Risk Stratification in Emergency Surgical Patients

Is the APACHE II Score a Reliable Marker of Physiological Impairment?

Thomas Koperna, MD, PhD; Dagmar Semmler, MD; Friedrich Marian, MD

**Hypotheses:** The APACHE II (Acute Physiology and Chronic Health Evaluation II) score used as an intensive care unit (ICU) admission score in emergency surgical patients is not independent of the effects of treatment and might lead to considerable bias in the comparability of defined groups of patients and in the evaluation of treatment policies. Postoperative monitoring with the APACHE II score is clinically irrelevant.

**Design:** Inception cohort study.

**Setting:** Secondary referral center.

**Patients:** Eighty-five consecutive emergency surgical patients admitted to the surgical ICU in 1999. The APACHE II score was calculated before surgery; after admission to the ICU; and on postoperative days 3, 7, and 10.

**Main Outcome Measures:** APACHE II scores and predicted and observed mortality rates.

**Results:** The mean ± SD APACHE II score of 24.2 ± 8.3 at admission to the ICU was approximately 36% greater than the initial APACHE II score of 17.8 ± 7.7, a difference that was highly statistically significant (P < .001). The overall mortality of 32% favorably corresponds with the predicted mortality of 34% according to the initial APACHE II score. However, the predicted mortality of 50% according to the APACHE II score at admission to the ICU was significantly different from the observed mortality rate (P = .02). In 40 long-term patients (≥10 days in the ICU), the difference between the APACHE II scores of survivors and patients who died was statistically significant on day 10 (P = .04).

**Conclusions:** For risk stratification in emergency surgical patients, it is essential to measure the APACHE II score before surgical treatment. Longitudinal APACHE II scoring reveals continuous improvement of the score in surviving patients but has no therapeutic relevance in the individual patient.


Many authors have proposed using the APACHE II (Acute Physiology and Chronic Health Evaluation II) score to compare the severity of disease in patients with intra-abdominal infection to make results of surgical treatment comparable in terms of severity of illness.

The APACHE II score contains not only physiological variables but also age points and chronic health evaluation points for the measurement of comorbid diseases. Nevertheless physiological variables have the utmost importance in the APACHE II score. On the other hand, the APACHE II score offers no opportunity to evaluate interventions, although interventions might significantly alter many of the physiological variables.

Because the APACHE II score was designed as an intensive care unit (ICU) admission score, many studies with surgical patients use it in that way regardless of the requirement that the score has to be independent of treatment. Although the APACHE II score is an admission score, longitudinal APACHE II scoring is also reported.

There is general agreement that mortality is mainly related to an overwhelming cytokine-mediated inflammatory response, which may lead to multiple organ dysfunction syndrome. On the other hand, the inability of elderly and frail patients to immunologically react against infection may contribute to an anergic state, which can also be recorded by the APACHE II score.

Our aim was to elucidate the matter of timing in the calculation of the APACHE II score in emergency surgical patients and the clinical relevance of ICU admission and longitudinal APACHE II scores.
PATIENTS AND METHODS

We examined 85 patients admitted to the surgical ICU after undergoing emergency surgical procedures in 1999. Elective operations were excluded so that the sample comprised only patients, in whom prediction of postoperative mortality is desirable in particular. Another exclusion criterion was an APACHE II score less than 10 at admission to the ICU so that the sample comprised only patients with a significant physiological derangement.

The 50 male and 35 female patients had a mean±SD age of 68.1±13.9 years (range, 19-91 years).

The population consisted of 35 patients who underwent abdominal surgical procedures (34 patients with secondary peritonitis, 11 with gastrointestinal tract hemorrhage, and 10 with intestinal obstruction); 9 patients who underwent vascular surgery, especially for ruptured abdominal aortic aneurysms (n=5); and 21 patients who underwent surgery for multiple trauma (n=17) or head trauma (n=4). Head trauma does not need to be the only injury, but this categorization better reflects the patient’s injuries for individual risk calculation.17

All patients in the multiple trauma group had multiple osseous injuries requiring emergency surgical treatment but no head trauma.

Time between admission to the hospital and surgical treatment was always less than 8 hours. The APACHE II score was calculated before surgery; after admission to the ICU; and on postoperative days 3, 7, and 10. The score was obtained using the worst laboratory values during 24 hours after admission to the ICU and on days 3, 7, and 10, and the initial APACHE II score was obtained using the worst laboratory levels from hospital admission to surgical treatment.

Individual risk was calculated from the APACHE II score using the equation and category weights from Knaus et al.17 The predicted death rate for groups of patients was calculated by dividing the sum of the individual risks by the total number of patients.

For statistical analysis, the χ² test was used for qualitative variables; for continuous risk factors, 1-factor analysis of variance and multiple logistic regression were applied. Differences were considered statistically significant at P<.05. Data are given as mean±SD.

Because the APACHE II score is not specific for intra-abdominal infection, we intended to examine not only patients with secondary peritonitis but also those who underwent various emergency surgical procedures. Another aim was to compare the predicted mortality rate with the observed mortality rate at different points to demonstrate the best timing for the measurement of the APACHE II score.

RESULTS

The APACHE II score at the beginning of treatment was 17.8±7.7 (range, 5-43), which was statistically significantly different from the APACHE II score of 24.2±8.3 (range, 10-42) calculated at admission to the ICU (P<.001). Only 10 of 85 patients had an initial APACHE II score greater than or equal to the score measured at admission to the ICU. Overall, the APACHE II score measured at ICU admission was about 36% greater than the initially calculated score, which showed a similar difference for each disease (Table 1).

For APACHE II scores measured before and after surgical treatment and on postoperative days 3, 7, and 10, multiple logistic regression analysis revealed statistically significant relations to patient survival (P=.003-.0001). Age also significantly affects outcome (P=.03). Overall, the APACHE II scores on days 7 and 10 showed a significant decrease compared with the score calculated before surgical treatment (P=.008 and P=.01, respectively), whereas the APACHE II score on day 3 was not different from the initial score (P=.32) (Table 2). In patients who died, there was no improvement in the APACHE II score.

Overall mortality was 32% and was highest for patients who underwent surgery for gastrointestinal tract bleeding and lowest for those who underwent surgery for trauma, and the observed mortality rates were considerably lower for peritonitis and bowel obstruction compared with the predicted mortality rates according to the APACHE II score (Table 1). Overall predicted mortality according to the APACHE II score was 34%, which did not significantly differ from the observed mortality rate for any disease and for any range of the APACHE II score (Figure 1). The APACHE II score measured at ICU admission was significantly higher than the initial APACHE II score (Tables 1 and 2). Therefore, overall predicted mortality of 50% differed significantly from observed mortality of 32%, especially in patients with an APACHE II score of 25 or greater (P=.02) (Figure 2).

In 40 patients treated for 10 days or more in the ICU, the APACHE II scores showed no improvement on days 3, 7, and 10 in patients who died but continuously decreased in survivors. A statistically significant difference was achieved on day 10 (P=.04) (Figure 3).

The duration of surgery, 129.2±111.4 minutes (range, 20-855), varied considerably because of the different diseases and related surgical treatment.

The duration of ICU treatment was 14.7±16.9 days (range, 1-97), and the hospital stay was 27.8±23.3 days (range, 1-164). Eight of the 58 surviving patients needed nursing care at home after discharge from the hospital.

COMMENT

The APACHE II score is suitable for comparing studies with patients with the same disease in terms of severity of physiological derangement and related mortality rates. It is also possible to predict mortality rates in defined groups of patients using the APACHE II score and diagnostic category weights.2,8,17,18 However, a roughly accurate prediction of outcome for the individual patient is neither possible nor justified.9,10,19-27 Nevertheless, the APACHE II score has been used to indicate which operation should be done for patients with perforated ulcers20 and for patients with necrotizing pancreatitis.29
Clinical assessment for predicting mortality risk in patients admitted to an ICU shows similar sensitivity and specificity compared with various commonly used scores.26,30,31

The original APACHE II system was designed to predict mortality risks for stratification by assessing the patient independent of the effects of treatment, which therefore has to be before surgery. Nevertheless, the APACHE II score was primarily designed to stratify groups of ICU admissions prognostically.17 That is why some authors7,18,20,32-37 calculated the score after admission to the ICU regardless of preceding surgical treatment. Other investigators2,5,6,9,24,28,38-47 calculated the APACHE II score using laboratory values and physiological data before surgical treatment. Especially in surgical patients, interventions might keep physiological variables stable, while these patients have undergone considerable physiological stress. On the other hand, emergency surgery might increase the postoperative APACHE II score because of the inability to effectively mask the physiological derangement caused by operative stress. These questions were previously addressed, but the difference was not measured before in earlier literature.21 We focused our attention on emergency surgical patients because pretreat-

Table 1. APACHE II Scores Measured Initially (APACHE1) and After Admission to the Intensive Care Unit (APACHE2) According to Various Diseases and Associated Mortality Rates*

<table>
<thead>
<tr>
<th>Disease</th>
<th>Patients, No.</th>
<th>Observed Mortality, %</th>
<th>APACHE II Score, Mean ± SD (Predicted Mortality, %)</th>
<th>P †</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peritonitis</td>
<td>34</td>
<td>38</td>
<td>19.9 ± 7.9 (50)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>GI tract bleeding</td>
<td>11</td>
<td>46</td>
<td>21.4 ± 8.7 (38)</td>
<td>.02</td>
</tr>
<tr>
<td>GI tract obstruction</td>
<td>10</td>
<td>30</td>
<td>16.0 ± 5.8 (38)</td>
<td>.002</td>
</tr>
<tr>
<td>Trauma</td>
<td>21</td>
<td>19</td>
<td>16.6 ± 6.4 (15)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Vascular surgery</td>
<td>9</td>
<td>22</td>
<td>10.0 ± 4.0 (10)</td>
<td>.001</td>
</tr>
<tr>
<td>Total</td>
<td>85</td>
<td>32</td>
<td>17.8 ± 7.7 (34)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*APACHE II indicates Acute Physiology and Chronic Health Evaluation II; GI, gastrointestinal.
† The significance of the difference is between APACHE1 and APACHE2 (analysis of variance).

Table 2. APACHE II Scores in 85 Patients After Emergency Surgery and 25 Patients Who Died Measured Initially (APACHE1); After Admission to the Intensive Care Unit (APACHE2); and on Postoperative Days 3, 7, and 10 (APACHE3, 7 and 10, Respectively)*

<table>
<thead>
<tr>
<th>Measurement Point</th>
<th>Overall APACHE II Score, Mean ± SD (Range)</th>
<th>P †</th>
<th>APACHE II Score for Patients Who Died, Mean ± SD (Range)</th>
<th>P †</th>
</tr>
</thead>
<tbody>
<tr>
<td>APACHE1</td>
<td>17.8 ± 7.7 (5.0-43.0)</td>
<td></td>
<td>22.0 ± 9.3 (9.0-43.0)</td>
<td></td>
</tr>
<tr>
<td>APACHE2</td>
<td>24.2 ± 8.3 (10.0-42.0)</td>
<td>&lt;.001</td>
<td>28.9 ± 9.3 (12.0-42.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>APACHE3</td>
<td>17.0 ± 7.1 (5.0-39.0)</td>
<td>.32</td>
<td>22.5 ± 9.2 (5.0-39.0)</td>
<td>.56</td>
</tr>
<tr>
<td>APACHE7</td>
<td>15.1 ± 6.5 (2.0-35.0)</td>
<td>.008</td>
<td>21.0 ± 5.6 (14.0-35.0)</td>
<td>.60</td>
</tr>
<tr>
<td>APACHE10</td>
<td>15.1 ± 4.9 (5.0-25.0)</td>
<td>.01</td>
<td>20.5 ± 3.2 (15.0-25.0)</td>
<td>.32</td>
</tr>
</tbody>
</table>

*APACHE II indicates Acute Physiology and Chronic Health Evaluation II.
† Significance of the difference is related to APACHE1 (analysis of variance).

Clinical assessment for predicting mortality risk in patients admitted to an ICU shows similar sensitivity and specificity compared with various commonly used scores.26,30,31

The original APACHE II system was designed to predict mortality risks for stratification by assessing the patient independent of the effects of treatment, which therefore has to be before surgery. Nevertheless, the APACHE II score was primarily designed to stratify groups of ICU admissions prognostically.17 That is why some authors7,18,20,32-37 calculated the score after admission to the ICU regardless of preceding surgical treatment. Other investigators2,5,6,9,24,28,38-47 calculated the APACHE II score using laboratory values and physiological data before surgical treatment. Especially in surgical patients, interventions might keep physiological variables stable, while these patients have undergone considerable physiological stress. On the other hand, emergency surgery might increase the postoperative APACHE II score because of the inability to effectively mask the physiological derangement caused by operative stress. These questions were previously addressed, but the difference was not measured before in earlier literature.21 We focused our attention on emergency surgical patients because pretreat-

Figure 1. Observed and predicted mortality rates according to APACHE II (Acute Physiology and Chronic Health Evaluation II) scores calculated before surgery.

Figure 2. Observed and predicted mortality rates according to APACHE II (Acute Physiology and Chronic Health Evaluation II) scores calculated after admission to the intensive care unit. There was a significant difference between overall observed and predicted mortality rates and between observed and predicted mortality rates at the APACHE II score of 25 or greater (P = .02).
ment stratification for comparison of different groups of patients and prediction of mortality risk is of paramount importance in these patients. Comprising this selected group of patients, we could demonstrate a significant increase in the APACHE II score after surgical therapy.

If the APACHE II score is measured after admission to the ICU a lower mortality rate should be found compared with the predicted mortality rate according to the APACHE II score. Nevertheless, the mortality rate might also be equal to or higher than expected in patients in whom the APACHE II score has been calculated after surgery, which may be explained by a larger proportion of elective surgical patients or a larger proportion of patients transferred after initial treatment elsewhere.

After elective surgery, patients might have a lower APACHE II score at admission to the ICU than before surgical treatment because of preceeding interventions. In many studies, patients are poorly defined and measurement of the APACHE II score differs considerably, making results of treatment difficult or nearly impossible to compare.

Overall, the APACHE II score has been found variously to underestimate or overestimate death, especially in high-risk patients. Criticism about the prognostic strength of the APACHE II score also arose when postoperative peritonitis and necrotizing pancreatitis were found to be associated with considerably higher mortality rates than expected, which, however, may be because laboratory values had been corrected, resulting in too low APACHE II scores.

Aggressive surgical treatment in patients with severe intra-abdominal infection may also significantly decrease postoperative mortality, whereas the impossibility of eradicating the source of infection initially significantly increases the postoperative mortality rate. For a superior prediction of mortality in patients with intra-abdominal sepsis, the combination of the multiple organ dysfunction syndrome score according to Goris et al or the Mannheimer Peritonitis Index together with the APACHE II score has been proposed. For a higher predictive value, the APACHE III score also includes values for preliminary treatment before admission to the ICU.

In trauma patients, the observed mortality rate was significantly higher than the predicted mortality rate. These results suggest that the APACHE II score may not be applicable in these patients. Indeed, in our patients, the observed mortality rate was higher than the expected mortality rate but did not show a significant difference. It was also found that in patients with ruptured abdominal aortic aneurysms, the predicted mortality rate is too low and precise prediction of outcome is impossible.

According to our results, a continuous decrease in the APACHE II score in long-term patients from day 7 on is associated with a favorable outcome, whereas an increase in the APACHE II score strongly predicts death. However, in the individual patient, clinical significance of longitudinal APACHE II scoring seems to be limited, especially in making decisions regarding relaparotomies. Nevertheless, also in the individual patient, a trend toward survival or death may be seen after 7 to 10 days of ICU treatment. In any case, the APACHE II score loses its predictive value for outcome when it is calculated during the ICU stay.

Corresponding author and reprints: Thomas Koperna, MD, PhD, Department of Surgery, Hospital Mistelbach, Liechtensteinstrasse 67, A-2130 Mistelbach, Austria.

REFERENCES

17. Knauß WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of dis-
18. Barie PS, Hyde LJ, Fischer E. Development of multiple organ dysfunction syn-
drome in critically ill patients with perforated vescic: predictive value of APACHE
Acute Physiology and Chronic Health Evaluation (APACHE II) scoring in a sur-
20. Cerra FB, Negro F, Abrams J. APACHE II score does not predict multiple organ fail-
21. Civetta JM, Hudson-Civetta JA, Nelson LD. Evaluation of APACHE II score for cost con-
22. Dellinger EP, Wertz MJ, Meakins JL, et al. Surgical infection stratification sys-
23. Griebel GD, Troidl H. Möglichkeiten und Grenzen von Scores: theoretische Über-
24. Levison MA, Zeigler D. Correlation of APACHE II score drainage technique and
172:89-94.
25. Ohmann C, Wittmann DH, Wacha H, and the Peritonitis Study Group. Prospec-
tive analysis of multiple intervention and outcome variables in 1238 patients. Crit Care
Evaluation (APACHE II) score and outcome in the surgical intensive care unit: an
analysis of multiple intervention and outcome variables in 1238 patients. Crit Care
27. Wahl W, Pelletier K, Schmidtmann S, Jungtinger T. Erfahrungen mit ver-
28. Schein M. Planned reoperations and open management in critical intra-
abdominal infections: prospective experience in 52 cases. World J Surg. 1991;
15:537-545.
29. Schildberg FW. Peritonitis: einzeitig geschlossen mit Drainage oder offen mit programmierter
30. Schmidtmann S, Jungtinger T, Mialkowskyj O, Stokstad P, Schildberg FW. Peritonitis-
management with the etappenlavage (EL): Prognosekriterien und Behandlungsver-
31. Garcia-Sabrido JL, Tallado JM, Christou NV, Polo JR, Valdecantos E. Treatment
of severe intra-abdominal sepsis and/or necrotic foci by an "open-abdomen" ap-
führer Peritonitis: geplant vs. on-demand. Langenbecks Arch Chir. 1996;
381:343-347.
Society Intra-abdominal Infection Study: prospective evaluation of manage-
versus relaparotomy on demand in the treatment of intra-abdominal infections.
35. Pacelli F, Doglietto GB, Affieri S, et al. Prognosis in intra-abdominal infections:
36. Sawyer RG, Rosenlof JK, Adams RB, May AK, Spengler MD, Pruitt TL. Perito-
nitis into the 1990s: changing pathogens and changing strategies in the criti-
37. Schein M. Planned reoperations and open management in critical intra-
abdominal infections: prospective experience in 52 cases. World J Surg. 1991;
15:537-545.
38. Tichmann W, Wittmann DH, Andreone PA. Scheduled reoperations (etappen-
39. Wittmann DH, Aparhamian G, Bergstein JM. Etappenlavage: advanced diffuse peri-
tonitis managed by multiple laparatomies utilizing zippers, slide fasteners and
Velcro analogue for temporary abdominal closure. World J Surg. 1990;14:218-
226.
40. Wittmann DH, Bansal N, Bergstein JM, Wallace JR, Wittmann MM, Aparhamian
C. Staged abdominal repair compares favorably with conventional operative therapy
for intra-abdominal infections when adjusting for prognostic factors with a lo-
41. Goris RJA, te Boekhorst TPA, Nuytinck KS, Gimbrere JSF. Multiple-organ fail-
ure: generalized autodestructive inflammation? Arch Surg. 1985;120:1109-
1115.
Mannheimer Peritonitis Index: ein Instrument zur intraoperativen Prognose der
43. Knauß WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system:
risk prediction of hospital mortality for critically ill hospitalized adults. Chest
1991; 100:1619-1636.
44. Cachecho R, Clas D, Gersin K, Grindlinger GA. Evolution in the management of