Prospectively validated prediction of physiologic variables and organ failure in septic patients: The Systemic Mediator Associated Response Test (SMART)

Gus J. Slotman, MD

Objective: **Conventional outcomes research provides only percentage risk of such end points as mortality rate, utilization of resources, and/or broad groupings of multiple organ system dysfunction. These prognostications generally are not applicable to individual patients. The purpose of the present study was to determine whether the Systemic Mediator Associated Response Test (SMART) methodology could identify interactions among demographics, physiologic variables, standard hospital laboratory tests, and circulating cytokine concentrations that predicted continuous and dichotomous dependent clinical variables, in advance, in individual patients with severe sepsis and septic shock, and whether these independent variables could be integrated into prospectively validated predictive models.**

Design: **Data review and multivariate stepwise logistic regression.**

Setting: **University research laboratory.**

Patients: **Three hundred three patients with severe sepsis or septic shock who comprised the placebo arm of a multiple-institution clinical trial, who were randomly separated into a model building training cohort (n = 200) and a predictive cohort (n = 103).**

Interventions: **None.**

Measurements and Main Results: **From baseline data and baseline plus serial input, including patient demographics, hospital laboratory tests, and plasma concentrations of interleukin-6, interleukin-8, and granulocyte colony stimulating factor, multiple regression models were developed that predicted clinically important continuous dependent variables quantitatively, in individual patients. Multivariate stepwise logistic regression was used to develop models that prognosticated dichotomous dependent end points. Data from individual patients in the predictive cohort were inserted into each predictive model for each day, with prospective validation accomplished by simple linear regression of individual predicted vs. observed values for continuous dependent variables, and by establishing the receiver operator characteristics area under the curve for logistic regression models that predicted dichotomous end points. Of SMART models for continuous dependent variables, 100 of 143 (70%) were validated at r values >.7 through day 3, and 184 of 259 (71%) above r = .5 through day 5. SMART predictions of dichotomous end points achieved receiver operator characteristics areas under the curve >.7 for up to 84% of the equations in the first week. Many SMART models for both continuous and dichotomous dependent variables were validated at clinically useful levels of accuracy as far as 28 days after baseline.**

Conclusions: **SMART integration of demographics, bedside physiology, hospital laboratory tests, and circulating cytokines predicts organ failure and physiologic function indicators in individual patients with severe sepsis and septic shock. (Crit Care Med 2002; 30:1035–1045)**

KEY WORDS: prognostication; severe sepsis; cytokines; shock; multiple organ failure

During the past two decades, predicting the clinical course of critically ill patients has been the subject of a large body of outcomes research. Most of these studies have investigated correlations of demographics and physiologic observations with subsequent intensive care unit or in-hospital death or survival. For example, the Sepsis Score (1), the Multiple Organ System Dysfunction Score (2), and the recently developed Brussels Sepsis Score (3, 4) have attempted to predict group percentage of risk of death among septic patients. In the general adult intensive care unit population, the Multiple Probability Model I and II (5, 6), the Simplified Acute Physiology Score I and II (7, 8), and the Acute Physiology and Chronic Health Examination I, II, and III (9–11) scores are the best known predictive engines (12). All of them predict mortality rates with relatively similar efficacy. Investigators who use these methods have attempted to achieve physiologic uniformity among patient populations in clinical trials of new drugs for sepsis (13) and to establish benchmarking between hospitals for quality assurance/quality improvement (11). However, because these scoring systems lump patients with disparate pathophysiologies together simply because they have a similar probability of dying, they ultimately forecast only grouped percentage risk of hospital death and possibly consumption of health care resources (14). Unfortunately, therefore, these methods do not predict the important pathophysiologic events, such as organ failure and shock, that actually determine the risk of mortality in the first place. Advance knowledge of these life-threatening complications might facilitate timely therapeutic intervention.

From the Department of Surgery, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School-Camden, Cooper Hospital/University Medical Center, Camden, NJ.

Address requests for reprints to: Gus J. Slotman, MD, Department of Surgery, Three Cooper Plaza, Suite 411, Camden, NJ 08103. E-mail: slotman@umdnj.edu

To be beneficial for individual patients, outcomes research must achieve a state of the art that predicts subclinically, in advance, the sequelae of systemic inflammation that lead to mortality. A key to achieving this new level of critical care prediction is recognizing temporal links between baseline data and subsequent events in the clinical course of individual patients. In septic patients, increased blood concentrations of inflammatory response mediators such as the eicosanoids and the cytokines, among others, have been associated with the development of shock, end organ dysfunction, and high mortality rates (15–17). It was hypothesized that clinical manifestations of sepsis and its sequelae could be predicted quantitatively and qualitatively in individual patients through analyzing interactions among circulating inflammatory response mediators, patient demographics, physiologic data, and standard hospital laboratory results. As a first step in developing this type of prognostic methodology, my colleague and I previously evaluated retrospectively whether circulating concentrations of the eicosanoids and cytokines, plus physiologic measurements and standard clinical laboratory results, could be integrated into multivariate predictive models, identified as the Systemic Mediator Associated Response Test (SMART). This initial foray was successful in predicting inflammation and organ failure quantitatively in patients with severe sepsis (18). Circulating eicosanoids and cytokines were predicted as well (19). Building on the results of that pilot project, the objective of the present study was to further develop SMART multivariate models that predicted physiologic measurements, results of hospital laboratory tests, circulating cytokines, and organ failure or shock, in advance, in individual patients with severe sepsis and septic shock, and to validate those models prospectively.

MATERIALS AND METHODS

Data of 303 patients with severe sepsis and septic shock who had comprised the placebo arm of a phase III clinical trial (20) were tabulated. These patients then were divided by a randomization program into a model-building training cohort (n = 200) and a prospective validation, predictive cohort (n = 103). Demographics including sex, race, age, admitting service (surgery or nonsurgical), and comorbidities were recorded at baseline for each patient. At baseline and on days 1–7, 14, 21, and 28, the physiologic variables and hospital laboratory tests listed on Table 1 were recorded for all patients who were alive at that observation point. In addition, at baseline and on days 1, 2, 3, and 4, plasma concentrations of interleukin (IL)-6, IL-8, and granulocyte colony stimulating factor (GCSF) were measured by enzyme-linked immunosorbent assay, by using commercially available kits and standard enzyme-linked immunosorbent assay laboratory methodology. Specific quality control data regarding these assays such as recovery and coefficients of variation were not available from the sponsors of the original clinical trial.

Table 1. Measures and outcomes in patients with severe sepsis

Age	Admitting service
Gender	Surgery
Race	Medicine
Albumin	$P(A\text{-}a)O_2$
Alkaline phosphatase	Base deficit
ALT	pH
AST	PaO_2
BUN	SaO_2
Calcium	FIO_2
Cholesterol	Fluids in
Creatinine	Fluids out
GGT	PaO_2/FIO_2
Glucose	Chloride
Hematocrit	Eosinophils
Hemoglobin	Lymphocytes
MCH	Mononuclear cells
MCHC	Metamyelocyte
MCV	Segmental neutrophils
Phosphorous	Band neutrophils
Platelet count	Basophils
Potassium	Granulocytes
Total protein	% Granulocytes
PT	% Lymphocytes
PTT	Eosinophils
RBC	Lactic acid
Sodium	PAOP
Total bilirubin	Cardiac index
Triglycerides	SVR
Uric acid	PEEP
WBC	Pressure support
IL-6	Respiratory rate
IL-8	Mechanical ventilation
GCSF	Trauma
P-r interval	Systolic BP
EKG	Diastolic BP
q-T interval	Heart rate
DIC	MAP
GCS	Temperature
Hepatobiliary baseline	Height
Shock	Weight
ARDS	
Renal failure	
Comorbidities	
Alcohol abuse, cirrhosis	
HIV	
Dialysis	
Neutropenia	
COPD	
Solid tumor,	
Hematologic malignancy	
Chronic renal failure	

ALT, alanine aminotransferase; AST, aspartate aminotransferase; BUN, blood urea nitrogen; GGT, glutamyl-glutamate aminotransferase; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; PT, prothrombin time; PTT, partial thromboplastin time; RBC, red blood cells; WBC, white blood cells; IL, interleukin; GCSF, granulocyte colony-stimulating factor; EKG, electrocardiogram; DIC, disseminated intravascular coagulation dysfunction; GCS, Glasgow Coma Scale; ARDS, acute respiratory distress syndrome; HIV, human immunodeficiency virus; COPD, chronic obstructive pulmonary disease; $P(A\text{-}a)O_2$, alveolar-arterial oxygen tension difference; SaO_2, arterial oxygen saturation; PAOP, pulmonary artery occlusion pressure; SVR, systemic vascular resistance; PEEP, positive end-expiratory pressure; BP, blood pressure; MAP, mean arterial pressure.

Table 2. SMART: Prediction of physiologic, respiratory, and metabolic variables in severe sepsis from baseline data only, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Heart rate	.429	.425	.310	.249	.360	.386	0.377	0.109	.183	.366
Temperature	.468	.411	.161	.243	.371	.295	.342	.033	.177	—
Cardiac index	.570	.445	.645	.437	.525	.440	—	—	—	—
SVR	.488	.304	.420	.014	.061	.265	.124	—	—	—
Glasgow Coma Scale	.601	.575	.458	.387	.287	.400	.325	.184	.213	.101
FIO_2	.443	.115	.078	.452	.517	.308	.409	.023	.218	.092
HCO_3	.571	.551	.562	.477	.500	.401	.350	.371	.421	.126
Pressure support	.893	.738	.763	.402	.421	.481	.167	—	.290	—
PEEP	.893	.716	.669	.372	.391	.270	.317	.168	.016	.071
Albumin	.881	.720	.770	.767	.767	.709	.647	.420	.373	.204
Cholesterol	.725	.832	.794	.722	.479	.395	.295	.356	.258	.055
Glucose	.217	.251	.247	.447	.472	—	.079	.197	.239	.313
Total protein	.785	.684	.701	.635	.587	.556	.483	.289	.229	.031
Triglycerides	.711	.922	.771	.403	.407	.313	.155	.343	.194	.120
Uric acid	.939	.910	.826	.740	.685	.593	.506	.283	.353	.512
Calcium	.696	.663	.424	.580	.611	.605	.510	.360	.450	.312

SVR, systemic vascular resistance; PEEP, positive end-expiratory pressure.

Table 3. SMART: Prediction of liver and renal function indicators in severe sepsis from baseline data only, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Alkaline phosphatase	.869	.550	.691	.679	.798	.710	.619	.421	.369	.105
ALT	.959	.844	.391	.485	.606	.242	.224	.354	.305	.108
AST	.786	.659	.231	.287	.153	.061	.093	—	—	.461
GGT	.943	.807	.717	.707	.671	.499	.578	.491	.456	.169
Total bilirubin	.965	.941	.832	.676	.770	.753	.824	.869	.815	.688
BUN	.970	.922	.881	.832	.816	.804	.767	.450	.337	.331
Creatinine	.896	.831	.741	.706	.657	.645	.567	.303	.384	.379

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, glutamyl-glutamate aminotransferase; BUN, blood urea nitrogen.

Table 4. SMART: Prediction of hematologic and coagulation indicators in severe sepsis from baseline data only, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Lymphocytes	.937	.982	.976	.994	.105	.158	.114	.995	.978	.980
Monocytes	.971	.998	.997	.994	.999	.161	.168	.988	.999	.997
Segmental neutrophils	.999	.999	.999	.999	.999	.999	.999	.989	.995	.999
Bands	.999	.991	.704	.511	.935	.994	.984	.020	.102	.331
Granulocytes (segs + bands)	.999	.999	.999	.859	.999	.878	.734	.637	.999	—
% Granulocytes	.999	.999	.999	.685	.999	.999	.999	.245	.999	—
Platelet count	0.921	.850	.777	.732	.670	.604	.438	.301	.450	.147
Prothrombin time	.932	.923	.926	.922	.917	.402	.928	.879	.809	.887

By using Statistical Analysis System (SAS) programs (21), the continuous dependent variables listed in Tables 2–4 were screened for Spearman cross-correlations with each independent variable at days 1–7, 14, 21, and 28 after baseline. The Spearman correlations were ordered by absolute magnitude, and those with the largest absolute correlations were entered into a stepwise selection model. The number of variables entered into the stepwise procedure was chosen to approximately maintain a ratio of one variable to ten patients in the model (22). Missing value patterns present in the data caused a different number of variables to be input to the stepwise procedure for each dependent variable. These multiple regression models then were validated prospectively by entering raw data from each of the patients in the predictive cohort into each model and plotting linear regression curves of the predicted value of each variable for each patient vs. the measurements actually observed. The extent of association between the quantitative predictions and observed data then was described by the Pearson product moment, or linear regression correlation coefficient (r) (23). The linear relationship of the independent variables was evaluated in model residual graphs. Examination of residuals revealed no evidence to conclude that the residuals were not normal in distribution. Outliers were examined also, and there was no evidence that any observations were an extreme influence in this model.

Again by using the training cohort, by stepwise logistic regression, I developed multivariate models that predicted the presence or absence of acute respiratory distress syndrome (ARDS), renal insufficiency, hepatobiliary dysfunction, and disseminated intravascular coagulation dysfunction (DIC), all defined according to established diagnostic criteria in the literature for these entities (20); cerebral dysfunction, defined as Glasgow Coma Scale score <11; and the number of lung quadrants on chest radiograph that were affected by pulmonary edema (0–4). Again, independent variables for each model were limited to ten, and all-ways elimination was used in the model-building process. Glasgow Coma Scale score <11 was chosen as the threshold for cerebral dysfunction because of the automatic absence of an appropriate verbal response for endotracheally intubated patients who otherwise might have intact cerebral function. The SMART multiple regression models derived for these dichotomous dependent variables then were validated prospectively by entering raw data from the 103 predictive cohort individual patients into the training cohort logistic regression formulas and then assessing predictive accuracy by calculating the area under the curve of receiver operating characteristic statistics (ROC AUC) (24, 25).

Stepwise multiple regression and stepwise multiple logistic regression models that predicted continuous and dichotomous dependent variables, respectively, 24 hrs after baseline, used only baseline data. For predictions beyond 24 hrs, SMART modeling was carried out in two ways for each variable at each time point: a) from baseline data only; and b) from serial data, where baseline measurements and/or subsequent determinations up to 24 hrs before the time being prognosticated were incorporated into the multiple regression and/or multivariate stepwise logistic regression modeling.

For differences in distribution of independent and dependent variables, the chi-square statistic was used (23).

To evaluate whether independent variables that made sense pathophysiologically but were not weighted predictors in the predictive models could improve SMART prognostications, baseline dichotomous models were reevalu-

ated for day 5. Specifically, independent variables that logically were associated with chest radiograph score (FIO_2, IL-8, GCSF, peak end-expiratory pressure [PEEP]), with ARDS (FIO_2, platelet count, IL-8, GCSF), with DIC (platelet count, prothrombin time, partial thromboplastin time), with hepatobiliary failure (activated clotting time, aspartate aminotransferase, IL-6, IL-8, GCSF), with renal insufficiency (shock, IL-6, IL-8, GCSF), and with shock (platelet count, IL-6, IL-8, GCSF, FIO_2), but were not captured mathematically as significant predictors, were "forced" into the SMART models that predicted each of these dichotomous dependent variables in day 5. Predicted vs. observed ROC AUC results then were compared with corresponding curves from models that did not include these variables.

RESULTS

Follow-up of subjects in the parent clinical trial was quite detailed and complete. Baseline data were sufficient for SMART equations of most of all the 103 validation patients. Only seven SMART models were validated with sample sizes <30: chest radiograph score at 14 (n = 21) and 21 (n = 11) days, HO_3 at 24 (n = 26) and 28 (n = 14) days, PEEP at 21 (n = 22) and 28 (n = 12) days, and pressure support ventilation at 28 days (n = 11).

Prospectively validated SMART predictions of physiologic, respiratory, and metabolic variables in patients with severe sepsis and septic shock, from baseline data only, are listed in Table 2. The highest linear regression correlation coefficients were for predictions of the level of pressure support ventilation, PEEP, serum albumin, cholesterol, total protein, triglycerides, and uric acid. Through 7 days, quantitative predictions of HCO_3, FIO_2, systemic vascular resistance, cardiac index, temperature, and heart rate also approached clinically useful levels of prospective validation, with $r \geq .500$. Predictions from baseline data of continuous dependent variables at 14 days and beyond were clinically significant only for HCO_3, serum albumin, cholesterol, total protein, uric acid, and calcium.

Results of prospectively validated SMART multiple regression predictions of liver and renal function indicators among patients with severe sepsis from baseline data only are displayed in Table 3. Clinically useful levels of correlation between SMART predictions and the values actually observed in individual patients, frequently exceeding r = .700, were achieved for alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, glutamyl-glutamate aminotransferase, total bilirubin, blood urea nitrogen, and creatinine. Many of the multiple regression models yielded clinically useful results at 14 days and beyond.

Prospectively validated SMART predictions of hematologic and coagulation indicators in patients with severe sepsis from baseline data only are tabulated in Table 4. Quantitative predictions of lymphocyte, monocyte, segmented neutrophil, band, and granulocyte counts; the differential percentage of granulocytes and lymphocytes; platelet count; and prothrombin time consistently resulted in linear regression correlations between predicted and observed values in individual patients in the clinically useful range above .9. SMART predictions of hematocrit, red blood cell count, white blood cell count, and partial thromboplastin time also were significant.

Prospectively validated SMART predictions of physiologic, respiratory, and metabolic variables in patients with severe sepsis from baseline data plus serial information, including maximum levels and change from baseline, are tabulated in Table 5. Plots of predicted vs. observed values in individual patients for Glasgow Coma Scale, HCO_3, pressure support ventilation, PEEP, albumin, cholesterol, triglycerides, and uric acid produced $r > .8$ during days 1–7. Predicted vs. observed correlations >.4 were recorded for heart rate, temperature, cardiac index, systemic vascular resistance, FIO_2, glucose, total protein, and calcium.

Prospectively validated SMART predictions of liver and renal function indicators from baseline plus serial data are displayed in Table 6. Clinically useful levels of accuracy, evidenced in Pearson product moments >.8, were achieved with alkaline phosphatase, alanine aminotransferase, glutamyl-glutamate aminotransferase, total bilirubin, blood urea nitrogen, and creatinine for up to 28 days of observation.

Prospectively validated SMART predictions of hematologic and coagulation indicators in patients with severe sepsis, from models derived from baseline plus serial data analysis, are displayed in Table 7. Clinically useful levels of accuracy were evidenced in r values >.9 for

Table 5. SMART: Prediction of physiologic, respiratory, and metabolic variables in severe sepsis from baseline plus serial data, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Cardiac index	0.570	.157	.404	.352	.298	.167	.007	—	—	
Glasgow Coma Scale	.601	.897	.804	.665	.377	.024	.164	—	.066	.079
HCO_3	.571	.277	.853	.375	.362	.211	.350	.112	.233	.218
Pressure support	.893	.877	.904	.674	.620	.481	.297	.258	—	.325
PEEP	.893	.877	.899	.674	.263	.291	.450	.167	.368	.188
Albumin	.881	.815	.937	.794	.819	.680	.622	.386	.227	.055
Cholesterol	.725	.832	.957	.633	.403	.303	.180	.287	.011	.058
Total protein	.785	.656	.638	.598	.588	.563	.520	.324	.047	.204
Triglycerides	.711	.846	.802	.415	.602	.454	.158	.457	.384	.117
Uric acid	.939	.910	.957	.720	.623	.545	.446	.304	.353	.517
Calcium	.696	.522	.346	.589	.551	.635	.142	.357	.553	.153

PEEP, positive end-expiratory pressure.

Table 6. SMART: Prediction of liver and renal function indicators in severe sepsis from baseline plus serial data, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Alkaline phosphatase	0.869	.594	.689	.055	.878	.720	.809	.699	.670	.818
ALT	.959	.865	.772	.506	.497	.175	.016	.041	.161	.572
AST	.786	.659	.605	.180	.134	.302	—	—	.138	.426
GGT	.943	.810	.837	.689	.701	.683	.736	.652	.443	.415
Total bilirubin	.965	.982	.983	.889	.895	.912	.822	.927	.949	.933
BUN	.970	.970	.946	.906	.811	.844	.792	.419	.553	.429
Creatinine	.896	.879	.815	.716	.603	.593	.568	.312	.384	.359

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, glutamyl-glutamate aminotransferase; BUN, blood urea nitrogen.

SMART predictions of lymphocyte, monocyte, segmented neutrophil, band and granulocyte counts, differential percentage of granulocytes and lymphocytes, platelet count, and prothrombin time. Pearson product moments >.5 were recorded also for hematocrit, red blood cell count, white blood cell count, and partial thromboplastin time.

The distribution of predicted vs. observed correlation coefficients for continuous dependent variable predictions from baseline data in patients with severe sepsis is tabulated in Table 8. Through day 3, more than half of the plots of individual patients had r ≥ .7. For days 4 and 5, most multiple regression models were validated at r ≥ .5. Among predictions beyond 14 days, approximately 20% of r values were ≥.6. Clinically useful levels of accuracy, reflected as a Pearson product moment ≥.8, were noted in 50% of SMART models for continuous dependent variables at day 1, 47% for day 2, and 25% for day 3. Thereafter, through day 28, 14% to 22% of quantitative predictions in individual patients generated predicted vs. observed plots of r ≥ .8 accuracy.

Distribution of correlation coefficients for prospectively validated SMART predictions of continuous dependent variables in individual patients with severe sepsis from baseline plus serial data is listed in Table 9. Through day 5 from baseline, more than half of predicted vs. observed r were values >.5, and 53% had r values >.8 at day 3. On days 4–28, between 17% and 31% of serial data multiple regression models generated Pearson product moments of ≥.8.

Prospectively validated predictions of the dichotomous dependent variables chest radiograph score, need for mechanical ventilation, ARDS, DIC, hepatobiliary failure, renal insufficiency, shock, and cerebral dysfunction are seen in Tables 10 and 11. SMART models derived from baseline data only (Table 10) were validated at clinically useful levels of accuracy, with 38 of 56 (68%) ROC AUC determinations up to 7 days after baseline >.700. When serial plus baseline data were included in the modeling process (Table 11), predictive accuracy was even higher, as 46 of 56 (84%) of these equations were validated at ROC AUC > .700 during the first week.

Comparative examples of SMART predictions from baseline data only vs. baseline plus serial data are seen in Table 12. For continuous dependent variables, the frequency of cross-validated r > .5 and r

Table 7. SMART: Prediction of hematologic and coagulation indicators in severe sepsis from baseline plus serial data, r (predicted vs. observed)

	Day									
	1	2	3	4	5	6	7	14	21	28
Lymphocytes	.937	.982	.975	.989	.132	.996	.970	.994	.986	.981
Monocytes	.971	.989	.989	.387	.999	.161	.139	.988	.999	.998
Segmental neutrophils	.999	.999	.999	.999	.999	.999	.999	.999	.999	.999
Bands	.999	.989	.038	.519	.956	.995	.980	.095	.102	.386
Granulocytes (segs + bands)	.999	.999	.999	.857	.999	.999	.748	.658	.999	—
% Granulocytes	.999	.999	.999	.704	.999	.999	.999	.209	.999	—
Platelets	.921	.894	.759	.754	.754	.789	.726	.382	.743	.581
Prothrombin time	.932	.932	.991	.885	.912	.911	.900	.866	.849	.865

Table 8. SMART: Predicted vs. observed correlation coefficients for continuous dependent variables in patients with severe sepsis from baseline data

Day After Baseline	r Value (%)				
	≥.5	≥.6	≥.7	≥.8	≥.9
1	29/36 (81)	25/36 (69)	22/36 (61)	18/36 (50)	13/36 (36)
2	26/36 (72)	23/36 (64)	20/36 (56)	17/36 (47)	12/36 (33)
3	22/36 (61)	22/36 (61)	19/36 (53)	9/36 (25)	7/36 (19)
4	18/36 (50)	16/36 (44)	12/36 (33)	6/36 (17)	4/36 (11)
5	21/36 (58)	16/36 (44)	10/36 (28)	7/36 (19)	6/36 (17)
6	13/36 (36)	11/36 (31)	8/36 (22)	5/36 (14)	3/36 (8)
7	13/36 (36)	9/36 (25)	7/36 (19)	5/36 (14)	4/36 (11)
14	7/36 (19)	7/36 (19)	6/36 (17)	6/36 (17)	4/36 (11)
21	8/36 (22)	8/36 (22)	8/36 (22)	8/36 (22)	6/36 (17)
28	6/36 (17)	6/36 (17)	5/36 (14)	5/36 (14)	4/36 (11)

Table 9. SMART: Predicted vs. observed correlation coefficients for continuous dependent variables in patients with severe sepsis from baseline plus serial data

Day After Baseline	r Value (%)				
	≥.5	≥.6	≥.7	≥.8	≥.9
1	29/36 (81)	25/36 (69)	22/36 (61)	18/36 (50)	13/36 (36)
2	26/36 (72)	23/36 (64)	21/36 (58)	21/36 (58)	11/36 (31)
3	26/36 (72)	26/36 (72)	21/36 (58)	19/36 (53)	13/36 (36)
4	22/36 (61)	13/36 (36)	13/36 (36)	7/36 (19)	4/36 (11)
5	19/36 (53)	17/36 (47)	12/36 (33)	11/36 (31)	6/36 (17)
6	17/36 (47)	14/36 (39)	10/36 (28)	9/36 (25)	7/36 (19)
7	14/36 (39)	12/36 (33)	11/36 (31)	7/36 (19)	5/36 (14)
14	10/36 (28)	9/36 (25)	6/36 (17)	6/36 (17)	5/36 (14)
21	13/16 (36)	11/36 (31)	10/36 (28)	8/36 (22)	7/36 (19)
28	11/36 (31)	10/36 (28)	7/36 (19)	7/36 (19)	5/36 (14)

Table 10. SMART: Prediction of shock and organ failure in severe sepsis from baseline data only, ROC AUC

	Day									
	1	2	3	4	5	6	7	14	21	28
Chest radiograph score[a]	.913	.798	.787	.832	.931	.702	.703	.624	.625	1.000
Mechanical ventilation	.828	—	—	—	.669	—	.765	.427	.596	.582
ARDS	.713	.770	.586	.708	.696	.849	.801	.860	.980	—
DIC	.912	.806	.797	.719	.681	.884	.767	.263	.484	
Hepatobiliary failure	0.816	.929	.686	.875	.838	.778	.817	.701	.744	.951
Renal insufficiency	.910	.802	.866	.894	.842	.732	.753	.699	.717	.848
Shock	.654	.607	.595	.626	—	.750	.655	.563	.319	—
GCS <11	.845	.768	.685	.712	.687	.715	.6060	.486	.469	.328

ROC AUC, area under the curve of the receiver operating characteristic plot; ARDS, acute respiratory distress syndrome; DIC, disseminated intravascular coagulation dysfunction; GCS, Glasgow Coma Scale.

[a]0–4 quadrants pulmonary edema (Murray JF, et al. *Am Rev Respir Dis* 1988; 138:720–723). Organ failure and shock definitions were from Fisher CJ, Slotman GJ, Opal SM, et al. *Crit Care Med* 1994; 22:12–21.

> .6 at 3 and 7 days was similar in baseline and serial SMART models. At 3 days, 53% of serial equations yielded r > .8 vs. 25% of baseline models. Performance of serial and baseline SMAR models for dichotomous dependent variables varied widely.

To test specifically the ability of the SMART modeling process to predict organ failure and shock subclinically in patients with severe sepsis, baseline data from patients in the predictive cohort who did not have ARDS at baseline were entered into the SMART models for predicting ARDS from baseline data on days 1–28. Similarly, data from patients who did not have DIC at baseline were entered into the models for DIC, and so on, as well, for individual patients who did not have hepatobiliary failure, renal insufficiency, shock, and Glasgow Coma Scale <11 at baseline. Results of these analyses are tabulated in Table 13. SMART multiple logistic regression models predicted the presence or absence of ARDS, DIC, hepatobiliary failure, renal insufficiency, shock, and cerebral dysfunction in patients without each of these conditions at baseline up to 28 days in advance, with 25 of 60 (42%) achieving ROC AUC values of ≥0.7. Conversely, predicted vs. observed analysis for shock and each type of organ dysfunction was performed by using baseline data from predictive cohort patients who had shock or each defined organ dysfunction at baseline. These results are displayed in Table 14. In 38 of 60 (63%) of these predictive models, the ROC AUC for predicted vs. observed plots was >.5, thus predicting the continued presence, or resolution, of shock and organ failure.

To evaluate whether biological constancy of continuous dependent variables was a factor in promoting accurate r values, the percentage of individual patients in whom each quantitatively predicted variable changed by >25% from baseline was calculated. These results are displayed in Table 15. Only two of these variables (6%), prothrombin time and band leukocyte count, varied >25% from baseline in <40% of the patients evaluated. On the other hand, 22 of 36 (61%) of the continuous dependent variables varied >25% from baseline in >80% of the patients evaluated. For completion, on day 3, the coefficient of variation (R^2) for training cohort SMART models that predicted each continuous dependent variable was plotted vs. the percentage of patients in whom that variable had varied >25%. No statistically significant direct

Table 11. SMART: Prediction of shock and organ failure and survival in severe sepsis from baseline plus serial data, ROC AUC

	Day									
	1	2	3	4	5	6	7	14	21	28
Chest radiograph score[a]	.913	.841	.832	.849	.864	.738	.745	.532	.803	1.00
Mechanical ventilation	.828	—	—	—	.797	.853	.937	.669	.771	.828
ARDS	.713	.808	.600	.643	.809	.886	.737	.881	1.000	1.000
DIC	.912	.912	.874	.677	.702	.889	.447	—	.491	—
Hepatobiliary failure	.816	.789	.680	.845	.858	.873	.883	.789	.955	.953
Renal insufficiency	.910	.864	.864	.965	.880	.805	.854	.886	.854	.688
Shock	.654	.494	.838	.795	.803	.626	.652	.625	.264	—
Survival status	.752	.754	.684	.665	.657	.657	.621	.701	.841	.839
GCS <11	.845	.873	.857	.793	.748	.682	.743	.651	.783	—

ROC AUC, area under the curve of the receiver operating characteristic plot; ARDS, acute respiratory distress syndrome; DIC, disseminated coagulation dysfunction; GCS, Glasgow Coma Scale.

[a]0–4 quadrants pulmonary edema (Murray JF, et al. *Am Rev Respir Dis* 1988; 138:720–723). Organ failure and shock definitions were from Fisher CJ, Slotman GJ, Opal SM, et al. *Crit Care Med* 1994; 22:12–21.

Table 12. Summary of baseline vs. baseline plus serial SMART model performance

r values for continuous dependent variables (% of total)			
Day 3, r	>.5	>.7	>.8
Baseline, %	61	53	25
Baseline plus serial, %	72	58	53
Day 7			
Baseline, %	36	19	14
Baseline plus serial, %	39	31	19
ROC AUC for dichotomous dependent variables (% of total)			
Day 3, ROC AUC	>.6	>.7	>.8
Baseline, %	63	37	12.5
Baseline plus serial, %	100	53	56
Day 7			
Baseline, %	100	75	25
Baseline plus serial, %	89	67	33

Table 13. SMART: Prediction of shock and organ failure in septic shock from baseline data among patients without baseline shock or organ dysfunction, ROC AUC

	Day									
	1	2	3	4	5	6	7	14	21	28
ARDS	.544	—	.400	.503	.408	.799	.723	.879	1.000	—
DIC	.875	.908	.754	.705	.700	.802	.852	.292	.482	—
Hepatobiliary failure	.698	.771	.420	.773	.798	.839	.748	.564	.789	.952
Renal insufficiency	.957	.610	.917	.928	.881	.621	.765	—	.611	.816
Shock	.500	—	—	—	—	.444	—	—	—	—
GCS <11	.846	.375	.622	.566	.616	.670	.462	.232	.688	.259

ROC AUC, area under the curve of the receiver operating characteristic plot; ARDS, acute respiratory distress syndrome; DIC, disseminated intravascular coagulation dysfunction; GCS, Glasgow Coma Scale.

Organ failure and shock definitions were from Fisher CJ, Slotman GJ, Opal SM, et al. *Crit Care Med* 1994; 22:12–21.

Table 14. SMART: Prediction of shock and organ failure in septic shock from baseline data among patients with baseline shock or organ dysfunction, ROC AUC

	Day									
	1	2	3	4	5	6	7	14	21	28
ARDS	.198	—	.478	.333	.833	.521	.464	.625	.857	—
DIC	.950	.667	.550	.300	.500	—	—	—	—	—
Hepatobiliary failure	.491	.768	.893	.931	.792	.908	.837	.788	.545	.974
Renal insufficiency	.545	.453	.373	.417	.595	.688	.579	—	.680	.861
Shock	.591	.519	.537	.599	—	.811	.653	.573	.253	—
GCS <11	.703	.785	.546	.769	.576	.733	.633	.767	—	—

ROC AUC, area under the curve of the receiver operating characteristic plot; ARDS, acute respiratory distress syndrome; DIC, disseminated intravascular coagulation dysfunction; GCS, Glasgow Coma Scale. Organ failure and shock definitions were from Fisher CJ, Slotman GJ, Opal SM, et al. *Crit Care Med* 1994; 22:12–21.

or inverse relationship was found between the degree of variation for each variable and the R^2 of the multiple regression predictive models. In 218 of 245 (89%) continuous baseline models for days 1–7, the baseline value of the dependent variable had the highest predictive weight compared with 68 of 105 (65%) of the equations for days 14–28 ($p < .01$). Similarly, for serial models, the baseline dependent variable was the strongest dependent variable in 187 of 245 (76%) of days 1–7, vs. 43 of 105 (41%) of days 14–28 ($p < .001$). Among dichotomous models, baseline dependent variables were highest predictors in 44 of 56 (79%) baseline equations for 7 days, vs. 7 of 24 (29%) on days 14–28 ($p < .01$), and for serial models 37 of 53 (59%) on days 1–7 vs. 5 of 27 (19%) on days 14–28 ($p < .001$).

Stepwise multiple regression prediction of IL-8, GCSF, and IL-6 from baseline data only is displayed in Table 16. These inflammatory response mediators were measured in plasma at baseline and at days 1, 2, 3, and 4 thereafter. Predicted vs. observed Pearson product moments as high as .988 were achieved by inserting data from individual patients from the predictive cohort into the multiple regression models for each mediator. IL-6, IL-8, or GCSF was correlated significantly in 103 of 360 (29%) of baseline models and in 112 of 360 (31%) of baseline plus serial data model matrixes for continuous dependent variables, and in 47 of 66 (71%) of baseline models and 54 of 64 (84%) of baseline plus serial data models for dichotomous dependent variables. Ninety-four of the SMART predictive models had two or more of these inflammatory response mediators correlated as independent variables.

Predicted vs. observed ROC AUC values from baseline models for day 5, when mathematically selected models were compared with those into which additional pathophysiologic variables were "forced," respectively, included the following: chest radiograph score, 0.931 vs. 0.862; ARDS, 0.696 vs. 0.742; DIC, 0.681 vs. 0.736; hepatobiliary failure, 0.838 vs. 0.764; and renal insufficiency, 0.842 vs. 0.727. Shock at day 5 was not predicted successfully by purely mathematical baseline modeling, but with forced variables had an ROC AUC of .608.

The equations developed to predict each dependent variable are available from the author for research and independent confirmation.

Table 15. Variation of continuous dependent variables in patients with severe sepsis and/or septic shock from baseline data

Variable	Patients, n	>25% Change from Baseline, %	Range of Variable
Albumin	269	62	0.3–5.8 g/dL
Alkaline phosphatase	279	91	7–2140 units/L
ALT	259	95	0–3590 units/L
AST	281	96	0–4410 units/L
BUN	286	92	1–115 mg/dL
Calcium	281	26	1–14 mg/dL
Cholesterol	255	87	5–837 mg/dL
Creatinine	289	82	0.1–12.2 mg/dL
GGT	233	92	0–1499 units/L
Glucose	287	91	22–787 mg/dL
Hematocrit	284	47	2.8–50%
Hemoglobin	287	44	1.1–17.0 mg/dL
Heart rate	287	62	26–188 bpm
Platelet count	287	94	7–1262 $\times 10^3$
Total protein	268	60	0.48–14.9 g/dL
PT	278	34	10–146 secs
PTT	276	63	18–240 secs
RBC	277	44	1.47–9.80 K/μL
Total bilirubin	281	95	0.06–60 mg/dL
Triglycerides	250	93	7–1640 mg/dL
Uric acid	267	88	0.4–19.2 mg/dL
WBC	287	95	800–138,000/mm^3
Glasgow Coma Scale	200	95	3–15
HCO_3	255	53	2–46 mEq/L
FIO_2	262	87	21–100%
Lymphocytes	252	98	0–2900/mm^3
Monocytes	246	100	0–1848/mm^3
Segmental neutrophils	239	64	0–36,064/mm^3
Bands	181	99	0–2002/mm^3
Granulocytes	206	39	0–36,432/mm^3
% Granulocytes	206	92	0–99%
% Lymphocytes	253	98	0–28%
CI	156	63	1.3–9.6 $L \cdot min^{-1} \cdot m^{-2}$
SVR	159	74	36–8125 dyne · sec/cm^5
PEEP	152	93	0–25 cm H_2O
Pressure support	81	99	0–70 cm H_2O

ALT, alanine aminotransferase; AST, aspartate aminotransferase; BUN, blood urea nitrogen; GGT, glutamyl-glutamate aminotransferase; PT, prothrombin time; PTT, partial thromboplastin time; RBC, red blood cell count; WBC, white blood cell count; CI, cardiac index; SVR, systemic vascular resistance; PEEP, positive end-expiratory pressure.

DISCUSSION

The results of the present study demonstrate that models developed with the SMART approach predict end-organ function indicators quantitatively, and prognosticate the presence or absence of shock and organ dysfunction up to 28 days in advance, in individual patients with severe sepsis and septic shock. Many of these SMART predictive engines were validated prospectively at clinically useful levels of accuracy. Multivariate logistic regression models predicting the dichotomous dependent variables of shock, chest radiograph score, DIC, ARDS, cerebral dysfunction, and liver and renal failure in severe sepsis also were validated prospectively in this study. Furthermore, some SMART models predicted the onset of shock and organ failure in patients without each of these systemic inflammatory conditions at baseline, as well as the continuation/resolution of shock and organ failure among patients who entered this study with these septic complications already established. Biological constancy of organ function indicators did not ap-

Table 16. SMART: Prediction of inflammatory response mediators in severe sepsis from baseline data only, r (predicted vs. observed)

	Day			
	1	2	3	4
IL-8	.614	.917	.015	.678
GCSF	.856	—	.804	.671
IL-6	.829	.988	.300	.117

IL, interleukin; GCSF, granulocyte colony stimulating factor.

pear to underlie SMART's predictive accuracy, because all but two continuous dependent variables varied by >25% from baseline in >40% of the patients studied. The importance of measuring inflammatory response mediators to these prognostications was demonstrated by the correlation of IL-6, IL-8, and/or GCSF in developing 29% to 31% of SMART models predicting continuous dependent variables, and in 71% to 84% of equations for dichotomous end points. Finally, plasma concentrations of IL-6, IL-8, and GCSF themselves were predicted accurately from baseline data up to 4 days in advance. A review of the literature indicates that prospectively validated prognostic models predicting continuous and dichotomous indicators of cardiopulmonary performance and end-organ function in individual patients have not been reported previously. These results are significant findings of this study.

One of the most clinically important aspects of this investigation was the prospectively validated, quantitative prediction of physiologic variables and standard hospital laboratory organ function indicators. This information was known days and sometimes weeks in advance, from both baseline and serial data input. Although the intrinsic value of knowing what hemodynamics, oxygen requirements, coagulation profiles, white blood cell counts and differential, or renal and liver function profiles will be days or even weeks in advance may be intuitively obvious, hypothetical practical application may clarify its potential uses. For example, therefore, if the intensivist at the bedside knows that the patient will require increasing F_{IO_2} and PEEP during the next 3 days, then more aggressive treatment of known pulmonary pathology and, particularly in the septic patient, a search for unrecognized infections or other sources of systemic inflammation will be instituted early, with possible beneficial therapeutic results. Similarly, advance prediction of increasing bilirubin and alkaline phosphatase could lead to early discovery and treatment of acute biliary pathology, whereas predicted increases in blood urea nitrogen and creatinine might lead physicians to optimize renal perfusion before irreversible tissue damage occurred. Early warning of impending thrombocytopenia could alert the physician to occult continued sepsis and inflammation-induced organ failure or to subclinical bleeding, for which early intervention could be lifesaving. Advance prediction of white blood cell count and, particularly, the differential cell count in the septic patient could give additional subclinical clues about the developing presence of secondary infections. Conversely, knowing that previously known pathologic concentrations of these variables were normalizing could help decision making regarding the length of antibiotic treatment, the need for ventilatory and hemodynamic support, or the adequacy of surgical drainage of infection. These examples are practical possibilities for the SMART models reported in this study.

Several of the SMART models were validated prospectively at levels of accuracy that approached clinical significance. Among continuous dependent variables, respiratory function variables, such as requirements of pressure support ventilation or PEEP, were predicted up to 3 days in advance, with $r > .700$ for baseline models and as high as .904 when serial data were included. Predictions of Glasgow Coma Scale, serum bicarbonate, albumin, cholesterol, triglycerides, and uric acid up to 7 days in advance had $r > .800$. Indicators of liver and renal function were predicted effectively, consistently yielding prospective $r \geq .700$ through 7 days and, for alkaline phosphatase and bilirubin, up to 28 days. This was true also for quantitative SMART predictions of hematologic and coagulation results, as counts of lymphocytes, monocytes, segmented neutrophils, bands, total granulocytes and platelets, and prothrombin time all had predicted vs. observed linear regression correlations >.9, with many at that level of accuracy through 28 days. Advance knowledge of this breadth and quantitative quality among individual septic patients has not been available previously. Its potential clinical usefulness has not yet been tested, but the possibilities are myriad.

Prospectively validated SMART predictions of shock and end-organ dysfunction, by using clinical definitions established in the literature, were another important finding of this study. For example, baseline logistic regression models predicted pulmonary edema scores, the need for mechanical ventilation, and the presence or absence of ARDS, shock, DIC, and acute hepatobiliary, renal, and central nervous system dysfunction accurately. Among SMART models for the first 7 days after the onset of severe sepsis/septic shock, 70% were validated prospectively at ROC AUC > .700. For pulmonary edema scoring, mechanical ventilation, ARDS, hepatobiliary dysfunction, and acute renal failure, SMART prognostic models were clinically reliable up to 28 days in advance. SMART also predicted 28-day survival with predicted vs. observed ROC AUCs as high as .839. Although some studies have correlated physiologic values and even some measurement of circulating inflammatory mediators with broad groupings of multiple organ failure (26, 27), the specific, prospectively validated predictions of shock, pulmonary edema, end-organ dysfunction, and survival demonstrated in the present study have not been reported previously. These results represent a significant methodological advance in outcomes research.

Another prognostic improvement manifested in the SMART models presented here is that their predictions are applicable to individual septic patients. Previously described illness severity scoring systems, including the Sepsis Score (1), Simplified Acute Physiology Score II (8), Multiple Probability Model II (6), Acute Physiology and Chronic Health Examination II and III (10, 11), and International Classification of Disease Illness Severity Score (28, 29), among others, generally have predicted only grouped risk of intensive care unit or hospital mortality. Even when serial physiologic assessments (30) were included in these scoring models, or Bayesian analysis was used to predict mortality rates (31), prognostication still has been limited to relative risk of mortality, and, therefore, has not been applicable to individual patients. Similarly, attempts at predicting multiple organ dysfunction (26, 27, 32, 33), ventilator dependence (34), or length of stay and hospital costs (28, 29) have faltered, by being insufficiently accurate for clinical usefulness, by lacking prospective validation, by not being applicable to individual patients, and/or by predicting only wide groupings of organ dysfunction or length of stay probabilities. In contrast, the SMART models presented here approach prognostications involving critically ill patients from the opposite direction, that is, by predicting continuous and dichotomous dependent variables specifically in individual patients. Many SMART quantitative and qualitative predictions of laboratory test results and clinical events occurring days (and sometimes even weeks) later were validated prospectively at clinically useful levels of accuracy.

The SMART modeling approach differs from conventional prognostic methods in several ways. Whereas physiologic scoring systems use clinically obvious information to predict mortality probability, SMART explores relationships among demographics, physiologic data, standard hospital laboratory tests, and subclinical measurement of inflammatory response mediators and analyzes these relationships to predict clinical events that determine the course of each patient. Mortality scoring methods group patients together who have like mortality risks but widely varying pathophysiologies; SMART predicts major clinical changes that may affect survival, based on results from patients with similar mechanisms of illness. Acute Physiology and Chronic Health Examination III, Multiple Probability Model II, Simplified Acute Physiology Score II, and others assess relative risk of death, while assuming, statistically, at least, that contributing factors remain constant and that outcome is not altered by treatment. SMART can predict conditions that may not yet be evident externally, and, therefore, it might improve outcomes by making possible early intervention and timely modification of the host inflammatory response.

In this study, circulating concentrations of inflammatory response mediators contributed to many SMART models. Plasma IL-6, IL-8, or GCSF was correlated significantly in the development of 29% of baseline predictions for continuous dependent variables and in 31% of equations derived from baseline plus serial data. The value of measuring inflammatory mediators was even greater in SMART predictions of shock and organ dysfunction, where 71% of baseline models and 84% of serial data model building had cytokines as correlated factors. Considering the predictive successes described here, one might speculate that the accuracy of SMART equations could be optimized further by measuring additional inflammatory response mediators that also have been associated with sepsis, shock, and organ failure. If IL-1β (33, 34), IL-2 (35), IL-8 (36, 37), tumor necrosis factor-α (35, 38, 39, 40), endothelin (41), E-, P-, and L-selectin (42, 43), intercellular adhesion molecule and other adhesion molecules (44), neutrophil elastase (43, 45, 46), antithrombin III (45), leukotrienes (41, 38), prostaglandins (17, 38), activated complement (17), and IL-6 (38, 47), among others, were all included simultaneously as independent variables in SMART models, it may be that clinically useful predictive accuracy could be achieved for many more dependent variables than was possible with the present database. Measurements of anti-inflammatory response reactants, like IL-10 and IL-1 receptor antagonist (48), similarly might improve SMART's performance.

Multiple regression models that predicted circulating IL-6, IL-8, and GCSF themselves in this study were validated prospectively, with r values as high as .988. These findings confirmed a previous retrospective report of SMART in which plasma prostaglandins, leukotrienes, IL-1, IL-6, and tumor necrosis factor were predicted in septic patients (19). Although confirmatory quality control data regarding the assays used to measure these substances are lacking, the available data nevertheless suggest that, by analyzing interactions among pathophysiologic observations, standard laboratory tests, and levels of inflammatory response reactants in blood, and possibly in other biological fluids, concentrations of mediators such as IL-1, tumor necrosis factor, and others can be known in advance in critically ill patients. If such SMART models can be matured to consistently high quantitative accuracy, then patients with life-threatening systemic inflammatory conditions can be matched subclinically with novel drugs directed specifically at the cytokine or other substance that drives each individual's pathophysiology. This kind of advance information could improve selecting patients for clinical trials to the point of facilitating determination of new drug efficacy for sepsis and other life-threatening systemic illnesses.

In predicting continuous dependent variables among septic patients, the question of biological constancy must be considered: That is, does SMART's predictive success merely reflect lack of change in the organ function indicators studied? This possibility was examined in two ways. First, the percentage of individual patients in whom each continuous dependent variable changed after baseline by >25% was calculated. Only two variables, 6% of the continuous dependent variables, varied by 25% in <40% of the database patients. On the other hand, 61% of the continuous dependent variables changed by >25% in >80% of the patients studied. Biological constancy, therefore, was rare. This concept was further evaluated by plotting the coefficients of variation (R^2) from day 3 baseline models against the percentage of patients in whom each predicted variable had varied >25%, reasoning that an inverse correlation between percentage variation and R^2 would indicate a prime influence of biological constancy. No such relationship existed, suggesting that SMART may have potential as a predictive clinical aid. However, that the baseline values of dependent variables are important factors in multiple-variant modeling was evidenced by their roles as prominent predictors, significantly more frequently in the first week than on days 14, 21, or 28.

The question of whether serial measurements of independent variables add to the prognostic accuracy of SMART models is not answered by the present data. Analysis of performance between the baseline only and baseline plus serial data prognostications revealed similar frequencies of r values up to .700 through day 7 from SMART models developed for baseline and from baseline plus serial data. Serial models achieved r values in the validation cohort >.800 in 53% of models at 3 days, compared with 25% for baseline equations. Results varied widely for SMART models that predicted dichotomous dependent variables. From these results, then, one might speculate that serial data should supplement baseline information in building reference databases, until the true value of repeated daily measurements to SMART predictive models can be determined.

Whether SMART mathematical models should have independent variables that make pathophysiologic sense "forced" into them, so as to improve prognostic accuracy, is not clear from the data. Although baseline variables that were forced into mathematical models modestly improved prediction of ARDS and DIC and enabled day 5 prognostication of shock, there was no benefit for the other dichotomous dependent variables. This suggests that clinically relevant independent observations that were not mathematically weighted predictors might help to optimize some SMART equations. More comprehensive assessment of these questions should be done in larger SMART databases.

Several limitations of this preliminary report should be considered. First, the number of patients in the training cohort and in the predictive/validation population was small, perhaps accounting for some of the variation in predictive accuracy. In the future, before SMART predic-

The Systemic Mediator Associated Response Test (SMART integration of demographics, bedside physiology, hospital laboratory tests, and circulating cytokines predicts organ failure and physiologic function indicators in individual patients with severe sepsis and septic shock.

tions can become useful at the bedside, they need to be refined on much larger databases to perform consistently at the 90% level of accuracy. Second, although the cross-validation in this study was prospective in that it was done with data that were not used to develop the equations, the acid test of the final SMART models will be to make real-time predictions in individual patients and then follow them clinically for outcomes. Third, in an effort to capture as many weighted values as possible, independent variables were screened for each dependent variable. As discussed previously, in the mature SMART sepsis, this wide-angle statistical approach will be supplemented by inserting into the models independent variables that make physiologic sense but are not screened as significant, so to add clinical judgment. Ultimately, only the minimum number of independent variables that are needed to achieve clinically useful accuracy will be included in each equation. Fourth, every result presented in this article, whether Pearson product moment or ROC AUC, describes the aggregate agreement between up to 103 predicted vs. observed pairs. The data were presented in this format consistent with the stated objective of validating the models. Because the ranges of the dependent and independent variables in septic shock are well known, including the measurements seemed superfluous and potentially confusing to this report. Finally, although the models presented here are a preliminary step in the development of SMART, these early equations may be difficult for other investigators to evaluate independently because key independent variables, such as the specific cytokines measured in the subject clinical trial, might not be common to other available databases.

The results of this study validate prospectively multiple regression models, by using baseline and/or baseline plus serial determinations, that predict physiologic variables and concentrations of circulating metabolic and organ function indicators, as well as cytokines, in severe sepsis and septic shock, quantitatively, up to 28 days in advance. Multivariate logistic regression equations were confirmed prospectively, also, that predicted the presence, occurrence, and resolution of shock and organ system dysfunction in septic patients. These prognostications are applicable to individual patients, frequently at clinically useful levels of accuracy. Plasma cytokine concentrations contributed significantly as weighted independent variables to many quantitative models and to most predictions of dichotomous dependent variables. Future prospective studies should examine further the value of serial independent data collection and should screen a more comprehensive list of inflammatory response reactants as possible independent predictors.

REFERENCES

1. Dominioni L, Dionigi R, Zanello M, et al: Sepsis score and acute-phase protein response as predictors of outcome in septic surgical patients. *Arch Surg* 1987; 122: 141–146
2. Marshall JC, Cook DJ, Christou NV, et al: Multiple organ dysfunction score: A reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995; 23:1638–1652
3. Vincent JL, Moreno R, Takala J, et al: The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* 1996; 22:707–710
4. Vincent JL, Mendonca 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 multicenter, prospective study. *Crit Care Med* 26:1793–1800
5. Teres D, Brown RB, Lemeshow S: Predicting mortality of intensive care unit patients. The importance of coma. *Crit Care Med* 1982; 10:86–95
6. Lemeshow S, Teres D, Klar J, et al: Mortality probability models (MPM II) based on an international cohort of intensive care unit patients. *JAMA* 1993; 270:2478–2486
7. Le Gall J, Loirat P, Alperovith A, et al: A simplified acute physiology score for ICU patients. *Crit Care Med* 1984; 12:975–977
8. Le Gall J, 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
9. Knaus WA, Zimmerman JE, Wagner DP, et al: APACHE—Acute Physiology and Chronic Health Evaluation: A physiologically based classification system. *Crit Care Med* 1981; 9:591–597
10. Knaus WA, Draper EA, Wagner DP, et al: APACHE II: A severity of disease classification system. *Crit Care Med* 1985; 13: 818–829
11. Knaus WA, Wagner DP, Draper EA, et al: The APACHE III prognostic system: Risk prediction of hospital mortality for critically ill hospitalized adults. *Clin Invest Crit Care* 1991; 100:1619–1636
12. Castella X, Artigas A, Bion J: 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
13. Knaus WA, Harrell FE, Fisher CJ: The clinical evaluation of new drugs for sepsis. A prospective study design based on survival analysis. *JAMA* 1993; 270:1233–1241
14. Seneff MG, Zimmerman JE, Knaus WA, et al: Predicting the duration of mechanical ventilation. The importance of disease and patient characteristics. *Chest* 1996; 110:469–479
15. Meduri GU, Headley S, Kohler G, et al: Persistent elevation of inflammatory cytokines predicts a poor outcome in ARDS. Plasma IL-1β and IL-6 levels are consistent and efficient predictors of outcome over time. *Chest* 1995; 107:1062–1073
16. Dinarello CA: Proinflammatory and anti-inflammatory cytokines as mediators in the pathogenesis of septic shock. *Chest* 1997; 112:321S–329S
17. Slotman GJ, Burchard KW, Yellin SA, et al: Prostaglandin and complement interaction in clinical acute respiratory failure. *Arch Surg* 1986; 121:271–274
18. Slotman GJ, Quinn JV: Multivariate regression modeling for the prediction of inflammation, systemic pressure, and end-organ function in severe sepsis. *Shock* 1997; 8:225–231
19. Slotman GJ, Quinn JV: Multivariate regression modeling for the prediction of plasma eicosanoid and cytokine concentrations in patients with severe sepsis. *J Surg Outcomes* 1998; 1:24–30
20. Fisher CJ, Dhainaut JFA, Opal SM, et al: Recombinant human interleukin-1 receptor antagonist in the treatment of patients with sepsis syndrome. *JAMA* 1994; 271:1836–1843
21. SAS Institute: SAS/STAT User's Guide, Version 6. Fourth Edition. Volumes 1 & 2. Cary, NC, SAS Institute, 1994
22. Afifi A, Clark C: Computer-Aided Multivariate Analysis. Second Edition. New York, Van Nostrand Reinhold, 1990

23. Kirkwood BR: Essentials and Medical Statistics. Oxford, UK, Blackwell Scientific, 1988
24. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29–36
25. Ferguson GA: Statistical Analysis in Psychology and Education. Third Edition. New York, McGraw Hill, 1971
26. Sauaia A, Moore FA, Moore EE, et al: Early predictors of post-injury multiple organ failure. *Arch Surg* 1994; 129:39–45
27. Sauaia A, Moore FA, Moore EE, et al: Multiple organ failure can be predicted as early as 12 hours after injury. *J Trauma* 1998; 45: 291–303
28. Osler TM, Rutledge R, Deis J, et al: ICISS: An International Classification of Disease-9 based Injury Severity Score. *J Trauma* 1996; 41:380–388
29. Osler TM, Rogers FB, Glance LG, et al: Predicting survival, length of stay, and cost in the surgical intensive care unit: APACHE II versus ICISS. *J Trauma* 1998; 45:234–237
30. Wagner DP, Knaus WA, Harrell FE, et al: 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
31. Millili JJ, Philiponis VS, Nusbaum M: Predicting surgical outcome using Bayesian analysis. *J Surg Res* 1998; 77:45–49
32. Barie PS, Hydo LJ, Fischer E: Utility of illness severity scoring for prediction of prolonged surgical critical care. *J Trauma* 1996; 40:513–519
33. Barie PS, Hydo LJ, Fischer E: Development of multiple organ dysfunction syndrome in critically ill patients with perforated viscus. *Arch Surg* 1996; 131:37–43
34. Sellers BJ, Davis BL, Larkin PW, et al: Early prediction of prolonged ventilator dependence in thermally injured patients. *J Trauma* 1997; 43:899–903
35. Gando S, Nakanishi Y, Tedo I: Cytokines and plasminogen activator inhibitor-1 in post-trauma disseminated intravascular coagulation: Relationship to multiple organ dysfunction syndrome. *Crit Care Med* 1995; 23: 1835–1842
36. Slotman GJ, Friedman B, Brathwaite C, et al: Interleukin-1 mediates increased plasma levels of eicosanoids and cytokines in patients with sepsis syndrome. *Shock* 1995; 4:318–323
37. Pinsky MR, Vincent JL, Deviere J, et al: Serum cytokine levels in human septic shock. Relation to multiple-system organ failure and mortality. *Chest* 1993; 103:565–575
38. Marty C, Misset B, Tamion F, et al: Circulating interleukin-8 concentrations in patients with multiple organ failure of septic and nonseptic origin. *Crit Care Med* 1994; 22: 673–679
39. Miller EJ, Cohen AB, Matthay MA: Increased interleukin-8 concentrations in the pulmonary edema fluid of patients with acute respiratory distress syndrome from sepsis. *Crit Care Med* 1996; 24:1448–1454
40. Függer R, Zadrobilek E, Götzinger P, et al: Perioperative TNFα and IL-6 concentrations correlate with septic state, organ function, and APACHE II scores in intra-abdominal infection. *Eur J Surg* 1993; 159:525–529
41. Mitaka C, Hirata Y, Nagura T: Circulating endothelin-1 concentrations in acute respiratory failure. *Chest* 1993; 104:476–480
42. Cowley HC, Heney D, Gearing AJH, et al: Increased circulating adhesion molecule concentrations in patients with the systemic inflammatory response syndrome: A prospective cohort study. *Crit Care Med* 1994; 22:651–657
43. Schlag G, Redl H: Mediators of injury and inflammation. *World J Surg* 1996; 20: 406–410
44. Boldt J, Wollbrück M, Kuhn D, et al: Do plasma levels of circulating soluble adhesion molecules differ between surviving and nonsurviving critically III patients? *Chest* 1995; 107:787–792
45. Nast-Kolb D, Waydhas C, Gippner-Steppert C, et al: Indicators of the post-traumatic inflammatory response correlate with organ failure in patients with multiple injuries. *J Trauma* 1997; 42:446–455
46. Davis JM, Meyer JD, Barie PS, et al: Elevated production of neutrophil leukotreine B_4 precedes pulmonary failure in critically ill surgical patients. *Surg Gynecol Obstet* 1990; 170:495–500
47. Meade P, Shoemaker WC, Donnelly TJ, et al: Temporal patterns of hemodynamics, oxygen transport, cytokine activity, and complement activity in the development of adult respiratory distress syndrome after severe injury. *J Trauma* 1994; 36:651–657
48. Parsons PE, Moss M, Vannice JL, et al: Circulating IL-1ra and IL-10 levels are increased but do not predict the development of acute respiratory distress syndrome in at-risk patients. *Am J Respir Crit Care Med* 1997; 155: 1469–1473