Chapter 7

Conclusions
In this study we evaluated the application of the ratio between observed and predicted mortality - standardised mortality ratio (SMR) - in the evaluation of the performance or relative effectiveness of care of the ICU.

According to this methodology we tested if the outcome prediction models used were able to adequately describe the population under analysis. This step is seen today as crucial. A recent Consensus Conference [1] stated that “Mortality prediction models are almost always overspecific for the patient samples upon which they were developed, and thus performance usually deteriorates when models are applied to different population samples…. For this reason, we recommend that mortality prediction models always be tested in patient samples distinct from those in which the models were developed”.

Our results demonstrate that none of the utilised models was able to accurately predict mortality in the two analysed populations. This was true for SAPS II and APACHE II in the Portuguese cohort and for SAPS II and MPM II in the EURICUS-I cohort. But what are the implications of these results for general ICUs?

One important implication at the ICU level is that a general outcome prediction model cannot be applied in a new cohort of patients without previous validation. The accuracy of the model should always be empirically tested before utilisation on a new patient population. Otherwise, differences between observed and expected mortality may be attributed to differences in ICU performance while they should be explained by the inadequacy of the utilised model.

Our results document also the inappropriateness of standardised mortality ratios for the evaluation of performance in intensive care, since most of the differences between predicted and observed mortality can be assigned to differences in performance of the prediction models across relevant subgroups. We tested formally this hypothesis and concluded that, even with better calibrated models, different ICU compositions in terms of patient baseline-characteristics have an important impact on the relation between predicted and observed mortality. Therefore, the use of prediction models for the evaluation of performance of the ICU should be restricted to experimental settings.

Does the presented above call for a moratorium in the use of general outcome models for performance evaluation? Can the problems be solved?

Efforts should be done to overcome the problems. In the classical development and application of general outcome prediction models, we must assume the independence of the observations (patient related data) and the non-variance of ICU related data. Otherwise, the positive correlation introduced by local factors (clinical and non-clinical) in the patients inside the same ICU will affect the stability of the predictive equations. New statistical techniques dealing with these questions have been proposed [2,3] and we may anticipate their increasing use in the future. Also, a recent trend can be observed, with models focusing more on the prediction of survival over time (survival curves) rather than responding to the dualistic question of survival or death at hospital discharge [4].

Maybe it is time to return to basic principles and redefine the main questions. Current
research in performance has focused only on input/output ratios: how many patients were admitted with a certain degree of physiologic derangement versus how many patients died. But is this performance? Is this quality of care?

Apart from the methodological questions involved in this process (e.g. quality of data, definitions of variables and procedures, verification of the assumptions), some other questions must be answered first. As Rosemary Butler (Department of Health, U.K.) pointed out recently while discussing a paper by Goldstein and Spiegelhalter [3]:

what aspects of performance are covered by the indicator;
what are the objectives of performance in the context of the indicator;
what standards of service are required to achieve the objectives;
how useful is the indicator likely to be;
what potential is there for follow-up action leading to a change.

For the present, a possible alternative is to focus more on patient stratification and description than on prediction, in other words to replace mortality with morbidity. This approach has been advocated in the last years [5] and efforts have been made by several groups to develop adequate instruments [6-8]. With this approach, we will lose our predictive capability but we will gain new views on the description and evolution of the diseases.

Is this a step backwards? Certainly not. Only with this information we will be able to describe and intervene in the process of care, in order to improve the outcome of our patients in terms of health and well-being. An this is what the Society expects from us.

“…accurate hospital statistics are much more rare than is generally imagined and at their best they only give the mortality which has taken place in the hospitals and take no cognisance of those cases which are discharged in a hopeless condition, to die immediately afterwards, a practice which is followed to a much greater extent by some hospitals than others…

Careful observers are now generally convinced that the origin and spread of fever in a hospital or the appearance and spread of hospital gangrene, erysipelas and pyaemia generally are much better tests of the sanitary state of a hospital than is mortality returns”

Florence Nightingale
Notes on Hospitals
1863
REFERENCES


