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

Flavio Lopes Ferreira, MD

Daliana 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.[1] 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.[2]

Currently available outcome prediction models (such as the APACHE [Acute Physiology and Chronic Health Evaluation],[3] SAPS [simplified acute physiology score],[4] and MPM [mortality probability models][5] 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,[6,7] and, as such, accounts for a high

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%.

JAMA. 2001;286:1754-1758   www.jama.com

proportion of the ICU budget.[7] Recently developed organ failure scores, such as the Sequential Organ Failure Assessment (SOFA) (**TABLE 1**)[8] 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 or-

Author Affiliation: Department of Intensive Care, Erasme University Hospital, Free University of Brussels, Belgium.
Corresponding Author and Reprints: Jean-Louis Vincent, MD, PhD, Department of Intensive Care, Erasme University Hospital, Route de Lennik, 808, B-1070 Brussels, Belgium (e-mail: jlvincent@ulb.ac.be).
Caring for the Critically Ill Patient Section Editor: Deborah J. Cook, MD, Consulting Editor, *JAMA*.
Advisory Board: David Bihari, MD; Christian Brun-Buisson, MD; Timothy Evans, MD; John Heffner, MD; Norman Paradis, MD.

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

Variables	SOFA Score				
	0	1	2	3	4
Respiratory Pao$_2$/Fio$_2$, mm Hg	>400	≤400	≤300	≤200†	≤100†
Coagulation Platelets ×10^3/µL‡	>150	≤150	≤100	≤50	≤20
Liver Bilirubin, mg/dL‡	<1.2	1.2-1.9	2.0-5.9	6.0-11.9	>12.0
Cardiovascular Hypotension	No hypotension	Mean arterial pressure <70 mm Hg	Dop ≤5 or dob (any dose)§	Dop >5, epi ≤0.1, or norepi ≤0.1§	Dop >15, epi >0.1, or norepi >0.1§
Central nervous system Glasgow Coma Score Scale	15	13-14	10-12	6-9	<6
Renal Creatinine, mg/dL or urine output, mL/d‖	<1.2	1.2-1.9	2.0-3.4	3.5-4.9 or <500	>5.0 or <200

*Norepi indicates norepinephrine; Dob, dobutamine; Dop, dopamine; Epi, epinephrine; and Fio$_2$, 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.

gan function and not to predict outcome, the obvious relationship between organ dysfunction and mortality has been demonstrated in several studies.[9-12] 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 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 medicosurgical 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-

dated patients, the assumed Glasgow Coma Score Scale was used to evaluate the neurological status.

The total SOFA was calculated as the sum of all daily SOFA scores during the ICU stay for each patient. The mean score was defined as the ratio of total score to the length of stay (LOS) in the ICU. The highest score recorded during the ICU stay was also noted. The Δ-SOFA score was defined as the difference between 2 subsequent scores; for example, the Δ-SOFA score 48-0 was the difference between the 48-hour SOFA score and the admission score.

Odds ratios with 95% confidence intervals were computed using a univariate logistic regression model with ICU outcome as the dependent variable. A χ^2 statistics test (with Yates correction when applicable) was used to evaluate the statistical significance of categorical variables. The results are presented as mean (SD). Comparisons of the areas under the receiver operating characteristic (ROC) curves were also performed with a test based on the difference between the 2 areas and the SE of the difference.[13] Using Statview (SAS Institute, Cary, NC) and Medcal (Medcal Software, Mariakerke, Belgium), all statistical tests were 2-tailed and a *P* value <.05 was considered significant.
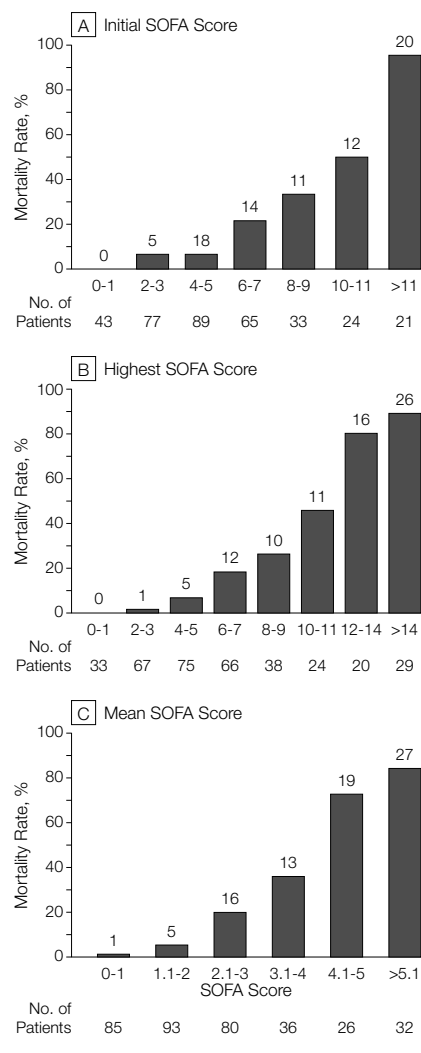
Table 2. Demographics of Study Population*

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

*ICU indicates intensive care unit.

RESULTS

The study included 352 patients with a mean (SD) age of 59 (17) years (**TABLE 2**). From the expected 13 620 variables, 267 were missing (215 bilirubin levels, 21 creatinine concentrations, 15 Pao$_2$/Fio$_2$ [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 1**A). 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

Figure 1. Mortality Rate in Relation to the Changes in Sequential Organ Failure Assessment (SOFA) Score During the First 48 Hours in the Intensive Care Unit

Numbers above the bars indicate number of deaths.

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 ini-

tial 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,[14-16] this process has never been validated. Derived measures from the APACHE III system have also been proposed for use on a daily basis,[17] 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.[18,19] Moreno et al[12] 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

Table 3. Univariate Logistic Regression Analysis of Length of Stay and Sequential Organ Failure Assessment (SOFA) Derived Parameters as Predictors of Mortality

Variables	Coefficient, Mean (SE)	Odds Ratio (95% Confidence Interval)	P Value
Mean SOFA score	1.12 (0.13)	3.06 (2.36-3.97)	<.001
Highest SOFA score	0.46 (0.05)	1.59 (1.43-1.76)	<.001
Δ-SOFA score, 48-0*	0.42 (0.08)	1.52 (1.29-1.78)	<.001
Initial SOFA score	0.37 (0.05)	1.45 (1.32-1.59)	<.001
SOFA score at 48 h	0.37 (0.06)	1.45 (1.30-1.61)	<.001
SOFA score at 96 h	0.33 (0.06)	1.39 (1.22-1.57)	<.001
Δ-SOFA score, 96-0†	0.21 (0.09)	1.24 (1.04-1.47)	.02
Length of stay	0.07 (0.02)	1.07 (1.03-1.11)	<.001
Total SOFA score	0.04 (0.01)	1.05 (1.03-1.06)	<.001

*Represents the difference between the 48-hour SOFA score and the admission score.
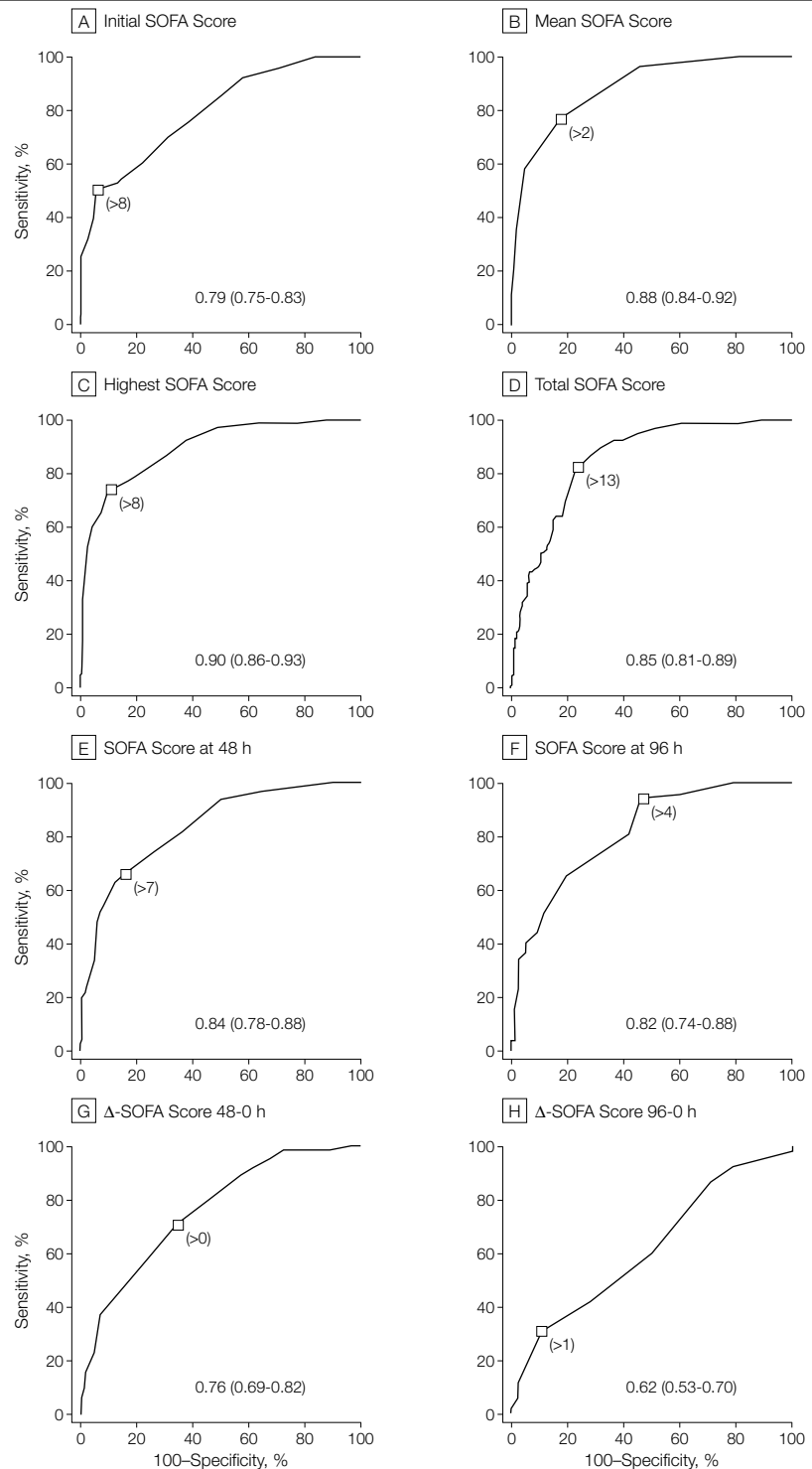†Represents the difference between the 96-hour SOFA score and the admission score.

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
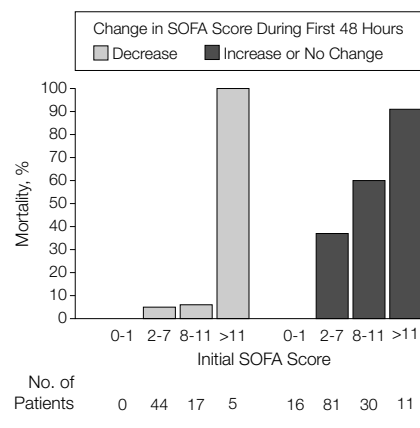

The area and the 95% confidence interval are presented in each panel. The Δ scores represent difference between the 48-hour SOFA score and the admission score and the difference between the 96-hour SOFA score and the admission score. Data markers are the optimal threshold for each SOFA score that discriminates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

Table 4. Changes in Sequential Organ Failure Assessment (SOFA) Score in Relation to Outcome*

First 48 Hours					Next 48 Hours				Average % of Deaths Over First 96 h
Evolution	No. at Risk	% of Deaths (SE)	95% CI	Mean LOS (SE)	Evolution	No. at Risk	% of Deaths	Total % of Deaths	
Increased	66	53 (6.1)	41-65	12.4 (6.4)	Increased	22	67	57	>50
					Unchanged	12	5	53	
					Decreased	32	44	50	
Unchanged	32	31 (8.2)	15-47	12.6 (9.5)	Increased	9	33	32	27-35
					Unchanged	9	11	27	
					Decreased	14	43	35	
Decreased	30	23 (7.7)†	8-38	10.9 (9.3)	Increased	12	33	26	<27
					Unchanged	6	33	25	
					Decreased	12	8	19	

*LOS indicates length of stay.
†$P = .01$, increased vs decreased.

Figure 3. Relation Between Mortality Rates and Sequential Organ Failure Assessment (SOFA) Scores

Change in SOFA Score During First 48 Hours
Decrease / Increase or No Change

| No. of Patients | 0 | 44 | 17 | 5 | 16 | 81 | 30 | 11 |

is a good prognostic indicator (especially the mean and the highest SOFA scores). Independent of the initial value, an increase in the SOFA score during the first 48 hours of ICU admission predicts a mortality rate of at least 50%.

Author Contributions: *Study concept and design:* Vincent.
Acquisition of data: Ferreira, Bota, Bross.
Analysis and interpretation of data: Ferreira, Bota, Mélot.
Drafting of the manuscript: Ferreira, Vincent.
Critical revision of the manuscript for important intellectual content: Bota.

REFERENCES

1. Shortell SM, Zimmerman JE, Rousseau DM, et al. The performance of intensive care units: does good management make a difference? *Med Care*. 1994; 32:508-525.
2. Cullen DJ, Chernow B. Predicting outcome in critically ill patients. *Crit Care Med*. 1994;22:1345-1348.
3. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-Acute Physiology and Chronic Health Evaluation: a physiologically based classification system. *Crit Care Med*. 1981;9:591-597.
4. Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270:2957-2963.
5. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care unit outcome prediction by using changing probabilities of mortality. *Crit Care Med*. 1988;16:470-477.
6. Tran DD, Groeneveld ABJ, Vander Meulen J, Nauta JJP, Strack Van Schijndel RJM, Thijs LG. Age, chronic disease, sepsis, organ system failure, and mortality in a medical intensive care unit. *Crit Care Med*. 1990; 18:474-479.
7. Deitch EA. Multiple organ failure: pathophysiology and potential future therapy. *Ann Surg*. 1992; 216:117-134.
8. Vincent JL, Moreno R, Takala J, et al, for the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707-710.
9. Regel G, Grotz M, Weltner T, Sturm JA, Tscherne H. Pattern of organ failure following severe trauma. *World J Surg*. 1996;20:422-429.
10. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicentric, prospective study. *Crit Care Med*. 1998;26:1793-1800.
11. Antonelli M, Moreno R, Vincent JL, et al. Application of SOFA score to trauma patients: Sequential Organ Failure Assessment. *Intensive Care Med*. 1999; 25:389-394.
12. Moreno R, Vincent JL, Matos A, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care: results of a prospective, multicentre study. *Intensive Care Med*. 1999;25:686-696.
13. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*. 1993;39:561-577.
14. Bion JF, Aitchison TC, Edlin SA, Ledingham IM. Sickness scoring and response to treatment as predictors of outcome from critical illness. *Intensive Care Med*. 1988;14:167-172.
15. Larvin M, McMahon MJ. APACHE-II score for assessment and monitoring of acute pancreatitis. *Lancet*. 1989;2:201-205.
16. Sawyer RG, Rosenlof LK, Adams RB, May AK, Spengler MD, Pruett TL. Peritonitis into the 1990s: changing pathogens and changing strategies in the critically ill. *Am Surg*. 1992;58:82-87.
17. Wagner DP, Knaus WA, Harrell FE, Zimmerman JE, Watts C. Daily prognostic estimates for critically ill adults in intensive care units: results from a prospective multicenter, inception cohort analysis. *Crit Care Med*. 1994;22:1359-1372.
18. Di Filippo A, De Gaudio AR, Novelli A, et al. Continuous infusion of vancomycin in methicillin-resistant staphylococcus infection. *Chemotherapy*. 1998;44:63-68.
19. Hynninen M, Valtonen M, Markkanen H, et al. Interleukin 1 receptor antagonist and E-selectin concentrations: a comparison in patients with severe acute pancreatitis and severe sepsis. *J Crit Care*. 1999;14: 63-68.
20. Goldhill DR, Sumner A. Outcome of intensive care patients in a group of British intensive care units. *Crit Care Med*. 1998;26:1337-1345.
21. Cryer HG, Leong K, McArthur DL, et al. Multiple organ failure: by the time you predict it, it's already there. *J Trauma*. 1999;46:597-604.