An Alarm System for Death Prediction

Rüdiger Brause, Computer Science Department, Goethe University, Frankfurt, Germany

Ernst Hanisch, Department of Surgery, Asklepios Klinik Langen, Academic Teaching Hospital of Goethe University,Frankfurt, Germany

ABSTRACT

The clinical treatment of sepsis is one of most severe issues in hospitals. Unfortunately, until now it has not been possible to significantly reduce the mortality rate of severe forms of sepsis like septic shock, which is as high as 50-60% worldwide. Often, the diagnosis and awareness for possible implications of sepsis can be facilitated by an automated online diagnosis. This contribution reports the development of a monitoring alarm system for the individual prediction of death based on the data of 382 patients with septic shock. The paper discusses the pros and cons of such a prediction system used in a medical environment, its principal usage issues and implementation.

Keywords: *Alarm System, Death Prognosis, Intensive Care Unit (ICU) Monitoring, Patient Bedside Monitoring, Septic Shock*

1. INTRODUCTION

The infection of men by pathogenic organisms in the bloodstream occurs in 1–2% of all hospitalizations and accounts for as much as 25% of Intensive Care Unit (ICU) bed utilization. The clinical treatment of sepsis is one of most severe issues in hospitals and is managed by several strategies. Nevertheless, more severe forms of sepsis, i.e. septic shock which occurs in about 5-7% of all septic cases are very difficult to handle and lead often to multi-organ failure and to death which is as high as 50-60% worldwide. Unfortunately, until now it has not been possible to significantly reduce the mortality rate of septic shock. Therefore, early treatment of complications in septic shock supported by an online prediction alarm system is an important issue in the hospital routine. In contrast to Personal Emergency Response Systems (PERS) which are triggered by the push button alarm of the person wearing a wireless transmitter, the ICU patients often are not aware of their state and do not have the possibility to push a button in the emergency case.

Therefore, the surveillance and alarm trigger of an automatic system is desirable. For this purpose, there do exist many alarm trigger established in the pulse monitor, respiration monitor, and other physiological devices. Unfortunately, due to the high number of different, not integrated devices the number of false alarms

DOI: 10.4018/ijmstr.2013040102

is very high. In a study in the John Hopkins Hospital, Baltimore (Cvach et al., 2012), they counted 771 alarm conditions per bed per day in the ICU. For instance in the case, of 157 apnea alarm conditions, 90% were thought to be false, for example, apnea alarm signals coming from patients on ventilators. Such frequent alarming causes desensitization due to a "cry wolf "effect. Most nurses don't know all of the ways that the alarm system can be customized to individual patients. They concluded that a coordinated, continuous effort is necessary for the staff in order to integrate all devices into a coordinated alarm system and set up and tune their alarm conditions to the appropriate case. By this, they managed to reduce the false alarms by 30%.

How can an alarm system be build which integrates the most valuable signals, can be adapted to the actual needs and has a minimal number of false alarms? This contribution describes the foundations, development and application issues of a septic shock alarm system which is able to predict death three days in advance individually. It is based on the data base of the MEDAN project (Hanisch, Brause, Paetz, & Arlt, 2011) of 382 patients out of 582 with septic shock (Hanisch et al., 2003). The data were collected in 102 German hospitals from 1998 to 2002. All handwritten patient records were transferred to an electronic database afterwards by a huge amount of man power to a consistent electronic database of 2.5 million data. We used programmed range and plausibility checks of different kinds to detect all faulty data in the electronic database. For this, static values (e.g. lower and upper bounds) and dynamic development (e.g. time sequence behavior) were checked (Paetz et al., 2004).

2. THE ALARM SYSTEM

The goal of the data analysis was the development of a prediction system for the individual mortality prognosis. Such systems can be used in the emergency case or for advances in treatment by automatic state monitoring. In our case, the analysis goal was two-fold: first, we liked to trace back the causes and influences of several clinical variables like coagulation or thrombocyte level to the patient outcome, and second, we aimed to build an alarm system which rings an alarm as soon as a bad state is reached. In this contribution, we focus on the latter case.

The main problem for a medical alarm system is the availability of data. Unfortunately, there is no clinical standard for bedside monitoring or patient data bases in Germany. For this reason, for our analysis of septic shock (which is a rare event) we had to initiate a multi-center study and concentrated on those 140 variables only, which are currently available in clinical routine. All gene tests or other special features were practically out of reach.

The next problem after obtaining the data is the question: What kind of analysis system should we use? It is well known that most metabolic processes are non-linear. Therefore, the usage of all linear methods like correlation analysis is not adequate. Instead, we used the method of formal neural networks.

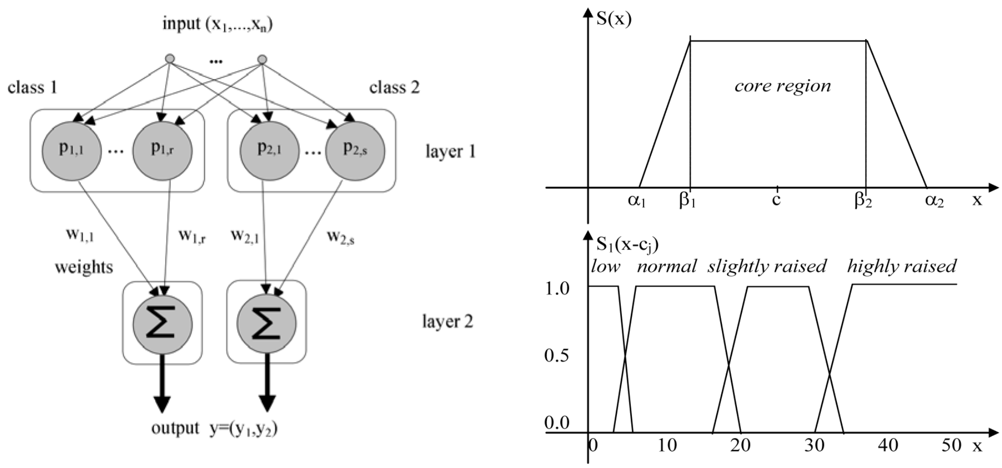
3. THE NEURAL NETWORK ALARM SYSTEM

It is well known that artificial neural networks can approximate every arbitrary function as close as desired (White 1992). Here, a two-layer network is sufficient. The main architecture is shown in Figure 1(a). As real valued input, the clinical variables are used. The output consists of two variable DEATH and ALIVE. Only one of these two outputs becomes exclusively TRUE depending on the maximum.

For the discrete choice of the network, both design choices have to be made: Which activation function and which learning function should we choose? Both decisions imply deep consequences for a network.

As activation function of the first layer neurons, a Radial Basis Function (RBF) $S_1(|x-c_j|)$ was chosen. For an input variable x, this function decreases monotonically with the distance from a center c. Here, it becomes only one (true) if the input represents a point which falls within

Figure 1. (a) The 2-layer architecture; (b) Activation functions of the first layer



(a) (b)

a certain interval in the neighborhood of a state c, see Figure 1(b). In detail, our RBF activation functions consist of an interval [β_1,β_2] (*core region*) where S(z) = 1 and an interval [α_1,α_2] (*support region*) where at least S(z) > 0. Thus, the *j*-th trapezoidal function of *z* might be:

$$S_1(z) = \begin{cases} 0 & z < \alpha_{1j} \\ \frac{z - \alpha_{1j}}{\beta_1 - \alpha_{1j}} & \alpha_{1j} \le z \le \beta_{1j} \\ 1 & \beta_{1j} \le z \le \beta_{2j} \\ \frac{{}_{\alpha 2j} - z}{\alpha_{2j} - {}_{2j}} & \beta_{2j} \le z \le \alpha_{2j} \\ 0 & z >_{\alpha 2j} \end{cases}$$

or purely rectangular:

$$S_1(z) = \begin{cases} 0 & z < \alpha_{1j} \\ 1 & \beta_{1j} \le z \le \beta_{2j} \\ 0 & z > \alpha_{2j} \end{cases}$$

A trapezoidal function can be seen like a fuzzy membership function where the function value varies between S(z) = 0 FALSE and S(z) = 1 TRUE. Contrary to classical logic, also values between those two (less or more TRUE) are possible. This can easily be used to transfer vague ideas and values of medical diagnosis from the doctors to the diagnostic system (the RBF network) and the results from the network to the doctors, describing the results in vague, medical terms like *slightly raised glucose level*, see Brause & Friedrich (2000). All human assumptions are formulated as rules and fed into the network; all network results are output and interpreted as rules. This formalism enables doctors to enter medical knowledge directly into the learning system and let the system output its results in medical terms which are better understood by them.

The input selection in our network is done by trapezoidal (RBF) functions of the first layer which cut out a certain area of the input space. The combination of several trapezoidals $S_1(|x-c_j|)$ gives the output of the second layer. Formally, the input-output function of the network for the *i*-th output y_i is given by:

$$y_i(x) = S_2\left(\sum_j w_{ij} S_1(|x - c_j|)\right) \qquad (1)$$

In the simple case, the activation function is linear:

$$S_2(z) = z$$

As learning algorithm for this network (Brause, Hamker, & Paetz, 2002) we used a variant of an algorithm (Huber & Berthold, 1995), developed by Paetz (2001). In our case, after training of the network we obtained 1284 rules for decease like:

```
IF organ_failure = YES
AND antiarrythmics = YES
AND haemodialysis = YES
AND peritoneal_lavage = YES
THEN class_deceased
WITH confidence = 0.8
AND frequency = 0.03
```

and 9976 rules for survival like:

```
IF peritoneal_lavage = NO
AND thrombocyte_concentrate = NO
AND haemodialysis = NO
THEN class_survived
WITH confidence = 0.98
AND frequency = 0.42
```

For an automatic computer diagnosis, this is ok, but it is difficult to compute all values in parallel manually.

4. THE SCORING ALARM SYSTEM

Although our neural network performed well for the task of death prediction, the neural network alarm system is not well accepted by medical persons. This non-rational fact is related to a more general human background: All doctors feel their profession as a kind of art; themselves they feel like artists. So, technical systems can only be accepted by doctors as (dump) assistance giving hints, not as valid diagnosis and never as prescription. For this, the medical diagnosis is disguised as a dummy alarm system, not as a profound diagnosis.

Additionally, the doctors want to understand the reasons for the alarm. Therefore, the first step was an output of the neural network system which can be understood very easy: the formulation as rules. This is good, but not preferable, because it needs a computer running the neural network application. In contrast to this, doctors are used to compute health indicators, called "scores" manually.

4.1. Neural Networks and Medical Scores

There are many famous medical scores used to predict the health of a septic patient: the SOFA (*Sepsis-Related Organ Failure Assessment*) score (Vincent et al., 1996,1998) which uses 12 different variables, the APACHE II (*Acute Physiological and Chronic Health Evaluation*) score (Knaus et al., 1985), using a scale of 0 to 71 of whole-number values based on 12 partial variables, the SAPS II (*Simplified Acute Physiology Score*) score of 15 variables (Le Gall et al., 1993), and the MODS (*Multiple Organ Dysfunction Score*) of only 6 variables (Marshall et al., 1995). All scores are computed by the same procedure: Using a weight table the doctors take the appropriate weights for the patient state variable values from the table and add them together to a score. The final score is then put into another table which maps the score to the patient's health state prediction.

Certainly, in order to propagate the usage of our results we should make it also available as a scoring system. But - how can we obtain our diagnostic networks results by a simple system like this? The solution comes up when we take a closer look to the scoring procedure and compare it to the neural network activity.

For computing a score, we have first to determine in which interval a measured variable falls, then assign a partial score value to it and then add all the values of the different partial scores together to the final score. We might mathematically formulate this by defining a function S(x) to be one within the *j*-th interval [β_{1j}, β_{2j}] if the measured value falls within the interval boarders β_{1j} and β_{2j}:

$$S_{ji}(x_i) = \begin{cases} 1 & \beta_{1ji} \le x_i \le \beta_{2ji} \\ 0 & \text{else} \end{cases}$$

and assign as score value the weight w_j to it. By multiplication, the *i*-th partial score value becomes:

$$\text{Partial score} = w_j S_{ji}(x_i)$$

As example the SOFA variable x_i ="Bilirubin" with its associated score values w_{ji} is shown in Figure 2.

Then the final score value is computed as sum of the partial scores only where S_{ji} =1:

$$\text{Score} = \sum_j \sum_i w_{ji} S_{ji}(x_i) = \sum_i w_{ji} S_{ji}(x_i) \quad (2)$$

using the variables x_i and their corresponding interval weights w_{ji} defined on each variable. Comparing the two expressions (1) and (2) we notice that they are formally equivalent under the condition that the activation functions S_{ji}(.) of the neural network is identical to the interval-shaping functions S_{ji}(.) of the scores. Thus, the scores can be seen as "score networks" and are special cases of trapezoidal networks which in turn are special cases of general artificial neural networks.

But scores and neural networks are not the same. In difference to scores which are statically defined and do not change, neural network parameters like weights are supposed to change. Special learning algorithms adapt the weights such that the network diagnostic performance becomes maximal in difference to the scores which are defined manually once by statistical and consensus considerations.

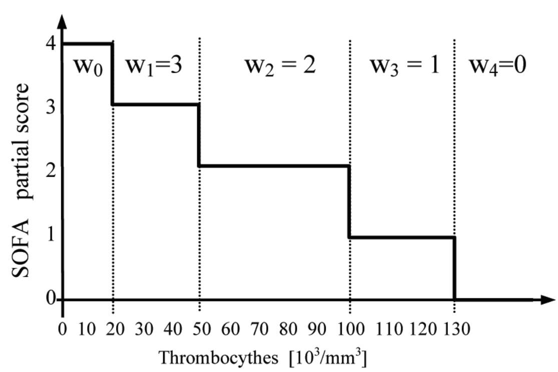
Now, given an optimally trained neural network, how can we construct the corresponding score?

4.2. Generating a Medical Score

The neural network obtained by training has two differences to the desired "score network": First, it contains many rules which are concurrently active, and second, it is based on trapezoidal basis functions, not on rectangular ones. Thus, in order to construct a score, we have to reduce the number of rules, i.e. the number of intervals of a variable, and they should not overlap. Additionally, all concurrently active rules have to be mapped to a single rule for one interval producing just one partial score value.

There is no canonical procedure for this mapping task. Therefore, we applied a random

Figure 2. The output function weights of the thrombocythes SOFA variable



mapping scheme with selection to this problem: a genetic algorithm. For the algorithm, a possible solution is represented by a tuple of interval boarders and number of intervals per variable. Starting with a set of random solutions (genotypes) the performance (fitness) of each solution can be obtained by using the score computation (phenotype) on the set of training samples and computing the score performance, see section 5. The best solutions can be selected, mutated and tested again. This process is visualized in Figure 3.

After the generation of 360 mutations and selections a score network was obtained (Paetz 2003) which did not change in several training sequences.

By our neural network analysis, we identified the systolic and diastolic blood pressure/ thrombocytes system as the most relevant variables for outcome prediction (Paetz & Arlt 2002). For clinical practice, the good performance of the neural network can be obtained by the MEDAN RRT score of the three variables. This score network can be described like an ordinary score by a table, see Table 1. It uses only three observed variables: the systolic blood pressure, the diastolic blood pressure and the concentration of thrombocythes. The values of the three variables systolic blood pressure RR_{sys}, diastolic blood pressure RR_{dia} and the amount of thrombocytes T_{hromb} listed in the table are each assigned to a single score value. The total RRT score is obtained as sum of the three 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.

By this definition, we get the following classification rule: "if RRT-Score < 6 then deceases" or „if RRT-Score ≥ 6 then survives". The new score can be easily implemented by a small piece of software which might be added to standard patient bedside monitoring devices. A pseudocode example of the necessary alarm software for computing the death prediction is shown in Figure 4.

The blood pressure can be monitored continuously, while the thrombocythe values have to be updated manually.

5. ALARM SYSTEM RESULTS

The performance of a new alarm system has to be compared with traditional approaches. As performance measure, we do not use the probability of successful diagnosis, because it does not include the characteristics of the diagnostic system at different working points (different thresholds and other parameters), but the resulting receiver operation characteristic (ROC) with its area under curve (AUC) value. Random diagnosis have an AUC of 0.5; in medicine, for an acceptable, serious diagnosis system the AUC value have to be higher than 0.8.

Figure 3. The iteration process of the genetic algorithm

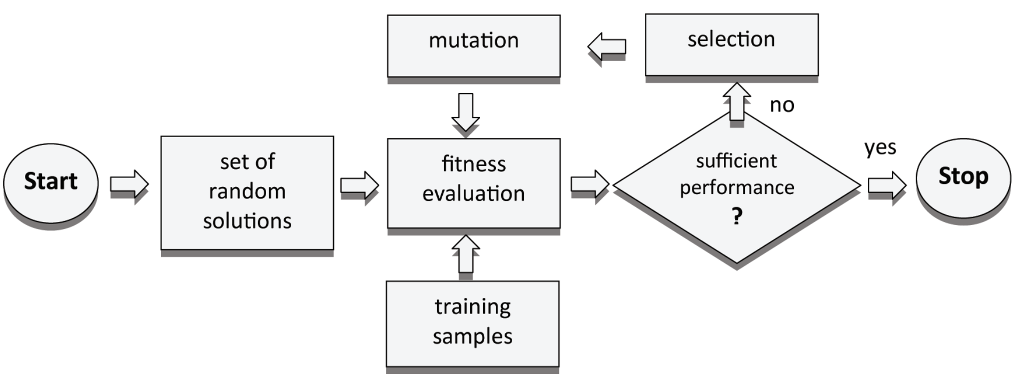


Table 1. The new MEDAN RRT score

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_{hromb}	≤112	>112	>202	>312	>371	>621	>770	-	-

Figure 4. Example pseudocode of the alarm system

```
public static boolean RRT_alarm(float RRTsys, float RRTdia, float Thromb){
    int score = 0, sc;
    boolean alarm;

    float[] Rs = {119,151,221,251,265};
    for (sc = 0; sc < Rs.length; ) {        // systolic blood pressure score
        if (RRTsys <= Rs[sc]) break;
        sc++;
    } score = score + sc;

    float[] Rt = {42,47,49,64,83,117,121,126};
    for (sc = 0; sc < Rt.length; ) {    // diastolic blood pressure score
        if (RRTdia <= Rt[sc]) break;
        sc++;
    } score = score + sc;

    float[] Tr = {112,202,312,371,621,770};
    for (sc = 0; sc < Tr.length; ) {    // thrombocythe concentration score
        if (Thromb <= Tr[sc]) break;
        sc++;
    } score = score + sc;

    alarm = (score < 6);
    return alarm;                    // return true for a score < 6
} // end RRT_score
```

5.1. The Performance Evaluation

For the neural network diagnosis, we get an AUC of 0.88. The metric variables hold most of the diagnostic information: After adding qualitative variables like treatment or medication the diagnosis augmented only slightly from AUC = 0,90 to 0,92 for a subset of 138 patients.

In the MEDAN score, the border between the two 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 border should be observed with special care.

In order to compute the score, the three score values have to be summed up. The threshold is $\theta = 6$: For a sum greater or equal to 6 the outcome prediction is favorable (85,7% correctly classified as "survived"), otherwise severe problems will arrive. In our experience, patients who stay several days below a score of 6 have a high probability to die. In Table 2, this is shown by the mortality associated to the score ranges.

How does the new score perform generally in comparison to the other standard scores used in ICUs? For a comparative analysis of the new MEDAN score we evaluated the traditional scores on our data. For our analysis, a score was calculated every time when the necessary variables were given without considering the Glasgow Coma Score (GCS). The GCS was not included in the scores since it was not always available for our data.

With an AUC = 0,89 the performance of the MEDAN score differs not much from

Table 2. Mortality related to MEDAN score ranges

MEDAN RRT Score	0..2	3..5	6..9	10..13
Mortality	98.41%	81.65%	13.68%	1.89%

the MEDAN neural network performance. Therefore, we might compare it directly to the other established scores. If we compare the diagnostic performance of the MEDAN score of our patients to their SOFA score (AUC = 0.89), their MODS score (AUC = 0.88), their SAPS II score (AUC = 0.85) and their APACHE II score (AUC = 0.79) we see that the classification power of the new score is equivalent or better, using much less variables (information) than all other scores.

We notice that the MEDAN RRT score is as performing as the best by experience-evolved score, the SOFA score. We might interpret that as if we short-cut the evolution of our score by an evolutionary algorithm based on the available data.

5.2. The Prediction Epoch

After selecting relevant variables for prediction, we have to determine the time period of the samples to analyze. This comes up to the question: Is the fate of the patient already determined after entering the ICU? At what day of the stay can we make valid predictions about the outcome? Is the outcome determined by the first three days of the stay? Or the first half of the stay? Or the second half of the stay? Or the last three days? For all the time periods, we analyzed the performance of all scores. An alarm message is given whenever input for the neural network generates high output for class "deceased." In Figure 5 we see the resulting alarm percentage for the first three days, for the first and second half of ICU stay and for the last three days, indicated separately for patients who either deceased or survived.

We can clearly see that predictions of the MEDAN score based on the start of the ICU stay are not reliable (AUC = 0.52) and are close to an arbitrary random decision. This is also true for the AUC of all other scores: SOFA = 0.54, APACHE II = 0.52, SAPS II = 0.52, MODS = 0.52, neural network = 0.52.

The usage of the score is not limited to the last three days: A bad score indicates a bad situation for the patient "as if he or she is in the state of being in the last three days before death" whenever the score is computed. Comparisons of individual histories and computed scores showed good correlations for the whole ICU stay. Only the 7% alarms stemming from the last three days can be interpreted as false alarms with respect to outcome prediction.

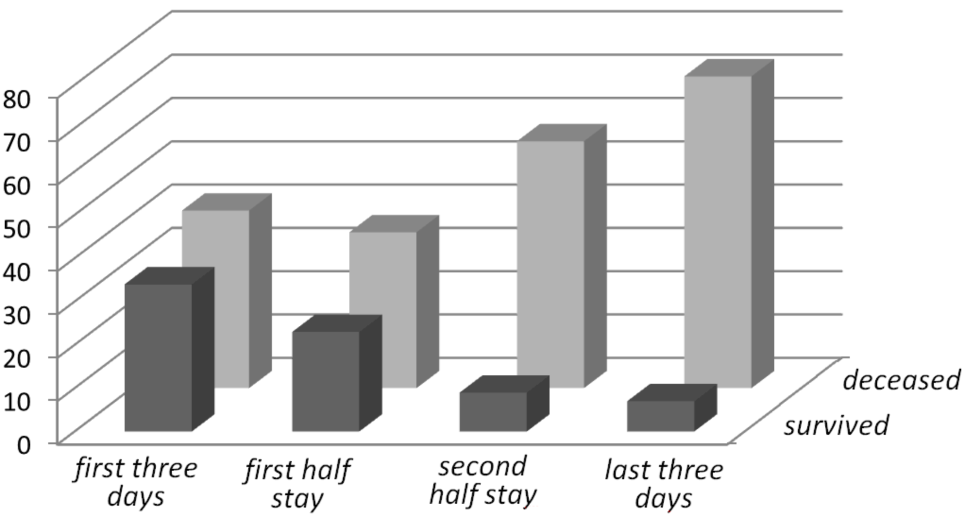
6. DISCUSSION

Most clinicians can recognize septic shock, but if you ask them, you get a hundred definitions (Rowe 1999), although consensus conferences should have resolved this issue (Levy et al., 2003). Different scoring systems have been developed, not only in order to document severity of illness, but also to estimate prognosis of critical ill patients. The best outcome predictor would be one that warns the physician on first day of ICU admission or when septic shock first appears (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 three days of the ICU period, scores reach acceptable AUC values, where by the SOFA score, based on ten variables, achieves the best AUC of all scores, together with the neural network and the MEDAN score which uses only three variables.

Does the failure of scores and of the neural network imply that it is impossible to predict the future state of the patient in advance than

Figure 5. MEDAN score alarms for different time periods



three days? Is it principally not possible to build an better alarm system? The answer is "No": We do not know whether a better prediction is possible for a certain subgroup of all patients, e.g. a group discriminated by a gene test. Or, other kinds of patient data may be available in future which may give better results.

Our main claim is that, given the standard patient data, there is no better alarm system prediction possible. Regarding Figure 5 we might even argue that after the first three days during the stay the fate of the patient is not determined. The decision seems to come only at the end of the stay: either by recovery or by death.

The resulting alarm system based on our analyses produces reliable alarms: in the last three days of the ICU stay there were ten times more alarms for deceased patients then for survivors. The alarm system that was trained with data of the last three days represents the patient conditions that lead to death or survival with a high probability. Only false alarms (7%) stemming from the last three days can be interpreted as "false alarms" with respect to outcome prediction, because on the other days one cannot retrospectively examine if the alarms are due to critical or uncritical states which might occur independently. Alarms in previous periods for survived patients might have not to be false; they can be seen as indicators for critical periods of ICU stay. Although the alarm system was trained with data of the last three 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 three days. If the patient is critical on his/her first day of ICU stay, the alarm system warns the physician, whether the patient will likely survive or die in the following days. If peripety happens later on, the alarm system will warn the physician at the right time.

Certainly, the bedside alarm application of the proposed alarm system is not dedicated to the direct use for the patient. It is a tool showing the involved probabilities; a not interpretive 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 a watch dog function and be integrated into other intensive care software. In this context it will serve as another indication for the supervising doctor, either causing additional diagnostic or treatment steps or confirming them.

ACKNOWLEDGMENT

We want thank all people who have supported us by their work, especially the MEDAN team for filling the data base with the 2.5 million data values and monitoring the clinical multi-center-study, B. Arlt for setting up the data base, transfer data from bedside systems to our data base and programming the alarm system, and Dr. Paetz for the data analysis. Without them, this work would not have been possible. This work was partially supported by the "Deutsche Forschungsgemeinschaft" DFG (German Research Foundation, project MEDAN: HA1456/7-1,2) by a grant to the authors.

REFERENCES

Brause, R., & Friedrich, F. (2000). A neuro-fuzzy approach as medical interface. In *Proceedings of the European Symp. on Art. Neural Networks (ESANN 2000),* D-Facto, Brussels, (pp. 201-206).

Brause, R., Hamker, F., & Paetz, J. (2002). Septic shock diagnosis by neural networks and rule based systems. In M. Schmitt, H.-N. Teodorescu, A. Jain, A. Jain, S. Jain, & L. C. Jain (Eds.), *Computational intelligence processing in medical diagnosis* (pp. 323–356). New York, NY: Springer Verlag. doi:10.1007/978-3-7908-1788-1_12

Cvach, M., et al. (2012). Using data to drive alarm system improvement efforts, The Johns Hopkins Hospital experience. *Healthcare Technology Safety Institute HTSI, AAMI Foundation*, Fairfax, Arlington. Retrieved January 10, 2013, from http://www.aami.org/htsi/SI_Series/ Johns_Hopkins_White_Paper.pdf

Hanisch, E., Brause, R., Arlt, B., Paetz, J., & Holzer, K. (2003). *The MEDAN data base*. Retrieved January 10, 2013, from http://medan.de/datenbank/download_database.htm

Hanisch, E., Brause, R., Paetz, J., & Arlt, B. (2011). Review of a large clinical series: Predicting death for patients with abdominal septic shock. *Journal of Intensive Care Medicine*, *26*(1), 27–33. doi:10.1177/0885066610384058 PMID:21262751

Huber, K.-P., & Berthold, M. R. (1995). Building precise classifiers with automatic rule extraction. In *Proceedings of the IEEE International Conference on Neural Networks* (Vol. 3, pp. 1263-1268).

Knaus, W. A., Draper, E. A., & Wagner, D. P. et al. (1985). APACHE II: A severity of disease classification system. *Critical Care Medicine*, *13*(10), 818–829. doi:10.1097/00003246-198510000-00009 PMID:3928249

Le Gall, Jr, Lemeshow, S., & Saulnier, F. (1993). A new simplified acute physiology score (SAPS II) based on a European / North American multicenter study. *Journal of the American Medical Association*, *270*, 2957–2963. doi:10.1001/jama.1993.03510240069035 PMID:8254858

Levy, M. M., Fink, M. P., & Marshall, J. C. et al. (2003). 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference. *Critical Care Medicine*, *31*, 1250–1256. doi:10.1097/01.CCM.0000050454.01978.3B PMID:12682500

Marshall, J. C., Cook, D. J., & Christou, N. V. et al. (1995). Multiple organ dysfunction score: A reliable descriptor of a complex clinical outcome. *Critical Care Medicine*, *23*(10), 1638–1652. doi:10.1097/00003246-199510000-00007 PMID:7587228

Paetz, J. (2001) Metric rule generation with septic shock patient data. In *Proc. of the 1st IEEE Int. Conf. on Data Mining* (pp. 637–638).

Paetz, J. (2003). Evolving score neural networks. In *Proc. of the 1st Indian Int. Conf. on Artificial Intelligence (IICAI 2003)*, Hyderabad, India (pp. 57-65). IICAI, ISBN 0-9727412-0-8.

Paetz, J., & Arlt, B. (2002). A neuro-fuzzy based alarm system for septic shock patients with a comparison to medical scores. In *Proc. of the 3rd Int. Symp. on Medical Data Analysis*, 42-52

Paetz, J., Arlt, B., Erz, K., Holzer, K., Brause, R., & Hanisch, E. (2004). Data quality aspects of a database for abdominal septic shock patients. *J Computer Methods and Programs in Biomed*, *75*(1), 23–30. doi:10.1016/j.cmpb.2003.09.001 PMID:15158044

Rowe, P. M. (1999). Septic shock – finding the way through the maze. *Lancet*, *354*(9195), 2058. doi:10.1016/S0140-6736(05)76802-1 PMID:10636381

Vincent, J. L., de Mendonca, A., & Cantraine, F. et al. (1998). Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: Results of a multicenter, prospective study. *Critical Care Medicine*, *26*(11), 1793–1800. doi:10.1097/00003246-199811000-00016 PMID:9824069

Vincent, J. L., Moreno, R., & Takala, J. et al. (1996). The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine*, *22*, 707–710. doi:10.1007/BF01709751 PMID:8844239

White, H. (1992). *Artificial neural networks - Approximation and learning theory*. Oxford, UK: Blackwell.