use truncated scores here!)

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.