Evaluation of Prognostic Scoring Systems in the Prediction of the Outcome in Critically Ill Patients with Perforative Peritonitis

Samir Delibegovic¹, Dragana Markovic¹, Samir Hodzic¹, Amir Nuhanovic²
University Clinical Center Tuzla, Tuzla, Bosnia and Herzegovina¹, Faculty of Electrical Engineering, University of Tuzla, Bosnia and Herzegovina²

Original paper
SUMMARY
Background: Patients with perforative peritonitis are among the most complex cases encountered in surgical practice. Early prognostic evaluation of these patients is desirable in order to make the correct therapeutic plan, selecting highly risky patients for less aggressive surgical procedures. Prospective evaluation of different prognostic scoring systems was performed in order to assess the possibility of prediction of outcome in these patients. Patients and methods: The prospective study of 145 patients with perforative peritonitis was performed. The main outcome of this study was peritonitis-related death. Variables necessary for calculation of the scoring systems were recorded at the initial admission to the hospital (during the first 24 hours) and the third and seventh day of hospitalization, except Mannheim Peritonitis Index, which was calculated during the first 24 hours after hospitalization, i.e. during laparatomy. Sensitivity and specificity are graphically shown for the different values of cut-off points. Results: ROC curve for TISS-28 and APACHE II is significantly more accurate in comparison with other scores. The area under the curve for the first postoperative day was 0.87 for TISS-28 score, 0.86 for APACHE II score, 0.83 for MOF; 0.83 for SAPS I, 0.72 for MPI score, 0.70 for Sepsis score. In addition, this discriminatory ability remained on the third and seventh postoperative day as well. The highest rate of correlation between the observed and the expected mortality rate was in APACHE II system, for the first (Kendall’s τ correlation 0.964) and the third (Kendall’s τ correlation 0.810) day. There was a decrease in the rate of correlation on the seventh day for all scoring systems except for MOF score. Conclusion: APACHE II is better in prediction of the outcome to other tested scoring systems.

Key words: peritonitis, APACHE II, Surgical Intensive Care Unit.

1. INTRODUCTION

Several scoring systems, based on logistic regression, have been developed in the last two decades. These scoring systems are made for assessment of severity of patients in Intensive Care unit (ICU) and assess general consequences of disease. Artificial neural network, as an alternative method of prediction of outcome, gave contradictory results (1,2). Scoring systems has been applied in surgical patients. Validity of APACHE II (Acute Physiology And the severity of Chronic illness) scoring system is confirmed in surgical patients with intraabdominal infection (3) and peritonitis (4). In literature we could not find any study about the use of SAPS I (Simplified Acute Physiology Score) in the group of patients with secondary peritonitis, although it is a simplified APACHE II score which could be applicable for the use in this group of patients.

Elebute and Stoner established the method of grading of sepsis (5); and validity of this sepsis score (SS) was confirmed (6). There are scoring systems specially developed to assess the impairment of organ function such as MOF (Multiple Organ Failure) score (7). In several studies MOF score was used in the group of peritonitis patients (8). Scoring systems based on intraoperative data have been developed such as Mannheim Peritonitis Index (MPI) and Peritonitis Altona Index. The validity of MPI has been shown in literature (9). It is one of the easiest scoring systems for risk prediction during the initial operation.

Perforative peritonitis is the most frequent form of intraabdominal infection—about 80% of cases of secondary peritonitis in large hospitals account for perforated peritonitis (10). These patients are among the most complex cases encountered in surgical practice (11). Early prognostic evaluation of these patients is desirable in order to make the correct therapeutic plan, selecting highly risky patients for less ag-
gressive surgical procedures. That is why we performed prospective evaluation of these five prognostic systems in order to assess the possibility of prediction of outcome in patients with perforative peritonitis.

2. PATIENTS AND METHODS

The prospective study involved 145 patients of both sexes with secondary perforative peritonitis. The average age of all patients was 58 (ranges from 21 to 81); the ratio of men: women was 91: 54, with no significant differences in the average age between the two genders. The patients hospitalized in surgical intensive care unit (SICU) longer than 24 hours were included in the study. The inclusion criterion was perforative peritonitis, as determined by laparatomy. The exclusion criterion was post-traumatic peritonitis. Patients were tracked either to discharge or death.

The main outcome of this study was peritonitis-related death. APACHE II, SAPS I, TISS-28, SS and MOF scoring systems were calculated on the admission (during the first 24 hours), the third and the seventh day after hospitalization, because we wanted to check the value of serial determination of scoring systems. MPI scoring system was calculated during the first 24 hours after hospitalization or during laparatomy. Quantitative estimation of the level of unconsciousness was performed using Glasgow coma scale (GCS). Data were collected in computer database made in the commercial program of Microsoft Access. The statistic analysis was performed using the commercial software (Microsoft Access and Microsoft Excel, SPSS 7.5 and Statistica 5.0).

Cut-off points were specified (26 points for APACHE II, 17 for SAPS I, 26 for MPI, 14 for SS, 39 for TISS-28) and all values greater than cut-off points were taken to predict death. Sensitivity and specificity are graphically shown for the different values of cut-off points. They are presented with ROC curve. The difference in the area below the curve between scoring systems was statistically compared. The test of the difference between areas under ROC curves was applied using trapezoidal rule to approximate areas, conservative estimation for the standard deviation, and Kendall’s τ as a measure of the correlation between the areas.

3. RESULTS

Relationship between prognostic scoring systems and probability of death is shown in figure 1-6. The values for all days were taken into account in order to check and compare validity of particular probabilities for each prognostic scoring system. The shape and position of the curve in APACHE II and TISS-28 score (starting and final part of curve are parallel with x-axis) demonstrates that these two scoring systems are superior at describing the probability of death.

3.1. Discriminatory ability of prognostic systems

ROC curve used for the prognostic scoring systems (relationship between sensitivity and the false positive rates (1-specificity) for different cut-off points) was given for the first, third and seventh postoperative days (Figures 7-9). APACHE II and TISS-28 curve demonstrated that their discriminatory ability was better than SAPS MPI, SS and MOF curve. The area under the curve for the first postoperative day was 0.87 for TISS-28 score, 0.86 for APACHE II score, 0.83 for MOF, 0.83 for SAPS, 0.72 for MPI and 0.70 for SS. This illustrated that TISS-28 and APACHE II are significantly better than the other systems (P<0.01) (Figure 1). In addition, this discriminatory ability remained on the third and seventh postoperative day as well (Fig. 2 and Fig. 3). In most cases, APACHE II (Table 1) gave the exact prediction (P<0.1) after that TISS-28 scoring system.

Although discriminatory ability of APACHE II is 0.86 and for TISS-28 0.87, it is far from perfect. Even with selection of optimal cut-off points it is not possible to get false-positive rate less than 20% and false-negative rate lesser than 20%.

3.2. Sharpness

The distribution of scoring systems, a measure for sharpness of

<table>
<thead>
<tr>
<th>Score</th>
<th>First day</th>
<th>Third day</th>
<th>Seventh day</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;0.1</td>
<td>0.1–0.9</td>
<td>&gt;0.9</td>
<td>&gt;0.1</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>

APACHE II 71 64 10 49 69 13 61 45 6 196 172 24
MOF 0 145 0 95 33 3 0 109 3 0 387 5
SAPS 30 110 5 62 61 8 77 29 6 142 230 20
SS 5 131 9 9 119 3 28 84 0 53 334 5
TISS-28 68 60 17 61 61 9 61 47 4 190 180 22
MPI 0 145 0 - - - - - - - -

Table 1. *Sharpness* of prediction of every scoring system

<table>
<thead>
<tr>
<th>The relation of observed and expected mortality</th>
<th>First day</th>
<th>Third day</th>
<th>Seventh day</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendall’s τ</td>
<td>p</td>
<td>Kendall’s τ</td>
<td>p</td>
<td>Kendall’s τ</td>
</tr>
<tr>
<td>APACHE II 0.96477 0.0001</td>
<td>0.81019 0.0011</td>
<td>0.52624 0.0342</td>
<td>0.80006 0.0013</td>
<td></td>
</tr>
<tr>
<td>MOF 0.91205 0.0002</td>
<td>0.68421 0.0059</td>
<td>0.92105 0.0002</td>
<td>0.76212 0.0022</td>
<td></td>
</tr>
<tr>
<td>SAPS 0.77652 0.0008</td>
<td>0.51210 0.0393</td>
<td>0.42502 0.0871</td>
<td>0.89661 0.0003</td>
<td></td>
</tr>
<tr>
<td>SS 0.96477 0.0001</td>
<td>0.81019 0.0011</td>
<td>0.52624 0.0342</td>
<td>0.80006 0.0013</td>
<td></td>
</tr>
<tr>
<td>TISS-28 0.93023 0.0002</td>
<td>0.70753 0.0044</td>
<td>0.57915 0.0198</td>
<td>0.76765 0.0020</td>
<td></td>
</tr>
<tr>
<td>MPI 0.90113 0.0007</td>
<td>- - - - - -</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Kendall’s τ – a measure of the correlation between the areas under ROC curve.

Table 2. Kendall’s τ and the significance (N=10) for the tested scoring systems.
evaluation of Prognostic scoring systems in the Prediction of the outcome in critically Ill Patients with Perforative Peritonitis

prediction, is shown in Table 1. APACHE II and TISS-28 predictions were sharper than those of other scoring systems.

3.3. Reliability of prognostic scoring systems
Testing Kendall’s τ as a measure of correlation (Table 2), we have found that the highest rate of correlation between the observed and the expected mortality rate was in APACHE II system, for the first (Kendall’s τ correlation 0.964) and the third (Kendall’s τ correlation 0.810) day. There was a decrease in the rate of correlation on the seventh day for all scoring systems except for MOF score.

4. DISCUSSION
There are several scoring systems available for the estimation of severity of the disease and prognosis in SICU, especially in peritonitis patients. Most studies have shown that among scoring systems based on physiological parameters, the most reliable system is APACHE II score (8,12).

APACHE II reliably assesses mortality in the group of surgical patients with systemic disarrangement, such as peritonitis (13).

APACHE score in our study was from 0 to 38, with the average of 25 points. No patients with a score higher than 28 survived. In other studies, different values of scores were reported for the dead patients. Chen et al. in their study cited that patients with a score higher than 40 did not survive (14), and Edwards et al. cited the score value of 22, so that value

---

**Figure 1.** APACHE II and probability of death.

**Figure 2.** TISS-28 score and probability of death.

**Figure 3.** MOF score and probability of death

**Figure 4.** SAPS score and probability of death

**Figure 5.** S5 and probability of death

**Figure 6.** MPI score and probability of death
can be used as an additional criterion for clinical decision not to operate (15). However, there are opposite opinions that this scoring system can be used in retrospective studies, but that it should not be used in a triage process nor as a predictor of the outcome in individual patients. The triage decision should be based on clinical estimation (16).

Mannheim peritonitis index (MPI) is based on intraoperative data and it has been developed specifically for abdominal infection (9). Although there are opinions that combination of APACHE II and MPI should be a standard classification system for grading severity of peritonitis and intraabdominal sepsis (17), in our study we show that MPI has no predictive power, while APACHE II does.

MOF score can be used in the prognosis of postoperative complications and deaths (8), but predictive power of APACHE II and TISS-28 score is far better. However, there is a characteristic good correlation between observed and expected mortality rate on the seventh day for MOF score, and a drop in the correlation rate for other tested scoring systems. MOF syndrome represents the process in surgical patients which cannot be manifested during the first 24 hours.

In literature we could not find any papers about the use of SAPS scoring system, in peritonitis patients, although it is a simplified APACHE II. However, in our study sensitivity and significance of this scoring system was lower than APACHE II. APACHE II in relation to Sepsis score, as well, is more objective and prospectively better validates large number of patients with different diagnosis (18), what was confirmed in our study.

In SICU TISS-28 reflects the amount of care and can provide useful information about severity of disease and prognosis (19). Usefulness of economic analysis supports its clinical application as well. So TISS-28 can be used as a valuable tool for the control of quality and the analysis of costs in ICU (20). In our study, high sensitivity and specificity of this system has been proven. Comparing observed and expected mortality we do not find significant difference in the first and third postoperative day for APACHE II and TISS-28 score, and the seventh day for MOF score. Overall, in the aspect of observed and expected mortality APACHE II exhibited the best predictive power.

Although some studies cited that there is a possibility of classification of severity of disease, defining risk groups of patients using scoring systems, especially APACHE II system (18), in this study it was not confirmed.

Despite of advancements in diagnostic techniques, the decision for reoperation in critically ill patients depends on medical estimation and can be the source of conflicts caused by broad spectrum of pressures centralized on surgeon, and costs which are generated by treatment of critically ill patients can be high, and its cost-benefit relation is not always acceptable. APACHE II is made for assessment of severity of patients in ICU and assesses general consequences of disease, respecting the age and previous medical conditions. Its practical use is pretty demanding, as it contains a large number of parameters that can be obtained only by invasive monitoring. On the other side, it can be expected, even in ICU in developing countries, to take over management of
these patents, because all parameters of APACHE II scoring system are a part of routine monitoring. It was proven as useful in defined population with syndrome of systemic inflammatory response (intraabdominal infection) in which global physiological disturbance is presented.

Quantification of the severity of the disease in critically ill patients is further controversial. However, scoring systems are a useful tool for controlling the quality of management and assessment of management itself. The value of APACHE II scoring system can alert doctor to a possible negative outcome and help him/her design a strategy for the patient's treatment.

REFERENCES