Reviewer's report

Title: Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups

Version: 2 Date: 14 December 2008

Reviewer: RICHARD PAUL MARSHALL

Reviewer's report:

This is a very useful and interesting paper.

The following comments all relate to potential Discretionary Revisions.

In the general discussion, it points out there is very little literature on quality measurement using indicators such as risk adjusted mortality. This, however, is partly because this indicator is particularly sensitive to factors beyond the control of the hospital. It can also become extremely politically sensitive and sometimes unfairly sensationalised by the media.

For example, because of their catchment populations, various hospitals can become the option of last resort. Such hospitals therefore default to a community assigned role of taking poorer, clinically more vulnerable, older or for other reasons higher risk cases even within particular DRGs. At the margins, higher mortality, even adjusted for statistical risk factors, in such hospitals is not necessarily “poor” performance. It may, however, measure poor performance by the referral or service planning system.

The use of Casemix Index (CMI) is a little fuzzy in the paper. CMI is really not an indicator of performance. It is certainly very relevant to this analysis because it can be a valid controlling factor when measuring efficiency. This is a point that could be made more explicitly in the paper.

On the other hand, importantly, CMI can certainly be an indicator of coding quality. Hospitals which are reporting a high level of detail in their statistical case summaries (in this study the FPMR) and coding it rigorously will tend to have higher CMI simply because of the detail they provide. In this regard a CMI of greater than 2 in even a high level tertiary hospital is quite exceptional and needs to be reviewed carefully. It could be that such a hospital is specialising in a very limited subset of cases and therefore cannot be validly compared in performance with other hospitals which are more open to take all comers. The authors acknowledge the early stage in data collection and the need to consolidate data standards to make the measures more valid and reliable over time.

The authors also acknowledge the issues with using charge data rather than actual cost data as a proxy for efficiency. This may not be such a problem because charge data does actually represent the cost to the payers (ultimately the patients and the community)! However, the point is well made that it is not a
direct measure of the hospital’s efficiency. Coding issues quality issues are also of importance in the accurate measurement of relative efficiency in the same way as outlined above.

The LOS comparison is well done and nicely presented as a proxy for efficiency as the control data on casemix comparability between the hospitals tightens up.

The method of comparing inpatient mortality in the low risk groups is very interesting and will benefit from the discussion that this article should stimulate. Among statisticians, the question should certainly be discussed as to whether the measured differences between the hospitals are truly statistically significant. Larger volumes of data may be needed to test for this. In very low incidence events such as mortality in low risk categories when larger volumes of data are available (eg 1 or 2 years, it may be useful to use a technique such as Poisson distribution analysis. For this reason, even though they are useful to demonstrate the idea of the approach, the findings on the mortality comparisons should probably be more carefully qualified.

As a very minor comment, the English language expression from the translation would benefit from a further edit for an international audience. Even though there are no real errors, the paper would be easier to read if the wording was simplified particularly in the introductory section.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.