# Review of A Large Clinical Series: Predicting Death for Patients With Abdominal Septic Shock

Ernst Hanisch, Rüdiger Brause, Jürgen Paetz and Björn Arlt

*J Intensive Care Med* 2011 26: 27
DOI: 10.1177/0885066610384058

The text of this article (DOI: 10.1177/0885066610384058) in the January/February 2011 Issue, *Journal of Intensive Care Medicine*, 26(1), 27-33, has been updated since its original online publication. This is the corrected and updated version.

Journal of Intensive Care Medicine
26(1) 27-33

DOI: 10.1177/0885066610384058
http://jicm.sagepub.com

# Predicting Death for Patients With Abdominal Septic Shock

**Ernst Hanisch, MD[1], Rüdiger Brause, PhD[2], Jürgen Paetz, PhD, and Björn Arlt, MSc[2]**

**Abstract**

This paper reports the result of the MEDAN project that analyzes a multicenter septic shock patient data collection. The mortality prognosis based on 4 scores that are often used is compared with the prognosis of a trained neural network. We built an alarm system using the network classification results. **Method.** We analyzed the data of 382 patients with abdominal septic shock who were admitted to the intensive care unit (ICU) from 1998 to 2002. The analysis includes the calculation of daily sepsis-related organ failure assessment (SOFA), Acute Physiological and Chronic Health Evaluation (APACHE) II, simplified acute physiology score (SAPS) II, multiple-organ dysfunction score (MODS) scores for each patient and the training and testing of an appropriate neural network. **Results.** For our patients with abdominal septic shock, the analysis shows that it is not possible to predict their individual fate correctly on the day of admission to the ICU on the basis of any current score. However, when the trained network computes a score value below the threshold during the ICU stay, there is a high probability that the patient will die within 3 days. The trained neural network obtains the same outcome prediction performance as the best score, the SOFA score, using narrower confidence intervals and considering three variables only: systolic blood pressure, diastolic blood pressure and the number of thrombocytes. We conclude that the currently best available score for abdominal septic shock may be replaced by the output of a trained neural network with only 3 input variables.

**Keywords**

septic shock, scores, neural network, alarm system

Received February 19, 2009, and in revised form October 22, 2009. Accepted December 3, 2009.

## Introduction

Since the description of sepsis by Schottmüller in 1914,[1] the knowledge concerning sepsis and its underlying pathophysiology has increased substantially. Epidemiologic examinations of patients with abdominal septic shock show the potential for a high risk resulting from the extensive therapy in the intensive care unit (ICU).[2] Unfortunately, until now it has not been possible to reduce significantly the mortality rate of patients with septic shock, which is as high as 50% to 60% worldwide. Nevertheless, a number of appropriate therapies are available[3]: early goal-directed therapy by volume resuscitation,[4] intensive insulin therapy,[5] and corticosteroids.[6] The latter 2 therapy options have recently been challenged.[7,8] The use of activated protein C[9] is controversial[10] and a recent meta-analysis demonstrated that antithrombin III has no effect.[11]

The heterogeneity of patient groups and the variations in therapy strategies is seen as one of the main problems in sepsis studies. Therefore, commonly available scoring systems are used for comparing critically ill patient groups. Moreover, 1 of the main objectives of scores is to help predict the outcome. Can we diagnose mortality very early at the beginning of the stay or only on the last day? How long before death or ICU dismissal can the outcome be predicted?

The task of the MEDAN project was to develop a means of diagnosing septic shock diagnosis based on a self-learning system, especially a neural network, to answer these questions. Moreover, its performance must be compared to those of several established scores (sepsis-related organ failure assessment [SOFA], Acute Physiological and Chronic Health Evaluation [APACHE] II, Simplified Acute Physiology Score [SAPS] II, multiple-organ dysfunction score [MODS]). For clinical purposes, the diagnostic outcome of "survival" or "death" should give a feedback to the treating doctor. If the

[1] Department of Surgery, Asklepios Klinik Langen, Academic Teaching Hospital at the J.W. Goethe University Frankfurt am Main, Germany
[2] Department of Informatics and Mathematics, Goethe University, Institute for Informatics, Frankfurt am Main, Germany

**Corresponding Author:**
Ernst Hanisch, Chirurgische Klinik, Asklepios Klinik Langen, Röntgenstr 20, D-63225 Langen, Germany
Email: E.Hanisch@Asklepios.com

diagnosis diverges from his/her appraisal, the treating doctor will be alerted to the need for additional measures.

For this purpose, a group of 382 patients made up exclusively of abdominal septic shock patients, for the first time in Germany, was investigated by using scores and a multivariate neural network analysis.

The classification results provided the basis for creating a reliable alarm system for patients with abdominal septic shock.

## Methods

For outcome prediction, the data of 382 patients out of 583 who met the consensus criteria for abdominal septic shock[12,13] were analyzed by using most of the commonly documented vital parameters and doses of medicine (metric variables). Of the 382 patients, 187 (48.9%) have died. Data were collected in 102 German hospitals from 1998 to 2002. The data were anonymized and coded by means of a paper-based mapping. The concept of data usage and privacy protection was approved by the data protection commission of the country.

Extensive manpower was required for subsequent transfer of all handwritten patient records into a consistent electronic database comprising 2.5 million data sets. We used programmed range and plausibility checks of different kinds to detect all faulty data in the electronic database. For this purpose, static values (eg, lower and upper bounds) and dynamic development (eg, time sequence behavior) were checked.[14]

The complete database is available at www.medan.de/datenbank/download_database.htm.

### Data Sets

Data from different periods of time were taken into account for the evaluation: the first 3 days of ICU stay (F3); the first 3 days after the septic shock occurrence (S3); all days of ICU stay (ALL); days 8, 7, and 6, counted backward from the last day of the ICU stay (D6-8), that is the last day of the ICU stay was day 0; days 4, 3, and 2 were counted from the last day of ICU stay (D2-4); and the last 5, 3, 2, 1 day(s) of ICU stay (L5, L3, L2, L1). All diagnosis results were characterized by their AUC value, the area under the ROC curve. Since the results on the day of admission and the day after were almost random (AUC = 0.5), we used a minimum of 3 days (S3) for AUC calculation.

Interestingly enough, none of the traditional indicators could achieve an acceptable level of successful diagnosis. For instance, neither the doses of catecholamine nor the fact of enforced respiration is significant—because nearly all patients with septic shock receive catecholamine and enforced respiration. This does not indicate anything significant about the chances of surviving. Other single variables such as base excess, lactate, or $O_2$ saturation also do not contribute to diagnosis. Even parameters with a good diagnostic value such as the central venous pressure or the diastolic blood pressure provide a pertinent diagnosis only in combination with other indicators. Therefore, the task was to find a small subset of the 140 variables with a good diagnostic power.

For this purpose, other data sets than those used by the scores are also taken into account:

**freq16:** (the 16 most frequently measured variables) heart rate (1/min), systolic blood pressure (mm Hg), diastolic blood pressure (mm Hg), temperature (°C), CVP (mm Hg), $O_2$ saturation (%), leukocytes (1000/µL), hemoglobin (g/dL), hematocrit (%), thrombocytes (1000/µL), PTT (s), sodium (mmol/L), potassium (mmol/L), creatinine (mg/dL), glucose (mg/d), urine volume (mL/h),

**coagulation:** leukocytes, erythrocytes (1000/µL), hemoglobin, hematocrit, thrombocytes, TPT (%), PTT (s), thrombin time (s), AT3 (%), fibrinogen (mg/dL), total protein (g/dL), glucose (mg/dL), RBC (mL), FFP (mL),

**heart:** heart rate, systolic blood pressure, diastolic blood pressure, CVP, cristalloids (mL), colloids (mL), adrenaline (µg/kg per minute), noradrenaline (µg/kg per minute), dopamine (µg/kg per minute), dobutamine (µg/kg per minute),

**lungs:** arterial $pO_2$ (mm Hg), arterial $pCO_2$ (mm Hg), base excess (−), bicarbonate (mmol), $O_2$ saturation, $O_2$ medication (l/min), PEAK (cm $H_2O$), I:E (−), respiratory frequency (1/min), $FiO_2$ (%), PEEP (mm Hg),

**bac:** (breathing and catecholamines) $FiO_2$, PEAK, respiratory frequency, adrenaline, noradrenaline, dopamine, dobutamine,

**bpt:** (systolic and diastolic blood pressure, thrombocytes) and the single variables systolic blood pressure, diastolic blood pressure, thrombocytes.

The main preprocessing steps are sampling of data within a 24-hour interval (mean values of 24 hours for each variable) and replacement of missing values with random values within the interval of the so-called interquartile range from a suitable normal distribution, done separately for every variable. Random missing value replacement is important not to simulate an increased performance.

### Computational Methods

First, we computed the conventional scores on the different data sets and evaluated their prognostic values on our database and its subsets. The scores for comparison were:

a. Sepsis-related organ failure assessment.[15,16] A maximum of 12 different variables are needed to calculate the score: respiration, $PaO_2/FiO_2$ (mmHg); coagulation, thrombocytes ($1/mm^3$); liver, bilirubin (mL/dL); cardiovascular, hypotension MAP (mm Hg) or amount of dopamine, dobutamine, epinephrine or norepinephrine; renal, creatinine (mg/dL) or urine output; central nervous system, Glasgow Coma Score.

b. Acute Physiological and Chronic Health Evaluation[17]: It uses a scale of 0 to 71 of whole-number values, based on 12 partial variables: rectal temperature (°C); cardiovascular, hypotension MAP (mm Hg), heart frequency (Cycl./min);

respiration, respiratory frequency (resp./min); oxygenation $FiO_2 \geq 0.5$: A-DO2 ($FiO_2 < 0.5$: $PaO_2$), arterial pH, serum sodium (Na+ mmol/L), serum potassium (K+ mmol/L), hematocrit (%), leukocytes ($\times 1000/mm^3$); renal, creatinine (mg/dL); central nervous system, Glasgow Coma Score.

c. Simplified acute physiology score[18]: The SAPS II score is another ICU score using 15 partial variables: age, heart frequency (Cycl./min), systolic blood pressure (mm Hg), rectal temperature (°C), oxygenation $PaO_2$ (mm Hg), urine output (mL/day), urea (mg/dL), leukocytes ($\times 1000/mm^3$), serum sodium (Na+ mmol/L), serum potassium (K+ mmol/L), bicarbonate; liver, bilirubin (mL/dL); central nervous system, Glasgow Coma Score, chronic disease, reason for admission.

   Originally, SAPS was introduced as a simplified APACHE score.

d. Multiple-organ dysfunction score[19]: The MODS score assesses organ states on a whole-number scale, using the 6 partial variables respiration: $PaO_2/FiO_2$ (mm Hg); coagulation, thrombocytes ($1/mm^3$); liver: bilirubin (mL/dL); renal: creatinine (mg/dL); cardiovascular: HR*CVP/MAP; central nervous system: Glasgow Coma Score.

A score was calculated every time the necessary variables were available. Mostly, we did not consider the Glasgow Coma Scale (GCS)[20] in the scores, because it was not always available in the database. A score was calculated every time the necessary variables were defined without considering the GCS.

The neural network was then trained on the data sets. Training was done with 50% of the samples and testing with the remaining 50%. In contrast to the predefined scores, the neural network algorithm[21] uses the class information of the training data in its training process to obtain its diagnostic power. The outcome labels "survived" and "deceased" are used as class information in the training procedure of the neural network for its parameters, the weights. There are two kinds of weights: those who determine the location of an input interval of a neuron for one input variable, and those who weight the output of the neuron in a second step. Thus, each neuron is sensitive to 1 multidimensional region within the input space. The weighted output of all neurons is summed up and compared to a threshold, giving rise to a class decision.

This kind of system adapts a nonlinear classification to the data by adapting the position and width of the input intervals in the input space, the weighting and the threshold for the classification decision.

The classification is trained for optimal class discrimination and learns automatically to use the best subset of input variables to perform its task, avoiding the time-consuming feature selection process. The result is similar to a nonlinear regression, but no regression model is needed a priori. For implementation details, see 22, 23.

### Experiment conditions and statistics

All (1-dimensional) score samples and the (multidimensional) samples of the data sets are classified by the neural network. For the 1-dimensional scores, the classification is nothing else than a linear classification using an optimal threshold. Training was done with 50% of the samples and testing with the remaining 50%. Data on training patients was not used for testing (disjoint patient sets). Finally, the samples of the test data sets are classified using the trained neural network. All experiments with 1 data set were repeated 20 times for robust estimation of mean and standard deviation. The area beneath the ROC curve (AUC) is used as criterion for comparison between the different performance results.

Additionally, we computed the 95% confidence intervals (CI) for the AUC values of our neural network diagnosis by assuming that AUC values in 1 data set calculated in repetitions of an experiment have a normal distribution. Exploratory statistics (Q-Q-plots) indicate that this is a reasonable assumption. Therefore, the CI bounds can be obtained by linearly transforming the CI bound of a normal distribution using our measured variances and mean values.[24]

## Results

We compare the diagnostic results from the different diagnostic approaches. The results of the analysis of 382 patients are very similar to intermediate results[25] obtained for only 138 patients.

### Epidemiology

Before comparing the performance of the neural network with score performance, an epidemiological overview of the data is given in Table 1.

All patients were diagnosed as having peritonitis from various sources: anastomotic leakages, bowel ischemia, perforations, or appendicitis.

### Neural Network Performance

With the data set freq16, a diagnostic performance of the neural network for mortality prognosis with an AUC = 0.56 (F3) and AUC = 0.59 (S3) was achieved Figure 1(a). The average AUC value is 0.65 (ALL) comprising the samples of all days. The best classification results are achieved on the basis of the last day L1 (AUC = 0.93). Since an outcome prognosis on the last day is not useful for building an alarm system, we consider L3 with a high AUC (AUC = 0.90) and a 3-day prognosis horizon as the most interesting prognostic period of time.

Taking this prognostic period, what are the most interesting prognosis parameters? What is required for a good diagnosis? In Figure 1(b), the corresponding AUC of different data sets for the last 3 days of the ICU stay (L3) are shown.

We conclude that the 3 variable systems heart, lungs, and the freq16 parameters taken very frequently are the most effective.

**Table 1.** Epidemiological Data of 382 Abdominal Septic Shock Patients

| | | All Patients 382 (100%) | | Male Patients 222 (58%) | | Female Patients 160 (42%) | |
|---|---|---|---|---|---|---|---|
| | All Patients | Survived | Deceased | Survived | Deceased | Survived | Deceased |
| Number of patients | 382 | 195 (51%) | 187 (49%) | 111 (50%) | 111 (50%) | 84 (53%) | 76 (47%) |
| Age (years) | 66.0 | 63.6 | 68.5 | 61.3 | 67.6 | 66.8 | 69.8 |
| ICU stay (days) | 18.6 | 20.6 | 16.5 | 22.2 | 18.0 | 18.5 | 14.4 |
| Artificial respiration[a] (days) | 13.3 | 13.0 | 13.8 | 15.1 | 14.9 | 12.0 | 9.6 |
| Weight (kg) | 75.5 | 77.4 | 73.4 | 83.1 | 77.0 | 69.9 | 68.2 |
| Height (m) | 1.70 | 1.71 | 1.70 | 1.76 | 1.74 | 1.64 | 1.62 |

[a] Artificial respiration duration was averaged only for patients that were respirated.

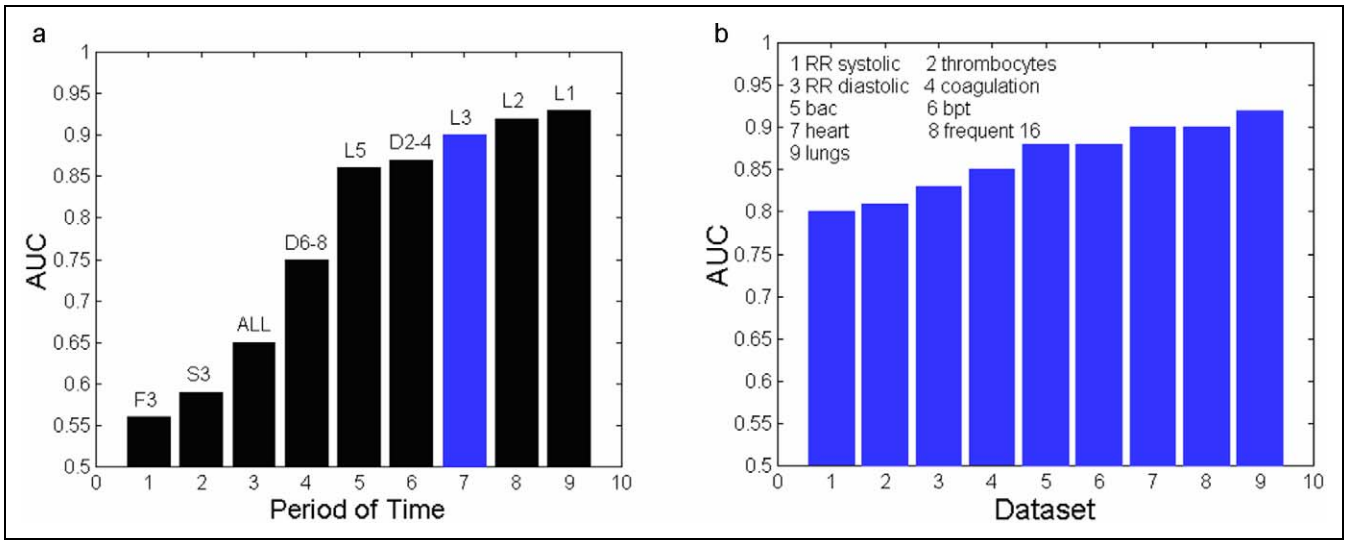


**Figure 1.** a. Area beneath ROC curves (AUC) for different periods of time of ICU stay for *freq16* data, b, AUC values for different data sets of the last 3 days of ICU stay L3.

## *Score Performance*

In this section, we compare these findings with the traditional approach using the scores that are often used for prediction purposes. For time period L3, the 3 scores MODS, SAPS II, and APACHE II perform differently (AUC = 0.88, 0.85, and 0.79, see Figure 2(a), with APACHE II performing worst. The SOFA score (AUC = 0.90) yields a clearly better classification. In comparison, considering the first 3 days (F3), the AUC for the SOFA score equals 0.54, for APACHE II = 0.52, SAPS II = 0.52, MODS = 0.52, Neural Network = 0.52. The CI for AUC values of scores are presented in Table 2.

We note that the 95% CI of AUC for the scores in Figure 2a are wide, for example, for MODS the CI is from 0.77 to 0.99. In contrast to this, the CI width for the neural network results is mostly less than the score's CI width, for example, the range for the bpt CI is (0.83, 0.92).

By definition, result A is significantly higher (95% CI) than result B if the lower bound LB of interval A's confidence interval is higher than the upper bound UB of interval B's confidence interval. For the results of freq16 datasets (Figure 1a and first column of Table 2), the time periods L1, L2, L3, D2-4, and L5 have a significantly higher AUC than F3, S3, and ALL. The systems lungs, heart, bpt, and SOFA have similar CIs and AUCs.

## *The Resulting Alarm System*

We used the diagnostic results of the neural network to create an alarm system,[26] using 138 patients. Here, we present the results for the extended group of 382 patients. An alarm message is given whenever the input for the neural network generates high output for class "deceased." We obtained the classification results using the lungs, heart, bpt, or SOFA data set (last 3 days). Since in the bpt system 3 parameters are only used as input, we use the bpt data for the alarm system, simplifying the bedside input for physicians. In Figure 2b, we see the resulting alarm percentage for the first 3 days, for the first and second half of ICU stay, and for the last 3 days, indicated separately for patients who either died or survived. In the time periods 1, 2, 3, and 4, there are 34%, 23%, 9%, and 7% alarms for surviving patients, respectively, and 41%, 36%, 57%, and 72% alarms for deceased patients, respectively (ie, 1.2, 1.5, 6.5, and 9.9 times more alarms for deceased patients, respectively).

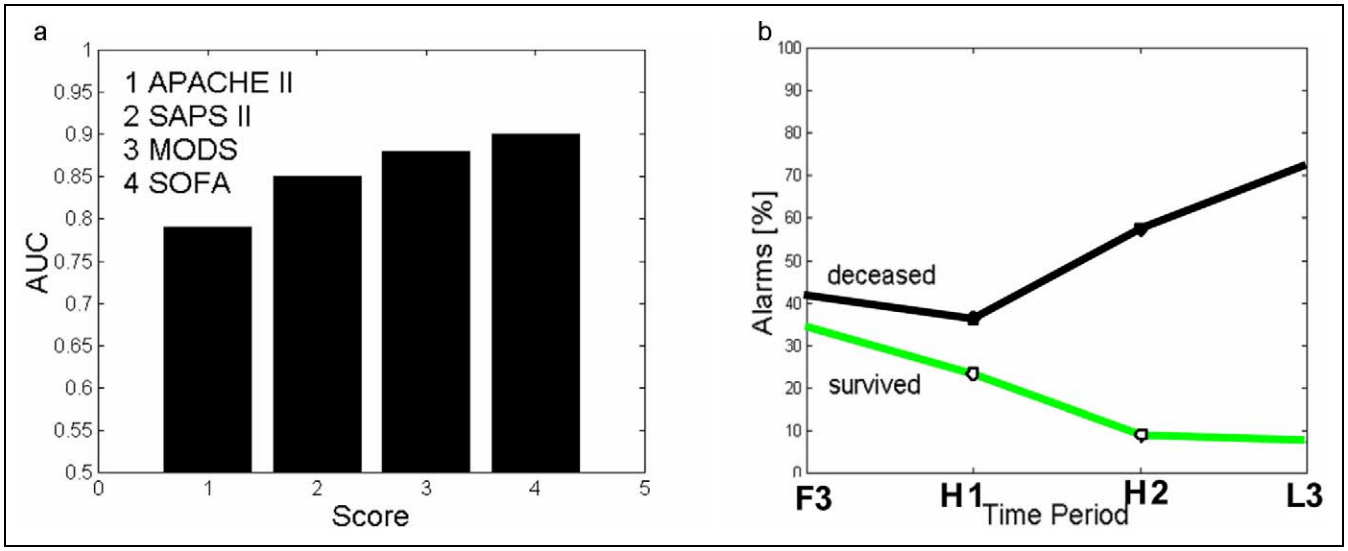


**Figure 2.** a. AUC for MODS, SAPS II, APACHE II, and SOFA in the L3 time period. b, Neural network alarm rate in percentage for different time periods (F3) the first 3 days; (H1) the first half of ICU stay; (H2) the second half of ICU stay; (L3) the last three days.

**Table 2.** 95% Confidence Intervals (CI) for AUC of all Data Sets

| Data Set Figure 1a | 95% CI | Data Set Figure 1b | 95% CI | Data Set Figure 2a | 95% CI |
|---|---|---|---|---|---|
| freq16/F3 | [0.50, 0.62] | $RR_{sys}$ | [0.70, 0.90] | MODS | [0.77, 0.99] |
| freq16/S3 | [0.53, 0.65] | Thrombocytes | [0.74, 0.89] | SAPS II | [0.77, 0.92] |
| freq16/ALL | [0.59, 0.70] | Bac | [0.84, 0.93] | APACHE II | [0.70, 0.89] |
| freq16/D6-8 | [0.70, 0.81] | $RR_{dia}$ | [0.78, 0.88] | SOFA | [0.83, 0.96] |
| freq16/L5 | [0.81, 0.92] | Lungs | [0.88, 0.96] | | |
| freq16/D2-4 | [0.82, 0.92] | Heart | [0.86, 0.94] | | |
| freq16/L3 | [0.86, 0.94] | Bpt | [0.83, 0.92] | | |
| freq16/L2 | [0.86, 0.97] | Coagulation | [0.76, 0.94] | | |
| freq16/L1 | [0.87, 0.99] | freq16/L3 | [0.86, 0.94] | | |

Abbreviations: APACHE II, Acute Physiological and Chronic Health Evaluation; AUC, area under the curve; $RR_{sys}$, systolic blood pressure; bac, breathing and catecholamines; $RR_{dia}$, diastolic blood pressure; bpt, systolic and diastolic blood pressure, thrombocytes, MODS; multipleorgan dysfunction score, SOFA; sepsis-related organ failure assessment, SAPA II, simplified acute physiology score.

For clinical practice, the good performance of the neural network can be obtained by the RRT score [25,27,28] of the 3 variables, see http://medan.de/scores/scores.htm. The score table was generated by a genetic algorithm and is defined by the following Table 3 (results with 282 patients).

The values of the 3 variables systolic blood pressure $RR_{sys}$, diastolic blood pressure $RR_{dia}$, and the amount of thrombocytes $T$ listed in the table are each assigned to a single score value. The total RRT score is obtained as the sum of the 3 values. Using this score, we might demonstrate the alarm system. Please note that higher values are associated with a less critical state of the patient.

The optimal threshold $\theta$ for the score is obtained for $\theta = 6$, see Table 4.

On the basis of this definition, we get the following classification rule:

> "if RRT-Score <6 then deceases or if RRT-Score $\geq$6 then survives." The evaluation shows that this is true for 85.70% of correctly classified data samples. If we compare the AUC values of the neural network (AUC = 0.88) and the SOFA score (AUC = 0.89) with the diagnostic performance of the new RRT-Score (AUC = 0.89), we see that the classification power is equivalent, using much less information than the SOFA score.

The boundary between the 2 classes is very sharp: patients with 5 points have a mortality of 68.3% whereas only 19.8% of the patients with 6 points died. Thus, patients with a score near the boundary should be observed with special care.

## Discussion

Most clinicians recognize septic shock. However, there are divergences in the applied definitions,[29] although consensus conferences should have resolved this issue long since.[12,13] In this article, we define the term "septic shock" stringently, the term "severe sepsis" is deliberately avoided, since we could demonstrate in a previous study that in a former study "severe sepsis" comprises almost identical patients in a state of abdominal septic shock.[30]

Different scoring systems have been developed, not only to document the severity of the illness but also to predict the prognosis of critically ill patients. The best outcome predictor

**Table 3.** The New RRT Score[a]

| Var.\Score | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $RR_{sys}$ | ≤119 | >119 | >151 | >221 | >251 | >265 | – | – | – |
| $RR_{dia}$ | ≤42 | >42 | >47 | >49 | >64 | >83 | >117 | >121 | >126 |
| T | ≤112 | >112 | >202 | >312 | >371 | >621 | >770 | – | – |

Abbreviations: $RR_{sys}$, systolic blood pressure; $RR_{dia}$, diastolic blood pressure.[a]For each single measurement, a score value is obtained. The resulting RRT score is the sum of all three values.

**Table 4.** Lethality at Different Medan RRT Score Values

| Score | 0-2 | 3-5 | 6-9 | 10-13 |
|---|---|---|---|---|
| # Samples | 126 | 267 | 351 | 53 |
| # "Deceased" | 124 | 218 | 48 | 1 |
| Mortality | 98.4% | 81.7% | 13.7% | 1.9% |

would be one that warns the physician on the first day of ICU admission or when septic shock is first manifested (this is usually the second day of the patient's ICU stay according to our analysis). Our results demonstrate that none of the scoring systems achieves this goal. Only in the last 3 days of the ICU period, scores reach acceptable AUC values, whereby the SOFA score, based on 10 variables, achieves the best AUC of all scores. Like the SOFA score, the data-driven neural network approach has a similar performance, using only 3 variables (bpt).

The failure of scores and of the neural network does not imply that it is impossible to predict the future condition of the patient in advance by more than 3 days. We do not know whether a better prediction is possible for a certain subgroup of all patients, for example, a group discriminated by a gene test. Other categories of patient data may become available which may provide better results in the future.

Given the standard patient data, no better prediction is possible. Although scores and the neural network under investigation provide relevant outcome prediction information only in the last 3 days of the ICU stay of patients (ie, often without clinical relevance), it is worthwhile to look at the data more closely. The CI values in Table 2 show that scores are difficult to use for individual patients. A score value does not indicate with a high confidence death or survival that results in long CIs. The neural network data concerning the nonscore data sets are more reliable since CI length is usually shorter. The SOFA score has the lowest interval length (0.13) of all the traditional scores, so that it is the best score for abdominal septic shock. But, SOFA's CI length is 0.13, whereas bpt's CI length is only 0.09. Considering all data sets (eg, lungs, heart, bpt, freq16), the results show the superiority of neural networks compared to scores with respect to the confidence of a classification of individual patients.

The resulting alarm system based on our analyses produces reliable alarms. In the last 3 days of the ICU stay, there were 10 times more alarms for deceased patients than for survivors. The alarm system that was trained with data of the last 3 days represent the patient conditions that lead to death or survival with a high probability. Only false alarms (7%) stemming from the last 3 days can be interpreted as "false alarms" with respect to outcome prediction, because on the other days one cannot retrospectively establish whether the alarms are due to critical or uncritical states which might have occurred independently. Alarms in previous periods for surviving patients are not necessarily false; they can be seen as indicators for critical periods of ICU stay. Although the alarm system was trained on the basis of data from the last 3 days, it can be used as an online bedside alarm system. Right from the start of the patients' ICU stay, physicians are warned when patients reach the same critical condition as deceased patients had within the last 3 days. If the patient is critical on his/her first day of ICU stay, the alarm system warns the physician whether the patient is likely to survive or to die within the next few days. If a dramatic change occurs later on, the alarm system will warn the physician at the right time.

Without doubt, the bedside alarm application of the proposed alarm system is not designed for direct use of the patient. It is a tool showing the probabilities involved; a noninterpretative usage might only lead to fatalistic or euphoric behavior without a benefit for the patient. Therefore, in clinical practice, the system should be regarded as having a watchdog function and be integrated into other intensive care software. It will serve as another indication for the supervising doctor, either prompting additional diagnostic or treatment steps or corroborating them.

Of course, the clinical usefulness of this Web-based alarm system must be checked in a prospectively randomized multicenter study. We started the trial after closing the database and analyzing the data. Only patients with abdominal septic shock are included in this study (protocol at www.medan.de) in accordance with the (not new) notion that it is time to reconsider the sepsis concept and to change the design of future trials.[31] The progress of the trial is supervised by a specialist for medical statistics who had participated in a data-monitoring committee. Unfortunately, the less divergent the outcomes of the 2 branches are (1 with prediction, 1 without), the more data had to be collected by the double-blind study to make a significant statement in either direction. At present, the randomized trial is still in progress.

## Authors' Note

The authors have been members of the MEDAN working group: Ernst Hanisch, 1998–today; Rüdiger Brause, 1998–today; Jürgen Paetz, 1998–2003; and Björn Arlt, 2000–2003.

## Acknowledgment

The authors would like to thank all the participants of the MEDAN working group, especially all the medical students involved in the project, the participating hospitals and all data typists. We thank Prof Dr A Encke for supporting the MEDAN working group.

## Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

## Funding

This study was partially supported by the "Deutsche Forschungsgemeinschaft" DFG (German Research Foundation, project MEDAN: HA1456/7-1,2).

## References

1. Schottmüller H. Wesen und Behandlung der Sepsis. *Inn Med*. 1914;31:257-280.
2. Hanisch E, Encke A. Intensive care management in abdominal surgical patients with septic complications. In: Faist E, ed. *Immunological Screening and Immunotherapy in Critically Ill Patients With Abdominal Infections.* Berlin: Springer-Verlag; 2001:71-138.
3. Hotchkiss R, Karl I. The pathophysiology and treatment of sepsis. *N Engl J Med*. 2003;348(2):138-150.
4. Rivers E, Nguyen B, Havstad S, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*. 2001;345(19):1368-1377.
5. Van den Berghe G, Wouters P, Weekers F, et al. Intensive insulin therapy in critically ill patients. *N Engl J Med*. 2001;345(19):1359-1367.
6. Annane D. Corticostereoids for septic shock. *Crit Care Med*. 2001;29(suppl):S117-S120.
7. Brunkhorst F, Engel Ch, Bloos F, et al. Intensive insulin therapy and pentastarch resuscitation in severe sepsis. *N Engl J Med*. 2008;358(2):125-139.
8. Sprung CH, Annane D, Keh D, et al. Hydrocortisone therapy for patients with septic shock. *N Engl J Med*. 2008;358(2):111-124.
9. Bernard GR, Vincent J-L, Laterre PF, et al. Efficacy and safety of recombinant human activated Protein C for severe sepsis. *N Engl J Med*. 2001;344(10):699-709.
10. Eichacker P, Danner R, Suffredini A, Cui X, Natanson C. Reassessing recombinant human activated protein C for sepsis: Time for a new randomized controlled trial. *Crit Care Med*. 2005;33(10):2426-2428.
11. Afshari A, Wetterlev J, Brok J, Moller A. Antithrombin III in critically ill patients: a systematic review with meta-analysis and trial sequential analysis. *BMJ*. 2007;335(7632):1248-1251.
12. Bone RC, Balk RA, Cerra FB, et al. American college of chest physicians/society of critical care medicine consensus conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med*. 1992; 20(6):864-875.
13. Levy MM, Fink MP, Marshall JC, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Crit Care Med*. 2003;31(4):1250-1256.
14. Paetz J, Arlt B, Erz K, Holzer K, Brause R, Hanisch E. Data Quality Aspects of a Database for Abdominal Septic Shock Patients. *Computer Methods Programs Biomed*. 2004;75(1):23-30.
15. 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(7):707-710.
16. Vincent JL, de 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*. 1998;26(11):1793-1800.
17. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13(10): 818-829.
18. Le Gall Jr, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270(24):2957-2963.
19. 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(10):1638-1652.
20. Jennett B, Teasdale G. Assessment of coma and impaired consciousness: a practical scale. *Lancet*. 1974;1(2782):81-84.
21. Paetz J. Metric Rule Generation with Septic Shock Patient Data. Proc. of the 2001 IEEE Int. Conf. on Data Mining (ICDM'01), San Jose, CA, USA: IEEE Computer Society Press; 2001:637-638.
22. Paetz J. Knowledge-based approach to septic shock patient data using a neural network with trapezoidal activation functions. *Artif Intell Med.* 2003;28(2):207-230.
23. Brause R, Hamker F, Paetz J. Septic shock diagnosis by neural networks and rule based systems. In Schmitt M, et al. eds. *Computational Intelligence Processing in Medical Diagnosis*. Heidelberg: Physica; 2002:323-356.
24. Hartung J. *Statistik*. 11th ed. Munich: Oldenbourg; 1998.
25. Brause R, Hanisch E, Paetz J, Arlt B. The MEDAN-Project: Results and their medical meaning. In: Nierhaus A, Schulte am Esch J, eds. *3rd Int. Symp, Sepsis, SIRS, Immune Response–Concepts, Diagnostics and Therapy*. Lengerich, Germany: PABST Science Publishers; 2003:122-129.
26. Paetz J, Arlt B. A Neuro-fuzzy Based Alarm System for Septic Shock Patients with a Comparison to Medical Scores. Lecture Notes in Computer Science, Berlin/Heidelberg, Germany: Springer Verlag 2002; (2526):42-52.
27. Paetz J. Evolving Score Neural Networks. In: Bhanu Prasad(Ed.): Proc. of the 1st Indian Int. Conf. on Artificial Intelligence IICAI 2003, Hyderabad, India, IICAI, ISBN 0-9727412-0-8, 2003: 57-65.
28. Paetz J. Finding optimal decision scores by evolutionary strategies. *Artif Intell Med*. 2004;32(2):85-95.
29. Rowe PM. Septic shock–finding the way through the maze. *Lancet*. 1999;354(9195):2058.
30. Wade S, Büssow M, Hanisch E. Epidemiologie von SIRS, Sepsis und septischem Schock bei chirurgischen Intensivpatienten. *Chirurg.* 69, 1998:648-655.
31. Carlet J, Cohen J, Calandra T, Opal S, Masur H. Sepsis: time to reconsider the concept. *Crit Care Med*. 2008;36:964-966.