Using Artificial Neural Networks for Classifying ICU Patient States

A Comparison Between Physician Methods and a Model-Based Approach

The rapid accurate diagnosis of critical disorders is an essential component of intensive care. Traditional diagnostic techniques have relied on physician experience, which is based on a data set chosen from his or her personal preferences, rather than from scientific merit. In this article, we aim to show that there are alternative methods of selecting clinical variables on which to base a diagnosis.

We suggest that a model-based technique utilizing artificial neural networks (ANNs) can be used to investigate alternative, objectively chosen data input sets. Traditionally, ANNs have been used for diagnosis or prediction tasks; however, this article introduces a novel method of exploring the inner structure of suitably trained ANNs to determine a set of key variables for each clinical state defined. Two different ANN techniques are proposed: self-organizing maps (SOMs) and back-propagation networks (BPNs). We do not claim that these techniques provide the optimal data set for decision making, but we do show that other combinations of data exist that may be an improvement over those currently used.

An unnoticed disturbance of sufficient oxygen delivery to any vital organ may lead to severe complications or even to death. Morgan, et al. [1], have defined

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAPn</td>
<td>systemic arterial pressure; n = m, mean (sometimes known as Mean Arterial Pressure, MAP); n = s, systolic; n = d, diastolic</td>
</tr>
<tr>
<td>PAPn</td>
<td>pulmonary arterial pressure: n = m, mean; n = s, systolic; n = d, diastolic</td>
</tr>
<tr>
<td>Tp, Tc</td>
<td>peripheral temperature, core temperature</td>
</tr>
<tr>
<td>HR</td>
<td>heart rate</td>
</tr>
<tr>
<td>ResF</td>
<td>respiratory frequency</td>
</tr>
<tr>
<td>CI</td>
<td>cardiac index</td>
</tr>
<tr>
<td>PCWP</td>
<td>pulmonary capillary wedge pressure</td>
</tr>
<tr>
<td>SVRi</td>
<td>systemic vascular resistance index</td>
</tr>
<tr>
<td>PFI</td>
<td>perfusion index (PaO2/FiO2)</td>
</tr>
<tr>
<td>PaO2</td>
<td>arterial oxygen tension</td>
</tr>
<tr>
<td>FiO2</td>
<td>fraction of inspired oxygen</td>
</tr>
<tr>
<td>aBpH</td>
<td>arterial pH</td>
</tr>
<tr>
<td>DO2, VO2</td>
<td>oxygen delivery, oxygen consumption</td>
</tr>
<tr>
<td>ExO2</td>
<td>oxygen extraction</td>
</tr>
<tr>
<td>aBBE</td>
<td>base excess</td>
</tr>
<tr>
<td>Lact</td>
<td>blood lactate concentration</td>
</tr>
</tbody>
</table>
The output from Node 8 is being investigated. The link and Node 8 with the largest modulus weight is from Node these nodes, therefore, the largest modulus weights a set of agreed upon four significant problems: library. They

gas-exchange abnormality. The IM
typical disorders of oxygen deliv ery and
ords for annotated 24-hour
val by Kuopio (evolved into high blood-flow
after
was granted Ethics Committee
risk
and subsequent publication or
application
by a physician, but where strict
the clinical definitions are given. The table shows that there are developing the above disorders.

As mentioned above, the diagnostic
process is classically based on the combination of "test and hypothesis" cycle and knowledge based on experience. The weakness of this method can be shown by appropriate use of the IMPROVE data library. Table 1 shows the correlation between clinical annotations by a physician and actual periods in a subset (which equals the cross-validation set used in the neural-network experiments described below) of the data library when the given patient states (NHC, HBFS, CF) occurred according to predefined clinical definitions. In these states (results of survey in the three ICUs described in [1]), HBFS is defined as either a state with abnormally high flow and the need of vasocative treatment to maintain perfusion pressure and low systemic vascular resistance index (SVRI), high flow with acceptable perfusion pressure and signs of tissue hypoxia, or high flow and acceptable perfusion pressure. Cardiac failure is defined as either a situation of inadequate flow and metabolic signs of tissue hypoxia, inadequate flow and no signs of tissue hypoxia, or acceptable flow and

a set of disorders for the IMPROVE data library. They studied the incidence of typical disorders of oxygen delivery and agreed upon four significant problems: hypovolaemia, cardiac failure, sepsis (evolved into high blood-flow state), and gas-exchange abnormality. The IMPROVE data library comprises a fully annotated 24-hour set of patient records for 60 patients who where at high risk of developing the above disorders. The monitoring, collection/storage of data, and subsequent publication or public presentation of this information was granted Ethics Committee approval by Kuopio University Hospital, after informed assent was given by the patients' relative(s). Subsequently, one patient withdrew consent and technical problems prevented the incorporation of another patient, leaving a database of 58 patients. Data to be used were generated every two minutes.

The database size, assuming that every variable is recorded exactly every 2 minutes during 24 hours, 24 hours times 30 measurements times 58 patients, is 41,760 samples for each variable (which are defined below). In reality, this number is less due to shorter recordings for different variables (see [2] for a more detailed discussion on this). To constrain the clinical problem, we investigate just two of the clinical disorders, high blood-flow state (HBFS) and cardiac failure (CF). We introduce a third class for convenience, which corresponds to the periods when patients are neither HBFS nor CF, denoted NHC (neither HBFS nor CF). By comparing the traditional physician approach with that of the ANNs, we hope to show that new data combinations are both possible and meaningful in a clinical context.

From a technical perspective, the objectives of this article are to provide an intracomparison between ANN techniques and an intercomparison between the ANN techniques employed and the physician. By doing this, we anticipate that clinical variables other than those already used routinely will emerge. Potentially, this may have a great impact on the diagnostic process employed in the care of the critically ill patient.

As mentioned above, the diagnostic process is classically based on the combination of "test and hypothesis" cycle and knowledge based on experience. The weakness of this method can be shown by appropriate use of the IMPROVE data library. Table 1 shows the correlation between clinical annotations by a physician and actual periods in a subset (which equals the cross-validation set used in the neural-network experiments described below) of the data library when the given patient states (NHC, HBFS, CF) occurred according to predefined clinical definitions. In these states (results of survey in the three ICUs described in [1]), HBFS is defined as either a state with abnormally high flow and the need of vasocative treatment to maintain perfusion pressure and low systemic vascular resistance index (SVRI), high flow with acceptable perfusion pressure and signs of tissue hypoxia, or high flow and acceptable perfusion pressure. Cardiac failure is defined as either a situation of inadequate flow and metabolic signs of tissue hypoxia, inadequate flow and no signs of tissue hypoxia, or acceptable flow and

<table>
<thead>
<tr>
<th>Annotations According to Definitions</th>
<th>NHC</th>
<th>HBFS</th>
<th>CF</th>
<th>Predictive Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHC</td>
<td>12057</td>
<td>1025</td>
<td>3940</td>
<td>71%</td>
</tr>
<tr>
<td>HBFS</td>
<td>1059</td>
<td>3114</td>
<td>59</td>
<td>74%</td>
</tr>
<tr>
<td>CF</td>
<td>1154</td>
<td>170</td>
<td>4833</td>
<td>78%</td>
</tr>
<tr>
<td>sensitivity/specificity</td>
<td>85%</td>
<td>72%</td>
<td>55%</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Correlation between physician annotations and clinical definitions (N=27411). The columns present presence of patient states as annotated by a physician. In the rows, patient states as indicated by strictly applying their clinical definitions are given. The table shows that there are 1025 periods in the DL where the patient state is annotated as HBFS by a physician, but where strict application of the clinical definition results in NHC.
2. Four examples of SOM weight vector component planes of one 12x16 node SOM. Each picture shows the magnitude in grayscale of one component of the weight vectors for all nodes in the SOM: white indicates a high value, black a low value. The nodes are labeled with patient states on the basis of the number of times observation vectors of that patient state were matched onto those nodes (nodes that do not have a label could not clearly be assigned to one specific patient state). HBFS states are covered by nodes in the upper part of the network, a large group of CF states in the right hand side, NHC in the rest of the network, with a relatively large number of CF states interspersed among them. Figure a) shows the distribution of weight components corresponding to Tp, b) to HR, c) to CI, and d) to PCWP. It can be seen that the magnitudes of CI, Tp, and PCWP clearly correlate with patient state; whereas, for HR there is no clear correlation.

Continuous need of exceptional support. Cardiac failure is defined separately for left and right ventricular failures, based on clinical judgment.

The numbers in the off-diagonal of Table 1, combined with the given predictive values and sensitivity to specificity ratio, demonstrate that the traditional methods of diagnosis are far from foolproof. Indeed, there are considerable periods in the data library when neither HBFS nor CF are correctly identified by the physician. This finding indicates that gradual changes in pathophysiological state of the patient are either not observed, or that the clinical definitions do not apply at that time. Alternatively, disorders may exist even though all the subcriteria for clinical definitions are not fulfilled. These findings highlight the need for a more detailed examination of the clinical variables used for diagnosing HBFS and CF. They also demonstrate the weakness of simple rule-based system approaches to this class of problem.

**Methods**

**Data-Set Generation**

Based on the descriptions given by physicians, 20 trend data variables were used to investigate HBFS and CF. There were too few hypovolaemia cases present in the data available to draw any significant statistical conclusions, and the gas-exchange abnormalities are very diverse, and, although important, the variety of patient states is too high for this phase of research.

The following parameters were used as input for the ANNs (see the Glossary for definitions): SAPs, SAPm, SAPd, PAPs, PAPm, PAPd, Tp, Tc, HR, ResF, CI, PCWP, SVRI, PFI, aBpH, DO2, VO2, ExO2, aBBE, Lact. The data vectors obtained were investigated by applying the trend data as feature vectors to two types of ANNs, Kohonen self-organizing maps (SOMs) and multilayer perceptrons trained with the BPNs.

**Self-Organizing Maps**

The main strength of SOMs lies in the possibilities they provide for modeling and analysis of complex experimental vectorial data such as process states, where even nonlinear relationships between variables may exist. The SOM is a nonlinear projection method going from a high-dimensional input space onto a lower-dimensional (often 2-D) regular lattice of neurons. The mapping is topology preserving; feature vectors close to each other in input space will be repre-
The main strength of SOMs lies in the possibilities they provide for modeling and analysis of complex experimental vectorial data.

sent as close to each other in the SOM [3]. The weights of the processing elements after training are indicative of how the data are represented in a “feature space,” and, in this particular application, can give information on the relative importance of the various variables for disorder detection.

The total set of trend data available comprised recordings from 58 patients. Of these, the data of 39 patients were selected (randomly) to be used for research and examination of the performance of “automatic disorder detection” algorithms to be developed, the data of the 19 remaining patients was kept separate to serve as a “final” test set for ensuing experiments.

The feature vectors were preprocessed before applying them as input to the SOM, so as to come to a standard range of values, in order to make the design of the patient-state assessment easier. Each variable was scaled into the range [-1, 1] by a scaling function that is continuous and piece-wise linear. This scaling function was formulated by clinicians on the basis of their experience of normal and abnormal ranges of recorded data [4]; from a clinical perspective, a function value of 1 indicates an extremely low value, 0 coincides with a regular normal value, and 1 is considered extremely high.

The experiments were performed with SOM_PAK software available from Helsinki University of Technology [5]. A range of 18 different networks, varying in sizes from 10x14 to 16x20, using a merged training set of 28039 feature vectors, (39 patients recorded for 24 hours at 1 sample per 2 minutes), and different values for learning rate, neighborhood and number of training iterations (in the order of 10^6). Training of each network configuration was repeated 10 times using a different subset of 35 patients each time (according to a 10-fold cross-validation scheme).

**Back-Propagation Networks**

A more appropriate technique for trying to maximize classification performance can be found by the use of BPNs. These ANNs use a supervised training method to optimize the weights between nodes of a predetermined ANN topology.

Initially, these internodal weights are assigned randomly, but through the iterative presentation of input patterns, with known outputs, the error between the ANN output and the actual output can be minimized [6]. During the training process, the weights associated with inputs that have the greatest influence on determining an outcome will increase in value. In contrast, inputs with little influence on an outcome will tend toward smaller values of internodal weight. By interrogating a suitably trained ANN to discover the path of largest weights for

3. Sammon mapping representation of weight vectors of the same SOM as depicted in Fig. 2. Each labeled dot represents the projection of one (20-dimensional) weight vector onto the 2-D plane. The relative distances between the projected weight vectors approximate the relative distances between the weight vectors in the 20-dimensional feature space. In this way, clusters in “weight space” (which approximates the actual “observations space” for an adequately trained SOM) can be visualized. Connecting lines indicate that nodes are topological “neighbors” in the network.
each outcome, key input variables for each condition can be identified.

A subset of the IMPROVE data set, with 23 patients, was used for this part of the study. The same 20 clinical variables that were used in the SOMs were used as network inputs. Six patient states were considered in these tests, CF, CF1, CF2, CF3 (each being a "substate" of CF in decreasing order of severity), HBFS, aO2_4c (oxygen-content-related problems) and NHC. These conditions were represented by binary values in the output of the ANNs. These data were randomly divided into BPN training and test sets.

A range of 38 different ANNs was developed. These networks varied in hidden-layer architecture and transfer functions used (linear, sigmoid, tanh). All networks were trained for 110,000 cycles, with testing every 100 cycles, during the last 10,000 iterations. For each ANN topology, the best performing network, in terms of RMS error, was further investigated with respect to its clinical predictive capability. The sensitivity, specificity, and positive and negative predictive values were calculated for each architecture on the test sets. The overall "best performing" network, from a clinical perspective, was then explored to discover the key clinical variables contributing to the decision process.

This analysis was based on the magnitude of weights on BPN interconnections. By tracing back paths from each output (patient state) to the inputs (clinical variable) along the modulus of the major weight links, and subsequently calculating the sum of the associated weights, an assessment could be made of the relative importance of each variable. For example, consider the simplified, fully trained three-layer ANN shown in Fig. 1. The output from Node 8 is being investigated. The link between the hidden layer and Node 8 with the largest modulus weight is from Node 6; this, in turn, has the largest modulus weights from input Nodes 1 and 3. These input parameters at these nodes, therefore, contribute the most to diagnosing the condition at Node 8. The same technique could then be applied to all output nodes to identify key patient variables for all states.

For a subjective comparison with the SOM results, the CF results were combined and the oxygen-content-state results were not included in this study.

Results
Self-Organizing Maps
Interpretation of the results is typically done by examining the network weights obtained. The contents of the high-dimensional (equal to the dimension of the input vectors) weight vectors of the SOM can be visualized by illustrations that contain “component planes.” These pictures show the relative distributions of the weight components associated with the measured variables. A dark color indicates a low value, whereas a light color indicates a high value (e.g., Fig. 2). The

4. Comparison of physicians, SOM, and BPN for CF. A subjective “weight” estimating the impact of the various numerical data on the ability to diagnose a particular disorder was applied to each variable by physicians from three different ICUs. These weights ranged from 0 (no impact) to 3 (very high impact), and they are present in the first three series of data points of Fig. 4. These comparisons can be related to the results from our ANN methods that are shown in the last two series of data points.

5. Comparison of physicians, SOM, and BPN for HBFS. The chart is similar to the one presented in Fig. 4, but here the investigated disorder is HBFS.
component planes show the correlation between different measurements and patient states. By examining a large number of component planes from the different networks trained, it was found that especially Tp, CI, PCWP, SVR, pH, and BE play an important role in indicating where in the feature space the disorders are located. From Fig. 2(a), for example, it can be seen that the area with node representing a low Tp matches the "CF area" on the right-hand side of the network. The area with node representing a high Tp matches the "PF area" on the left-hand side of the network. Conversely, HR (Fig. 2(c)) does not show a clear coincidence with any of the "patient state areas." It can also be noted that nodes with CF are present inbetween areas with NHC. Although different SOMs were trained using a wide range of parameter variations, as described above, similar patient state "groupings" were obtained from all networks trained.

Another way to visualize the weights obtained by training the SOM can be done by using a nonlinear mapping technique, called Sammon mapping [7]. This technique can be used to map the 2D-dimensional weight vectors to points on a 2-D plane, whereby the distances between image vectors approximate the Euclidean distances between the weight vectors. Since the distribution of weight vectors in a suitably trained SOM resembles that of the input vectors, in this application a Sammon mapping can reveal information about the distribution of patient states among the weight vectors and their relative distances in feature space. It can be seen from Fig. 3 that HBF is well separated from weights representing other patient states (HBFS is mainly present in the lower part of Fig. 3), but that separation between CF and NHC is likely to be much more difficult (a "cloud" of intertwined CF and NHC states in the center of Fig. 3).

Back-Propagation Networks

For the experiments with the BPNs, the results are very promising, with correct classification rates higher than 90% for all patient states. Results from experiments with parameters mentioned above are presented in Table 2. In clinical terms, the overall best-performing network had a 20-15-13-15 topology using a tanh transfer function.

Discussion

The IMPROVE data library is a unique library in the sense that it is fully annotated. This integration has allowed us to gain significant insight into how clinical data can be combined for the best effectiveness. We have reported in this article a novel method for objective clinical decision-making, which we show to be an improvement over traditional methods that rely on the relative experience of the physician in attendance.

Although different SOMs were trained using a wide range of parameter variations, as described above, similar patient state "groupings" were obtained from all networks trained.

Another way to visualize the weights obtained by training the SOM can be done by using a nonlinear mapping technique, called Sammon mapping [7]. This technique can be used to map the 2D-dimensional weight vectors to points on a 2-D plane, whereby the distances between image vectors approximate the Euclidean distances between the weight vectors. Since the distribution of weight vectors in a suitably trained SOM resembles that of the input vectors, in this application a Sammon mapping can reveal information about the distribution of patient states among the weight vectors and their relative distances in feature space. It can be seen from Fig. 3 that HBF is well separated from weights representing other patient states (HBFS is mainly present in the lower part of Fig. 3), but that separation between CF and NHC is likely to be much more difficult (a "cloud" of intertwined CF and NHC states in the center of Fig. 3).

Table 2. Back-Propagation Network Classification. Patient states studied are given in the columns (CFx indicating substates of CF), different performance assessments in the rows.

<table>
<thead>
<tr>
<th></th>
<th>NHC</th>
<th>LCF1</th>
<th>LCF2</th>
<th>LCF3</th>
<th>HBFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensitivity</td>
<td>99.72%</td>
<td>99.78%</td>
<td>98.18%</td>
<td>98.29%</td>
<td>100.00%</td>
</tr>
<tr>
<td>specificity</td>
<td>99.46%</td>
<td>99.98%</td>
<td>99.89%</td>
<td>99.33%</td>
<td>100.00%</td>
</tr>
<tr>
<td>pos. pred. value</td>
<td>99.53%</td>
<td>99.89%</td>
<td>97.30%</td>
<td>98.61%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

This integration has allowed us to gain significant insight into how clinical data can be combined for the best effectiveness.

Caution must be employed in the interpretation of these findings, as the clinical definition employed evolved from the diagnosis "sepsis" to the disorder "HBFS." These are not the same entities, but nevertheless are sufficiently related to draw some tentative conclusions. Here again, similarities between physicians and ANNs exist with respect to the importance of variables such as SVRI and Tc. However, there are many other variables that are considered important in the ANN models (especially the BPN representation) but are not important according to the physicians (e.g., ResF, HR, PFI).

It is important to note that while we have been able to gain some insight into the classification of patient states, we are not suggesting that ANNs replace physicians in the clinical setting. Rather, we believe that ANNs can be used as an aid to physicians in making decisions about patient care. This has important implications for the future of clinical decision-making, as ANNs have the potential to significantly improve the accuracy and speed of clinical diagnosis.
Other features will include the application of further model-based approaches: a neuro-fuzzy approach and principal factor analysis will allow a more quantitative description of the comparisions and show potential to replace the subjectiveness of the “weights” given to the clinical variables shown in Figs. 4 and 5.

This innovative model-based approach to determine the most important clinical variables for the stated disorders has the potential to make significant scientific and patient-care impact. It points a way toward more objective clinical decision making of benefit to the care process. However, we are not advocating that this type of method be used as a replacement for the physician, but rather as an aid to interpretation of clinical disorders. Indeed, the purpose of this study was to investigate the appropriateness of data sets for clinical decision-making by looking at the internal “weights” assigned to SOMs and BPNs. This approach provides for an overall classification of clinical disorders. Inclusion of time representation will allow this purpose to be extended to prediction of future patient states, giving knowledge on patient-state trajectory (e.g., [8]), and giving the physician an opportunity to react before a more morbid state is reached.

Conclusions

On the basis of investigation of the fully annotated IMPROVE data library, there seems to be a mismatch between actual annotations by the physician regarding onset of clinical disorders in the 58 patients and periods in the data library defined by application of their clinical definitions. The fact that the physicians are perfectly correct in some cases to disregard the clinical rules confirms that medicine is indeed an art rather than a science. ANNs in the form of SOMs and BPNs have conclusively demonstrated their utility and potential as an aid in clinical decision making. This article presented some interesting results, which have extended knowledge and understanding about the value of combining certain variables that physicians would not normally consider. This result can serve as a basis for development of biosignal interpretation (BSI) methods, which aim to detect disorders as early as possible by using multivariable information.

Acknowledgments

The authors would like to thank Dr. Aarno Kari and the medical and nursing staff of the intensive-care unit at Kuopio University Hospital as well as the annotators who shared responsibility for work on the data library.

Mark van Gils was born in Bergen op Zoom, The Netherlands, in 1964. He received his M.Sc. degree in Applied Physics from the Technical University of Eindhoven in 1990. In 1995 he obtained a Ph.D. degree from the same university for research on the use of artificial neural networks for processing auditory-evoked potentials. He worked as a visiting scientist at the NIH/NIA Gerontology Research Center in Baltimore, MD, in 1995. He is currently employed as research scientist at VTT Information Technology, Tampere, Finland, as well as a lecturer in biomedical signal processing at Helsinki University of Technology. His research interests include processing, analyzing, and visualizing information present in biomedical signals.

Kari Nieminen was born in 1964 in Espoo, Finland. He has been working as a research scientist since 1995 and as an intern training clinician since 1996 in the Department of Anaesthesiology and Intensive Care in Kuopio University Hospital.

Ron Summers was born in Billingham, Cleveland, UK, in 1957. He is currently a Senior Lecturer in Systems Science at City University, London, UK. He obtained a B.Sc. in Biophysical Science at the North East London Polytechnic (now University of East London) in 1982. His Masters degree in Information Engineering was taken in 1985 at City University, where his dissertation focused on the development of an expert system for the diagnosis of jaundice. Always interested in matters biomedical, he completed his Ph.D. in 1992 for which the theme was the design, implementation, and evaluation of an intelligent knowledge-based system for management of patients who required ventilatory therapy. He is an active member of both IMEKO and IFAC international committees and is committed to the development and use of medical informatics as a provider for enhanced patient care.

Peter Weller was born in Sussex, England, in 1959. He has a background in designing medical equipment. He received a Mathematics degree from the Open University in 1988 and a Masters degree in Medical Informatics from City University in 1994. He is currently working as a Research Assistant at City University and working toward a Ph.D. in the use of artificial neural networks for the monitoring and control of a complex, critical system. His research interests include the applications of artificial neural networks, genetic algorithms, and fuzzy systems in modeling and controlling complex systems.

Address for Correspondence: Mark van Gils, Ph.D., VTT Information Technology, P.O. Box 1206, FIN-33101 Tampere, Finland.

References