Author's response to reviews

Title: Variations in hospital standardised mortality ratios (HSMR) as a result of frequent readmissions

Authors:

Wim F van den Bosch (w.bosch@antoniusziekenhuis.nl)
Peter Spreeuwenberg (p.spreeuwenberg@nivel.nl)
Cordula Wagner (c.wagner@nivel.nl)

Version: 2 Date: 14 October 2011

Author's response to reviews: see over
Reviewer's report

Title: How the hospital standardised mortality ratio (HSMR) may be biased as a result of frequent readmissions. A retrospective database study

Version: 1 Date: 17 August 2011

Reviewer: david ben-tovim

Reaction of authors

Dear Mr Ben-Tovim,

You have provided valuable comments on our manuscript. Please find below our point by point response to your comments which we have incorporated in the new version of the manuscript. We think the quality of our manuscript improved by this substantially.

Also on behalf of my two co-authors I like to thank you for your excellent comments.

Best regards,
Wim F. van den Bosch

Reviewer's report:

This is a follow-up article to a previously published article on the same topic, the previous article having been based on a limited sample of hospitals. The articles address an interesting and important topic, how to deal with the impact of previous hospital admissions in calculations of comparative hospital mortality, the most frequently used measure of which is the HSMR.

There is a subtle issue here. In a data set that uses linked information (ie in which it is possible to trace the admission history of individual patients over time and over multiple admissions) then it is clear that, with access to retrospective information, patients who were admitted on a number of occasions were not in fact equally liable to die on each occasion, because they survived a number of admissions, and only died once. How, and whether, to deal with that issue is a matter of some interest.

The risk adjustment process used to generate HSMRs commonly adjust for characteristics of patients present at the point of admission (eg age, sex, principle diagnosis) or during admission (length of stay) but if the patient level characteristic that leads to multiple admissions is not a function of those risk factors, then some mortality relevant influence may be at work in the multiple admission patient which is not captured in the usual HSMR risk adjustment process. Does this matter in relation to the utility of HSMRs?

HSMRs are relative mortality measures- they indicate the extent to which a hospital's mortality outcome varies in relation to a standard hospital, comprised of all patients treated in the total sample of hospitals of interest. They allow individual hospitals to compare their outcomes against this standard, and locate themselves within peer groups. They are best seen as screening tools to promote hospitals to ask hard questions about the care they provide- they are not of themselves diagnostic of good or poor quality care.

The influence of multiple re-admissions only becomes a source of concern in relation to the hard questions a hospital should ask if there are non-random
variations in the frequency of re-admissions across hospitals that relate to matters other than the care provided in hospitals. If hospital X has 5% of patients who are readmissions at any one time, and hospital y has 10%, but the mortality rate in that 5% in hospital X is sufficiently different to the mortality rate in hospital y (after standard risk adjustment for age, sex, diagnoses, co-morbidity and transfer status and length of stay) to cause a substantial difference in HSMR, then that is certainly a matter about which hard questions might be asked, and a measure that can narrow the search for areas of differences such as a patient view methodology presented here, is of substantial value. But if the frequency of readmission is influenced by easily identified administrative rather than clinical practice, then that is an issue to be dealt with by appropriate adjustment prior to presentation of HSMR values.

In my view, the first publication from this group (BMC health service research 2011, 11:57), failed to adequately discuss this issue, and I am concerned that this publication fails to deal with that issue either. A close examination of the data in the first publication from this group seems to me to indicate that the important variations in mortality attributed to differences between a patient view class between hospitals could mostly be attributed to the differences between hospitals in relation to providing chemotherapy on a day or in-patient basis. When chemotherapy is provided on a day basis, separations for chemotherapy become large in number, each separation is of low mortality risk. But the patients are from a diagnostic group that otherwise may be considered to be a high mortality risk population. On a patient view level, hospitals who use day chemotherapy may appear at lower mortality risk than hospitals who provide similar treatments for cancer patients as inpatients.

The current article only addresses inpatient admissions (plus deaths of day-cases which are very limited)

Variations in practice of this kind may produce statistical variability, but are of limited practical utility in relation to using HSMRs in a constructive manner. They can be dealt with by a number of administrative strategies-by defining day chemotherapy as an out-patient procedure and excluding it from HSMR calculation, by providing a specific exclusion based on a procedure code, etc. Screening tools need to be calibrated to be useful, and this is an obvious calibration issue, and should be discussed in those terms. I note that a substantial number of the CCS codes (12-44) in which comparisons between Model 1 and Model 2 generated changes to high SMRs (table 1) in this paper were cancer diagnoses-the influence of administrative issue in relation to cancer treatments should be looked at carefully before this article is published, and the issue discussed at greater length.

We have investigated to what extent both models differ in terms of statistical distance. As a result substantial differences may occur across most diagnostic groups (not just neoplasms), particularly chronic diseases. Some heart diseases, COPD and other show larger distances compared to neoplasms. We have added these data to table 1. Furthermore we have expanded the discussion to show that there are multiple reasons for this effect, not just differences in administering chemotherapy.
There is a more general point to be made- the concept of the patient view class is an interesting one, and provides an interesting perspective on mortality measurement. But using emotive terms such as ‘wrong’ or ‘right’ is inappropriate. There are no gold standards to use as the basis for value judgements in this area. HSMRs are empirical measures-does the number of previous admissions a patient has had to a hospital produce a systematic impact on mortality measurement that should provoke hard questions about the care provided within that hospital, and if so, what is the best way to measure and report that? Those are important issues, and deserve clear and objective study-the use of term such as better and worse and wrong and right in relation to small differences in c-statistics betrays a potential bias on behalf of the authors; a tendency to look for ‘gotcha’ results to invalidate the HSMR measure, rather than promoting an objective scientific discussion. Such language does not contribute to the debate about the validity and utility of HSMRs.

In addition to the c-statistics, we have enriched the manuscript with other quality metrics for the two models: calibration using Hosmer Lemeshow tests and explanatory power using Nagelkerke R square. We also tested the behaviour of the metrics for various lengths of the look-back-period (1, 2 and 5 years). In all cases the metrics of model 2 were in favour compared to model 1 and the sensitivity to adjustment increased going from 1 to 2 years and going from 2 to 5 years. Thank you for providing us with the idea that ‘better’ is not the appropriate term for model 2. We propose instead to use ‘more favourable’.

In our view the more fundamental issue is not so much the choice of model but the fact that two more or less comparable models may produce such divergent results. Consequently there may be (much) more uncertainty in presented (H)SMR-outcomes than suggested by the confidence intervals. This reasoning has been included in the new manuscript.

**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 received grant fund support to analyze hospital mortality data from Australian hospitals.
Reviewer’s report

Title: How the hospital standardised mortality ratio (HSMR) may be biased as a result of frequent readmissions. A retrospective database study

Version: 1  Date: 20 August 2011
Reviewer: Mette Nørgaard

Reviewer’s report:

Reaction of authors

Dear Mrs Nørgaard,

You have provided valuable and comprehensive comments on our manuscript. Please find below our point by point response to your comments which we have incorporated in the new version of the manuscript. We think the quality of our manuscript improved by this substantially.

Also on behalf of my two co-authors I like to thank you for your excellent comments.

Best regards,
Wim F. van den Bosch

The aim of the present paper is to assess a possible bias due to use of a less optimal prediction model when computing hospital standardised mortality ratio (HSMR). The HSMR has emerged as a potentially universal system-level indicator of quality in a hospital and improving the method is thus of interest. The authors have in a recent study published in BMC Health Services Research shown that the existing model in the Netherlands may favour hospitals with relatively many readmissions. This current paper is thus an extension of the former study in which the authors try to quantify this potential bias. However, the study has some limitations that should be addressed.

Major compulsory revisions:
1. The HSMR is based on indirect standardization in which each hospital is assessed on the patients that it has actually treated. You could therefore have two hospitals (A and B) with identical sizes and identical mortality rates which both have a considerably higher than the average mortality rate in one disease group and average mortality rates on the remaining disease groups. If hospital A treats a much higher number of patients within the disease group with higher than expected mortality than hospital B, hospital A will have a higher HSMR than hospital B even though they have the exact same mortality rates and thereby may have identical quality of care. This dilemma is not solved by a better case-mix adjustment and although improving the prediction model may give you a more accurate estimate of how a hospital performs compared with the national average it will not make comparisons between hospitals more correct. Therefore HSMR is most suitable for comparison over time within hospitals. This should be discussed.

Our article does not aim to discuss the usage (comparing) of HSMRs but rather to address an opportunity to improve as you state the accuracy of the model to reduce
potential uncertainty in case no adjustment is done for readmission. Achieving that
goal in our article is already a challenge and we prefer to focus on the subject and
not to burden the article with HSMR-applications for which already many other
articles provide. That potentially would confuse the reader.

2. The prediction model in HSMR is used to give an estimate of the expected
number of deaths in a given hospital had this hospital had mortality rates as the
national average. Prediction models are characterised by at least two aspects;
discrimination and calibration. C-statistics measures only discrimination. The
calibration is however of importance in estimating the expected number of
deaths. The authors should therefore include a test for the calibration of the
model (e.g. Hosmer lemeshow statistics).

In addition to the c-statistics, we have enriched the manuscript with other quality
metrics for the two models: calibration using Hosmer Lemeshow tests and
explanatory power using Nagelkerke R square. We also tested the behaviour of the
metrics for various lengths of the look-back-period (1, 2 and 5 years). In all cases the
metrics of model 2 were in favour compared to model 1 and the sensitivity to
adjustment increased going from 1 to 2 years and going from 2 to 5 years.

3. A more thorough discussion on the clinical interpretations of readmissions is
missing. It is correct that readmissions may lower HSMR because a patient can
only die once. However, readmissions may on one hand be a marker of poor
quality and on the other hand they may be a marker of high disease severity.
These aspects should be included.

We have enriched the discussion, addressing a variety of things that cause the effect.

4. The authors define readmissions as more than one admission within a 5 year
period regardless of diagnosis. However they do not provide an argumentation
for this choice of definition, Since HSMR is usually measured for a shorter period
than 5 year one could argue that a 5 year period is not relevant since admissions
5 admissions 4 year ago would not able to affect the HSMR calculation. Did the
authors do any sensitivity analyses to estimate how changing their definition
would change the results?

Please see our comments under 2.

5. Also, one could discuss whether admissions with different diagnoses should
be classified as admissions with the same diagnosis.

I’m not sure we understand the suggestion. There is probably no need to do that
because we added an extra correction factor independent of diagnoses.

6. According to the c-statistics the model without readmission is actually better to
discriminate between those who die and those who survive for patients with
cancer of the breast and prostate. How do the authors explain this finding?
There is no explanation for this. It is just how the model reacts to certain adaptations and sometimes it is going in the wrong direction. The most remarkable thing is that these two diagnoses are the only sex dependent diagnoses of the 50.

Minor essential revisions:
7. Is the description on how HSMR fluctuates in different countries needed?

We have investigated to what extent both models differ in terms of statistical distance. As a result substantial differences may occur across most diagnostic groups particularly chronic diseases. Neoplasms, some heart diseases, COPD and some others show substantial distances. We have added these data to table 1. We do not see a reason why this phenomenon would not occur in other countries.

8. Could the questions raised by the existing literature regarding reliability of the HSMR model be more thoroughly explained?

Please see our first comment regarding keeping focus in the article.

9. The figures could be left out.

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
Reviewer's report
Title: How the hospital standardised mortality ratio (HSMR) may be biased as a result of frequent readmissions. A retrospective database study
Version: 1 Date: 25 August 2011
Reviewer: Paul Aylin

Reaction of authors

Dear Mr Aylin,

You have provided valuable comments on our manuscript. Please find below our point by point response to your comments which we have incorporated in the new version of the manuscript. We think the quality of our manuscript improved by this substantially.

Also on behalf of my two co-authors I like to thank you for your excellent comments.

Best regards,
Wim F. van den Bosch

Reviewer's report:
On balance, as the authors already acknowledge, the addition of previous emergency admissions is already used in other countries' models (namely the UK), so I am not sure how original this work is, or even how internationally relevant, however it is certainly relevant to the Netherlands.

The impact on the current HSMR in the UK is not clear, since the UK-model does not adjust for non-emergency cases and because admissions with different primary diagnoses are not counted as readmissions. Furthermore the UK-model is restricted to a look-back-period of one year maximum.

This study attempts to answer the question "How the hospital standardised mortality ratio (HSMR) may be biased as a result of frequent readmissions." It is a retrospective analysis looking at 5 years of Dutch hospital data, comparing the current Dutch model for risk adjustment with a slightly more complex model which has the addition of one further variable based on the number of emergency admissions in patients over the previous 5 years. The overall c-statistic improved marginally from 0.852 to 0.867.

Please see comment above.

Is the question posed by the authors new and well defined?
I believe the general question is not really new, and has already been addressed by the group in the UK, as the paper already makes it clear that an adjustment for emergency admissions has been incorporated into the English HSMR risk model.

Please see comment above.
The findings are relevant to the Netherlands, but as the data are not yet available for all Dutch hospitals, it is not practical to include this variable in the casemix adjustment model just yet.

Being not practical does not mean the impact is not there. Since we demonstrated the effect for 80% of the Dutch hospitals, we think there is a substantial uncertainty in the HSMR-figure that even may exceed the uncertainty of the 95% confidence interval.

The authors may want to cite more recent papers that discuss the limitations of the HSMR and details some sensitivity analyses of the effect of modifying the casemix adjustment models, and also the attempt in the UK to garner a consensus on a hospital mortality measure.


We have included the second article in our list of references.

Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

The methods are well described, but would suggest that the c-statistic is only one way of measuring a model's 'accuracy', and that the inclusion of other measures such as Brier's score and the R-statistic might provide additional useful information on the models.

In addition to the c-statistics, we have enriched the manuscript with other quality metrics for the two models: calibration using Hosmer Lemeshow tests and explanatory power using Nagelkerke R square. We also tested the behaviour of the metrics for various lengths of the look-back-period (1, 2 and 5 years). In all cases the metrics of model 2 were in favour compared to model 1 and the sensitivity to adjustment increased going from 1 to 2 years and going from 2 to 5 years.

The authors also seem to think that having a higher c-statistic in one model invalidates the previous model. At one point the authors even announce that "current HSMR-model indicated the wrong label with respect to higher-than-expected mortality”. Unfortunately, the only way to find out whether an HSMR model wrongly indicates a hospital as having higher than expected mortality is to compare it against some gold standard, and unfortunately, no such gold standard exists at the moment.

Thank you for suggesting better terminology here. Indeed, ‘better’ is not the appropriate term for model 2. We propose instead to use ‘more favourable’.

Are the data sound and well controlled?
Yes they would seem to be.

Does the manuscript adhere to the relevant standards for reporting and data
deposition?
Yes I think so.
Are the discussion and conclusions well balanced and adequately supported by the data?
I think the discussion and conclusion could be phrased more usefully to offer an improvement to the current Dutch HSMR model, and be less geared towards suggesting that the current model is wrong. The idea of the HSMR is more of a screening tool, to identify hospitals with potential issues, which may or may not be due to quality of care. Just because a new screening tool comes along (albeit in this case with a very small improvement in discrimination), that does not invalidate the previous test, particularly if the new test is not practical at the moment using Dutch data.

Please see our various comments on these issues above.

The more fundamental issue here is not so much the choice of model but the fact that two more or less comparable models may produce such divergent results. Consequently there may be (much) more uncertainty in presented (H)SMR-outcomes than suggested by the confidence intervals. This reasoning has been included in the new manuscript.

The UK model already uses this variable, and where practical, other countries ought to incorporate it as well.
Do the title and abstract accurately convey what has been found?
Might change the title to include the words "how might the current Dutch HSMR model be improved", as the c-statistic is not a measure of bias.

We have changed the title as follows:
Variations in hospital standardised mortality ratios (HSMR) as a result of frequent readmissions.

Is the writing acceptable?
The English is perfect, but see above for other comments.

Declaration of competing interests:
I do have to declare an interest as a member of the UK group developing HSMRs.