The Logistic Organ Dysfunction System

A New Way to Assess Organ Dysfunction in the Intensive Care Unit

Jean-Roger Le Gall, MD; Janelle Klar, MS; Stanley Lemeshow, PhD; Fabienne Saulnier, MD; Corinne Alberti, MD; Antonio Artigas, MD; Daniel Teres, MD; for the ICU Scoring Group

Objective.—To develop an objective method for assessing organ dysfunction among intensive care unit (ICU) patients on the first day of the ICU stay.

Design and Setting.—Physiological variables defined dysfunction in 6 organ systems. Logistic regression techniques were used to determine severity levels and relative weights for the Logistic Organ Dysfunction (LOD) score and for conversion of the LOD score to a probability of mortality.

Patients.—A total of 13 152 consecutive admissions to 137 adult medical/surgical ICUs in 12 countries from the European/North American Study of Severity Systems.

Outcome Measure.—Patient vital status at hospital discharge.

Results.—The LOD System identified from 1 to 3 levels of organ dysfunction for 6 organ systems: neurologic, cardiovascular, renal, pulmonary, hematologic, and hepatic. From 1 to 5 LOD points were assigned to the levels of severity, and the resulting LOD scores ranged from 0 to 22 points. Model calibration was very good in the developmental and validation samples (P=.21 and P=.50, respectively), as was model discrimination (area under the receiver operating characteristic curves of 0.843 and 0.850, respectively).

Conclusion.—The LOD System provides an objective tool for assessing severity levels for organ dysfunction in the ICU, a critical component in the conduct of clinical trials. Neurologic, cardiovascular, and renal dysfunction were the most severe organ dysfunctions, followed by pulmonary and hematologic dysfunction, with hepatic dysfunction the least severe. The LOD System takes into account both the relative severity among organ systems and the degree of severity within an organ system.

JAMA. 1996;276:802-810

NEW SYSTEMS to assess organ dysfunction are proposed almost every year, and each system differs from the others in large or small ways. As early as 1980, Fry et al[1] proposed a system of 4 organ failures for surgical patients: pulmonary, hepatic, gastrointestinal, and renal failure. In 1983, Stevens[2] described the Sepsis Severity Score comprising 7 failures, each with 5 severity levels. The system of Marshall et al[3] contained metabolic failure and took anergy into account. The widely used Organ System Failure (OSF) score was published in 1983 by Knaus et al,[4] and in 1989 hepatic failure was added.[5] Fagon et al[6] added infection to the assessment of organ dysfunction and called their system ODIN (Organ Dysfunctions and/or Infection). Hebert et al[7] published a multiple organ failure scoring system for patients who have sepsis syndrome. Recently, Marshall et al[8] proposed the Multiple Organ Dysfunction Score (MODS) based on a review of 30 reports in the literature.

The aim of our study was to create an objectively derived system from a large database of intensive care unit (ICU) patients, using the statistical technique of multiple logistic regression. Although based on sophisticated statistical methods, our goal was to develop a system that was as simple as possible to apply in the ICU. In developing a statistically based system, ranges and weights of the variables defining levels of organ dysfunction can be determined objectively, the significance of severity levels for each organ can be identified, and the levels of dysfunction can be weighted according to their relative prognostic significance.

In the resulting Logistic Organ Dysfunction (LOD) System, the points for individual severity levels of each organ system reflect both the relative severity of the levels within an organ system and the relative severity of the levels among organ systems. The LOD score

From the Intensive Care Unit, Hôpital Saint-Louis, Paris, France (Drs Le Gall and Alberti); the School of Public Health, University of Massachusetts, Amherst (Ms Klar and Dr Lemeshow); Intensive Care Unit, Hôpital Calmette, Lille, France (Dr Saulnier); Servei de Medicina Intensiva, Consultation Hospitalari Parc Tauli, Sabadell, Spain (Dr Artigas); and the Critical Care Division, Baystate Medical Center, Springfield, Mass, and Tufts University School of Medicine, Boston, Mass (Dr Teres).

A complete list of members of the ICU Scoring Group was previously published in *JAMA* (1993;270:2957-2963).

Reprints: Jean-Roger Le Gall, MD, Service de Réanimation Médicale, Hôpital Saint-Louis, 1 avenue Claude Vellefaux, 75475 Paris Cedex 10, France.

Concepts in Emergency and Critical Care section editor: Roger C. Bone, MD, Consulting Editor, *JAMA*.

Advisory Panel: Bart Chernow, MD, Baltimore, Md; David Dantzker, MD, New Hyde Park, NY; Jerrold Leiken, MD, Chicago, Ill; Joseph E. Parrillo, MD, Chicago, Ill; William J. Sibbald, MD, London, Ontario; and Jean-Louis Vincent, MD, PhD, Brussels, Belgium.

is a global score that can be calculated to summarize the combined effect of dysfunction among several organs. In addition, the LOD model is a logistic regression equation that can be used to translate the score into a probability of mortality based on organ dysfunctions.

METHODS

The database for this analysis was assembled as part of the European/North American Study of Severity Systems (ENAS) that was used to develop the Simplified Acute Physiology Score (SAPS II) system for estimating the probability of mortality among ICU patients.[9] Data on 14 745 consecutive ICU admissions were collected in 137 medical, surgical, or mixed ICUs in 12 countries. Eligible patients were aged 18 years or older; burn patients, coronary care patients, and cardiac surgery patients were excluded. To develop and validate the LOD System, 80% of the patients in the database were randomly selected to constitute the developmental sample, and the remaining 20% composed the validation sample. As with the development of SAPS II, differences by site were not a consideration in the development of the system. A detailed description of the data collection procedures is given in Le Gall et al.[9]

Variables were extracted from the database to define organ dysfunction, based on a combination of 12 variables for 6 organ systems: neurologic system (Glasgow Coma Score [GCS]), cardiovascular system (heart rate and systolic blood pressure), renal system (urea and creatinine levels, and urine output), pulmonary system (ventilation/CPAP [continuous positive airway pressure] status and Pao_2/Fio_2 [fraction of inspired oxygen] ratio), hematologic system (white blood cell and platelet counts), and hepatic system (bilirubin level and prothrombin time). The variables had been recorded as the worst value in the first 24-hour period in the ICU. The worst value was defined as the value that would have been assigned the greatest number of SAPS points in the original SAPS system. For sedated patients, the GCS was ascertained either from interviewing the physician who ordered the sedation, or by reviewing the patient's medical record before sedation. If a variable was not measured for a patient, it was assumed to be within the range of normal. All variables except platelet counts and prothrombin time (PT) were continuously scaled. Platelet counts were recorded as being less than 50×10^9/L, and PT was recorded as being more than 3 seconds over standard or less than 25% of standard.

Similar to the methodology for developing SAPS II, the analysis was designed to first identify cut points that defined variable ranges associated with changes in mortality rate. That is, the continuum of measurements for a variable was divided into discrete categories, 1 of which was defined as being within normal limits. The other categories were each defined by upper and lower cut points, and each category represented a range of values that had a higher association with mortality than the range considered to be within normal limits. Subsequent analyses associated a point value with each range.

To identify the cut points that would define the ranges of severity for each variable, the continuous independent variables were plotted against the dependent variable (vital status at hospital discharge), and the LOWESS smoothing function, using locally weighted least squares, was used to suggest the cut points.[10] The plot of this smoothed function shows values of the independent variable that are associated with changes in the predicted value of the dependent variable. For example, if the plot of the smoothed function showed an increased association with mortality when the daily urine output declined to less than 0.5 L, a cut point of 0.5 L could be chosen to define a boundary for a range of severity.

Once the cut points for each variable were identified, dummy variables were created for each range outside the normal range. The dummy variable for each range took on the value of 1 for patients whose value for the variable fell within the range and the value of 0 for all other patients. Patients with values in the normal range composed the referent group and took on the value of 0 for the dummy variables.

The variables were entered simultaneously into a multiple logistic regression model. Since the variables all were valued at 0 or 1, multiple logistic regression modeling produced an equation in which each variable had an associated coefficient (β) that was a measure of the relative weight of that level of the variable while controlling for all other variables in the model. The βs for the variables were grouped by magnitude into levels of increasing severity. The number of severity levels was not preestablished, but was based on the observed range of the coefficients.

In defining the severity levels by the size of the coefficients, comparable severity levels for the 2 or 3 individual continuous variables defining an organ dysfunction (eg, white blood cell [WBC] and platelet counts for hematologic dysfunction) were grouped together. Then, as was done initially for the individual continuous variables, a dummy variable for each level of organ dysfunction was newly defined. For each organ, a patient was coded as being in the referent group only if the values were within normal limits for each variable defining the organ dysfunction.

Each new dummy variable represented a level of severity for organ dysfunction, and the βs associated with each variable were the weights that formed the basis for the assignment of points. To obtain the point value for a level, the β for the dummy variable for each level of organ dysfunction was multiplied by a factor of 10 and rounded to obtain a whole number. When the points for each severity level were known, the LOD score was calculated by summing the points associated with each of the involved organ systems. The LOD score was then used as the single variable in a multiple logistic regression equation of the form: $logit=\beta_0+\beta_1(\text{LOD score})$.

The logit containing the LOD score was then converted to a probability of hospital mortality as $\Pr(y=1|\text{ logit})=e^{\text{logit}}/(1+e^{\text{logit}})$, where y equaled 1 for patients who died, y equaled 0 for patients who lived, Pr indicated probability, and e indicated a mathematical constant 2.7182818, which represented the base of the natural logarithm.

The assessment of model performance was the final stage of the analysis. To evaluate model calibration, Hosmer-Lemeshow goodness-of-fit tests, comparing observed with expected mortality, were performed.[11] To evaluate discrimination, area under the receiver operating characteristic (ROC) curve was calculated.[12]

RESULTS

Of the 14 745 patients in the ENAS database, 13 152 met the inclusion criteria. As a result of the random number generation used to create the developmental and validation samples, 10 547 patients were included in the developmental sample, and 2605 patients were included in the validation sample. The results of the data quality analyses indicated that the variables selected to compose the LOD System demonstrated good reliability. Intraclass correlations ranged from more than 0.87 for systolic blood pressure to more than 0.95 for urinary output, serum urea, and WBC count.

The results are presented in 4 parts corresponding to the successive steps of the analysis: initial analysis of individual variables in relation to mortality at hospital discharge, defining organ dysfunction using the combined information from 2 or more variables, simplifying the LOD scoring system, and developing and validating the LOD model to estimate the probability of hospital mortality.

Table 1.—Cut Points for Ranges for Several Levels of Increasing Severity for Individual Variables*

Variable	Severity Level 0	1	2	3	4
Glasgow Coma Score	14-15	9-13	6-8	4-5	3
Pao_2/Fio_2 ratio on MV or CPAP, mm Hg (kPa)	No ventilation	≥250 (≥33.2)	150-249 (19.9-33.1)	50-149 (6.6-19.8)	<50 (<6.6)
Heart rate, beats/min	30-139	140-159	≥160	<30	...
Systolic blood pressure, mm Hg	90-239	240-269 or 70-89	≥270 or 40-69	<40	...
Serum urea, mmol/L (g/L)	<6 (<0.36)	6-9.9 (0.36-0.59)	10-19.9 (0.60-1.19)	≥20 (≥1.20)	...
Creatinine, μmol/L (mg/dL)	<106 (<1.2)	106-141 (1.2-1.6)	>141 (>1.6)	...	...
Urine output, L/d	0.75-9.99	0.5-0.74 or ≥10	<0.5	...	...
White blood cell count, $\times 10^9$/L	2.5-49.9	1.0-2.4 or ≥50	<1.0	...	...
Bilirubin, μmol/L (mg/dL)	<34.2 (<0.6)	34.2-68.3 (0.6-4.0)	≥68.4 (>4.0)	...	...
Platelets, $\times 10^9$/L	≥50	<50	...	...	...
Prothrombin time, % of standard	≥25%	<25%	...	...	...

*Ellipses indicate data not applicable. Fio_2 indicates fraction of inspired oxygen; MV, mechanical ventilation; and CPAP, continuous positive airway pressure.

Initial Analysis of Individual Variables

The initial analysis used the LOWESS procedure to identify cut points that defined severity levels for each variable, shown in Table 1. Four levels of increasing severity outside the normal range were identified for the GCS and Pao_2/Fio_2 (fraction of inspired oxygen) ratio, 3 levels each for heart rate, systolic blood pressure, and serum urea, and 2 levels each for creatinine, urine output, WBC count, and bilirubin. Platelet count and PT each had 1 level of severity outside the normal range.

The cut points were used to define dummy variables for individual severity levels of each variable, a total of 27 dummy variables. These variables were used as the independent variables in a logistic regression equation that calculated the relative weight of each level in relation to hospital mortality. The weights were determined by the βs, and differences in the strength of the association with mortality were evidenced by the increasing size of the βs, from the least severe to the most severe level, within the levels of a continuous variable. For example, the βs ranged from 0.22 for the least severe level of bilirubin to 0.52 for the most severe level of bilirubin. The β for the least severe level of the GCS was 0.50, and it was 2.58 for the most severe level. Therefore, the most severe level of bilirubin had the same association with mortality as the least severe level of the GCS.

To take into account the relative severity of the organ systems involved in creating the LOD score, a grid was created in which variable levels were clustered according to their βs, or weights, for each level, as shown in Table 2. This table shows how the 4 levels of GCS and the Pao_2/Fio_2 ratio, the 3 levels of heart rate, systolic blood pressure, and serum urea, the 2 levels of creatinine, urine output, bilirubin, and WBC count, and the single level of platelet count and PT were distributed in the grid. Although the maximum number of ranges of severity for a single variable was 4, the relative strength of the association of the levels with mortality resulted in a grid with a total of 6 levels of severity, each level being associated with increasingly higher βs. The first or lowest severity level was composed of dummy variables with βs from 0.22 to 0.44. The sixth and highest level of severity was composed solely of the most severe level of the GCS, which had an associated β of 2.58.

Defining Organ Dysfunction Using Combined Information

When all the dummy variables were assigned to a cell in the grid, the organ dysfunction levels were defined according to the levels of the multiple variables for an organ system. That is, for the 4 organ systems involving more than 1 variable in the definition of dysfunction, the cut points for the 2 or 3 variables that were grouped together in the grid were used to define a single level of organ dysfunction. For example, both heart rate and systolic blood pressure were used to define levels of cardiovascular dysfunction. It can be seen from the grid that the second level of severity for heart rate (β=0.55) was similar to the lowest level of severity for systolic blood pressure (β=0.57), and the cut points for those levels were used in combination to define a single dummy variable for that level of severity for cardiovascular dysfunction. Heart rate and systolic blood pressure each had only 3 severity levels, but since they were distributed in 4 cells of the grid, there were 4 levels of cardiovascular dysfunction defined. Following this procedure for each of the organ systems defined by more than 1 variable, 3 levels of renal and hematologic dysfunction and 2 levels of hepatic dysfunction were defined. Neurologic and pulmonary dysfunction, based on a single variable for each organ, still had 4 severity levels each.

The next step in the analysis was to redefine the relative weights of the organ dysfunction levels, since 4 of them were now based on the combined information from 2 or more variables. New dummy variables were created for each level of dysfunction of the organ systems. The new dummy variables for dysfunction in the 4 organ systems, plus the original dummy variables for neurologic and pulmonary dysfunction, a total of 20 variables, became the terms in a new logistic regression equation to determine their relative weight with respect to hospital mortality.

The redistribution of the severity levels based on the new logistic regression coefficients is shown in Table 3. This table is similar to Table 2, but each coefficient now applies to a level of organ dysfunction, which may be defined by cut points for 1, 2, or 3 variables. Once the levels of organ dysfunction were distributed in the grid and the 6 severity levels were established, points for each level of severity were determined by multiplying the average of the coefficients by 10 and rounding to obtain a whole number. In increasing order of severity, the points assigned to the 6 levels at this stage of the analysis were 3, 7, 11, 13, 20, and 25. For each organ system, 0 points were assigned to values within the range of normal.

Simplifying the LOD Scoring System

The analysis to generate the final set of points for the LOD score was designed to use the optimum number of severity levels. This analysis resulted in a reduction in the number of levels from 6 to 3. Keeping the number of severity levels at 6 to this point permitted a more sensitive analysis of the association with mortality of the several levels of organ dysfunction than would have been achieved using only 3 levels from the beginning.

To define 3 levels of organ dysfunction instead of 6, the coefficients for the dummy variables for organ dysfunction were distributed into 3 groups by increasing size. The original 2 least severe levels

Table 2.—Coefficients for 6 Levels of Severity Based on Individual Variables, From Lowest to Highest Degree of Association With Hospital Outcome*

	Severity Level					
Organ System and Variable	1	2	3	4	5	6
Neurologic						
Glasgow Coma Score	...	0.50	1.14	...	2.08	2.58
Cardiovascular						
Heart rate	0.24	0.55	...	1.57	...	...
Systolic blood pressure	...	0.57	1.02	1.70	...	...
Renal						
Creatinine	0.31	0.47	...	...	...	...
Urea	0.44	0.80	1.13	...	...	...
Urine output	...	0.46	0.94	...	...	...
Pulmonary						
PaO_2/FIO_2	0.33	0.66	1.01	1.26	...	...
Hematologic						
White blood cell count	...	0.69	1.07	...	...	...
Platelets	0.28	...	...	...	...	...
Hepatic						
Bilirubin	0.22	0.52	...	...	...	...
Prothrombin time	0.30	...	...	...	...	...

*Ellipses indicate data not applicable; FIO_2, fraction of inspired oxygen.

Table 3.—Six Levels of Increasing Severity With Corresponding Coefficients for Each Organ System*

	Severity Level					
	1	2	3	4	5	6
	Points					
Organ System	3	7	11	13	20	25
	Coefficients					
Neurologic	...	0.53	1.16	...	2.09	2.61
Cardiovascular	0.22	0.72	...	1.30	...	2.41
Renal	...	0.61	1.14	...	1.94	...
Pulmonary	0.33	0.67	1.06	1.31	...	...
Hematologic	0.35	0.88	...	1.39	...	...
Hepatic	0.30	0.59	...	...	...	...

*Levels defined using combined variable information. Data not applicable indicated by ellipses.

were combined into 1 group, the 2 middle levels of severity were combined into 1 group, and the 2 most severe levels were combined into 1 group. Multiplying the average of the coefficients in each of the 3 levels by 10 and rounding to obtain severity points as whole numbers for each level resulted in the 3 levels of increasing severity being associated with 5, 12, and 23 points, respectively. To simplify the points, they were divided by 5 to bring the lowest point value to 1, which brought the other point values to 2.4 and 4.6. These were rounded up to the whole numbers 3 and 5, respectively, so that organ dysfunction in the LOD System was scored with 1, 3, or 5 points for increasing levels of severity. The point system was weighted by relative severity of levels of organ dysfunction, with a 5-fold difference between the lowest level (1 LOD point) and the highest level (5 LOD points) of dysfunction.

The LOD points that can be scored for each level of organ dysfunction are shown for each organ system in Table 4. From the Table, it can be seen that neurologic, cardiovascular, and renal dysfunction score the maximum of 5 LOD points for the most severe level of dysfunction. Pulmonary and hematologic system dysfunction score a maximum of 3 LOD points, and hepatic dysfunction scores a maximum of 1 LOD point. The LOD score can range from 0 to 22 points. Figure 1 shows the distribution of the LOD score from 0 to 22 points in the developmental sample. An LOD score of 0 indicates no organ dysfunction. An LOD score of 1 is the score for the lowest level of severity for 1 organ system dysfunction, and an LOD score of 22 points is the score for the highest level of severity for all 6 organ dysfunctions.

Of the 10 547 patients in the developmental sample, 1293 (12.3%) had no organs in dysfunction, 2723 (25.8%) had 1 organ dysfunction, 2615 (24.8%) had 2 organs in dysfunction, and 3916 (37.1%) had 3 or more organs in dysfunction. Regardless of the number of organs in dysfunction, the LOD score varied widely by the severity of the dysfunction. Depending on the involved organs and the level of severity, the LOD score can be as low as 1 or as high as 5 with 1 organ dysfunction. It can be as low as 6 or as high as 22 with 6 organs in dysfunction.

Table 4.—The Logistic Organ Dysfunction (LOD) System: Three Levels of Increasing Severity With Corresponding Points for Each Organ System*

	Severity Level			
Organ System	0	1	2	3
	LOD Points			
Neurologic	0	1	3	5
Cardiovascular	0	1	3	5
Renal	0	1	3	5
Pulmonary	0	1	3	...
Hematologic	0	1	3	...
Hepatic	0	1	...	...

*Ellipses indicate points not applicable at this severity level.

The final scoring system for the LOD score is presented in Table 5. For each organ dysfunction defined by more than 1 variable, only 1 of the variables needs to be in the abnormal range for the LOD points to be assigned. All of the variables defining an organ dysfunction must be within the normal range to receive 0 LOD points for that organ dysfunction (Table 6). To calculate the LOD score for a patient, the points for each organ dysfunction are summed.

The application of the LOD score in the ICU can be illustrated using data for a hypothetical patient as an example. Consider a patient admitted to the ICU for septic shock with extreme oliguria. The WBC count is 2.0×10^9/L, the systolic blood pressure is 60 mm Hg, and there is no evidence of pulmonary, hepatic, or neurologic dysfunction. The creatinine level is 88 μmol/L (1 mg/dL). Calculating the LOD score, being oliguric contributes 5 points to the LOD score for renal dysfunction, the low WBC count contributes 1 LOD point for hematologic dysfunction, and the systolic blood pressure contributes 3 LOD points for cardiovascular dysfunction, for a total LOD score of 9 points.

Developing and Validating the LOD Model

The LOD score was first calculated for each of the 10 547 patients in the developmental sample by summing the points for each organ system based on the recorded levels of each variable included in the system. The LOD score then was used as the only term, along with a constant term, in a new logistic regression equation, resulting in a model that provided an estimate of the severity of organ dysfunction as defined by the probability of hospital mortality. The equation for the logit was $logit = -3.4043 + 0.4173(\text{LOD score})$.

The logit was then converted to a probability of hospital mortality for each patient: $\Pr(y=1|\text{ logit}) = e^{-3.4043+0.4173(\text{LOD score})} / 1+e^{-3.4043+0.4173(\text{LOD score})}$.

The goodness-of-fit and area under the ROC curve for this model were both excellent in the developmental sample.

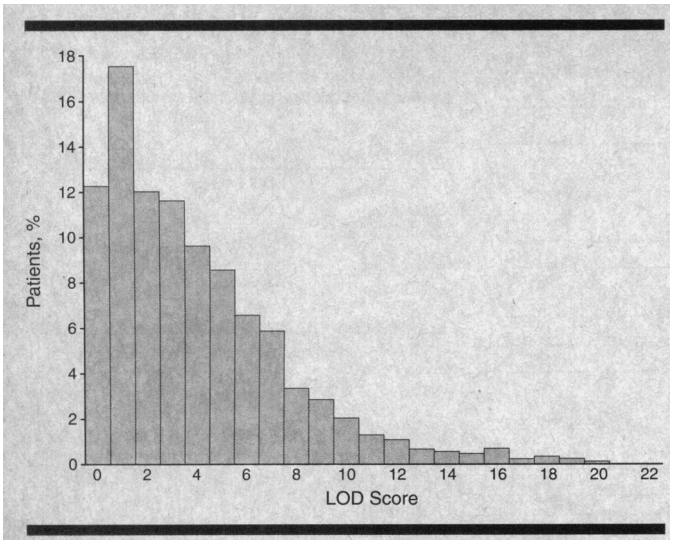

Figure 1.—Distribution of the Logistic Organ Dysfunction (LOD) score among 10 547 patients in the developmental sample.

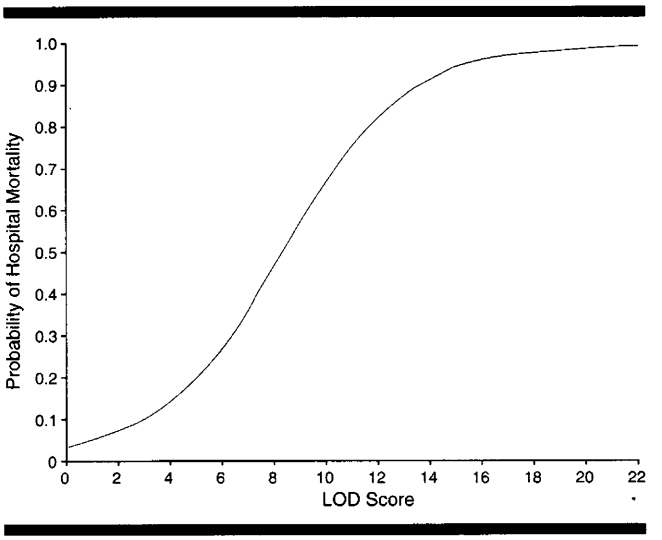

Figure 2.—Curve showing the relationship between the Logistic Organ Dysfunction (LOD) score and the probability of hospital mortality based on the LOD System.

The full goodness-of-fit table for the model in the developmental sample is shown in Table 7. The Table shows that there was very good agreement between observed and expected mortality across all risk groups, which was supported by the formal goodness-of-fit test. The value of the Hosmer-Lemeshow statistic was 10.86, with an associated P of .21 with 8 *df*. The area under the ROC curve in the developmental sample was 0.843.

Good model performance was also demonstrated among the 2605 patients in the validation sample. Excellent calibration was indicated by a value of the Hosmer-Lemeshow statistic of 9.32, with P=.50 with 10 *df*. The area under the ROC curve was 0.850, indicating excellent discrimination in the validation sample.

The probability of hospital mortality for each value of the LOD score is presented in Table 8, and Figure 2 provides a graphical presentation of the function. The figure was derived using the LOD model calculations for the logit and probability shown above. Once the LOD score has been calculated for a patient, it can be readily converted to the probability of mortality using the information in the Table.

The steepest increases in the probability of mortality occur for LOD scores from 5 to 11, with an approximate 10% increase in risk for each point increase in the score. For an LOD score of 12 or more, the risk is over 80%; the risk stays high but increases less rapidly as the score increases to the maximum of 22 points, which has an associated probability of mortality of 99.7%. There are several scenarios by which a patient could receive an LOD score of 12 points or more: either by several organs being involved at a moderate to severe level of dysfunction or by the severity level of fewer organs being very high. In any such scenario, the mortality risk is very high. Intermediate risk using the LOD System appears to occur in the range between approximately 5 and 10 points, and there are numerous combinations of organs and severity levels that would result in such a score.

The hypothetical patient described above had an LOD score of 9 points. From Table 8, it can be seen that the probability of hospital mortality for that patient would be 58.7%. While it is obvious that 3 organs were involved, the LOD System weights the severity of dysfunction for the specific organ systems and provides a corresponding estimate of the probability of hospital mortality. The relationship between the number of organs in dysfunction, the severity of dysfunction, and the LOD score for overall organ dysfunction is shown in Table 9. For each possible number of organ systems in dysfunction (from 1 to 6), 2 examples are shown to illustrate the range in the LOD score depending on the particular combination of organ systems involved, if each system is at its most severe level of dysfunction. For 3 organ systems at their maximum severity level, the LOD score could range from 7 points (for maximum hepatic, hematologic, and pulmonary dysfunction) to 15 points (for maximum neurologic, cardiovascular, and renal dysfunction). From Table 8, this translates into a corresponding range of estimated probability of hospital mortality from 38.2% to 94.6%. If the 3 involved organ systems were at their least severe level of dysfunction (1 LOD point each), the LOD score would be 3 regardless of which organs were involved, and the corresponding probability of mortality would be 10.4%.

COMMENT

An assessment of the severity of organ dysfunction in the ICU is a critical tool for conducting clinical trials, especially sepsis trials. The evaluation of new therapies cannot be successfully achieved without controlling for the degree of organ dysfunction. Our analysis demonstrated that it is not adequate to assess severity, or to describe a patient's condition, by simply counting the number of dysfunctioning organ systems. Using the LOD System, 6 organs could be dysfunctioning at levels that scored 1 point each (an LOD score of 6 points), and the associated risk of mortality would be 28.9%, compared with a mortality risk of 68.3% if 2 organs were dysfunctioning at a level that scored 5 points each (an LOD score of 10 points). Even 1 point in the LOD System is associated with an increased risk of mortality.

In many scoring systems,[2,7,8] each organ dysfunction is graded from 1 to 4 points, or from 1 to 6 points, and a score is produced by adding the points. Our results indicate that these systems cannot adequately reflect patient severity. Not only are the ranges defining the levels different from those we found using statistical methods, but weighting each organ system in the same way does not take into account the differential prognostic significance of the involved organs. The LOD score measures both the importance of the organ system relative to the others and the degree of severity within that system. Most organ dysfunction systems are scored with the worst severity level for each organ assigned the same number of points, but giving the

Table 5.—Scoring for the Logistic Organ Dysfunction (LOD) System*

	LOD Points						
	Increasing Severity/ Decreasing Values			Organ Dysfunction Free	Increasing Severity/ Increasing Values		
Organ System Measures	5	3	1	0	1	3	5
Neurologic							
Glasgow Coma Score	3-5	6-8	9-13	14-15	...	...	...
Cardiovascular							
Heart rate, beats/min	<30 or	...	...	30-139 and	≥140 or	...	...
Systolic blood pressure, mm Hg	<40	40-69	70-89	90-239	240-269	≥270	...
Renal							
Serum urea, mmol/L (g/L) or	...	...	...	<6 (<0.36)	6-9.9 (0.36-0.59)	10-19.9 (0.60-1.19)	≥20 (≥1.20)
Serum urea nitrogen, mmol/L (mg/dL)	...	...	...	<6 (<17) and	6-9.9 (17-<28) or	10-19.9 (28-<56) or	≥20 (≥56)
Creatinine, μmol/L (mg/dL)	...	...	...	<106 (<1.20) and	106-140 (1.20-1.59)	≥141 (≥1.60) or	...
Urine output, L/d	<0.5	0.5-0.74	...	0.75-9.99	...	≥10	...
Pulmonary							
Pao_2 (mm Hg)/FIo_2 on MV or CPAP		<150	≥150	No ventilation; no CPAP	...	...	...
(Pao_2 [kPa]/FIo_2)		(<19.9)	(≥19.9)	no IPAP	...	...	...
Hematologic							
White blood cell count, $\times 10^9$/L	...	<1.0	1.0-2.4 or	2.5-49.9 and	≥50.0	...	...
Platelets, $\times 10^9$/L	...	...	<50	≥50	...	...	...
Hepatic							
Bilirubin, μmol/L (mg/dL)	...	...	...	<34.2 (<2.0) and	≥34.2 (≥2.0) or	...	...
Prothrombin time, s above standard (% of standard)	...	...	(<25%)	≤3 (≥25%)	>3	...	...

*To calculate the LOD score, each organ system receives points for the single variable associated with the most points. For example, if the worst heart rate of the day was 25 beats/min (5 LOD points), but the systolic blood pressure remained at 50 mm Hg (3 LOD points), then 5 LOD points are assigned. The points are not added to obtain 8 LOD points for the organ dysfunction: the maximum number of points for an organ is 5, and the maximum LOD score is 22. Ellipses indicate data not applicable; FIo_2, fraction of inspired oxygen; MV, mechanical ventilation; CPAP, continuous positive airway pressure; and IPAP, intermittent positive airway pressure.

same number of points for a low GCS (5 LOD points) as for a high bilirubin level (1 LOD point) does not correctly reflect severity and so cannot correctly characterize patient condition.

Of the 6 organ systems described by the LOD System, neurologic, cardiovascular, and renal dysfunction were the most severe and received the maximum of 5 LOD points for the most severe level of dysfunction. Pulmonary and hematologic dysfunction both received 3 points for the most severe level of dysfunction. Hepatic dysfunction received 1 point. It is notable that Fagon et al[6] found that cardiovascular, renal, respiratory, and neurologic system dysfunction were the most severe, while hematologic and hepatic system dysfunction were less severe.

The most severe level of neurologic dysfunction, receiving 5 LOD points, was defined by a GCS less than 6. Neurologic dysfunction was measured by the actual GCS in patients who were not sedated and by the estimated GCS in sedated patients. The criteria and weights for neurologic dysfunction proposed for the MODS[8] are somewhat similar to those for the LOD System, although 4 levels of dysfunction are defined, rather than 3.

In the LOD System, cardiovascular system dysfunction could also be very severe, with a state of severe shock adding 5 points to the LOD score. Adding therapeutic measures such as the use of vasoactive drugs was not included in the LOD definitions. The LOD score was developed using data from the first ICU day, and the physiological measurements represented patient condition prior to therapy. The worst recorded values are those that receive the highest number of LOD points. For example,

Table 6.—Variables and Definitions for the Logistic Organ Dysfunction (LOD) System

All variables must be measured at least once. If they are not measured, they are assumed to be within the normal range for scoring purposes. If they are measured more than once in the first 24 h, the most severe value is used in calculating the score.

Neurologic System

Glasgow Coma Score: Use the lowest value; if the patient is sedated, record the estimated Glasgow Coma Score before sedation. The patient is free of neurologic dysfunction if the estimated Glasgow Coma Score is 14 or 15.

Cardiovascular System

Heart rate: Use the worst value in 24 h, either low or high heart rate; if it varied from cardiac arrest (5 LOD points) to extreme tachycardia (3 LOD points), assign 5 LOD points.

Systolic blood pressure: Use the same method as for heart rate (eg, if it varied from 60 to 250 mm Hg, assign 3 LOD points). The patient is free of cardiovascular dysfunction if both heart rate and systolic blood pressure are scored with 0 LOD points. This principle is the same for all organ dysfunctions that may be defined by more than 1 variable.

Renal System

Serum urea or serum urea nitrogen level: Use the highest value in mmol/L or g/L for serum urea, in mmol/L (mg/dL) of urea for serum urea nitrogen.

Creatinine: Use the highest value in μmol/L (mg/dL).

Urinary output: If the patient is in the ICU for less than 24 h, make the calculation for 24 h (eg, 1 L/8 h = 3 L/24 h). If the patient is on hemodialysis, use the pretreatment values.

Pulmonary System

If ventilated or under continuous positive airway pressure (CPAP), use the lowest value of the Pao_2/FIo_2 (fraction of inspired oxygen) ratio (whether Pao_2 is mm Hg or kPa). A patient who has no ventilation or CPAP during the first day is free of pulmonary dysfunction.

Hematologic System

White blood cell count: Use the worst (high or low) white blood cell count that scores the highest number of points.

Platelets: If there are several values recorded, find the lowest value and assign 1 LOD point if the lowest value is less than 50×10^9/L.

Hepatic System

Bilirubin: Use the highest value in μmol/L (mg/dL).

Prothrombin time (seconds or %): If there are several values recorded, assign 1 LOD point if the prothrombin time was ever more than 3 s above standard or less than 25% of standard during the day.

Table 7.—Goodness-of-Fit of the Logistic Organ Dysfunction (LOD) Model Among 10 547 Patients in the Developmental Sample*

Probability	Survived, No.		Died, No.	
	Observed	Expected	Observed	Expected
0.00-0.032	1242	1251.4	51	41.6
0.032-0.048	1781	1765.0	73	89.0
0.048-0.071	1172	1184.3	103	90.7
0.071-0.104	1108	1100.1	120	127.8
0.104-0.150	853	866.2	166	152.8
0.150-0.211	733	720.2	180	192.8
0.211-0.382	891	889.9	442	443.2
0.382-0.587	314	319.8	368	362.2
0.587-0.833	131	132.0	388	387.0
0.833-1.000	29	25.1	402	405.9

*For this model, df=8; $\hat{C}$=10.86; P=.21; and area under the receiver operating characteristic (ROC) curve is 0.843.

Table 8.—Conversion of the Logistic Organ Dysfunction (LOD) Score to a Probability of Hospital Mortality Using the LOD Model

LOD Score	Probability of Hospital Mortality, %
0	3.2
1	4.8
2	7.1
3	10.4
4	15.0
5	21.1
6	28.9
7	38.2
8	48.4
9	58.7
10	68.3
11	76.6
12	83.3
13	88.3
14	92.0
15	94.6
16	96.4
17	97.6
18	98.4
19	98.9
20	99.3
21	99.5
22	99.7

if at different times on the first ICU day a patient has tachycardia of 150 beats per minute (1 LOD point) and bradycardia of 25 beats per minute (5 LOD points), 5 points are added to the LOD score. After the first ICU day, when a patient is receiving continuous therapy, the problem of scoring cardiovascular variables is, indeed, a difficult one. The variable proposed for the assessment of cardiovascular dysfunction in the MODS[8] is pressure-adjusted heart rate (product of heart rate multiplied by the ratio of the central venous pressure to the mean arterial pressure). This variable depends on resuscitation and the use of blockers and pressors. The central venous pressure is not recorded in all patients, which limits the value of this variable. Although hypertension and bradycardia have not classically been regarded as part of the multiple organ system syndrome, they nevertheless reflect an abnormality in the functioning of the cardiac system and were associated with a worse outcome than was the case for patients without these factors in our study. This result suggests that previous definitions of early cardiovascular dysfunction need to be modified.

Renal dysfunction, as manifested by low urine output (oliguria) or high serum urea levels, also receives 5 LOD points for the most severe level of dysfunction, which has been noted in other studies of renal dysfunction in intensive care. There is no distinction made between chronic and acute renal dysfunction in the LOD scoring, as the focus is on the relevant physiological measurements without having to rely on diagnostic assessments. Again, the decision as to what constitutes the worst value is based on the number of points assigned. For example, if a patient has oliguria of 0.4 L/d (5 LOD points), 5 points are added to the LOD score, regardless of the level of creatinine. To rely only on creatinine could actually postpone the confirmation of renal dysfunction, since it may take several days to observe a rise in creatinine. In several assessment systems,[8,13] serum creatinine concentration is the only component of renal dysfunction measurement. Serum urea or serum urea nitrogen, as well as daily urinary output, are measured in many countries and have a prognostic weight independent of creatinine. The coefficients for both urea and urinary output demonstrated a stronger association with hospital mortality than the coefficients for creatinine, and when the variables were considered in combination to define renal dysfunction, the association with mortality was even stronger.

Pulmonary dysfunction receives only 3 LOD points for the most severe level. Patients who have been assisted with neither ventilation nor CPAP are considered to be free of pulmonary dysfunction and receive 0 points towards the LOD score. The Pa_{O_2}/FI_{O_2} ratio was also used in the MODS[8] calculations to define levels of pulmonary dysfunction; however, it was not clear whether all of their population of 692 surgical patients were receiving mechanical ventilation, which was not the case for the consecutive admissions that composed the ENAS database.

Hematologic dysfunction also scores a maximum of 3 LOD points, with the most severe level defined by a WBC count less than 1.0×10^9/L. This suggests that a very low WBC count is not as strongly associated with mortality as the most severe levels of dysfunction of other organs, all other things being equal. The data for platelet counts were collected as a binary variable indicating only whether platelet counts were low (less than 50×10^9/L), and this level of measurement resulted in a severity level that receives only 1 LOD point. The MODS uses only platelet counts, measured on a continuous scale, in the assessment of hematologic system dysfunction. Platelet counts less than 50×10^9/L showed a strong association with mortality in that study, consistent with the LOD System categorization.

Hepatic dysfunction scores a maximum of 1 LOD point. This suggests that early hepatic dysfunction by itself is not strongly associated with mortality, but its occurrence in association with the dysfunctioning of the other organ systems worsens the prognosis in an ICU patient. Unlike in the MODS, hepatic dysfunction contributed the least to the scoring of multiple organ dysfunction in the LOD System, allowing a maximum of 1 LOD point. Using PT to assess hepatic dysfunction incorporated the measurement of a variable that may be abnormal even when the bilirubin is within normal limits. Since our analysis was restricted to the first 24 hours in the ICU, it would be expected that hepatic dysfunction would be more heavily weighted later in the ICU stay than during the first ICU day. In future LOD research, data for platelet counts and PT should be collected as a continuous measurement to confirm whether the current cut points are best suited to reflect the association with mortality.

Although developed using the same database, there are important differences between the SAPS II and the LOD systems. The former takes into account not only several physiologic parameters, but also includes age, the type of patient admission, and several comorbidities. The LOD System was designed to characterize 6 distinct organ systems and uses only physiological measurements to do so. The information from the physiological measurements is grouped in a manner that permits the characterization of organ dysfunction, both as to the number of affected organs and the degree of dysfunction for each organ. In the LOD System, 1 abnormal element is

sufficient for the classification of organ dysfunction.

Multiple organ dysfunction is not necessary for the application of the LOD System, as it applies to both single or multiple organ dysfunctions. This makes the LOD System more broadly applicable, since less than one third of patients in an ICU may have 2 or more organs in dysfunction, with the majority having only 1 organ dysfunction.[14] In our database, which comprised several tertiary care units, 26% of patients had 1 organ dysfunction, and 62% had 2 or more organs in dysfunction. The concept of multiple organ dysfunction implies the involvement of multiple organs, rather than a single organ, but the LOD System grades organ dysfunction in such a way that severity due to organ dysfunction can be quantified, whether 1 to 6 organs are involved.

Patients with 1 organ dysfunction, especially if it is severe (ie, 5 LOD points), are not comparable with patients free of organ dysfunction. With this concept, patients with acute myocardial infarction, pulmonary embolism, or irreversible brain injury would be classified as having 1 organ system dysfunction with an associated LOD score and probability of hospital mortality that is higher than patients with no organ dysfunction. Like others,[6,8] we have purposely used the term organ dysfunction rather than organ failure, since dysfunction encompasses from moderate to severe impairment.

The aim of our study was to develop a new system to quantify organ dysfunction based on objective criteria of severity. Other scoring systems (APACHE [Acute Physiology and Chronic Health Evaluation] III, MPM [Mortality Probability Model] II, and SAPS II) were created using multiple logistic regression, while earlier versions of some systems (APACHE II, SAPS I) were based on more subjective methods. The objective models in the most recent generation were demonstrably superior to previous ones.[15,16] Until now, all of the systems that proposed to estimate organ dysfunction were also based on subjective methods. These systems varied widely, and there was no statistical means for assessing how well they reflected severity of illness. While mortality generally increases with the number of involved organs and the degree of organ dysfunction,[7,8,14] the systems do not provide an estimate of the risk of mortality, with the exception of the ODIN system.[6] It is, therefore, difficult to compare or validate the systems, since it is not possible to evaluate their calibration.

The proposed LOD System, which was developed using statistical methods that determined the relative weights of the several organ systems and of the levels of severity within each organ system, also produces an estimate of the risk of mortality that demonstrated excellent calibration and discrimination. The LOD System can be used for the assessment of organ dysfunction in an objective manner. Many organ dysfunctions develop later in the ICU stay, particularly among elective surgery patients who are often free of organ dysfunction when admitted to the ICU. The large ENAS database from which the system was developed, however, comprised a mix of medical patients, emergency surgical patients, and elective surgical patients who manifested measurable levels of organ dysfunction on the first ICU day. Many of the ICUs in the ENAS database were tertiary care units, and patients entered them at a relatively advanced stage of disease, as reflected by the 62% of patients with 2 or more organs in dysfunction on the first day in the ICU.

Any system to assess organ dysfunction that uses first ICU day variables must be validated for use at other time periods, including the LOD System, and future studies must be designed for that purpose. The validity of the estimate of the probability of mortality from the conversion of the LOD score to a probability using the LOD model on subsequent ICU days has not been tested. Studies of the association with mortality of an LOD score that is collected daily in the ICU need to be undertaken. Also, further studies that take into account the duration of dysfunction will be needed to estimate the probability of hospital mortality at later points in the ICU stay. Duration of dysfunction is commonly associated with a worsening prognosis, even if a patient's condition is unchanging, since absence of improvement is a negative sign.

In addition to studying the applicability of the LOD System over time, another issue that can be considered is the question of which measurements to use on the first day or subsequent days to characterize patients. As with other severity systems for ICU patients, the LOD System uses the worst value recorded for the patient in the 24-hour period. Organ dysfunction generally occurs following treatment for some stress, trauma, or infection, and if the effort is completely successful, the physiological parameters will be normalized. If not, they will remain outside the normal ranges. Organ dysfunction is more likely to occur in the latter case, which suggests a justification for using the worst values. Using the best measurement or a measurement collected at the same time every day are alternatives to using the worst value, and the issue of which values to use merits further study.

A potential criticism of the LOD System is its use of hospital mortality, rather than ICU mortality, as the outcome of interest. This is a debatable point, and using ICU mortality as did Fagon et al[6] and Marshall et al[8] also has appeal. Again, this is an issue that should be studied further. Our choice was to use hospital mortality as a more objective criterion, especially in light of the multinational nature of the database. ICU discharge practices may vary considerably among units in different countries or regions, with patients who have no chance of improvement being discharged to floors in some units but not in others, and this would tend to distort the comparability of ICU mortality across units. Patients who die in the ICU are still included in the outcome measure of hospital mortality, as well as patients who are discharged from the ICU for further hospital care and who subsequently die prior to hospital discharge.

It may be argued that an organ dysfunction system is best used to describe

Table 9.—Logistic Organ Dysfunction (LOD) Score, by Number of Organ System Dysfunctions for 2 Examples With Involved Organ Systems at Their Highest Severity Level *

	Example 1		Example 2	
No. of Organs	Score	Involved Organ Systems (LOD Points)	Score	Involved Organ Systems (LOD Points)
1	1	HP (1)	5	N (5)
2	4	HP, HE (1, 3)	10	N, CV (5, 5)
3	7	HP, HE, P (1, 3, 3)	15	N, CV, R (5, 5, 5)
4	12	HP, HE, P, R (1, 3, 3, 5)	18	N, CV, R, P (5, 5, 5, 3)
5	17	HP, HE, P, R, CV (1, 3, 3, 5, 5)	21	N, CV, R, P, HE (5, 5, 5, 3, 3)
6	22	HP, HE, P, R, CV, N (1, 3, 3, 5, 5, 5)	22	N, CV, R, P, HE, HP (5, 5, 5, 3, 3, 1)

*HP indicates hepatic system; HE, hematologic system; P, pulmonary system; R, renal system; CV, cardiovascular system; and N, neurologic system. Ellipses indicate data not applicable.

severity of illness rather than to estimate the risk of mortality. There are other components of severity, such as pain, anatomical extent of injury, or therapeutic requirements, but it is the physiological markers that are most objectively quantifiable in the ICU. Severity can describe gradations of dysfunction, but risk quantifies the measurable components of severity into an estimate of the probability of mortality that can be objectively tested by comparing it with the observed outcome.

Some severe morbidities do not, of course, necessarily result in mortality, but the proposed system can still be useful. For example, if a patient enters the ICU with oliguria and no other organ system is in dysfunction, that patient has 1 dysfunctioning organ and would be scored as 5 LOD points. Even if the patient undergoes dialysis, the LOD score for the patient would be 5 and the probability of hospital mortality would be 21.1%, which is still a statistically measurable risk of mortality. The same would apply to a patient who received mechanical ventilation for exacerbation of chronic obstructive pulmonary disease. If the Pa_{O_2}/FI_{O_2} ratio were 140, with no other dysfunctioning organ, the LOD score would be 3, and the probability of hospital mortality would be 10.4%. So, although mortality may not be a perfect endpoint for the assessment of organ system dysfunction, any degree of organ dysfunction is associated with a risk of mortality, and it does provide an objective way to evaluate observed vs expected outcome in the conduct of clinical trials.

Having a tool to quantify the severity of organ dysfunction is necessary in order to evaluate the effectiveness of treatment, not only on mortality but on the resolution of organ dysfunction. Successful resolution of organ dysfunction, however, has not been clearly defined. Some researchers have proposed the number of days free from organ dysfunction as a marker of resolution for an outcome measure.[13,17] For such a purpose, the number of days with a zero for the LOD score could be calculated.

As with all models designed for use in a dynamic and changing environment, the LOD System must be kept up-to-date and applicable in the face of changing case mix and ICU therapies. In its present form, the LOD System we have proposed is based on objectively derived coefficients that weight the severity of organ dysfunction differentially both among the 6 organ systems and within each organ system. The results of our analysis, by refining the relative association of levels of severity with hospital mortality, suggest that the LOD System has great potential as a tool with which to assess the real severity of organ dysfunction among general medical and surgical ICU patients.

References

1. Fry DE, Pearlstein L, Fulton RL, Polk HC Jr. Multiple system organ failure: the role of uncontrolled infection. *Arch Surg.* 1980;115:136-140.
2. Stevens LE. Gauging the severity of surgical sepsis. *Arch Surg.* 1983;118:1190-1192.
3. Marshall JC, Christou NV, Horn R, Meakins JL. The microbiology of multiple organ failure: the proximal gastrointestinal tract as an occult reservoir of pathogens. *Arch Surg.* 1988;123:309-315.
4. Knaus WA, Draper EA, Wagner DP, Zimmermann JE. Prognosis in acute organ system failure. *Ann Surg.* 1985;202:685-693.
5. Knaus WA, Wagner DP. Multiple systems organ failure: epidemiology and prognosis in critical care clinics. In: Pinsky MR, Matuschak GM, eds. *Multiple Systems Organ Failure.* Philadelphia, Pa: WB Saunders Co; 1989:221-232.
6. Fagon JY, Chastre J, Novara A, Medioni P, Gibert C. Characterization of intensive care unit patients using a model based on the presence or absence of organ dysfunction and/or infection: the ODIN model. *Intensive Care Med.* 1993;19:137-144.
7. Hebert RC, Drummond AJ, Singer J, Bernard GR, Russel JA. A simple multiple organ system failure scoring system predicts mortality of patients who have sepsis syndrome. *Chest.* 1993;104:230-235.
8. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple Organ Dysfunction Score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995;23:1638-1652.
9. 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.
10. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc.* 1979;74:829-836.
11. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons Inc; 1989.
12. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29-36.
13. Sibbald WJ, Vincent JL. Round table conference on clinical trials for the treatment of sepsis. *Crit Care Med.* 1995;23:394-399.
14. Rauss A, Knaus WA, Patois E, Le Gall JR, Loirat P, and the French Multicentric Group of ICU Research. Prognostic for recovery from multiple organ system failure: the accuracy of objective estimates of chances for survival. *Med Decis Making.* 1990;10:155-162.
15. Castella X, Artigas A, Bion J, Kari A, and the European/North American Severity Study Group. A comparison of severity of illness scoring systems for intensive care unit patients: results of a multicenter, multinational study. *Crit Care Med.* 1995; 23:1327-1335.
16. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients: a system update. *JAMA.* 1994;272:1049-1055.
17. Rangel-Frausto MS, Pittet D, Costigan M, Hwang T, Davis CS, Wenzel RP. The natural history of the systemic inflammatory response syndrome (SIRS): a prospective study. *JAMA.* 1995; 273:117-123.