Author's response to reviews

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

Authors:

Paola D'Errigo (paola.derrigo@iss.it)
Maria E Tosti (mariaelena.tosti@iss.it)
Danilo Fusco (fusco@asplazio.it)
Carlo A Perucci (dir@asplazio.it)
Fulvia Seccareccia (fulvia.seccareccia@iss.it)

Version: 4 Date: 25 May 2007

Author's response to reviews: see over
Dear Referees,

We appreciate your comments. The following is a response to all of those comments. Also, your suggestions have been incorporated into the manuscript, as noted below.

Response to David Spiegelhalter comments:

P 3: Thank you for appreciating our efforts to better describe the reasons leading to the use of the multilevel approach. We agree that the presence of “dependent patients within hospitals” can not represent the only reason to justify the use of this methodology. To this extent, in the “Background” section of the previous revised version we sent, we widely described some issues concerning the “regression to mean” effect and the small samples problem. Now we tried to further clarify these topics also in the “Discussion” section by rephrasing some sentences and re-arranging the first paragraphs. Now the first part explains as follows:

“This study was conducted to build a hierarchical logistic model using data from the ‘Italian CABG outcome study’. The Italian CABG outcome study collected data from 2002 to 2004. Thirty-day overall mortality was 2.61%, comparable to the mortality observed by similar studies in other countries[5,6,29-31], but with a great variability among surgery centres (range 0.33 - 7.63). To investigate this heterogeneity, an empirical algorithm with a single-level multiple logistic regression procedure was used[7]. Standard single-level models, usually adopted in these kinds of studies, treat all patients as independent observations, when developing the risk-adjustment equation. Actually, patients undergoing a surgical intervention are not randomly allocated but nested in hospitals on the basis of reasons that lead them to make the same choices (place of residence, trust in a particular surgeon’s skill, hospital’s reputation, etc), thus violating one of the basic assumptions of traditional regression analysis . Hierarchical (or multilevel) models consider the hospitals involved in the study as a random sample from a population of hospitals and partition the random variability of the data into two parts: that between different patients and that between different hospitals. The hospital-specific random error component is interpreted as representing differences in hospital quality. Moreover, hierarchical models account for regression-to-the-mean by providing more accurate assessments of standardized mortality rates, giving more robust estimates to small sample sizes, with the most improvement for smaller centres[11-14]. Therefore, hierarchical modelling represents the most appropriate statistical method for dealing with outcomes data when individual patients are clustered within hospitals and in particular when there is a great heterogeneity in sample size[8-14]. In spite of these characteristics, the scientific literature still lacks outcome studies that have actually employed the hierarchical methodology[20, 23-27]. One possible cause could be represented by the high level of technology (software and hardware) required to implement the multilevel methodology[28], but the most convincing reason is the well-known risk of under/overestimating the importance of clinical centres in determining the outcome variability if the adjustment for confounding factors is not exhaustive.” (PAGE 10)
Page 11: We have modified the sentence you mentioned and it now explains as follows: “Standard single-level models, usually adopted in these kinds of studies, treat all patients as independent observations when developing the risk-adjustment equation.”

**Figure 1:** Concerning your comments about figure 1, we regret realizing only now we did not answer your question in the previous revised version. Now we try to explain the reasons leading to arrange figure 1 as it is. Differently from other authors and being inspired by the work of Merlo J et al. (J Epidemiol Community Health 2005), in the multilevel approach we identified the outliers using the group-level residuals instead of the RAMRs. Therefore, in this particular case, we preferred to graph the group-level residuals as more informative. Anyway, being the vertical scale relabeled, it is now more informative and the figure legend in this version is more clear and legible.

**Response to Patrick S Romano comments:**

Concerning your comment we have modified the sentence you mentioned and it now explains as follow:

“Differently from the classical approach, the hierarchical model identifies the outliers using the group-level residuals. Trying to simplify the concept, it may be stated that the group-level residual, for each centre, represents the distance between its estimated mortality and the overall estimated mortality.

In order to compare clinical centres’ performances with the overall mean, group-level residuals were ordered from the smallest to the largest and graphically presented together with their 95% confidence intervals (CI) [17,20-22].”

Moreover, we submitted our paper to a new English revision and some language corrections have been made.