10-15-2017

A Two-Biomarker Model Predicts Mortality in the Critically Ill with Sepsis.

Carmen Mikacenic

Brenda L Price

Susanna Harju-Baker

D Shane O'Mahony
Department of Pulmonary and Critical Care Medicine, Swedish Medical Center, Seattle, Washington

Cassianne Robinson-Cohen

See next page for additional authors

ORIGINAL ARTICLE

A Two-Biomarker Model Predicts Mortality in the Critically Ill with Sepsis

Carmen Mikacenic[1], Brenda L. Price[2], Susanna Harju-Baker[1], D. Shane O'Mahony[3], Cassianne Robinson-Cohen[4], Frank Radella[1], William O. Hahn[5], Ronit Katz[4], David C. Christiani[6,7], Jonathan Himmelfarb[4], W. Conrad Liles[5]*, and Mark M. Wurfel[1]*

[1]Division of Pulmonary and Critical Care Medicine, Department of Medicine, [2]Department of Biostatistics, [4]Kidney Research Institute, and [5]Division of Allergy and Infectious Diseases, Department of Medicine, University of Washington, Seattle, Washington; [3]Department of Pulmonary and Critical Care Medicine, Swedish Medical Center, Seattle, Washington; [6]Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; and [7]Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts

Abstract

Rationale: Improving the prospective identification of patients with systemic inflammatory response syndrome (SIRS) and sepsis at low risk for organ dysfunction and death is a major clinical challenge.

Objectives: To develop and validate a multibiomarker-based prediction model for 28-day mortality in critically ill patients with SIRS and sepsis.

Methods: A derivation cohort (n = 888) and internal test cohort (n = 278) were taken from a prospective study of critically ill intensive care unit (ICU) patients meeting two of four SIRS criteria at an academic medical center for whom plasma was obtained within 24 hours. The validation cohort (n = 759) was taken from a prospective cohort enrolled at another academic medical center ICU for whom plasma was obtained within 48 hours. We measured concentrations of angiopoietin-1, angiopoietin-2, IL-6, IL-8, soluble tumor necrosis factor receptor-1, soluble vascular cell adhesion molecule-1, granulocyte colony–stimulating factor, and soluble Fas.

Measurements and Main Results: We identified a two-biomarker model in the derivation cohort that predicted mortality (area under the receiver operator characteristic curve [AUC], 0.79; 95% confidence interval [CI], 0.74–0.83). It performed well in the internal test cohort (AUC, 0.75; 95% CI, 0.65–0.85) and the external validation cohort (AUC, 0.77; 95% CI, 0.72–0.83). We determined a model score threshold demonstrating high negative predictive value (0.95) for death. In addition to a low risk of death, patients below this threshold had shorter ICU length of stay, lower incidence of acute kidney injury, acute respiratory distress syndrome, and need for vasopressors.

Conclusions: We have developed a simple, robust biomarker-based model that identifies patients with SIRS/sepsis at low risk for death and organ dysfunction.

Keywords: sepsis; systemic inflammatory response syndrome; biomarkers; IL-8; tumor necrosis factor receptor

(Received in original form November 18, 2016; accepted in final form April 14, 2017)

*These authors contributed equally.

Supported by the NHLBI (K23HL120896, P50HL073996, and R01HL060710), the National Institute of Allergy and Infectious Diseases (T32AI007044), the Parker B. Francis Fellowship, and the University of Washington Department of Medicine. The funding sources had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. C.M., M.M.W., and W.C.L. have a pending patent broadly related to this work.

Author Contributions: Study concept and design, C.M., B.L.P., J.H., W.C.L., and M.M.W. Acquisition, analysis, or interpretation of data, all authors. Drafting of the manuscript, C.M. Critical revision of the manuscript for important intellectual content, all authors. Statistical analysis, B.L.P., C.R.-C., and R.K. Obtained funding, W.C.L., D.C.C., and M.M.W. Administrative, technical, or material support, C.M., S.H.-B., D.S.O'M., C.R.-C., F.R., W.O.H., D.C.C., and J.H. Study supervision, R.K., D.C.C., J.H., W.C.L., and M.M.W.

Correspondence and requests for reprints should be addressed to Carmen Mikacenic, M.D., Box 359640, 325 9th Avenue, Seattle, WA 98104. E-mail: cmikacen@uw.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 196, Iss 8, pp 1004–1011, Oct 15, 2017

Originally Published in Press as DOI: 10.1164/rccm.201611-2307OC on April 18, 2017
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: The accurate identification of critically ill patients with systemic inflammatory response syndrome or sepsis at low risk for death is an impediment to optimal allocation of resources and triage in this population.

What This Study Adds to the Field: This study develops and validates a prediction model for death in critically ill patients with systemic inflammatory response syndrome based on plasma levels of two biomarkers: IL-8 and soluble tumor necrosis factor receptor-1, measured early in their intensive care unit admission. This model identifies patients at low risk of death and who have a lower incidence of acute kidney injury, acute respiratory distress syndrome, and shock.

Sepsis is a dysregulated response to infection that is one of the most common reasons for intensive care unit (ICU) admission in the United States and a major cause of mortality in critically ill populations worldwide (1–5). The systemic inflammatory response syndrome (SIRS) criteria, designed to encapsulate the physiologic response to critical illness, have been part of the sepsis definition for more than two decades (6, 7). Recent proposed changes to this definition reflect the need to better stratify patient risk in a heterogeneous syndrome (8). The Acute Physiology and Chronic Health Evaluation (APACHE) III score is a well-validated predictor of hospital mortality in ICU patients, but calculation of the score requires extensive clinical data collected over the first 24 hours of admission, thereby limiting its utility for bedside decision-making (9). Identification of patients at low risk for mortality or organ dysfunction in sepsis may improve resource use and allow for more targeted approaches for rational clinical trials.

Endothelial dysfunction, inflammation, and apoptosis have been implicated in the pathogenesis of sepsis. Circulating biomarkers may indicate the level of activity of these pathways and have prognostic utility for the early identification of patients at risk for multiorgan dysfunction and death. Prior work has shown that angiopoietin (Ang)-1 and Ang-2, angiogenic factors that regulate endothelial stability and permeability, respectively, are predictive of mortality in patients with severe sepsis (10, 11). Circulating levels of inflammatory pathway proteins IL-6, IL-8, and soluble tumor necrosis factor-α receptor 1 (sTNFR-1) have been shown to predict mortality or organ dysfunction in patients with established septic shock (12–16). Apoptosis has also been implicated in the pathophysiology of sepsis-related organ failure (17, 18). We have previously shown that early measurements of circulating biomarkers of these pathways are associated with mortality in patients admitted to the ICU meeting SIRS criteria and in patients with bacteremia (19, 20). We hypothesized that a multiple-biomarker-based model could estimate the probability of 28-day mortality in a diverse group of critically ill patients presenting with SIRS and/or sepsis.

Some of the results of these studies have been previously reported in the form of an abstract (21).

Methods

Derivation and Internal Test Cohort

Patients were recruited between 2006 and 2010 from ICUs at Harborview Medical Center (Seattle, WA) (19, 22). Adult patients who met two out of four SIRS criteria were prospectively enrolled on admission to the ICU. We obtained plasma specimens within 24 hours of admission to the ICU. The study was approved by the University of Washington Human Subjects Research Committee and granted a waiver of consent.

External Validation Cohort

Patients were enrolled from ICUs at Massachusetts General Hospital (Boston, MA) between 1999 and 2010 (23). Adult patients were enrolled on admission to the ICU if they had a defined risk factor for acute respiratory distress syndrome (ARDS) including sepsis, trauma, multiple transfusion, or aspiration. Plasma specimens were obtained within 48 hours of ICU admission (24). This study was approved by the Massachusetts General Hospital Human Subjects Research Committee and signed consent was obtained from each patient or legal surrogate.

Biomarker Measurement

Plasma concentrations of Ang-1, Ang-2, IL-6, IL-8, sTNFR-1, soluble vascular cell adhesion molecule-1 (sVCAM-1), granulocyte colony–stimulating factor, and soluble Fas were measured using an immunoassay-based method (Meso Scale Discovery, Rockville, MD) (19). We assigned the lower limit or upper limit of detection to samples that fell below or above the range of detection, respectively. The detection limits and number of samples below or above the limits of detection are available in the online supplement METHODS section and Table E1.

Model Selection and Development

Models were constructed from the $\log_{10}$-transformed biomarkers and the covariates of age, sex, source of ICU admission (medical/surgical), diabetes, chronic kidney disease, cirrhosis, and body mass index (Figure 1).

Our primary outcome of interest was inpatient mortality at Day 28 after enrollment. The initial model-building methodology used least absolute shrinkage and selection operator (LASSO) (25). The LASSO is a model-building strategy that performs variable selection by constraining the sum of the regression coefficients, thereby selecting variables for the model and penalizing less predictive variables to prevent overfitting. We subsequently identified the best two-biomarker model using all subsets selection logistic regression. We also considered the Sequential Organ Failure Assessment score (SOFA) as a continuous variable to predict mortality (26). Discrimination power was quantified using C statistics (area under the receiver operating characteristic curve [AUC]) and appropriate goodness of fit was verified via Hosmer-Lemeshow chi-square statistics (27).

We compared the capability of these models with APACHE III to predict 28-day mortality using the 95% confidence interval (CI) for the difference in the AUC (28, 29). Akaike information criterion values were calculated to assess relative goodness of fit of the models (30). We performed a sensitivity analysis to see how the models performed when limited to

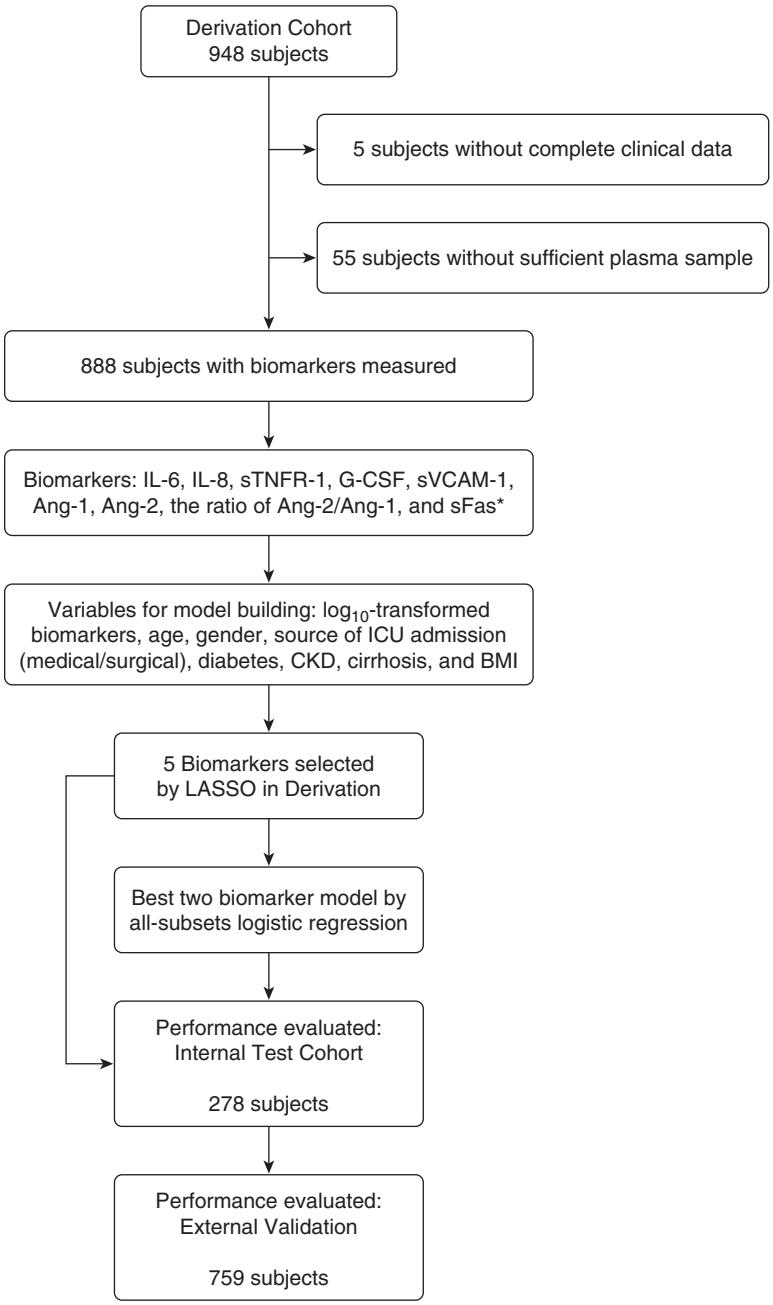

Figure 1. Sample and statistical flow diagram for mortality prediction model. Subject numbers for the development and validation of the least absolute shrinkage and selection operator or all subsets selection logistic regression models for 28-day mortality. *Lower limit or upper limit of detection applied to samples falling below or above the range of detection. Ang = angiopoietin; BMI = body mass index; CKD = chronic kidney disease; G-CSF = granulocyte colony–stimulating factor; ICU = intensive care unit; LASSO = least absolute shrinkage and selection operator; sVCAM = soluble vascular cell adhesion molecule; TNFR = tumor necrosis factor receptor.

patients with sepsis. Sepsis was defined as "Sepsis-2" for patients who met two or more SIRS criteria with suspected infection (7) or as "Sepsis-3" for patients who had a Day 0 SOFA score of greater than or equal to 2 with suspected infection (8).

We identified a threshold score for the mortality prediction model to maximize positive predictive value (PPV) defined as the probability of those with a score above the threshold dying within 28 days. We also identified a threshold score that gave a negative predictive value (NPV) of 0.95, where NPV is defined as the probability of those with a score below the threshold not dying within 28 days.

Analysis of Organ Failure

We compared the extent of persistent organ dysfunction (Day 4) (31) or length of stay among survivors in those falling above or below the NPV threshold using a Wilcoxon rank sum test. We compared proportions of ARDS (Berlin criteria within the first 7 d) (32), severe acute kidney injury (AKI; Acute Kidney Injury Network score $\geq$2) (18), or cardiovascular dysfunction (need for vasopressors within first 72 h), or any combination of these three endpoints.

Two sided P values of P less than 0.05 were considered significant. All statistical analyses were performed in R (R Core Team 2015) (33). Additional details regarding patient enrollment, biomarker measurement, and model development are detailed in the online supplement.

Results

Patients in the derivation (n = 888), internal test (n = 278), and external validation cohorts (n = 759) were predominantly male and middle aged (Table 1). The derivation cohort comprised white persons previously used in genetic studies from our group (19), whereas the internal test cohort was nonwhite and the external validation cohort was of diverse race. The derivation cohort was comprised of a mixed medical (n = 480; 54%) and surgical population (n = 408; 46%), whereas the internal test cohort and external validation cohort were predominantly medical ICU patients (n = 215, 77% and n = 718, 95%, respectively). Most of the patients in all three cohorts met the definition of Sepsis-2 or Sepsis-3 (7, 8). The 28-day mortality was 12% in the derivation cohort (n = 104), 9% in the test cohort (n = 25), and 13% in the external validation cohort (n = 97). The distribution of the biomarker concentrations (median and interquartile range) are listed in Table E2.

Mortality Prediction Model Derivation

We first developed the prediction model in the derivation cohort. We included age, sex,

Table 1. Subject Characteristics

	Derivation (*n* = 888)	Internal Test (*n* = 278)	External Validation (*n* = 759)
Age, yr, mean ± SD	56 ± 16	54 ± 16	61 ± 18
Male, n (%)	564 (64)	199 (72)	487 (64)
Race, n (%)			
White	888 (100)	0 (0)	696 (92)
African American	0 (0)	145 (52)	21 (2.8)
Asian/Pacific Islander	0 (0)	90 (32)	29 (3.8)
Native American	0 (0)	43 (15)	1 (0.13)
BMI, kg/m^2, mean ± SD	30 ± 10	31 ± 23	28 ± 8
Comorbidities, n (%)			
Diabetes	221 (25)	104 (37)	200 (26)
Cirrhosis	75 (8.4)	17 (6.1)	28 (3.7)
Chronic kidney disease	63 (7.1)	36 (13)	—
Source of ICU admit, n (%)			
Medical	480 (54)	215 (77)	718 (95)
Surgical	408 (46)	63 (23)	41 (5.4)
Day 0 SOFA ⩾2, n (%)*	703 (79)	229 (82)	749 (99)
Source of critical illness, n (%)†			
Sepsis-2	698 (79)	206 (74)	614 (81)
Sepsis-3	566 (64)	177 (64)	604 (80)
Pneumonia	179 (20)	72 (26)	340 (45)
Other	299 (34)	95 (34)	145 (19)
APACHE III, mean ± SD	51 ± 26	51 ± 26	67 ± 25
Day 4 SOFA score, mean ± SD	3.0 ± 2.9	3.1 ± 3.2	—
ICU LOS among survivors‡, d, mean ± SD	7.7 ± 11	7.4 ± 10	6.5 ± 6.9
28-d mortality, n (%)	104 (12)	25 (9)	97 (13)

Definition of abbreviations: APACHE = Acute Physiology and Chronic Health Evaluation; BMI = body mass index; ICU = intensive care unit; LOS = length of stay; SOFA = Sequential Organ Failure Assessment.
*SOFA scores greater than or equal to 2 on Day 0.
†Source of critical illness is not mutually exclusive.
‡ICU LOS among patients who survived 28 days.

source of ICU admission (medical/surgical), diabetes, chronic kidney disease, cirrhosis, and body mass index, and $\log_{10}$-transformed IL-6, IL-8, sTNFR-1, granulocyte colony–stimulating factor, sVCAM-1, Ang-1, Ang-2, the ratio of Ang-2/Ang-1, and soluble Fas as candidate predictors of mortality. In the derivation cohort, a LASSO model including the combination of IL-8, sTNFR-1, Ang-2, Ang-2/Ang-1, and sVCAM-1 most accurately predicted mortality (AUC, 0.80; 95% CI,0.75–0.84) (*see* Figure E1). Notably, clinical variables were excluded from the LASSO model. The performance of this model was comparable with APACHE III (AUC, 0.75; 95% CI, 0.70–0.80) (Table 2). There was no significant difference between the LASSO model and APACHE III (95% CI for difference, −0.03 to 0.11). The LASSO model combined with APACHE III significantly improved the discrimination (0.84; 95% CI, 0.80–0.88) above that of the APACHE III score alone (95% CI for difference, 0.01–0.16); however, this model was not a good fit on the basis of the Hosmer-Lemeshow test and may not be reliable (Table 2). Table E3 shows the predictive capacity for each of the individual biomarkers.

We then sought the most parsimonious model using one, two, three, or four biomarkers. A two-biomarker model of IL-8 and sTNFR-1 demonstrated an AUC of 0.79 (95% CI, 0.74–0.83) (Figure 2A) and improved on a single biomarker model (AUC, 0.77; 95% CI, 0.73–0.82) (*see* Figure E2) but addition of other biomarkers added only small incremental value (*see* Figure E2, Table E4). The two-biomarker model of IL-8 and sTNFR-1 performed similarly to APACHE III score (AUC, 0.75; 95% CI, 0.70–0.80) (Table 2) and better than Day 0 SOFA score as a continuous variable (AUC, 0.71; 95% CI, 0.66–0.76). In a sensitivity analysis using inpatient mortality rather than 28-day inpatient mortality the AUCs for APACHE III score and our two-biomarker model were unchanged. The two-biomarker model was carried forward for internal and external validation.

Mortality Prediction Model Testing and Validation

We tested the ability of the two biomarker model with IL-8 and sTNFR-1 to discriminate 28-day mortality in an internal test and external validation cohort. The AUC in the internal test cohort was 0.75 (95% CI, 0.65–0.85) (Figure 2B) and in the external validation cohort was 0.77 (95% CI, 0.72–0.83) (Figure 2C). This performed similarly to APACHE III in the internal test cohort (AUC, 0.77; 95% CI, 0.68–0.86) (Figure 2B) and in the external validation cohort (0.83; 95% CI, 0.79–0.87) (Figure 2C). Adding our two-biomarker model to APACHE III slightly improved the performance above APACHE III alone in the derivation (AUC, 0.83 [95% CI, 0.80–0.87] vs. AUC, 0.75 [95% CI, 0.70–0.80]; difference of 0.08 [95% CI, 0.004–0.15]), but not in the internal test (AUC, 0.79 [95% CI, 0.70–0.88] vs. AUC, 0.77 [95% CI, 0.68–0.86]; difference of 0.03 [95% CI, −0.14 to 0.20]) or external validation cohort (AUC, 0.82 [95% CI, 0.77–0.87] vs. 0.83 [95% CI, 0.79–0.87]; difference of 0.03 [95% CI, −0.05 to 0.11]) (Table 2). For comparison, a Day 0 SOFA score of 2 or greater, recently proposed as a new criteria for sepsis in the ICU (34) and shown to predict mortality in a general ICU population, has a maximum AUC in these cohorts of 0.60 (*see* Figure E1). When SOFA was considered as a continuous variable for mortality prediction, it performed similarly to the two-biomarker model in the internal test cohort (AUC, 0.75; 95% CI, 0.66–0.84). SOFA was not as predictive as the two-biomarker model in the external validation cohort (AUC, 0.67; 95% CI, 0.61–0.73) (Table 2) and was not a good fit on the basis of the Hosmer-Lemeshow test ($P < 0.001$).

We performed sensitivity analyses to test our ability to discriminate mortality when limiting to patients with sepsis defined as Sepsis-2 and Sepsis-3 in all cohorts. When limited to patients with either definition of sepsis, we found that the combination of IL-8 and sTNFR-1 predicted mortality with similar AUC values (*see* Figure E3).

Based on our model using IL-8 and sTNFR-1, we observed increasing mortality for each quartile of model score in the

Table 2. Model Performance in Derivation, Test, and Validation Cohorts

	Derivation			Internal Test			External Validation		
	AUC	HL*	Difference in AUC (95% CI)	AUC	HL*	Difference in AUC (95% CI)	AUC	HL*	Difference in AUC (95% CI)
LASSO	0.80 (0.75 to 0.84)	0.31	0.04 (−0.03 to 0.11)	0.77 (0.66 to 0.87)	0.01	0.01 (−0.15 to 0.16)	NA	NA	NA
Two-biomarker	0.79 (0.74 to 0.83)	0.46	0.03 (−0.04 to 0.10)	0.75 (0.65 to 0.85)	0.16	−0.03 (−0.18 to 0.13)	0.77 (0.72 to 0.83)	0.09	−0.04 (−0.12 to 0.04)
Day 0 SOFA†	0.71 (0.66 to 0.76)	0.36	0.04 (−0.04 to 0.12)	0.75 (0.65 to 0.85)	0.53	0.02 (−0.13 to 0.17)	0.67 (0.61 to 0.73)	<0.001	−0.16 (−0.08 to −0.24)
APACHE III	0.75 (0.70 to 0.80)	0.32	—	0.77 (0.68 to 0.86)	0.37	—	0.83 (0.79 to 0.87)	0.11	—
LASSO + APACHE III	0.84 (0.80 to 0.88)	0.03	0.08 (0.01 to 0.16)	0.80 (0.69 to 0.90)	0.2	0.05 (−0.12 to 0.22)	NA	NA	NA
Two-biomarker + APACHE III	0.83 (0.80 to 0.87)	0.09	0.08 (0.004 to 0.15)	0.79 (0.70 to 0.88)	0.57	0.03 (−0.14 to 0.20)	0.82 (0.77 to 0.87)	0.13	0.03 (−0.05 to 0.11)

Definition of abbreviations: APACHE = Acute Physiology and Chronic Health Evaluation; AUC = area under the receiver operator characteristic curve; CI = confidence interval; HL = Hosmer-Lemeshow; LASSO = least absolute shrinkage and selection operator; NA = biomarkers not available to calculate the model; SOFA = Sequential Organ Failure Assessment.
*Chi-square value (*P* value) from HL test with 10 bins. The model is rejected if *P* value is <0.05.
†SOFA score on Day 0.

discovery, internal test, and external validation cohort (*see* Table E5). We derived PPV and NPV for mortality (Table 3). A model score threshold of 0.54 maximized the PPV in the derivation cohort at 0.6 (Table 3). A model score below the threshold of 0.135 had an NPV of 0.95 in the derivation cohort that was identical in the internal test dataset (NPV, 0.95) and similar in the external validation cohort (NPV, 0.93).

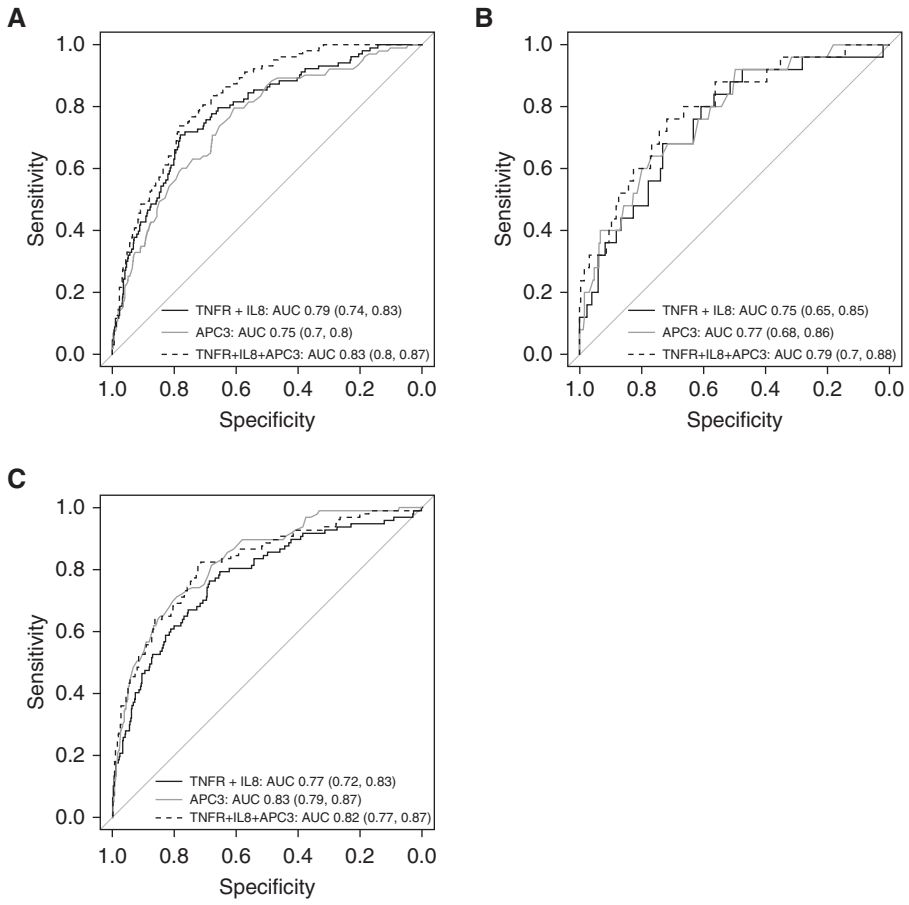

Figure 2. Receiver operating curves for mortality prediction models. Area under the receiver operating characteristic curve (AUC) for mortality in the (*A*) derivation cohort, (*B*) internal test cohort, and (*C*) external validation cohort. AUC values are presented with 95% confidence intervals in parenthesis. The models are shown as TNFR + IL8 to represent the model of two biomarkers alone, Acute Physiology and Chronic Health Evaluation III (APC3), and TNFR + IL8 + APC3 for the combined model of biomarkers and APC3. TNFR = tumor necrosis factor receptor.

Other ICU Outcomes

Although mortality is a relevant endpoint, other outcomes including organ dysfunction and ICU length of stay are clinically important. Need for vasopressors, AKI, and ARDS are clinical factors that often warrant admission to the ICU. We wanted to characterize whether patients above or below the model score threshold derived to maximize NPV for mortality had differences in these outcomes. In the derivation cohort, 246 patients had scores less than 0.135, whereas 642 patients had model scores greater than 0.135. Patients below the threshold had lower median SOFA scores at 4 days (2 vs. 4; $P < 0.001$) (*see* Figure E4A) and among survivors, lower median length of ICU stay (4 d vs. 7 d; $P < 0.001$) (*see* Figure E4B). The internal test cohort patients (n = 78) below the threshold also had lower median Day 4 SOFA scores (4 vs. 5; $P < 0.001$) and lower median ICU length of stay (4 d vs. 5 d; $P < 0.001$). The external validation cohort patients below the threshold (n = 229) had lower ICU length of stay (3 d vs. 6 d; $P < 0.001$). Day 4 SOFA scores were not available for the external validation cohort.

We also examined the proportions of patients above and below the model threshold with either AKI, need for vasopressors within 72 hours, or ARDS within 7 days of ICU admission. Among patients with scores below the threshold (0.135) in the derivation, internal test, and external validation cohorts, there was a lower proportion of each ICU-defining

Table 3. Negative and Positive Predictive Values for the Two-Biomarker Model

Cohort	Threshold Goal	Model Score Threshold*	Patients above/below Threshold (*n*)	PPV†	NPV‡
Derivation	NPV	0.135	642/246	0.29	0.95
	PPV	0.540	868/20	0.60	0.89
Internal test	NPV	0.135	200/78	0.18	0.95
	PPV	0.540	272/6	0.50	0.92
External validation	NPV	0.135	530/229	0.29	0.93
	PPV	0.540	747/12	0.83	0.88

Definition of abbreviations: NPV = negative predictive value; PPV = positive predictive value; sTNFR-1 = soluble tumor necrosis factor receptor 1.
*Model score threshold obtained from model fit to the discovery dataset using IL-8 and sTNFR-1 [logit (probability of death) = $-12.8 + 2.36 \times \log_{10}$(sTNFR-1 concentration) + $0.80 \times \log_{10}$(IL-8 concentration)].
†PPV is the probability of those with a score above the threshold dying within 28 days.
‡NPV is the probability of those with a score below the threshold not dying within 28 days.

illness compared with patients with scores above the threshold (Table 4). The most dramatic example is the proportion of patients with severe AKI (Acute Kidney Injury Network score ⩾2), which was 0.05 for patients below the threshold, but was 0.37 for patients above the threshold.

Discussion

In this study, we show that a simple predictive model based on early plasma measurements of IL-8 and sTNFR-1 predicts mortality in a diverse group of critically ill patients meeting criteria for SIRS at two academic medical centers. To our knowledge, this is the largest and most inclusive study of a predictive biomarker model in SIRS/sepsis.

Prior reports in smaller groups of patients have suggested the potential utility of sTNFR-1 and IL-8 for prediction of mortality in patients with sepsis. A small single center study (n = 52) in adult patients with septic shock showed that circulating sTNFR-1 levels were modestly predictive of mortality and that sTNFR-1 levels remained elevated in nonsurvivors compared with survivors for 48 hours (14). In patients with ARDS inclusion of circulating sTNFR-1 levels in a multibiomarker prediction model was predictive of death (35). IL-8 was also included in a multiple-biomarker decision tree model designed to predict mortality in adult and pediatric patients with septic shock (13, 36) and adult patients with ARDS (35). Low levels of IL-8 in a pediatric septic shock cohort predicted a high likelihood of survival (37). In contrast, our study used sTNFR-1 and IL-8 in a simple two-biomarker predictive model in two large cohorts of ICU patients with SIRS/sepsis. IL-8 and sTNFR-1 were among biomarkers shown by latent class analysis to distinguish a subphenotype of ARDS that has differential mortality and responsiveness to fluid management strategies (35, 38). This suggests that these biomarkers could be considered to better understand subphenotypes or predictive enrichment of treatment responsiveness in sepsis (39).

Our model is comparable with APACHE III in terms of mortality prediction in the three cohorts. Notably, the APACHE III score requires measurement of multiple variables over a 24-hour time period, limiting its utility for early clinical decision-making. Thus, our simple model may represent a good alternative for early mortality prediction in this patient population. In addition, our biomarker model could be developed for rapid point-of-care testing moving closer toward a goal of precision medicine.

Table 4. Proportions of Severe Organ Dysfunction among Patients below or above the Model Threshold

	Model Score below 0.135			Model Score above 0.135			All Subjects: Total Observations
	Patients with Outcome	Total with Observation*	Prop†	Patients with Outcome	Total with Observation‡	Prop†	
ARDS in 7 d§							
Derivation	116	642	0.18	52	246	0.21	888
Internal test	39	203	0.19	25	75	0.33	278
External validation	30	460	0.07	41	299	0.14	759
AKI (AKIN ⩾2)‖							
Derivation	34	639	0.05	88	241	0.37	880
Internal test	13	200	0.06	41	75	0.55	275
External validation	3	198	0.02	48	179	0.27	377
Shock¶							
Derivation	109	642	0.17	97	246	0.39	888
Internal test	33	203	0.16	31	75	0.41	278
External validation	150	373	0.40	153	248	0.62	621

Definition of abbreviations: AKI = acute kidney injury; AKIN = Acute Kidney Injury Network; ARDS = acute respiratory distress syndrome; Prop = proportion.
*Number of patients with score below 0.135 with organ failure data available for analysis.
†Proportion of subjects with outcome.
‡Number of subjects with score above 0.135 with organ failure data available for analysis.
§Presence of ARDS within 7 days as defined by Berlin criteria.
‖Severe AKI as defined by an AKIN score ⩾2.
¶Vasopressors required in first 72 hours of admission.

The clinical definition of sepsis is in a state of evolution. Proposed changes have moved from SIRS criteria (Sepsis-2) to an approach based on manifestations of organ dysfunction (Sepsis-3) (6–8). Part of the rationale is that SIRS criteria may not be effective because they define a heterogeneous population that may not be at elevated risk for adverse outcomes (40, 41). Notably, in contrast to the recent publication supporting the Sepsis-3 criteria, in our cohorts a SOFA score on Day 0 of 2 or greater was a poor predictor of mortality (34). Using SOFA score as a continuous variable increased its predictive capacity, but the biomarker model still outperformed the SOFA in the derivation and external validation cohorts. This may be explained by differences in underlying diagnoses and our inclusion criteria that mandated presence of SIRS criteria on enrollment resulting in a more severely ill population or decreased specificity of the Sepsis-3 definition. Regardless, the strong performance of our model using both definitions of sepsis suggests that it will be robust when applied to other sepsis populations.

The most straightforward application of our model would be in the identification of patients at low risk of death and organ dysfunction. We identified a model score threshold that had an NPV for mortality in the derivation/test cohorts of 0.95 and remained high in the validation cohort (0.93). Patients below the same threshold had less persistent organ dysfunction measured by Day 4 total SOFA score, proportionally less AKI, need for vasopressors, ARDS, and among survivors, shorter ICU length of stay. Our model built for maximizing NPV for mortality distinguishes certain organ failure, such as AKI, better than shock. Future models should be built directly for organ failure to maximize the utility of these biomarkers as a clinical tool. Taken together, these findings suggest that our model might be helpful to clinicians in early triage decisions in patients with sepsis and to assist researchers in rational enrollment for clinical trials.

Our study has several limitations. First, we studied only patients who met SIRS criteria and were admitted to the ICU at academic medical centers. This limits the generalization of our results to a broader set of critically ill patients. Second, we did not include several clinical measurements commonly used for assessment of severity including arterial lactate, procalcitonin, or C-reactive protein (42–46). However, lactate measurements were not included in the recently proposed sepsis definition for the lack of improvement in predictive capacity above the proposed model and procalcitonin has not shown strong evidence of predictive utility as a stand-alone biomarker (8, 34, 47). Third, our external validation cohort was recruited over an extended period of time. This may lead to heterogeneity in clinical treatment over that time frame. Nonetheless, our model performs well despite that heterogeneity. Finally, although in this study we aimed for a parsimonious and easily clinically applicable model, other biomarkers (IL-6, Ang-1, Ang-2, sVCAM-1) did have some, albeit limited, incremental prognostic value. Future studies need to explore whether other biomarkers might provide additional value, particularly for specific forms of organ dysfunction.

In summary, we have developed and validated a robust biomarker-based model that predicts mortality in a large heterogeneous population of critically ill patients with SIRS/sepsis. This algorithm could potentially be incorporated in either laboratory-based or point-of-care diagnostics to guide management of critically ill patients. Future studies will explore whether our model can improve early clinical decision-making in patients with sepsis. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1. Gaieski DF, Edwards JM, Kallan MJ, Carr BG. Benchmarking the incidence and mortality of severe sepsis in the United States. *Crit Care Med* 2013;41:1167–1174.
2. Stoller J, Halpin L, Weis M, Aplin B, Qu W, Georgescu C, Nazzal M. Epidemiology of severe sepsis: 2008-2012. *J Crit Care* 2016;31: 58–62.
3. Vincent J-L, Marshall JC, Namendys-Silva SA, François B, Martin-Loeches I, Lipman J, Reinhart K, Antonelli M, Pickkers P, Njimi H, *et al*.; ICON investigators. Assessment of the worldwide burden of critical illness: the Intensive Care Over Nations (ICON) audit. *Lancet Respir Med* 2014;2: 380–386.
4. Fleischmann C, Scherag A, Adhikari NKJ, Hartog CS, Tsaganos T, Schlattmann P, Angus DC, Reinhart K; International Forum of Acute Care Trialists. International Forum of Acute Care Trialists. Assessment of global incidence and mortality of hospital-treated sepsis. current estimates and limitations. *Am J Respir Crit Care Med* 2016;193: 259–272.
5. Vincent J-L, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, Moreno R, Lipman J, Gomersall C, Sakr Y, *et al*.; EPIC II Group of Investigators. EPIC II Group of Investigators. International study of the prevalence and outcomes of infection in intensive care units *JAMA* 2009;302:2323–2329.
6. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM, Sibbald WJ; The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Chest* 1992;101: 1644–1655.
7. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, Cohen J, Opal SM, Vincent J-L, Ramsay G; SCCM/ESICM/ACCP/ATS/SIS. SCCM/ESICM/ACCP/ATS/SIS. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference *Crit Care Med* 2003;31: 1250–1256.
8. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche J-D, Coopersmith CM, *et al*. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801–810.
9. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A, *et al*. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619–1636.
10. Ricciuto DR, dos Santos CC, Hawkes M, Toltl LJ, Conroy AL, Rajwans N, Lafferty EI, Cook DJ, Fox-Robichaud A, Kahnamoui K, *et al*. Angiopoietin-1 and angiopoietin-2 as clinically informative prognostic biomarkers of morbidity and mortality in severe sepsis. *Crit Care Med* 2011;39:702–710.
11. Parikh SM, Mammoto T, Schultz A, Yuan H-T, Christiani D, Karumanchi SA, Sukhatme VP. Excess circulating angiopoietin-2 may contribute to pulmonary vascular leak in sepsis in humans. *PLoS Med* 2006; 3:e46.
12. Mat-Nor MB, Md Ralib A, Abdulah NZ, Pickering JW. The diagnostic ability of procalcitonin and interleukin-6 to differentiate infectious from noninfectious systemic inflammatory response syndrome and to predict mortality. *J Crit Care* 2016;33:245–251.

13. Wong HR, Lindsell CJ, Pettilä V, Meyer NJ, Thair SA, Karlsson S, Russell JA, Fjell CD, Boyd JH, Ruokonen E, *et al*. A multibiomarker-based outcome risk stratification model for adult septic shock. *Crit Care Med* 2014;42:781–789.
14. de Pablo R, Monserrat J, Reyes E, Diaz-Martin D, Rodriguez Zapata M, Carballo F, de la Hera A, Prieto A, Alvarez-Mon M. Mortality in patients with septic shock correlates with anti-inflammatory but not proinflammatory immunomodulatory molecules. *J Intensive Care Med* 2011;26:125–132.
15. Ware LB, Koyama T, Zhao Z, Janz DR, Wickersham N, Bernard GR, May AK, Calfee CS, Matthay MA. Biomarkers of lung epithelial injury and inflammation distinguish severe sepsis patients with acute respiratory distress syndrome. *Crit Care* 2013;17: R253.
16. Wong HR, Cvijanovich NZ, Anas N, Allen GL, Thomas NJ, Bigham MT, Weiss SL, Fitzgerald J, Checchia PA, Meyer K, *et al*. A multibiomarker-based model for estimating the risk of septic acute kidney injury. *Crit Care Med* 2015;43:1646–1653.
17. Albertine KH, Soulier MF, Wang Z, Ishizaka A, Hashimoto S, Zimmerman GA, Matthay MA, Ware LB. Fas and fas ligand are up-regulated in pulmonary edema fluid and lung tissue of patients with acute lung injury and the acute respiratory distress syndrome. *Am J Pathol* 2002;161:1783–1796.
18. Robinson-Cohen C, Katz R, Price BL, Harju-Baker S, Mikacenic C, Himmelfarb J, Liles WC, Wurfel MM. Association of markers of endothelial dysregulation Ang1 and Ang2 with acute kidney injury in critically ill patients. *Crit Care* 2016;20:207.
19. Mikacenic C, Hahn WO, Price BL, Harju-Baker S, Katz R, Kain KC, Himmelfarb J, Liles WC, Wurfel MM. Biomarkers of endothelial activation are associated with poor outcome in critical illness. *PLoS One* 2015;10:e0141251.
20. Hahn WO, Mikacenic C, Price BL, Harju-Baker S, Katz R, Himmelfarb J, Wurfel MM, Liles WC. Host derived biomarkers of inflammation, apoptosis, and endothelial activation are associated with clinical outcomes in patients with bacteremia and sepsis regardless of microbial etiology. *Virulence* 2016;7:387–394.
21. Mikacenic C, Price B, Harju-Baker S, Katz R, Himmelfarb J, Liles W, Wurfel M. 1372: Development of a robust biomarker predictive model for death in ICU patients with SIRS and sepsis. *Crit Care Med* 2016; 44 (Suppl 1):418.
22. Ko DC, Shukla KP, Fong C, Wasnick M, Brittnacher MJ, Wurfel MM, Holden TD, O'Keefe GE, Van Yserloo B, Akey JM, *et al*. A genome-wide in vitro bacterial-infection screen reveals human variation in the host response associated with inflammatory disease. *Am J Hum Genet* 2009;85:214–227.
23. Gong MN, Thompson BT, Williams P, Pothier L, Boyce PD, Christiani DC. Clinical predictors of and mortality in acute respiratory distress syndrome: potential role of red cell transfusion. *Crit Care Med* 2005;33:1191–1198.
24. Ahasic AM, Zhai R, Su L, Zhao Y, Aronis KN, Thompson BT, Mantzoros CS, Christiani DC. IGF1 and IGFBP3 in acute respiratory distress syndrome. *Eur J Endocrinol* 2012;166:121–129.
25. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B* 1996;58:267–288.
26. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med* 1996;22:707–710.
27. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965–980.
28. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.
29. Demler OV, Pencina MJ, D'Agostino RB Sr. Misuse of DeLong test to compare AUCs for nested models. *Stat Med* 2012;31:2577–2587.
30. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected papers of Hirotugu Akaike. New York, NY: Springer; 1998. pp. 199–213.
31. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA* 2001;286:1754–1758.
32. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS; ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;307: 2526–2533.
33. The R Project for Statistical Computing [accessed 2015 Feb]. Available from: http://www.r-project.org/
34. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, Rubenfeld G, Kahn JM, Shankar-Hari M, Singer M, *et al*. Assessment of clinical criteria for sepsis: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315: 762–774.
35. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA; NHLBI ARDS Network. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med* 2014;2:611–620.
36. Wong HR, Cvijanovich NZ, Anas N, Allen GL, Thomas NJ, Bigham MT, Weiss SL, Fitzgerald J, Checchia PA, Meyer K, *et al*. Prospective testing and redesign of a temporal biomarker based risk model for patients with septic shock: implications for septic shock biology. *EBioMedicine* 2015;2:2087–2093.
37. Wong HR, Cvijanovich N, Wheeler DS, Bigham MT, Monaco M, Odoms K, Macias WL, Williams MD. Interleukin-8 as a stratification tool for interventional trials involving pediatric septic shock. *Am J Respir Crit Care Med* 2008;178:276–282.
38. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, Calfee CS; ARDS Network. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. *Am J Respir Crit Care Med* 2017;195:331–338.
39. Prescott HC, Calfee CS, Thompson BT, Angus DC, Liu VX. Toward smarter lumping and smarter splitting: rethinking strategies for sepsis and acute respiratory distress syndrome clinical trial design. *Am J Respir Crit Care Med* 2016;194:147–155.
40. Churpek MM, Zadravecz FJ, Winslow C, Howell MD, Edelson DP. Incidence and prognostic value of the systemic inflammatory response syndrome and organ dysfunctions in ward patients. *Am J Respir Crit Care Med* 2015;192:958–964.
41. Kaukonen K-M, Bailey M, Pilcher D, Cooper DJ, Bellomo R. Systemic inflammatory response syndrome criteria in defining severe sepsis. *N Engl J Med* 2015;372:1629–1638.
42. Casserly B, Phillips GS, Schorr C, Dellinger RP, Townsend SR, Osborn TM, Reinhart K, Selvakumar N, Levy MM. Lactate measurements in sepsis-induced tissue hypoperfusion: results from the Surviving Sepsis Campaign database. *Crit Care Med* 2015;43:567–573.
43. Bakker J, Coffernils M, Leon M, Gris P, Vincent JL. Blood lactate levels are superior to oxygen-derived variables in predicting outcome in human septic shock. *Chest* 1991;99:956–962.
44. Dellinger RP, Carlet JM, Masur H, Gerlach H, Calandra T, Cohen J, Gea-Banacloche J, Keh D, Marshall JC, Parker MM, *et al*.; Surviving Sepsis Campaign Management Guidelines Committee. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. *Crit Care Med* 2004;32:858–873.
45. Ryu J-A, Yang JH, Lee D, Park C-M, Suh GY, Jeon K, Cho J, Baek SY, Carriere KC, Chung CR. Clinical usefulness of procalcitonin and C-reactive protein as outcome predictors in critically ill patients with severe sepsis and septic shock. *PLoS One* 2015;10:e0138150.
46. Ruiz-Alvarez MJ, García-Valdecasas S, De Pablo R, Sanchez García M, Coca C, Groeneveld TW, Roos A, Daha MR, Arribas I. Diagnostic efficacy and prognostic value of serum procalcitonin concentration in patients with suspected sepsis. *J Intensive Care Med* 2009;24: 63–71.
47. Tang BMP, Eslick GD, Craig JC, McLean AS. Accuracy of procalcitonin for sepsis diagnosis in critically ill patients: systematic review and meta-analysis. *Lancet Infect Dis* 2007;7:210–217.