Reviewer’s report

Title: Use of hierarchical models to evaluate performance of cardiac surgery centres in the Italian CABG outcome study

Version: 1 Date: 5 March 2007

Reviewer: David Spiegelhalter

Reviewer’s report:

General

This is a well-written description of an interesting study. My concerns are about the motivation of the methods and the interpretation of the results.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Page 3. The justification for the hierarchical model is here presented through allowing for ‘non-independent’ observations. This does not seem an adequate argument. The results within centres are only considered marginally independent if a pooled model is used without any allowance for a centre effect, and such a model is clearly of no interest when comparing performance. If a fixed effects model is used or the centre effect, as in the standard approach such as in NY State, then the observations are no longer marginally independent (of course under any model they are still conditionally independent given centre and risk factors). The crucial question is why a random-effects rather than a fixed effects model should be used: traditionally random-effects (mixed) models were chose when the specific units were not of interest, but that is not the situation here. I think the authors should refer to the literature that justifies these models as appropriate for estimating ‘long-term’ performance, and essentially allowing for ‘regression-to-the-mean’. See for example Burgess et al (2000) and Christiansen and Morris (1997).

Page 7. The interest seems to lie in assessing whether centres are above or below the mean. But the hierarchical model explicitly says the centres are different and distributed around the mean. Thus the hypotheses being tested are simply whether the centres are in the top or bottom half of this distribution, which does not seem of great interest. Certainly the test is not whether the centres are ‘extreme’, since any large centre will be confidently placed in the top of bottom half with enough data.

Page 8. When fitting centre-level covariates, it does not make sense to also have a factor for clinical centre. With such a factor it is inevitable that the covariates will show no additional predictive power.

Page 9. There needs to be some discussion of the actual results- what are the risk-adjusted mortality rates of the extremes, and so on.

Page 10. ‘standard single level model …. assumes independence of observations’ this is simply untrue (except of course in the sense of conditional independence given centre effects, which the hierarchical model also assumes); the NY model, for example, certainly does not assume marginal independence of observations since it fits a centre effect.


Christiansen, CL and Morris, CN (1997) Improving the statistical approach to health care provider profiling, Annals of Internal Medicine, 127, 764—768

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Figure 1. This plot gives no idea of what the actual estimated risk-adjusted mortality rates are and the vertical scale is essentially meaningless: the vertical scale could be relabelled to be much more informative.

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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