Reviewer's report

Title: Variations and inter-relationship in outcome from emergency admissions in England. A retrospective analysis of Hospital Episode Statistics from 2005-2010

Version: 3  Date: 18 March 2014

Reviewer: Anthony Laverty

Reviewer's report:

I enjoyed reading this paper, and overall I think it is sound. I have a few points. Really these are all discretionary, but I do think they should improve the paper and make it more accessible to the general reader:

• I think it would be helpful in the abstract to mention that you have used the 20 most common reasons for admission, as the question most people will ask when you list various conditions is “how were they chosen”

• The intro gives the impression that the rationale for this study is in the context of how well a rating could be designed, but the abstract gives the impression it is more about how to improve care

• The methods seem sound to me, a few points here
  o I don’t think I have heard it called the RCS Charlson score (?) I think the general reader would certainly benefit form description of what Charlson represents and how it is derived
  o As it wasn’t explicit I was a little unclear about whether the analysis was at the hospital (site) level, or trust level. This should be clarified
  o I also was not clear how you made the division into medical vs surgical procedures, and I think there are a few ways you could do this. This should also be clarified

• I think language such as “High-risk diagnoses such as sepsis” should be avoided in the results (unless this is a defined category of high-risk – I don’t think it is)

• You refer in the discussion to your use of superspells, but you don’t say this in the methods. Again, I think you should just clarify this for readers not very familiar with HES

• I found the discussion of developing a rating interesting, but as with much discussion of ratings, there isn’t any mention of why we would do this. You say it could be supported as there is evidence of variation, but do we expect ratings to change this? Perhaps the question here is do we expect ratings to improve care through some vague notion of patient choice, or by simply pushing providers to improve.

  o There is discussion of the trade-off between a whole hospital rating and using more detailed data, but if the idea is to rely on patient choice, perhaps this would confuse patients more
• You say that using “a group of commonly encountered medical and surgical pathologies and ensured a plausible link between mortality and quality by including patients with conditions amenable to salvage” and also that this means that there is a relationship between process and outcome. I am not quite convinced about this – just because the conditions are common it does not mean that they are “amenable to salvage” I think it is a strength that you have included common conditions because they make up a large number of admissions, and this gives the study good power to detect differences, and in practical terms affect a lot of people. I think the discussion of process vs. outcome is more complex than you mention here, but you have used what is probably the best data around at the moment and used some important outcomes

• Also, I think a big strength is the linkage to out of hospital mortality, but you don’t seem to have mentioned this here (!)

• In table 1 I would remove the capital letters at the top of each column, and add SDs after those means and %. Also, for Charlson, the final category (I presume) is 3+ rather than just 3? I would also mark which of your deprivation categories is the most/least deprived

• I think tables 3 to 5 could be made clearer by replacing the +/- in favour of parentheses. Also, I think the general reader would find SDs more informative than SE

• Finally, I am not sure if these 3 very similar tables would be better if some of them were moved to the appendix. Perhaps one of your eFigures would be better in the main paper and the numbers behind them in the appendix. I appreciate that this may come down to space requirements etc however.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I have no competing interests to declare