Reviewer's report

Title: A predictive model for the early identification of patients at risk for a prolonged intensive care unit length of stay

Version: 2 Date: 19 February 2010

Reviewer: Elisa Estenssoro

Reviewer's report:

General Comments
The authors have tried to develop a linear regression model to identify patients at risk for prolonged ICU stay (ICU LOS). The number of patients used for the construction of the model is impressive, and the methods used (construction, internal validation, external validation) are, naturally, correct--the authors have a great experience on the subject. They found that use of MV, or sedation, and PaO2/FIO2 relationship, physiological variables (not stated) and other variables at day 5 very accurately predict LOS.

Somehow this is not unexpected. At day 5, severely ill patients that promptly die, and so have short stay, are no longer included. These patients generate “noise”. In this way, it is interesting how some coefficients change the sign. For example APS is negative on day 1, but positive on day 5.

The authors should be credited for their constant efforts to shed some light in this relevant but very complex subject.

Major compulsory revisions:
First, there are some factors that have been consistently associated with a prolonged LOS in ICU that are not included in the model. I understand that this is a limitation of working with a database. One of these factors is addressed by the authors: ICU-acquired infections, especially ventilation-associated pneumonia, which has been consistently related to a prolonged ICU stay. This is a problem of registration of the database and the authors mention it. The other very important risk factor that has been left aside is ICU-acquired paralysis (critical-illness polyneuropathy and myopathy). Is there any chance to have data on this very important risk factor for prolonged ICU LOS? Other risk factors that are associated to prolonged mechanical ventilation, which will subsequently augment ICU LOS are: interruption of sedation-analgesia practices (I guess that it is impossible to have data on this), and shock on admission. At least these should be mentioned, if not analyzed. The thing is-and I am impressed by it- that the model still works well WITHOUT all these risk factors!

Second, and most important- I don’t agree with cutting ICU stay at 30 days. That 1% of patients still represents thousands of them in the database. Do they add uncertainty to the model? What if you include them? Is their ICU LOS so unpredictable? In many cases this 1% might be greater. For example, in Europe
and Latin America LTACS are not widely available, so these extreme outliers must stay at the ICU until death or discharge. You include these patients in Figure 2, where is clearly seen that they use a disproportionate percentage of bed days. To have a model able to predict these outliers (if there is one< maybe not!) would of great value for clinicians.

I agree with the authors’ statement “…the infrequent availability of advanced health information technology …represents major barrier to the model’s widespread use”. However, had models like the authors’ been developed with scores universally disseminated like APACHE II, SAPS 2 and 3, this would contribute a lot to stimulate the use of these tools.

The final model: day 5 PaO2FIO2 and the inability to assess GCS because of sedation- paralysis are clear confounders of mechanical ventilation. The model might work with these factors, but I wonder if there is any clinical sense to have all together. What are “other physiological variables”? Apparently, to be mechanically ventilated at day 5 is the great predictor.

I would like to see the Kaplan-Meier survival curve of both groups.

The discussion should be shortened.

Table 5. Mechanical ventilation days and ICU LOS are clearly skewed, as the authors say: so why aren’t these presented as median and IQRs?

Figure 3. Chronic heart failure, which is the most frequent risk category for regular ICU stay, and second for prolonged ICU stay should be included.

Minor essential revisions
Sometimes APACHE score is referred as “APACHE IV” and sometimes just “APACHE”. I understand it should be always APACHE IV?

Table 1 What means rescaled Glasgow value?
Table 3 What means “visit”?

There are many typing errors, especially in the tables.

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 have no competing interests in relation to this study.
Elisa Estenssoro, MD