Trends in Anaesthesia and Critical Care 4 (2014) 170–174

Contents lists available at ScienceDirect

Trends in Anaesthesia and Critical Care

journal homepage: www.elsevier.com/locate/tacc

REVIEW

Outcome prediction models in end-of-life decison making

G. Villa*, P. Di Maggio, M. Baccelli, S. Romagnoli, A.R. De Gaudio

University of Florence, Department of Human Health Science, Section of Anesthesia and Intensive Care, Florence, Italy

Keywords:
End-of-life
Scoring system
Decision making
Score discrimination
Score calibration

SUMMARY

In the modern intensive care units, advanced life support systems may unreasonably prolong the dying process in end-of-life patients. Outcome prediction models may assist physicians to identify these patients, allowing the integration of palliative care in intensive care treatments. General and disease-specific models show the necessary discrimination and calibration to be applied in the daily medical practice for clinical and research purposes. However, clinical limitations and other general limitations, such as those related to the user, to the patient or to the model used, reduce their prospective applicability. The actual reliability of the estimates produced by these probabilistic models is one of the main limitations. Despite their potential role in recognizing end-of-life patients, none of the current outcome prediction models is routinely applied for supporting the clinicians' decision making process in critically-ill patients.

1. Introduction

In modern intensive care medicine, critically-ill patients are often subjected to advanced and invasive diagnostic and therapeutic interventions. In the effort to keep patients alive, the aggressiveness of the procedures performed is mostly related to the severity of the illness.[1] However, this approach is not always successful and mortality in the Intensive Care Unit (ICU) remains very high.[2] In these settings, advances in pharmacology, biotechnology and life support systems may only unreasonably postpone the time of death. When patients are maintained alive via administration of specific therapy to counteract critical illness, related adverse effects of treatment may involve patients and their family. The paradigm of life-sustaining or death-prolonging therapy has led to active debates[3] focusing on the concern that these invasive and expensive cares may be futile, ethically questionable and not beneficial for the patient. Therefore, nowadays, improving the quality of care received by dying patients still remains an ongoing challenge for every ICU team member.

The inappropriate use of ICU resources in end-of-life patients is mainly identified retrospectively, and one of the most challenging issues for ICU physicians is to prospectively identify patients who will not survive despite ICU care, and for whom palliative care associated with intensive care treatment could be an acceptable treatment option.[4–6]

Although a number of ICU outcome prediction models have been identified in the published literature, none of them have actually provided physicians with enough information on the suitability of intensive care treatments for individual patients. The ideal model should be a well-validated, calibrated and discriminated prediction tool, which would generate timely data and support clinical decisions, creating a greater awareness of the patient prognosis and the potential benefit of palliative care. However, development of a clinically applicable and scientifically accurate critical care prediction tool is a hard task, especially when considering problems concerning complex data collection and inconsistent analysis methods.

2. Characteristics of outcome prediction models

The entry data of outcome prediction models should be objectively measured, constantly collected and recorded with a standardized format, free of bias. Amongst them, *pre-treatment* variables such as age, underlying chronic health issues, ICU admission diagnosis and severity of illness, as well as *post-treatment* variables such as type and timing of therapies and subsequent response should be included.

Considering the different variables taken into account and the type of population for which the score system is designed, generic and disease-specific models can be identified.

Generic models predict the ICU outcome in a heterogeneous population of patients treated in a particular setting. Specifically

* Corresponding author. Department of Human Health Science, Section of Anaesthesia and Intensive Care, University of Florence, Italy. Largo Brambilla 3, Florence, Italy. Tel.: +39 3207615547.
E-mail address: gianlucavilla1@gmail.com (G. Villa).

http://dx.doi.org/10.1016/j.tacc.2014.10.003
2210-8440/

designed for adult ICU patients, they were introduced into clinical practice in the mid-1980s.[7,8] These generic models, which are usually developed from a suitable database or selected from existing databases, are applied to all patients in the ICU during a certain period.[7] Examples of generic models designed to evaluate in-hospital mortality are the first version of the Acute Physiology and Chronic Health Evaluation (APACHE I)[9] and the Mortality Probability Model (MPM)[10–16] for adult ICU patients, and the Pediatric Risk of Mortality Score (PRISM) for pediatric ICU patients[17,18] (Table 1).

Disease-specific models predict ICU outcome in more homogeneous groups of patients who are categorized by clinical syndrome or by primary diagnosis. Glasgow Coma Scale (GCS) for stroke[19] and Pneumonia Severity Index (PSI) or Ranson's criteria[19,20] for acute pancreatitis are examples of disease-specific models. The second and the third generation of APACHE score (APACHE II and ACAPHE III)[8,21] could also be considered as disease-specific models according to the different weights associated with the independent variables related to the specific disease (Table 1). However, many critical care providers consider them as generic models assuming that only the weight of the independent variables change between diseases, not the particular data.[19] It is generally believed that, for a specific population of patients, the generic models are less accurate than the disease-specific models in which the input variables are intentionally related to the specific disease (e.g., occurrence of cardiogenic shock in patients with acute myocardial infarction or GCS in stroke patients).

3. Current role and potential value of outcome prediction models

In critically-ill patients, decisions on withholding or withdrawal of life-sustaining treatments[22–24] or do-not-resuscitate procedures[25] are daily applied, with or without the support of suitable outcome prediction models. However, these models could potentially guide clinicians towards a more appropriate use of ICU resources basing the clinical judgment on the patient's likelihood of benefiting from therapy. In fact, in many critical ICU cases, the burdens of some treatments may outweigh the benefits or treatments may not achieve the desired, anticipated effects. Therefore, by orientating physicians towards the withdrawal or withholding of unnecessary treatments, these tools may have the potential to reduce the burdens of stress and suffering in end-of-life patients and family members.[19]

The outcome predicted by the model should be clinically relevant and clearly defined in terms of timing and methods of ascertainment.[26]

Several authors have retrospectively explored the possibility to extend the role of these outcome prediction models from primarily administrative systems to reliable tools, available in the daily clinical practice to help bedside decision making in critically-ill patients.[27–31]

On the contrary, very few prospective studies have been aimed at directly evaluating the relationship between outcome prediction models and the limitation of unnecessary cares.[19] In a prospective study conducted in 17 French ICUs on heterogeneous patients, Knaus et al.[32] investigated the role of predicted survival information, obtained by the multiple organ system failure (OSF), on clinical decision making. Although a small but significant increase in therapy, withholding was observed, this was limited to those patients with very low survival chances (with three or more OSFs), and a very poor impact on clinical decision emerged from the application of these models. However, although there was no evidence in this study suggesting that outcome prediction models may actually modulate clinical judgment in the everyday practice, the explicit provision of prognostic data may lead to a sense of therapeutic futility.[32]

In a study conducted on 1025 patients in four British neurosurgical units, Murray et al.[33] reported that clinicians could reduce the intensity of treatment in severely ill, head-injured patients on the basis of the objective estimation of the prognosis derived by outcome prediction models. In particular, via the introduction of a computer-based outcome prediction model, the authors observed a 39% reduction in the use of specific intensive care treatments in patients who were prognosticated to have the worst outcome. However, during the study, no evidence suggested that the prediction model affected the overall outcome, length of stay, or the recording of explicit decisions to limit unnecessary treatment.[33–35]

Finally, in the Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment (SUPPORT), no change in clinical behaviours (e.g. ICU length of stay, length of mechanical ventilation, length of coma before death or used ICU resources) was observed amongst patients randomized to receive their regularly updated predicted mortality statistics as outlined in their medical chart.[36]

Measures to assess model predictive performance include overall, discrimination and calibration measures.[26]

4. Statistical consideration

An outcome model is generally fitted on a “derivation set” and validated on a “validation set”. In particular, in order to measure the model performance, the predicted probability of death obtained in

Table 1
The main generic and disease-specific outcome prediction models used in the intensive care unit.[15,16]

	Number of variables	Time of assessment	Predicted outcome	Discrimination (ROC-AUC)	Calibration (Hosmer–Lemeshow C statistic)
APACHE-I[63,64]	34	First 32 h after admission	ICU mortality	NA	NA
APACHE-II[63,64]	12	First 24 h after admission	Hospital mortality	0.85	209.20, $p < .01$
APACHE-III[63–65]	17	First 24 h after admission	Hospital mortality	0.90	48.71, $p < .01$
APACHE-IV[63,64,66]	21	First 24 h after admission	Hospital mortality	0.88	16.9, $p = .08$
SAPS 1[67]	14	First 24 h after admission	ICU mortality	NA	NA
SAPS 2[67]	17	First 24 h after admission	Hospital mortality	0.86	219.83, $p < .01$
SAPS-3[67,68]	20	Prior to and within 1 h of ICU admission	Hospital mortality	0.84	NA
MPM_0-I[10]	7	Prior to and within 1 h of ICU admission	Hospital mortality	NA	NA
MPM_0-II[10]	15	Prior to and within 1 h of ICU admission	Hospital mortality	0.837	47.61, $p < .01$
MPM_0-III[10,11]	16	Prior to and within 1 h of ICU admission	Hospital mortality	0.823	NA
Ranson's Criteria[69]	11	First 48 h after admission	Hospital mortality	NA	NA
PRISM[70]	14	First 24 h after admission	Hospital mortality	0.851	1.746, $p = .627$
PIM[70]	8	First 24 h after admission	Hospital mortality	0.838	10.866, $p = .028$

Abbreviations: APACHE, acute physiology and chronic health evaluation; SAPS, simplified acute physiology score; MPM, mortality prediction model; PRISM, paediatric risk of mortality; PIM, Pediatric Index of Mortality; NA, not available.