DOI: http://dx.doi.org/10.21088/ijsn.2277.467X.7118.4

Severity Scoring Systems in Intensive Care: A Clinical Review

IJSN
Volume 7, Number 1
© Red Flower Publication Pvt. Ltd

Neethu Jose

Abstract

A large number of severity of illness scoring systems have been developed and they are widely used in intensive care practice. However, they are complex systems with their basis in mathematics. To use such systems effectively, it is important to appreciate what factors influence their performance so that they can be compared fairly and used most appropriately. The purpose of this review is to describe the methods commonly used to assess the various facets of performance in severity of illness scoring systems. The performance of the most frequently used scoring systems in adult intensive care practice are presented.

Keywords: Intensive Care; ICU Scoring; APACHE; Prognosis; Risk Assessment; Severity of Illness Index.

Introduction

The use of scoring systems to predict risk of mortality and evaluating outcome in critically ill patients is important in modern medicine. The first such system in widespread use was the APGAR score introduced in 1953 to asses the vitality of the newborn. The Glasgow Coma Scale (GCS) and Ranson score are other examples of systems that have gained widespread use. Within intensive care, a large number of scoring systems aimed either at the general intensive care unit (ICU) patient or defined subgroups have been developed during the last two decades. Prognostic or general severity scoring systems such as the Acute Physiology and Chronic Health Evaluation (APACHE) and Simplified Acute Physiology Score (SAPS) estimate risk based on data available within the first 24h of ICU stay. The standard mortality ratio (SMR), a key element in ICU benchmarking, can be calculated using these systems. Disease-specific scoring systems have been developed for several important subgroups treated in the ICU, such as pancreatitis, hepatic failure and adult respiratory distress syndrome. Because the ICU treats patients with one or more organ dysfunction (OD), several organ failure scoring systems have also been developed in the last 10 years. Scoring systems are also important in clinical trials and in the monitoring of quality-of-care.

Author Affiliation: Assistant Professor, Jubilee Mission College of Nursing, Thrissur, Kerala 680005, India and Assistant Nursing Superintendent cum Nurse Educator, Jubilee Mission Medical College & Research Centre, Thrissur, Kerala 680005, India.

Corresponding Author: Neethu Jose, Assistant Professor, Jubilee Mission College of Nursing, Thrissur, Kerala 680005, India and Assistant Nursing Superintendent cum Nurse Educator, Jubilee Mission Medical College & Research Centre, Thrissur, Kerala 680005, India.
E-mail: agnesfeb9@gmail.com

Received on 12.01.2018, **Accepted on** 09.02.2018

Classification of scoring systems [3]

There is no agreed classification of the scoring systems that are used in critically ill patients. Scores can be applied either to a single set of data or repeated over time. The available methods include;

1. *Anatomical scoring*. These depend on the anatomical area involved. Anatomical scoring systems are mainly used for trauma patients [e.g. abbreviated injury score (AIS) and injury severity score (ISS)].
2. *Therapeutic weighted scores*. These are based on the assumption that very ill patients require a greater number of interventions and procedures that are more complex than patients who are less ill. Examples include the therapeutic intervention scoring system (TISS).

3. *Organ-specific scoring*. This is similar to therapeutic scoring; the underlying premise is the sicker a patient the more organ systems will be involved, ranging from organ dysfunction to failure [e.g. sepsis-related organ failure assessment (SOFA)].
4. *Physiological assessment*. It is based on the degree of derangement of routinely measured physiological variables [e.g. acute physiology and chronic health evaluation (APACHE) and simplified acute physiology score (SAPS)].
5. *Simple scales.* It is based on clinical judgement (e.g. survive or die).
6. *Disease specific* [e.g. Ranson's criteria for acute pancreatitis, subarachnoid haemorrhage assessment using the World Federation of Neurosurgeons score, and liver failure assessment using Child-Pugh or model for end-stage liver disease (MELD) scoring].

Again ICU scoring systems can be divided into four major groups-

1. General risk-prognostication scores(severity of illness scores).
2. Disease-speciûc riskprognostication scores,
3. Trauma scoring and
4. Organ dysfunction (failure) scoring.

1. General risk-prognostication systems [1,2]

The basis for development of both the APACHE system and the SAPS in 1982 was the assumption that the severity of acute disease could be measured by quantifying the degree of abnormality of physiologic variables. These ûrst versions were soon replaced by more sophisticated models using prospective sampled patient data and advanced logistic regression analysis.

a. APACHE II (Acute Physiology And Chronic Health Evaluation)

The APACHE II model, published in 1985, was developed due to the complexity of the original model and it has become the most frequently used general mortality prediction model (MPM). The original number of physiologic variables was reduced from 34 to 12 and some were re-weighted. Patients under the age of 16 were not included.

In addition to the acute physiology variables, age, operative status and the presence of severe chronic Organ Dysfunction or immune suppression were incorporated. The final APPACHE II score is the sum of the acute physiology, age and chronic health points, calculated from the worst values during the first 24 hours of intensive care.

b. APACHE III

APACHE III was developed as a further reûnement of APACHE II. This comprises the three subscores age (0–24 points), acute physiology (0–252 points) and chronic health evaluation (0–23 points). A second objective of the developers was to reûne mortality prediction by correcting for risk in individually deûned patient groups. This could also be used to improve ICU discharge decisions. Finally, further development of the APPACHEIII score provides equations to estimate the length of stay in ICU,the amount and type of therapy required and the intensity of nursing care.

c. APACHE IV

Published in 2006, the APACHE IV system is made up of the acute physiology score (APS), age and admission circumstances, totalling 142 variables of which 115 are admission diagnoses. In contrast to SAPS III, the APS was found to be the most important factor, followed by disease group and age. As in earlier APACHE models, the APS was based on the most abnormal values registered during the ûrst 24h after ICU admission. APACHE IV also includes a separate scoring system for coronary bypass patients.

d. MPM II (Mortality Probability Model)

It was published in 1985, the MPM was the first general severity model to assess risk of death at ICU admission. Prediction models for assessment at admission and 24h were developed originally but models for assessment at 48 and 72h. MPM_0 includes a total of 15 variables collected at ICU admission; MPM_{24} consists of eight variables collected at 24h, as well as ûve variables obtained from the MPM_0. As the models consist mainly of dichotomous variables, scoring is very simple. The strength of the MPM II models lies in their simplicity of scoring and the possibility of sequential assessment of mortality risk throughout the ICU stay.

e. SAPS II (Simplified Acute Physiology Score)

It was developed and validated in France in 1984, used 13 weighted physiological variables and age to predict risk of death in ICU patients. This was published in 1994 . Like the APACHE scores, SAPS was calculated from the worst values obtained during the first 24 hours of ICU admission. The developers focused on maintaining a scoring system based mainly on physiological variables. Twelve physiologic variables were included in addition to age, admission type and the presence of metastatic or haematological cancer or AIDS.

f. SAPS III

The SAPS III Outcomes Research Group published their new scoring system in 2005. It was realized that a mainly physiology-based scoring system (SAPS II) had serious shortcomings facing case-mix and lead-time bias. Three subscores, namely patient characteristics before admission (five variables), circumstances of admission (five variables) and acute physiology (10 variables) are summed up to produce the SAPS III score. The patients' worst physiologic parameters at ICU admission (1h) are recorded. Probability of mortality is calculated using the total SAPS III score in a general or customized equation based on the location of the hospital.

2. Disease- and organ-specic prognostic scores

Scores to quantify single-organ failure or a specific disease are often used outside the ICU and knowledge of these scores may be valuable when communicating within the ICU. They have seldom been developed using large prospectively collected data and logistic regression analysis. Their use is often not validated for ICU patients with concomitant organ failure, but they continue to be used to guide treatment and prognostication.

a. GCS (Glassgow Coma Scale)

The GCS was developed as a method for assessing depth and duration of impaired consciousness and is one of the most widespread clinical scores in medicine. Motor response, verbal response and response to pain are noted, producing a total score from 3 to 15. A score of 14–15 indicates mild injury, 9–13 moderate injury and 3–8 severe injury. Its strengths lie in the ease of calculation and reproducibility. The GCS has become a standard method of assessing unconsciousness and coma, but its use outside the setting of trauma and traumatic brain injury is problematic.

Its use is not encouraged in patients with other reasons for unconsciousness such as intoxication and epileptic activity. It has no place in assessing the depth of sedation in the ICU. The importance of the GCS in the ICU, with the exception of neuro-intensive care, is probably its inclusion in more complex scoring systems.

b. Ranson score

The Ranson score was originally developed from a cohort of 100 patients with pancreatitis from a single centre. After univariate analysis, 11 variables were found to be associated with morbidity and mortality. Patients with severe pancreatitis are often admitted to the ICU and the Ranson score is still widely used despite the lack of formal validation and several complaints concerning the development of the score.

c. Child–Pugh (CP)

Child and Turcotte ûrst proposed a classiûcation system of liver failure in 1964 , later modiûed by Pugh in 1973. The CP classiûcation system grades the patients into three groups. When developing the CP score, empirical methods were used to select variables. Inclusion of two subjective variables (ascites and encephalopathy) may weaken inter-observer reliability and they are often altered by therapy. The CP is in common use and has been extensively validated outside the ICU.

d. Risk, injury, failure, loss and end-stage kidney (RIFLE) classification

Acute kidney failure is a frequent and important predictor of mortality in the ICU population. To establish a uniform classiûcation of acute kidney injury, the RIFLE classification was proposed by the acute dialysis initiative in 2004. Three severity levels of acute kidney injury (risk, injury and failure) and two outcome classes (loss and end-stage) were proposed. Characterization of acute kidney injury is based on urine output and the elevation of serum creatinine compared with baseline. In a validation study of the risk, injury and failure criteria in the ICU setting, patients in the injury and failure groups were shown to have a signiûcantly increased risk of mortality even after the correction for non-renal organ failure and other confounding factors.

3. Trauma scoring systems [4]

a. Therapeutic Intervention Scoring System(TISS)

TISS was originally designed to measure the severity of illness by quantifying the type and intensity of the treatment provided. TISS points are dependent on local, or even individual, treatment strategies, and the therapeutic capability of the unit, as well as the appropriateness of an intervention, the score cannot be used to compare the efficacy of intensive care in different units. It provide an accurate assessment of the level of care and monitoring and can be performed daily. It is therefore valuable administrative tool. TISS can be used to calculate the workload required by each patient to establish Nurse –Patient ratio, to determine a hospital's requirement for ICU beds and to calculate the costs.

b. Injury Severity Score(ISS) and Combined Trauma and Injury Severity Score(TRISS)

The ISS and TRISS system was developed to provide a standard approach for the evaluation of trauma care. Mortality following traumatic injury depends on the degree of physiological derangement, the extent of the anatomical injury, the age of the patient and whether the trauma was blunt or penetrating. The TRISS methodology combines these factors – the RTS, the ISS, age, blunt or penetrating injury- to provide a measure of the probability of survival.

4. OD Scoring Systems [1,2]

Multiple OD syndrome is the leading cause of death for patients admitted to the ICU. The general severity scoring systems, with the exception of $MPM_{48h-72h,}$ do not consider OD that develops after the ûrst 24h of ICU stay. Deûnitions of multi-organ failure do not take into account the fact that the development and resolution of organ failure is a continuum of alterations and severity rather than a definite event.

a. The Sepsis-Related Organ Failure Score(SOFA)

Developed in a conference initiated by the European Society of Intensive Care Medicine in 1994. During development, there was focus on keeping the score objective and independent of therapy, making the collection of variables uncomplicated in most ICUs. The SOFA score uses routinely collected data for the calculation of a score of 0–4 for each organ, the higher number meaning more severe failure. SOFA comprises separate daily scores for respiratory, renal, cardiovascular, CNS, coagulation and hepatic failure. The scores can be used in several ways, as individual scores (each organ), as the sum of scores on one single ICU day or the sum of the worst scores during the ICU stay.

b. Multiple-Organ Dysfunction Score (MODS)

Published in 1995, the MODS had similar goals as SOFA, in recognition of the need for a classiûcation and prognosis system that could quantify the effect of multiple-organ failure on outcome. This score uses variables which reflects physiological derangement, rather than therapeutic interventions used to support failing organs. Only post-resuscitation values are used in the calculation of MOD score, values are recorded at the same time each day and missing or unobtainable values are presumed to be normal. The MOD score provide a measure of admission severity of illness, intensity of therapeutic intervention and global ICU morbidity, and may be useful as an outcome measure in clinical trials.

c. LODS(Logistic Organ Dysfunction Score)

The European/North American Study of Severity Systems provided data for the LODS in 1996. It was the ûrst OD score to be developed with the use of multivariate regression analysis of a large database. Twelve variables for six organ systems (neurologic, cardiovascular, renal, pulmonary, haematologic and hepatic) were chosen to deûne OD. These variables were recorded as the worst value during the ûrst 24h in the ICU and do not include therapeutic interventions (except mechanical ventilation) or physiologic variables not readily available in all ICU patients. Four severity levels were identiûed assigning the scores 0, 1, 3 or 5 for each organ system according to the severity of failure. LODS was developed for the evaluation of OD on the ûrst day of ICU stay and not as a tool for monitoring disease progression, although there are modiûcations of LODS where scoring is performed on a daily basis.

Why Predict Outcome?

a. prognosis
b. cost-benefit analysis
c. withdrawal of treatment
d. comparison between different centres
e. monitoring/assessment of new therapies
f. population sample comparison in studies

Requirements for a Good Scoring System [3]

a. simple, reliable, easily obtainable
b. wide patient applicability - different diagnoses - all age groups - all levels / types of ICU's.
c. high sensitivity/specificity - ie. should be a good discriminator
d. stimulates improvement in outcomes
e. independent of treatment
f. physiological parameters

g. optimal time is unclear
h. number of criteria is unclear

Advantages of scoring systems

- Past experiences are taken into account in an unbiased manner, where as, with human decisions, recent experiences has a disproportionate influence.
- Objective outcome predictions should be more reliable because they are based on reproducible data.
- The database supporting the risk estimate is substantially larger than any one clinician's experience.
- The risk estimates are based solely on the patient's response to treatment.

Conclusion

General illness severity scores are widely used in the ICU to assess resource use, predict outcome, and characterize disease severity and degree of organ dysfunction. All the scores were developed to be used in mixed groups of ICU patients and their accuracy in subgroups of patients can be questioned; disease-specific scoring systems are increasingly being developed. As ICU populations change and new diagnostic, therapeutic and prognostic techniques become available, all the scoring systems will need to be updated. Importantly, the different scoring systems have different purposes and measure different parameters; Different forms of scoring systems are frequently used in the ICU. They have become a necessary tool to describe ICU populations and to explain differences in mortality. As there are several pitfalls related to the interpretation of the numbers supplied by the systems, they should not be used without knowledge on the science of severity scoring.

Reference

1. K. Strand, H. Flaatten. Severity scoring in the ICU: a review: Acta Anaesthesiol Scand 2008;52:467–478
2. Knaus WA, Draper EA, Wagner DP et al. APACHE II: a severity of disease classiûcation system. Crit Care Med 1985;13:818–29.
3. Jean-Louis Vincent, Rui Moreno. Clinical review: Scoring systems in the critically ill, Critical Care 2010;14:207 doi:10.1186/cc8204.
4. C.J. Hinds, D Watson. Intensive care; A concise text book, 3rd edition, Saunder's Elsevier publication; 21-30.