Reviewer's report

Title: Improved risk adjustment for surveillance of healthcare-associated bloodstream infections: a retrospective cohort study

Version: 1 Date: 13 May 2009

Reviewer: Alistair W Stewart

Reviewer's report:

This paper is a muddle of some detailed statistical method and an absence of description of other statistical detail and some BSI results thrown in. This isn’t the right balance for a paper in BMC Infectious Diseases. The main outcome, the reallocation of risk adjustment is only briefly covered. The paper reads like an extract from a masters thesis rather than a paper suitable for publishing.

Major Compulsory Revisions

All comments below need to be considered in any revision.

I am concerned about the circular nature of this whole project. Using data from hospitals to create a rule to assess those same hospitals is circular. Even the training and validation process doesn’t solve this as the same hospitals are used.

To have made this paper of interest and value the authors should have told us which services were associated with what hospitals so that the reallocation of risk level can be better understood. There is no presentation of the risk reallocation outcome in the results and only brief mention in the discussion.

There is much detail on aspects of the modelling method and yet there are figures on Bayesian shrinkage plots and 2/3 sigma control limits with no explanation at all.

Even with the detailed modelling there is insufficient information about how the training dataset and the validation set were used. It needs to be clearly described that the infection data is available by hospital only and not by service within hospitals (as I initially assumed). It is described that to achieve parsimony non significant medical services were dropped from the model. Prediction is the goal of this modelling exercise and so all information should be used, parsimony is not a desirable feature in a model when prediction is the goal. The hospitals with risk scores of zero presumably only had services that were eliminated from the model yet these services could have summed to a non trivial score.

The authors report the association of the medical services to the outcome, really a different issue from that on which they embarked (a different modelling strategy). If association is a purpose of the modelling, I would want to see what the effects are for each of the services even if their effects are not shown in the model to be significantly different from zero.
The method for calculating the risk score is referenced to a paper on the development of the Framingham equations which is not at all like the data being used in this paper. The referenced paper uses quite complex modelling and so doesn’t help in understanding what has been done in this paper. A non-technical brief description of the risk score development is necessary, then, perhaps, followed by a suitable reference that supplies the detail that some readers may require.

This paper needs extensive overhaul. The authors need to be clear what the objectives of this paper are and ensure this clear to the readers. Then, the paper needs to be written to achieve those objectives bearing in mind their audience.

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

**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'