Reviewer's report

Title: Stratification of the severity of critically ill patients with classification trees

Version: 2 Date: 28 May 2009

Reviewer: Michael Racz

Reviewer's report:

Major Compulsory Revisions

1) Please check and compare values in Table 6 in the revision with Table 5 in the original manuscript. In the development section of the original manuscript, the SMRs for internally developed models (a logistic regression model and 3 classification trees) are all 1. This is as one would expect. In Table 6 of the revision, the SMRs for logistic regression and C4.5 are 1.22 and 0.99, respectively. The ratio of observed over expected should be one when applying the model to the entire data set on which it was developed so I would think the original is correct.

2) I still have concerns about the way the severity scores are presented in Table 6. When you look across the development to validation sections for CART, for example, the SMRs went from 1.00 to 1.04 with the validation SMR not significantly different than one. Part of the reason the validation SMR is greater than one is that the mortality rate for the validation set is greater than that of the development set. For Apache II, the SMRs went from 1.30 to 1.36. This only had a small increase but was significantly different than 1 to begin with in the development set. If the severity scores were recalibrated, this comparison of development to validation would make sense. If recalibration can't be done, please make some indication in Table 6 that the severity scores are not developed in the development phase. Maybe a comment in the Comparison of model properties section is also appropriate.

3) If Table 5 in the original manuscript is correct, as referred to in comment 1, the SMRs for logistic regression went from 1.00 to 1.29. Algebraically this is hard to believe. The model was developed on a set of 1,880 cases with a mortality rate of 30.7%. Using this model on the 1,880, therefore, the expected mortality rate is 30.7%. If the SMR when the model is applied to the validation set is 1.29 and the mortality rate for this set is 32.8%, then the expected mortality rate must be 25.4%. Given the balance indicated in Table 2 between the characteristics of the development and validation sets, I find it hard to believe that the model predicted 30.7% on one set and only 25.4% on the other. If this is true, can you comment on what is driving this lower expected mortality in the validation set?

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