

NO PICO

PAPER ID: PMC1269432

TITLE:

Bench-to-bedside review: Outcome predictions for critically ill patients in the emergency department

ABSTRACT:

The escalating number of emergency department (ED) visits, length of stay, and hospital overcrowding have been associated with an increasing number of critically ill patients cared for in the ED. Existing physiologic scoring systems have traditionally been used for outcome prediction, clinical research, quality of care analysis, and benchmarking in the intensive care unit (ICU) environment. However, there is limited experience with scoring systems in the ED, while early and aggressive intervention in critically ill patients in the ED is becoming increasingly important. Development and implementation of physiologic scoring systems specific to this setting is potentially useful in the early recognition and prognostication of illness severity. A few existing ICU physiologic scoring systems have been applied in the ED, with some success. Other ED specific scoring systems have been developed for various applications: recognition of patients at risk for infection; prediction of mortality after critical care transport; prediction of in-hospital mortality after admission; assessment of prehospital therapeutic efficacy; screening for severe acute respiratory syndrome; and prediction of pediatric hospital admission. Further efforts at developing unique physiologic assessment methodologies for use in the ED will improve quality of patient care, aid in resource allocation, improve prognostic accuracy, and objectively measure the impact of early intervention in the ED.

Introduction:

The landscape of critical care delivery in the emergency department (ED) is rapidly changing. The phenomena of hospital and ED overcrowding are increasing in severity and remain unresolved. In the USA there are more than 110 million ED visits per year [1]. The proportion of critically ill patients presenting to the ED and admitted to the intensive care unit (ICU) has also risen. In California alone there was a 59% increase in the number of visits of critically ill patients to the ED from 1990 to 1999 [2]. Inpatient telemetry and ICU beds continue to be fully occupied for a significant amount of the time in many hospitals and is a primary cause of overcrowding in the ED [3,4]. As hospital census approaches 100%, the ED unavoidably becomes a surrogate ICU. Unfortunately, resources are often limited, and critical care delivery in the ED setting is fraught with inadequate space and medical equipment and lack of staffing. Increasingly stringent nurse-patient ratios are being mandated and enforced on the inpatient ward, consequently worsening the overcrowding problem, with ED nurses often far extended over their patient care capacity. ED physicians are often over-extended as well, and adequate critical care is often difficult to provide and sometimes overlooked in a busy ED. Early disease recognition and prognostication of outcome with the aid of physiologic scoring systems is a potentially valuable tool for the multitasking ED physician, and may result in improved critical care when intensive care expertise is not yet available.

In addition to the increasing focus on critical care in the ED, the framework of critical care within the ICU is evolving. The evolution of scoring systems has extended beyond just prognostication. Scoring systems now encompass critical care illness as a continuum that extends from the inciting event and treatment (often begun in the ED) to the post-ICU recovery and rehabilitation processes. Physiologic scoring systems are being utilized by clinicians and medical researchers in decision support, outcomes and evaluation research, quality care analysis, and internal and competitive benchmarking. This is the new face of ICU care and supports ongoing development of scoring systems in the ED setting as well [5,6].

We review existing physiologic scoring systems designed for application in critically ill patients, and examine how these systems have been applied in the ED. We also focus on scoring systems developed specifically for prognosticating outcome in ED patients.

Scoring systems in the intensive care unit:

Intensivists have used a variety of physiologic scoring systems in clinical decision making over the past few decades. There is currently increased emphasis regarding their use in continuous quality improvement processes, as entry criteria in clinical research trials, and even as indicators of the

efficacy of drug therapy [7]. Furthermore, in an era of rising health care expenditure, prognosticating outcome permits earlier detection of patients who will benefit most from early and aggressive therapeutic intervention. Numerous physiologic scoring systems have been developed and used widely in the ICU. Because these scoring systems are well known in the intensive care literature, we review them only briefly here.

The Acute Physiology and Chronic Health Evaluation (APACHE) II score is one of the first physiologic scoring systems developed as a mortality prediction model. It is a point scoring system that determines the severity of disease based on the worst measurements of 12 physiologic variables during the first 24 hours of ICU admission, prior health co-morbidities, and age. A high numeric score closely correlates with increased risk for in-hospital death [8]. APACHE II has been subjected to the most validation studies, which show that mortality prediction is accurate, and it is currently the most widely used scoring system in the ICU setting. It has been shown to predict outcome accurately in a variety of medical illnesses, including pancreatitis [9], cirrhotic liver disease [10], infective endocarditis [11], medical complications of oncologic patients [12], chronic obstructive pulmonary disease [13], gastrointestinal hemorrhage [14], myxedema coma [15], acute myocardial infarction requiring mechanical ventilation [16], and septic abortion [17]. APACHE II has even been shown to be superior to the American Society of Anesthesiologists classification in preoperative prediction of postoperative mortality [18]. The latest APACHE III scoring system was shown to be reliable in predicting outcome of surgical ICU patients as well [19,20].

Other scoring systems such as the Simplified Acute Physiology Score (SAPS) II [21], Sequential Organ Failure Assessment score [22], Multiple Organ Dysfunction Score (MODS) [23], Mortality Probability Models [24,25], and the Pediatric Risk of Mortality score [26,27] have been shown to be beneficial in predicting resource utilization, organ failure, and mortality in patient populations such as those with cardiovascular disease [28], adult [29] and pediatric [30] trauma, obstetric patients [31], surgical ICU patients [32,33], and nonsurgical ICU patients [34].

Although these systems were originally designed to predict mortality, their use is being progressively expanded to compare clinical trials [35-37] and for criteria to initiate drug therapy; for example, an APACHE II score of 25 or greater is often used as an indication for drotrecogin alfa (activated) in severe sepsis. Hence, there is difference between how scoring systems were derived and how they are being used clinically.

Scoring systems in trauma:

Trauma scoring systems have also been used in the triage of trauma patients and to predict their outcome. Trauma scores have been used to characterize severity of injury and physiologic derangements quantitatively.

The Glasgow Coma Scale (GCS) assesses the severity of head trauma based on three response parameters: eye opening, motor, and verbal response. Compared with other more extensive scoring systems, the GCS has been shown to be superior in predicting outcome, which it does with high sensitivity and specificity [38]. It is also simple to use and readily applied at the bedside. However, inter-rater reliability of GCS scoring was recently shown to be less adequate than was previously believed [39]. Furthermore, the three individual component scores of GCS have similar areas under the receiver operating characteristic (ROC) curve to that of the total GCS score for predicting ED intubation, neurosurgical intervention, brain injury, and mortality [40].

The Therapeutic Intervention Scoring System (TISS) evaluates the need in staffing, monitoring, and therapeutic intervention rather than stratifying severity of illness. Patients are assigned to a class from I to IV, ranging from those who do not require intensive therapy to those patients who are considered physiologically unstable. TISS has been shown to be effective in stratification and prediction of ICU cost [41]. With the new TISS-28, it may be possible to predict post-ICU outcome and identify those high-risk patients who would benefit from further observation [42].

The Trauma Score provides a numerical assessment of central nervous system and cardiopulmonary function. Prediction of survival was shown to be reliable [43]. The Revised Trauma Score is probably the more widely used scoring system currently in trauma and is an accurate predictor of outcome. However, its usefulness as a triage tool was recently questioned [44].

Other trauma scores have been designed using various combinations of physiologic parameters, mechanism, age, GCS, and systemic inflammatory response syndrome (SIRS). Examples of these scoring systems include the Injury Severity Score, Trauma and Injury Severity Score (TRISS), International Classification Injury Severity Score, and the Physiologic Trauma Score. These

scoring systems have been used in a variety of trauma scenarios, including motor vehicle accidents, blunt and penetrating trauma, and even in pediatric polytrauma [43,45-49].

Existing scoring systems applied to the emergency department:

ED scoring and outcome prediction are innovative but relatively novel concepts. As a result, few scoring systems are specific to the ED setting. Most scoring systems are applicable upon ICU admission and throughout the first 24 hours after admission. These systems usually do not take into account the ED length of stay and course of therapy. Several authors have taken existing physiologic scoring systems, originally designed for application in the non-ED setting, and applied them in the ED and prehospital patient population.

For example, TRISS was used to determine the effectiveness of ground versus air transport for major trauma victims [50]. TRISS accurately predicted 15 out of 15 deaths of the 110 patients transported by ground, but only 33 out of the 46 predicted deaths occurred in the 103 patients transported by air. Even though the study did not randomize patients to receive ground versus air transport, the authors concluded that air transport resulted in better outcome because only 72% of patients predicted to die actually died following air transport. Irrespective, the study suggests that current trauma scoring systems can be applied successfully in prehospital and ED settings. Another study used three physiologic scoring systems – APACHE II, SAPS II, and MODS – to assess the impact of ED intervention on morbidity and in-hospital mortality [51]. In that prospective, observational cohort study, patients were enrolled and their scores were computed at ED admission, ED discharge, and at 24, 48 and 72 hours in the ICU. The authors applied these scoring systems at specific time points in order to observe the trend in scores over a 72-hour period. Length of ED stay was approximately 6 hours. The hourly decreases in APACHE II, SAPS II, and MODS scores were noted to be most significant during the ED stay, as compared with scores computed during the subsequent 72 hours in the ICU. The APACHE II and SAPS II scores both exhibited notable decreases in predicted mortality during the ED stay. The nontraditional use of these scores allowed the authors to show that the highest scores and predicted mortalities occurred during the ED stay, and that traditional scoring during the first 24 hours after ICU admission (and after initial resuscitation) may not account for the actual severity of disease in the pre-ICU period. Although the study reemphasizes the significant impact that ED intervention has on critically ill patients, it also suggests that existing scoring systems such as APACHE II either are limited to their original design (which is prognosticate to outcome based only on the first 24 hours in the ICU) or need to be recalibrated to include physiologic parameters in the ED [51]. SIRS, part of the definition of sepsis, has been used as a predictor of outcome in patients admitted to the ICU from the ED [52]. SIRS in combination with an elevated lactate (≥ 4 mmol/l) in the ED was found to be 98.2% specific for admission to the hospital and the ICU, and 96% specific for predicting mortality in normotensive patients [53,54]. SIRS and elevated lactate (≥ 4 mmol/l) have also been used successfully in the ED as screening variables for initiation of invasive hemodynamic monitoring and early goal-directed therapy in severe sepsis or septic shock patients, resulting in significantly improved outcomes [35]. Because SIRS has been the limiting factor to a better definition of sepsis [55], the addition of lactate in the triaging of patients with a suspected infection may allow ED physicians to identify normotensive patients at high risk for septic shock.

The Pneumonia Severity Index [56] is a measure of severity of community-acquired pneumonia, taking into account physiologic parameters, age, medical co-morbidities, and laboratory studies. Even though it was designed as an outcome prediction tool, the Pneumonia Severity Index is widely used as a determinant for site of care in conjunction with clinical judgment [57] and as a quality assessment tool [58-60].

Scoring systems developed for use in the emergency department:

There are a number of physiologic scoring systems designed for use in the ED setting, some of which are discussed below and summarized in Table 1. These systems require several unique characteristics that are inherent to the ED, such as ease of use and bedside availability, accuracy of prediction within a shorter time frame of data collection, and comparability with current ICU scoring systems on hospital admission.

The Mortality in Emergency Department Sepsis Score (MEDS) is a recent scoring system developed from independent variables and univariate correlates of mortality. It was designed to predict patients in the ED who are at risk for infection and to stratify them into risk categories for mortality [61]. A prediction model was developed based on independent multivariate predictors of death, including terminal illness, tachypnea or hypoxia, septic shock, platelet count below

150,000/mm³, band proportion above 5%, age above 65 years, lower respiratory infection, nursing home residence, and altered mental status. Based on the MEDS score, patients in the developmental group were assigned to very low, low, moderate, high, and very high risk categories for mortality. MEDS as a valid outcome prediction model was established in a validation group, with an area under the ROC curve of 0.76 in this group [61]. MEDS is among the first scoring systems to be examined over the natural course of sepsis beginning in the ED. However, the mortality in the study patients of 5.3% is exceedingly low compared with the more familiar sepsis mortality range (16–80%) [62,63]. Thus, studies are needed to validate MEDS before it may be clinically applicable in other ED settings.

The Rapid Acute Physiology Score (RAPS) is an abbreviated version of the APACHE II scoring system. It was developed to predict mortality before, during, and after critical care transport. Limited physiologic parameters available on transport (i.e. pulse, blood pressure, respiratory rate, and GCS) were used and scored numerically [64]. RAPS correlated well with APACHE II score in a comparison analysis ($r = 0.85$; $P < 0.01$) [64]. RAPS, when initiated in the prehospital setting and extended into the full APACHE II score upon admission, is highly predictive of mortality [65,66]. RAPS is an efficient scoring system for use in the prehospital setting, but it is probably too abbreviated. Because most of the variables included in the score are vital signs, it may be too sensitive as a prediction tool. For example, patient anxiety during transport, leading to an elevated heart rate or respiratory rate, will easily increase the RAPS score over a very short time interval. The Rapid Emergency Medicine Score (REMS) is a modification of RAPS, with age and peripheral oxygen saturation added to the RAPS score. Its predictive value is superior to that of RAPS for in-hospital mortality when applied to patients presenting in the ED with common medical issues [67]. The area under the ROC curve is 0.85 for REMS, as compared with 0.65 for RAPS ($P < 0.05$) [67]. REMS has also been shown to have predictive accuracy similar to that of APACHE II [68]. A clinician can easily expand a REMS score into the full APACHE II score. Thus, an APACHE II score can be quickly calculated by the intensivist with a few additional parameters once the patient is admitted to the ICU. Although studies have examined its application in the ED, these studies are limited to the nonsurgical patient population.

The Mainz Emergency Evaluation Systems (MEES) was developed in Germany to assess prehospital therapeutic efficacy. It is based on seven variables: level of consciousness, heart rate, heart rhythm, arterial blood pressure, respiratory rate, partial arterial oxygen saturation, and pain. A MEES score is obtained before and after prehospital intervention to assess patient improvement or deterioration. Although it does not allow outcome prediction, it does provide an easy and reliable assessment of prehospital care [43,69]. A recent study [70] showed that adding end-tidal carbon dioxide capnometry to MEES has significantly greater value than MEES alone in predicting survival after cardiopulmonary resuscitation in nontraumatic cardiac arrest.

In Taiwan, severe acute respiratory syndrome (SARS) screening scores were developed specifically for prediction of this syndrome in febrile ED patients. Recently, two of these SARS screening scores, the four-item symptom score and the six-item clinical score, were tested and validated in different cohorts in Taiwan and were found to have good sensitivity and specificity for predicting SARS [71]. The study suggests that these scores could be used as a tool for mass screening in case of future outbreaks. However, they would not be applicable for screening on a case-by-case basis outside endemic regions.

The Pediatric Risk of Admission score includes nine physiologic variables, three medical history components, three chronic disease factors, two therapies, and four interaction terms. This score provides a probability of admission from the ED for pediatric patients. It was shown to be reliable in predicting admission and providing a measure of illness severity [72–74]. Although the score was not designed specifically for outcome prediction, it is an example of the use of scoring systems to risk stratify and triage patients in the ED.

Conclusion:

Emergency physicians have the opportunity to have a significant impact on the initial evaluation and treatment of the critically ill patient. Application of outcome prediction models in the form of physiologic scoring systems allows early recognition of illness severity and initiation of evidence-based therapeutic interventions. In the presence of overcrowded, under-staffed EDs, the utility of efficient and bedside physiologic scoring systems can be of tremendous value to the multitasking ED physician. As technology advances, immediate access to patient data and the availability of ED scoring systems on hand-held computers will further facilitate outcome prediction. However, the current development, implementation, and verification of these systems in the ED setting are limited.

Unique physiologic assessment tools and outcome prediction models should be developed for use in the ED setting. Physiologic scoring systems such as APACHE II, SAPS II, and MODS were developed to measure illness severity objectively, to provide mortality risk probabilities, and to evaluate the performance of ICUs. When these models are applied in the ED setting, lead-time bias may result because these systems were not originally designed to account for pre-ICU illness severity [51]. Thus, similar models specific to the ED should include the following: variables that reflect prehospital severity of illness and are commonly obtained in the ED; use of practical time-indexed variables that reflect response to treatment delivered in dynamic resuscitation during ED care; creation of an independent, multicenter database to establish adequate sample size and power for the development and validation of the model [21,75-79]; analysis of the relationships among the predictive variables and actual patient outcome for overall calibration and reliability of the model; establishment of outcomes other than mortality, such as patient disposition, number of return visits to the ED, lengths of ED and ICU stay, length of mechanical ventilation, and functional status at hospital discharge [80]; and the ability to be correlated with more established scoring systems already in place in ICUs.

Outcome prediction science is not considered synonymous to physician clinical judgment. However, the intent of prediction models is to reduce clinician variability and improve the overall accuracy of prognostic estimates. An ED patient-specific prediction model can assist clinicians by providing greater certainty in the effects of interventions provided in the ED; improving the understanding of existing physiologic measurements and their influence on outcomes; reducing variations in individual clinical judgment on the severity of patient illness at ED presentation; allowing for comparison of probability thresholds to guide important clinical decisions; and providing a common measurement tool with which to compare performance among EDs [80,81]. Physiologic assessment tools can also identify outliers by comparing actual outcomes with expected outcomes, and thus provide opportunities for quality improvement if inadequacies of care are identified in case reviews. However, it must be recognized that physiologic scoring systems are typically developed to provide estimates of outcome for a group of patients, and not to predict individual patient outcome. In addition, they should not be used to make end-of-life decisions in emergency situations.

Most EDs are staffed for short-term stabilization of critically ill patients. Because of overcrowding and prolonged ED lengths of stay, the care provided to patients with such high acuity may vary and is limited by available equipment, training, and staff-patient ratios. Methodologies such as physiologic scoring systems to assess the quality and quantity of critical care delivered will serve as tools to help remedy the varying care delivered in the ED setting. Thus unique physiologic assessment methodologies should be developed to examine and improve the quality of patient care, enhance the precision of clinical research, aid in resource allocation, improve the accuracy of prognostic decisions, and objectively measure the impact of clinical interventions and pathways in the ED.

Abbreviations:

APACHE = Acute Physiology and Chronic Health Evaluation; ED = emergency department; GCS = Glasgow Coma Scale; ICU = intensive care unit; MEDS = Mortality in Emergency Department Sepsis Score; MEES = Mainz Emergency Evaluation Systems; MODS = Multiple Organ Dysfunction Score; RAPS = Rapid Acute Physiology Score; REMS = Rapid Emergency Medicine Score; ROC = receiver operating characteristic; SAPS = Simplified Acute Physiology Score; SARS = severe acute respiratory syndrome; SIRS = systemic inflammatory response syndrome; TISS = Therapeutic Intervention Scoring System; TRISS = Trauma and Injury Severity Score.

Competing interests:

The author(s) declare that they have no competing interests.