Serial Evaluation of the SOFA Score to Predict Outcome in Critically Ill Patients

Flavio Lopes Ferreira, MD
Duliana Peres Bota, MD
Annette Bross, MD
Christian Mélot, MD, PhD, MSciBiostat
Jean-Louis Vincent, MD, PhD

Outcome prediction is important both in clinical and administrative intensive care unit (ICU) management. Although outcome prediction and measurement should not be the only measure of ICU performance, outcome prediction can be usefully applied to monitor the performance of an individual ICU and possibly to compare ICUs. Outcome prediction can also be useful in providing information on likely patient outcomes for relatives of critically ill patients and potentially for therapeutic decision making and guiding resource allocation. Outcome prediction models currently available have not been validated for use in directing individual patient management.

Currently available outcome prediction models (such as the APACHE [Acute Physiology and Chronic Health Evaluation], SAPS [simplified acute physiology score], and MPM [mortality probability models] systems) calculate a prediction on values taken within the first 24 hours of an ICU stay. However, these scores ignore the many factors that can influence patient outcome during the course of an ICU stay. Being able to evaluate changes in patient status over time thus represents an improvement on standard models.

Organ dysfunction is associated with high rates of ICU morbidity and mortality, and, as such, accounts for a high proportion of the ICU budget. Recently developed organ failure scores, such as the Sequential Organ Failure Assessment (SOFA) can help assess organ dysfunction or failure over time and are useful to evaluate morbidity. Although these scoring systems were developed to describe and quantify outcomes, they have not been validated for use in directing ICU management.

Context  Evaluation of trends in organ dysfunction in critically ill patients may help predict outcome.

Objective  To determine the usefulness of repeated measurement the Sequential Organ Failure Assessment (SOFA) score for prediction of mortality in intensive care unit (ICU) patients.

Design  Prospective, observational cohort study conducted from April 1 to July 31, 1999.

Setting  A 31-bed medicosurgical ICU at a university hospital in Belgium.

Patients  Three hundred fifty-two consecutive patients (mean age, 59 years) admitted to the ICU for more than 24 hours for whom the SOFA score was calculated on admission and every 48 hours until discharge.

Main Outcome Measures  Initial SOFA score (0-24), Δ-SOFA scores (differences between subsequent scores), and the highest and mean SOFA scores obtained during the ICU stay and their correlations with mortality.

Results  The initial, highest, and mean SOFA scores correlated well with mortality. Initial and highest scores of more than 11 or mean scores of more than 5 corresponded to mortality of more than 80%. The predictive value of the mean score was independent of the length of ICU stay. In univariate analysis, mean and highest SOFA scores had the strongest correlation with mortality, followed by Δ-SOFA and initial SOFA scores. The area under the receiver operating characteristic curve was largest for highest scores (0.90; SE, 0.02; P<.001 vs initial score). When analyzing trends in the SOFA score during the first 96 hours, regardless of the initial score, the mortality rate was at least 50% when the score increased, 27% to 35% when it remained unchanged, and less than 27% when it decreased. Differences in mortality were better predicted in the first 48 hours than in the subsequent 48 hours. There was no significant difference in the length of stay among these groups. Except for initial scores of more than 11 (mortality rate >90%), a decreasing score during the first 48 hours was associated with a mortality rate of less than 6%, while an unchanged or increasing score was associated with a mortality rate of 37% when the initial score was 2 to 7 and 60% when the initial score was 8 to 11.

Conclusions  Sequential assessment of organ dysfunction during the first few days of ICU admission is a good indicator of prognosis. Both the mean and highest SOFA scores are particularly useful predictors of outcome. Independent of the initial score, an increase in SOFA score during the first 48 hours in the ICU predicts a mortality rate of at least 50%.

*Context*  Evaluation of trends in organ dysfunction in critically ill patients may help predict outcome.

*Objective*  To determine the usefulness of repeated measurement the Sequential Organ Failure Assessment (SOFA) score for prediction of mortality in intensive care unit (ICU) patients.

*Design*  Prospective, observational cohort study conducted from April 1 to July 31, 1999.

*Setting*  A 31-bed medicosurgical ICU at a university hospital in Belgium.

*Patients*  Three hundred fifty-two consecutive patients (mean age, 59 years) admitted to the ICU for more than 24 hours for whom the SOFA score was calculated on admission and every 48 hours until discharge.

*Main Outcome Measures*  Initial SOFA score (0-24), Δ-SOFA scores (differences between subsequent scores), and the highest and mean SOFA scores obtained during the ICU stay and their correlations with mortality.

*Results*  The initial, highest, and mean SOFA scores correlated well with mortality. Initial and highest scores of more than 11 or mean scores of more than 5 corresponded to mortality of more than 80%. The predictive value of the mean score was independent of the length of ICU stay. In univariate analysis, mean and highest SOFA scores had the strongest correlation with mortality, followed by Δ-SOFA and initial SOFA scores. The area under the receiver operating characteristic curve was largest for highest scores (0.90; SE, 0.02; P<.001 vs initial score). When analyzing trends in the SOFA score during the first 96 hours, regardless of the initial score, the mortality rate was at least 50% when the score increased, 27% to 35% when it remained unchanged, and less than 27% when it decreased. Differences in mortality were better predicted in the first 48 hours than in the subsequent 48 hours. There was no significant difference in the length of stay among these groups. Except for initial scores of more than 11 (mortality rate >90%), a decreasing score during the first 48 hours was associated with a mortality rate of less than 6%, while an unchanged or increasing score was associated with a mortality rate of 37% when the initial score was 2 to 7 and 60% when the initial score was 8 to 11.

*Conclusions*  Sequential assessment of organ dysfunction during the first few days of ICU admission is a good indicator of prognosis. Both the mean and highest SOFA scores are particularly useful predictors of outcome. Independent of the initial score, an increase in SOFA score during the first 48 hours in the ICU predicts a mortality rate of at least 50%.
gan function and not to predict outcome, the obvious relationship between organ dysfunction and mortality has been demonstrated in several studies. We were interested in evaluating whether repeated measurement of the SOFA score, by including alterations over time, could help refine outcome prediction.

**METHODS**

Following approval by the ethical review board of Erasme University Hospital, Free University of Brussels, Belgium, which waived informed consent on the basis that this was an epidemiological study without intervention, all patients (>18 years) admitted to the 31-bed medicsurgical department of intensive care for more than 24 hours during a 4-month period (April 1–July 31, 1999) were included in the study.

Demographic, laboratory, and clinical data were collected, and the SOFA score (0-24, Table 1) was calculated, on admission and every 48 hours until discharge. In the calculation of the score, the worst values for each parameter in the 24-hour period were used. For a single missing value, a replacement was calculated from the mean of the sum of the results immediately preceding and following the missing value. In se-

---

### Table 1. The Sequential Organ Failure Assessment (SOFA) Score*

<table>
<thead>
<tr>
<th>Variables</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respiratory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pao2/FIO2, mm Hg</td>
<td>&gt;400</td>
<td>≤400</td>
<td>≤300</td>
<td>≤200</td>
<td>≤100</td>
</tr>
<tr>
<td><strong>Coagulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelets ×10^9/μL‡</td>
<td>&gt;150</td>
<td>≤150</td>
<td>≤100</td>
<td>≤50</td>
<td>≤20</td>
</tr>
<tr>
<td><strong>Liver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubin, mg/dL‡</td>
<td>&lt;1.2</td>
<td>1.2-1.9</td>
<td>2.0-5.9</td>
<td>6.0-11.9</td>
<td>&gt;12.0</td>
</tr>
<tr>
<td><strong>Cardiovascular</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypotension</td>
<td>No hypotension</td>
<td>Mean arterial pressure &lt;70 mm Hg</td>
<td>Dop ≥5 or dob (any dose)§</td>
<td>Dop &gt;5, epi ≤0.1, or norepi ≤0.1§</td>
<td>Dop &gt;15, epi &gt;0.1, or norepi &gt;0.1§</td>
</tr>
<tr>
<td><strong>Renal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine, mg/dL or urine output, mL/d</td>
<td>&lt;1.2</td>
<td>1.2-1.9</td>
<td>2.0-3.4</td>
<td>3.5-4.9 or &lt;500</td>
<td>&gt;5.0 or &lt;200</td>
</tr>
</tbody>
</table>

* Norepi indicates norepinephrine; Dob, dobutamine; Dop, dopamine; Epi, epinephrine; and FIO2, fraction of inspired oxygen.
† Values are with respiratory support.
‡ To convert bilirubin from mg/dL to µmol/L, multiply by 17.1.
§ Adrenergic agents administered for at least 1 hour (doses given are in µg/kg per minute).
||To convert creatinine from mg/dL to µmol/L, multiply by 88.4.

---

### Table 2. Demographics of Study Population*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>352</td>
</tr>
<tr>
<td>Age, mean (SD) [range], y</td>
<td>59 (17) [18-95]</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>230</td>
</tr>
<tr>
<td>Women</td>
<td>122</td>
</tr>
<tr>
<td>Type of admission, No. (%)</td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>195 (55.4)</td>
</tr>
<tr>
<td>Surgical</td>
<td>157 (44.6)</td>
</tr>
<tr>
<td>Length of ICU stay, d</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>6.5</td>
</tr>
<tr>
<td>Median</td>
<td>4.0</td>
</tr>
<tr>
<td>Range</td>
<td>1-56</td>
</tr>
<tr>
<td>No. (%) of deaths</td>
<td>81 (23)</td>
</tr>
</tbody>
</table>

*IICU indicates intensive care unit.

---

**RESULTS**

The study included 352 patients with a mean (SD) age of 59 (17) years (Table 2). From the expected 13620 variables, 267 were missing (215 bilirubin levels, 21 creatinine concentrations, 15 Pao2/Fio2 [fraction of inspired oxygen] ratios, and 16 platelet counts).

As expected, the initial SOFA score was significantly related to vital status. An initial SOFA score up to 9 predicted a mortality of less than 33% while an initial SOFA score of greater than 11 predicted a mortality rate of 95% (Figure 1A). The highest SOFA score was also correlated with mortality: highest scores of 10 correlated with a mortality rate of 40% while those higher than 11 were associated with a mortality rate of...
greater than 80% (Figure 1B). The mean SOFA score over the entire ICU stay was also correlated with mortality (Figure 1C). The predictive value of the mean SOFA score for mortality was similar regardless of the LOS.

By univariate logistic analysis, the mean SOFA score correlated most closely with mortality (Table 3), followed by the highest score, the Δ-SOFA 48-0 score, and the initial score. The highest SOFA score presented the largest area under the ROC curve (0.90, SE 0.02) compared with the other SOFA-derived variables, followed by the mean SOFA score (area under ROC curve 0.88, SE 0.03). The area under the ROC curve was significantly larger for the highest SOFA score than for the initial SOFA score (P <.001, Figure 2).

Trends in SOFA scores during the first 48 hours were also analyzed. Regardless of the initial score, the mortality rate was 50% or higher when the score increased, 27% to 35% when it did not change, and less than 27% when it decreased (Table 4). Differences in mortality were predicted better during the first 48 hours than in the subsequent 48 hours. There was no significant difference in LOS among these groups. When we analyzed this trend, taking into account the initial SOFA score for values of 11 or lower, a decreasing value was associated with a mortality rate of 6% or less (Figure 3). However, when the mean SOFA score increased or remained unchanged, the mortality rate averaged 37% when the initial SOFA scores ranged from 2 to 7, 60% when the initial SOFA scores ranged from 8 to 11, and 91% when the initial SOFA score was higher than 11.

**COMMENT**

In developing a scoring system, such as SOFA, for assessing and monitoring organ dysfunction, several important features need to be considered. First, organ failure is not an all-or-nothing phenomenon; rather, it is a continuum of alterations in organ function from normal function, through varying degrees of dysfunction, to organ failure. Second, the description of organ dysfunction needs to be based on simple, easily repeatable variables specific to the organ in question and readily available in all institutions. Third, organ dysfunction is not static. It will alter over time, and a scoring system needs to be able to take this time factor into account. In using the SOFA for outcome prediction, the ability to perform serial SOFA scores allow a more effective representation of the dynamics of illness including the effects of therapy compared with traditional outcome prediction models at the time of ICU admission. Although some investigators have used the APACHE II score over time, this process has never been validated. Derived measures from the APACHE III system have also been proposed for use on a daily basis, but APACHE III is not available in the public domain, and its daily use has again not been validated.

The SOFA score is a useful tool to stratify and compare patients in clinical trials. Moreno et al recently demonstrated that the initial SOFA score can be used to quantify the degree of organ dysfunction or failure present on admission, that the Δ-SOFA score can demonstrate the degree of dysfunction or failure developing during an ICU stay, and that the total maximum SOFA score can represent the cumulative organ dysfunction experienced by the patient. They also demonstrated a strong correlation of all these parameters with mortality outcome.

In our study, we have moved a step further, presenting selected SOFA
parameters, the mean and the highest SOFA scores, as reliable predictors of outcome throughout the ICU stay. The mean SOFA score gives an indication of the average degree of organ failure over time and could also be a useful tool for stratifying patients in clinical trials, according to the total score or the scores for individual organs. The highest SOFA score can identify the critical point at which patients exhibit the highest degree of organ dysfunction during their ICU stay. With these 2 variables, we can thus define the peak and the total amount of organ impairment for any patient or group of patients during their ICU stay. The equivalence of the areas under the ROC curve for these 2 parameters suggest that they are similarly effective in predicting outcome.

The Δ-SOFA score could be used to reflect patient response to therapeutic strategies and allow the physician to monitor daily progress, offering an objective evaluation treatment responses. For example, knowledge of the trend in SOFA score over time could facilitate decision making regarding the appropriateness of instituting organ support. Knowing that a decreasing SOFA score is associated with an improved outcome should prompt aggressive early therapy, which may reduce mortality.20 Others have shown that the development of organ failure may occur early during an ICU stay,21 and a scoring system that allows regular surveillance of organ function is thus needed. Trends in the SOFA score over the first 48 hours of an ICU stay could provide such a system and be a sensitive indicator of outcome, as reflected in the fact that a decreasing value was associated with a decrease in mortality rates from 50% to 27%.

Interestingly, the LOS was not related to outcome prediction. Indeed, the mean SOFA score had a better prognostic value than the other SOFA derived variables. This may be because patients who present with a limited degree of organ dysfunction and have a long ICU stay still have a high likelihood of survival.

In conclusion, evaluation of the SOFA score throughout the ICU stay

---

**Figure 2.** Comparisons of the Areas Under the Receiver Operating Characteristic (ROC) Curves for Prediction of Mortality

The area and the 95% confidence interval are presented in each panel. The Δ scores represent difference between the 48-hour SOFA score and the admission score and the difference between the 96-hour SOFA score and the admission score. Data markers are the optimal threshold for each SOFA score that discriminates between survival and nonsurvival.
**REFERENCES**


