Reviewer's report

Title: Impact of Introducing Multiple Evidence-Based Clinical Practice Protocols in a Medical Intensive Care unit: A Retrospective Cohort Study

Version: 1 Date: 22 May 2007

Reviewer: Maurizia Capuzzo

Reviewer's report:

The aim of this single-center study was to assess the impact of introducing four evidence-based protocols on the outcome of ICU admitted patients. The complete implementation of the protocols took 25 months and the comparison was performed between the Standardized Mortality Ratio (SMR) of the patients admitted to ICU in the 24 months before and those admitted in the 14 months after it. SMR, computed according to the APACHE III predictions, was significantly lower in the period after that in that before the introduction of the protocols. The Authors conclude that the protocols improved outcome.

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

I have some concerns about the general framework of this kind of studies. From a methodologic point of view, as a researcher, I am interested in knowing the impact of each intervention on outcome. I appreciate that only a large sample allows identifying significant but small differences as those we are talking about.

To detect a small difference, we need a sample collected by many ICUs in a short period of time or, alternatively, by one ICU over a long period of time. In the latter case, as in the present study, some changes in organization and practices over the period studied cannot be avoided. In fact, the Authors report that the bed capacity of the ICU increased from 15 to 19 in August 2002, and from 19 to 24 in December 2002 (page 5). At the same time, some change in the intensivist-to-bed ratio were performed as reported in the study published in Chest (correctly referenced as Ref. n. 26). Moreover, when data are collected over a long period of time for comparison with historical controls, we cannot avoid the influence of the continuous (unavoidable and frequently imperceptible) changes in practice, which represent a confounding variable. As a consequence, those changes could represent a relevant noise making difficult to identify the sound (the effect of a single intervention).

The most recent methodological evolution (the so called French “escamotage”) to overcome the problem seems to be the evaluation of the effect of a group of interventions: the underlying hypothesis is that, if each intervention has a small effect, the sum of their effects will show statistically relevant differences. As a result, we cannot see what is truly relevant for our patients’ outcome.

To identify the effect of each intervention, I suggest to the author a different analysis: the Variable Life-Adjusted Display (VLAD), as proposed by Lovergrove J et al (Lancet. 1997; 350(9085): 1128-30) to assess the performance of a single surgeon. This method measures the performance as number of lives saved and it has been used by Hans Flaatten (Qual Saf Health Care 2005; 14: 270-272) to show the decrease of ICU performance due to the decrease in the number of beds available. Of course, the authors could give the measure as the number of lives saved per ICU bed, solving the problem of the changes in the number of ICU beds over time.

The VLAD curve resulting from the cumulative survival will rise when performance grows and will fall when the performance deteriorates. The different slopes of the various part of the curve could suggest the efficacy of the different interventions implemented.

The authors report that they used the 28-day ICU free days to avoid the confounding effect of mortality (page 8, second paragraph). Such a complicated computation is not really useful for the reader, who usually knows the ICU length of stay (LOS) of his patients. I strongly suggest to use the ICU LOS.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Results, page 9 second paragraph. The text emphasizes that APS, APACHE III and predicted mortality were higher in the protocol than in the pre-protocol period, but in Table 1, predicted mortality is not reported as significantly different (and it appears roughly similar). Check the statistical significance.

Discussion, page 11, first paragraph, sentence before the last. As far as Activated Protein C is concerned, there are recent studies suggesting cautiousness (Intensive Care Med. 2007 Mar;33(3):435-43 and Intensive Care Med. 2007 Mar;33(3):426-34). This point should at least mentioned.

Discussion, page 12, second paragraph. The possible Hawthorne effect is not mentioned among limitations.

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 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:

NO