The Comparison of Reliabilities of Nutritional Risk Scoring and Physiological Severity Score in Critical Care Patients

Correspondence Address

Hatice KAYA ÖZDOĞAN
Adana Numune Training and Research Hospital, Department of Anesthesiology and Reanimation, Adana, Turkey
E-mail: hat_ka33@hotmail.com

ABSTRACT

Objective: APACHE II is most commonly used severity prognostic model in ICU patients. Nutritional risk screening 2002 (NRS-2002) score is widely suggested for screening of nutritional risk in general hospitalized patients. In this study, we aimed to compare the reliability of NRS-2002 with APACHE II on outcome in critical care patients.

Material and Methods: Discrimination and calibration characteristics of the scoring systems were evaluated.

Results: APACHE II on admission had moderate power of discrimination and calibration for mortality and complication prediction in critical care patients. On the other hand, NRS-2002 had insufficient discrimination statistics.

Conclusion: APACHE II is superior in outcome prediction compared to NRS-2002 in ICU patient population.

Key Words: APACHE II, NRS-2002, Mortality, Complication, Intensive care

INTRODUCTION

There are a number of scoring systems for outcome prediction in critical care patients. A common use of most of these systems is illness severity scoring to make comparisons between patient groups or intensive care units (ICU) (1, 2). However, they have also been used to assess or predict the risk for specific patient groups.
There are many scoring systems for risk prediction in critical care patients. These systems have been originally designed to predict disease severity and probability of hospital mortality and complications in general ICU patients. Acute Physiology and Chronic Health Evaluation II and III (APACHE II and III), Simplified Acute Physiology Score II (SAPS II), and Mortality Probability Model (MPM II) are examples. APACHE II has been the most commonly used one among these scores. Although these scoring systems take into account the chronic health problems, one of the drawbacks of all of the above-mentioned systems is that they ignore the nutritional status of the patient.

Malnutrition has been known to be associated with increased rates of complications, length of hospital stay (LOS), and mortality (3-7). It is well known that appropriate nutritional treatment decreases complication and infection rates, and reduces LOS (8, 9). The Nutritional risk screening 2002 (NRS-2002) score was introduced by the European Society for Clinical Nutrition and Metabolism (ESPEN) as a useful method for screening of nutritional risk in hospitalized patients (10). When the NRS-2002 score is ≥3, patients are accepted to be nutritionally under risk and nutritional support is advised (10-13).

True validity of a risk screening or predicting tool can only be discussed in the context of its impact on clinical outcome. The aim of this study was to compare the reliabilities of NRS-2002 and APACHE II in general ICU patients in predicting the outcome.

**MATERIAL and METHODS**

The Ethical Committee of the Adana Numune Training and Research Hospital approved the study (Chair: Ö. Keşkek, Assoc. Prof of Internal Medicine, 76/2017). This study was conducted in 239 consecutive critical patients admitted to the Anesthesia ICU of the Adana Numune Training and Research Hospital between June 2016 and June 2017. Patient records and APACHE II and NRS-2002 scores calculated on ICU admissions were retrospectively evaluated.

The outcome measures of the study were the reliability of the evaluated scoring systems in predicting the mortality and complication rates. Complications that have been taken into consideration included pulmonary (atelectasis, respiratory failure, adult respiratory distress syndrome, and pneumonia), cardiac (arrhythmias, ischemic events), thrombo-embolic, urinary, neurological, and infectious complications.

**Data Analysis and Statistics**

Discrimination and calibration define the overall predictive power of a model. “Discrimination” refers to ability of a model to distinguish patients who experienced an event from those who did not. Discrimination was measured by the receiver operating characteristic (ROC) curves. The area under the curve (AUC) represents the probability that a patient who experienced the event had a higher predicted probability of having that event than a patient who did not (14). The higher the true-positive rate is relative to the false-positive rate, the greater is the AUC. An AUC of 0.5 indicates that the model does not predict better than chance. The discrimination power of a model is considered perfect if AUC=1, good if AUC>0.8, moderate if AUC is between 0.6 and 0.8, and poor if AUC<0.6.

“Calibration” refers to the agreement between the “predicted probabilities” and the “true probabilities”. Calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test and the corresponding calibration curves. Small $P$ value, which means a significant difference exists between the observed and predicted event, indicates a lack of fit of the model (15).

Continuous variables were presented as means (min-max) and were compared using the Mann-Whitney $U$ test. Categorical values were analyzed using the Chi-square test. $P<0.05$ was considered as statistically significant. Statistical evaluation was performed by using the SPSS 15.0 statistical package.

**RESULTS**

During the study period, 239 patients were admitted to the Anesthesia ICU. A hundred and thirty-one patients were male and 108 were female. In this group, 53 patients were trauma patients, 67 patients were admitted following major surgery, and other patients were medical critical care patients. A total of 62 patients died in the ICU (Table I). Table II shows overall scores and the differences between hospital survivors and non-survivors. Non-survivors had

<table>
<thead>
<tr>
<th>Table I: The study population.</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>Patients with complications</td>
</tr>
</tbody>
</table>

NS: Non-significant
significantly higher APACHE II scores on admission. There were no statistically significant differences regarding NRS-2002 scores.

Discrimination and calibration statistics for mortality and complication prediction were presented in Table III. None of the systems had perfect discrimination power based on the finding that all AUC values are below 0.8. APACHE II was the best performer for mortality and complication prediction. Calibration characteristics were statistically adequate for both models.

**DISCUSSION**

Many illness severity scoring systems have been produced but few are currently in clinical use. The most commonly used systems include APACHE II and III, SAPS II and MPM II (2). These systems consider objective physiological criteria for scoring so that standardized comparisons can be performed between patient groups and between ICUs. However, to some extent, they can be used to predict risk or to assess a probability of mortality for general ICU patient groups (1, 2). There have been controversies in the use of these systems in the general critical care patient population. However, attempts have been made to apply and evaluate the performances of these scoring systems in patients with cancer (16), trauma (17), postoperative complications (18), emergency surgery (19), and elective surgery (20, 21). For example, in a critical review, den Boer et al. demonstrated that APACHE II had poor to good discrimination performance, and generally underestimated the mortality risk in solid organ cancer patients (16). The most widely used one of these systems, APACHE II, has been shown to be able to predict mortality (21), and even the increasing levels of both local and systemic complication rates in elective surgery (20, 21). Our results show that APACHE II is useful in mortality and complication prediction in general critical care patients. The use of APACHE II on admission to ICU can enable a reliable prediction of mortality in this patient group.

One of the major drawbacks of the APACHE II model is that it does not specifically take the nutritional status of the patient into account. Numerous tools for the screening of malnutrition and nutritional risk have been proposed. The NRS-2002, endorsed by ESPEN, consists of a nutritional score, a disease severity score, and an age adjustment for patients aged > 70 years. Total score is calculated and patients are classified as at no risk to high risk (10, 12). Although NRS-2002 has not been specifically constructed for use in surgical patients, it was also found to be a sensitive screening tool in patients undergoing elective surgery (13, 22). A high NRS-2002 score is significantly associated with increased complication rates and prolonged LOS in the general hospital population (11, 13, 22, 23). The purpose of nutritional screening is to predict the probability of better or worse outcome due to nutritional factors, and whether nutritional treatment is likely to influence it (10). Therefore, an ideal nutritional screening tool is supposed to predict mortality and complication rates so that nutritional intervention can be provided in high risk patients.

**Table II:** The scores of evaluated systems in patients.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Survivors</th>
<th>Non-survivors</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>APACHE II</strong></td>
<td>21.95 (5-52)</td>
<td>13.72 (5-33)</td>
<td>27.14 (11-52)</td>
<td>NS</td>
</tr>
<tr>
<td><strong>NRS-2002</strong></td>
<td>3.3±0.5</td>
<td>3.1±0.3</td>
<td>3.9±0.6</td>
<td>NS</td>
</tr>
</tbody>
</table>

**APACHE:** Acute Physiology and Chronic Health Evaluation  
**NRS-2002:** Nutritional risk screening 2002  
**NS:** Non-significant

**Table III:** Discrimination and calibration statistics of evaluated systems for outcome parameters.

<table>
<thead>
<tr>
<th></th>
<th>Mortality</th>
<th>Complication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>H-L P</td>
</tr>
<tr>
<td><strong>APACHE II</strong></td>
<td>0.67</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>APACHE II predicted outcome</strong></td>
<td>0.65</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>NRS-2002</strong></td>
<td>0.52</td>
<td>0.11</td>
</tr>
</tbody>
</table>

**APACHE:** Acute Physiology and Chronic Health Evaluation  
**NRS-2002:** Nutritional risk screening 2002  
**AUC:** Area under curve  
**H-L:** Hosmer-Lemeshow
The present study showed that NRS-2002 had insufficient discrimination power for mortality and complication in general ICU patients. In a previous study from our center, it was found that the NRS-2002 score had moderate to good discrimination power for complications and LOS in ICU prediction in general trauma patients admitted to the Trauma ICU (24). However, it was unreliable for mortality prediction. We also demonstrated that there was an insufficient discrimination power for NRS-2002 score by ROC analysis.

One of the drawbacks of the present study is that it consisted of a heterogeneous group, including trauma, postoperative, and medical critical care patients. The differences in demographic characteristics and medical conditions of trauma, major elective surgery, and medical critical care patients might have resulted in inevitable bias in statistics.

Overall, the critical care risk formulas are, in their current form, useful mainly for 1) clinical research to quantify the degree of risk in a study sample or to evaluate whether two study groups are comparable; and 2) quality improvement programs. The present study, in conclusion, showed that APACHE II is reliable in predicting mortality in general ICU patients. On the other hand, NRS-2002 is not supposed to predict mortality and complication in this patient population.

REFERENCES