Reviewer's report

Title: Computerized prediction of intensive care unit discharge after cardiac surgery: development and validation of a Gaussian processes model

Version: 1 Date: 21 April 2011

Reviewer: Thomas Higgins

Reviewer's report:

This study developed a predictive model for ICU discharge after cardiac surgery via a Gaussian processes model. Most prior work has used logistic regression. Limitations include a single center study, and a relatively small n of 461 in the development set. The authors acknowledge this limitation in potential generalizability (page 12). The ROC is moderate at 0.758 (realizing that ROC for LOS is generally lower than for mortality predictions) Median ICU LOS is fairly long by US standards, which raises the question whether patients might have left earlier. It would be helpful to report the standard deviation of LOS, or better yet display a histogram.

Major compulsory changes:

Add a histogram of ICU length of stay, at least by days if not hours

Page 5, discussion of missing numerical data: please cite the % of missing elements in the results section

I was unsure of what was meant by the statement on page 11 that "Validated applications to examine these databases are currently lacking".. Certainly kappa analysis can be employed, as well as other measures of intra-observer variability.

Page 11 - I am aware of at least two studies that have looked specifically at ICU length of stay in ICU patients. (Higgins TL et al, Predicting prolonged intensive care LOS following CABG. Clinical Intensive Care 1999; 10:175-182; Becker et al J Cardiovasc Surg 1995; 36:1; plus others referenced in the Kramer review). To my recollection, these studies also reported aROC metrics on performance, which may have been better than the EuroSCORE. Consider quoting these other sources. As you may know, APACHE has different equations for general ICU use versus customized equations for cardiac surgery, to accommodate the physiologic differences and differing independent variables of interest in the cardiac surgical population.

Minor Essential"

I would move the discussion of machine learning techniques from background to discussion.

Either a reference or statistical review should justify the statement of 0.25 as an acceptable cut-off for the Brier score.

On page 8, I did not see HIS (health information system) previously spelled out -
if this is the case, the abbreviation needs to be spelled out at first mention.
Student t-test should be capitalized (page 8)
The multiple statistics on page 10 would be easier to absorb in tabular form - could refer to the table here.

Discretionary:

Reference 1 is a review of other studies - a very good review, but consider adding two or three other primary references (see comments re: page 11, above). Reference 2 is also a review, and I would consider separately referencing the scores cited (EuroSCORE, STS, Parsonnet, etc.) rather than the review, especially since readers of this journal may be less familiar with the original work, and may need access to the detailed statistical descriptions in the original articles.

On Page 6, the paragraph on dynamic data is very long, and the last two sentences may be too much information.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I hold 100 shares of stock in Cerner Corporation, which licenses the APACHE system, which is referred to in the paper. I do not consider this a competing interest, but mention it for completeness.