DECISION SUPPORT SYSTEMS BASED ON ARTIFICIAL NEURAL NETWORKS IN THE DIAGNOSTICS AND TREATMENT OF ACUTE PANCREATITIS: LITERATURE REVIEW

A. A. Litvin¹, O. Y. Rebrova²

¹Gomel Regional Hospital, Gomel State Medical University, Gomel, Belarus
²Russian National Research Medical University n. a. N. I. Pirogov, Moscow, Russia

Key words: decision support system, artificial neural networks, diagnostics and treatment, severe acute pancreatitis, infected pancreatic necrosis

It is known that upon making medical decisions problems are lack of knowledge, lack of time resources, lack of ability to attract a large number of competent experts, incomplete information about the state of the patient and others [13, 17]. Considerable percentage of medical errors is being preserved currently, followed by a further loss of the patient’s health or death due to an error in the assessment of his/her condition, complicating the process of making the right decisions in view of the multiplicity of factors and symptoms of diseases and their interactions. The condition of a patient may be affected by certain factors not only individually, but also in certain combinations. Clinician often can’t grasp a large amount of information and identify regularities. Only modern computer techniques with appropriate software can cope with it [13]. At present, there is a need for more extensive use of various methods of computer decision support [6, 17]. Decision support systems (DSS) allow the doctors not only to check their own prognostic and diagnostic assumptions, but use artificial intelligence in complex clinical situations [7].

In emergency surgery when making medical decisions lack of time, insufficient information about the history and clinical manifestations of the disease, high dynamics of the disease, high price of medical errors are even more common. Surgeon on the basis of the analysis of a huge number of factors, based on the knowledge and previous experience, must often quickly decide on the optimal treatment strategy, the process of operational manuals and completing the operation
Acute pancreatitis (AP) among emergency abdominal diseases in frequency of hospitalization is at the second place (after acute appendicitis), while the development of pancreatic necrosis leads to the greatest number of adverse outcomes [12]. Therefore, the actual problem is the computer support decision-making in the diagnosis and treatment of AP with the development and improvement of information technology.

Systems that use software implementations of artificial neural networks are the example of DSS [1, 4, 14]. Artificial neural networks (ANN) are mathematical models that allow usually somewhat better classify objects than stochastic models [2, 9, 11]. A method of ANN constructing is based on certain principles of information processing in biological systems. The basis of each ANN is relatively simple, in most cases, the same elements that mimic the work of brain neurons. Each artificial neuron is characterized by its current state, by analogy with the nerve cells of the brain, which can be excited or inhibited. Artificial neuron has a group of "synapses" — unidirectional input links connected to the outputs of other artificial neurons, and also has an axon — output relationship of the neuron from which the signal (excitation or inhibition) arrives at synapses of the next artificial neurons [2]. For ANN principle of parallel processing of the signals is typical, which is achieved by combining a large number of neurons in the so called certain layers and compounds of neurons of the different layers. The strength of synaptic connections is modified in the process of extracting knowledge from the training data set (training mode), and then it is upon getting the results of the new data (performance mode) [9]. Neural networks allow you to make decisions on the basis that they identify hidden patterns in multidimensional data.

Neural networks have been used in many fields of technology, where they solve numerous applications [3, 9]. Examples of the use of computer prediction based on the ANN technology in medicine are also numerous enough, including the fields of medical and surgical gastroenterology. Thus, the technology of intelligent data analysis is applied by N. Horowitz et al. (2007) in the development of diagnostic profiles to detect gastroesophageal reflux disease [15]. 132 patients were examined,
computer model was constructed on the basis of the data, that allowed to identify the informative complex of signs of the disease: heartburn, regurgitation, the positive effect of antacid therapy and deterioration of health after acute fatty foods. The sensitivity and specificity of this method amounted to 75% and 78%. Y. C. Lee et al. (2007) used data analysis techniques to predict weight loss after bariatric surgery [24]. Data of 249 patients, who had been operated using two alternative methods (bypass or banding), were retrospectively analyzed: 177 women and 72 men. Two hundred-eight patients (83.5%) within two years after the operation successfully reduced the body weight, while 41 patients (16.5%) had no significant effect. Created by the authors predictive computer model based on ANN allows even before the operation to identify whether the expected effect would be after it (P=0.000). High preoperative levels of triglycerides, low levels of HbA1c and bypass surgery predict successful weight loss within two years after surgery [24].

DSS have recently become widespread in pancreatology [16, 18, 27], including the diagnostics and treatment of AP [8, 10, 19, 20, 21, 22, 23, 25, 26]. The first attempt to predict the severity of the AP based on ANN was taken by S. C. Kazmierczak et al. (1993), by analyzing the activity of pancreatic enzymes in blood serum [20]. Pancreatic lipase was the best predictor of severe AP, its accuracy was 82% (95% CI: 77-87). Prediction accuracy when using values of serum amylase was quite low — 76% (71-81), as the combination of such indicators as lipase and amylase did not allow to significantly increase the accuracy of the created ANN — predictive accuracy was 84% (79-89).

W. E. Pofahl et al. (1998) used ANN to predict the length of stay of patients with AP in the hospital [25]. In their study, they compared the ANN with Ranson, APACHE II systems criteria at sensitivity, specificity, positive predictive value and accuracy. The input features of neural network (input neurons) were 71 indices of clinical, laboratory and instrumental methods of examination, estimated in 156 randomly selected patients on admission to hospital. The results at the output of ANN consisted of indicator "length of stay" with two possible values ">7 days" (heavy AP) and "<7 days" (not heavy AP). The trained ANN was tested in 39 patients with AP.
The authors provide information about high sensitivity (75%), specificity (81%) and accuracy (79%) of ANN in determining the severity of AP, detectable by the predictable length of hospital stay [25].

In the study by M. T. Keogan et al. (2002) neural network model was used to predict the duration of the treatment of AP longer than average and being 8.4 days [21]. Since a large number (23) of the input variables towards relatively small number of patients (n=92) posed a risk for the so called ANN retraining, the number of variables was reduced with the help of step discriminant analysis from 23 to 6: fluid deficit; serum creatinine level; presence of comorbidity; blood pressure; degree of inflammation by CT; calcium levels in blood serum. The above data were subjected to neural network and linear discriminant analysis. ANN showed the best predictive accuracy (AUC=0,83±0,05) compared with the scales of Ranson (AUC=0,68±0,06; P<0,02) and Balthazar (AUC=0,62±0,06; P<0,003). However, as compared with the linear discriminant function, significant differences in the prognosis of duration of AP treatment weren’t identified (AUC=0,82±0,05; P=0,53). The disadvantage of this ANN model and the study at all was the fact that it had been trained and tested on the same data set, so the results should be verified in other studies.

K. I. Halonen et al. (2003) on the basis of retrospective analysis of case histories of 234 patients with severe AP developed two forecasting models to predict the possible death in necrotizing pancreatitis [22]. In the first model, called «LR4», death was predicted by logistic regression analysis based on four indicators — age, presence of cardiovascular disease, respiratory failure, and the highest level of blood creatinine within 72 hours after admission. The second model based on ANN («ANN8») in addition to the above-mentioned variables included four ones: need for vasopressor support, gender, body mass index and the lowest level of hemoglobin within the first 72 hours after admission. The predictive accuracy of the different models was compared using ROC-analysis. The greatest predictive accuracy was shown by «LR4» (AUC=0,862) and «ANN8» (AUC=0,847). The rest of the scales showed the following results in the prediction of mortality in severe AP: the scale of
Glasgow — AUC=0.536, Ranson — AUC=0.655, MODS — AUC=0.781, APACHE II — AUC=0.817. A simpler model «LR4», which consists of four conventional indices, had the highest accuracy [22].

In 2007, R. Mofidi et al. developed a neural network model for classifying the severity of AP, predicting organ failure and death [19]. To do this, the authors conducted a retrospective analysis of 664 case histories of patients with AP, including 181 patients with severe AP (27.3%). For ANN developing multi-layer perceptron (four layers, two of them — invisible intermediate layers) with error backpropagation was used. 60% of the data were made available for ANN training, 15% — testing, 25% — validation. Dimension feature space had been previously reduced from 29 to 10 characters. As a result, designed ANN was based on 10 clinical parameters (age, presence of hypotension, two or more symptoms of SIRS, PaO₂, LDH, glucose, urea, calcium, hematocrit levels, blood leukocytes), determined upon admission and in 48 hours. This model showed significantly better results than the system APACHE II and Glasgow: ANN was more accurate than APACHE II and Glasgow in predicting severe AP (P<0.05 and P<0.01, respectively), predicting the development of multiple organ insufficiency (P<0.05 and P<0.01) and in the prediction of death (P<0.05). This work differs from the above-mentioned by inclusion of a large number of patients (n=664), as well as the fact that the training and validation were carried out on different groups of patients. No less important advantage is that all 10 input variables are available for the doctor on duty within the first 6 hours after admission.

B. Andersson et al. (2011) conducted a study which purpose was to develop and test the effectiveness of ANN model for early prediction of AP severity [23]. The authors conducted a retrospective analysis of the results of treatment of 208 patients with AP (from 2002 to 2005, n=139, from 2007 to 2009, n=69). The severity of AP was determined in accordance with the criteria suggested by the conference on AP in Atlanta. Of the 23 potential indicators of AP severity, authors selected six the most informative criteria using ANN: duration of pain attack, blood creatinine, hemoglobin, alanine aminotransferase, heart rate and blood leukocytes. The area
under the ROC-curve in the neural network model was 0.92 (95% CI: 0.85-0.99), 0.84 (0.76-0.92) — in logistic regression (P=0.030, χ²), and 0.63 (0.50-0.76) — in assessing the severity of AP with the use of APACHE II (P<0.001, χ²). The authors concluded that developed ANN on the basis of the data obtained upon admission to hospital data was accurate enough to predict the severity of AP [23].

P. I. Mironov et al. (2011) also assessed the possibility of ANN in determining the severity of state and predicting outcomes AP [10]. The study included 100 patients with severe AP, who were under the supervision of the authors from 2004 to 2010. To construct the ANN, they selected 33 parameters of 5 categories (demographics, physiological variables, laboratory tests, temporary variables, patient outcomes). Subsequently, the number of input data was reduced by stepwise logistic regression analysis to 6. They used the standard three-layer perceptrons with a compound of each of the hidden and output neurons with all the elements of the previous layer. Neural networks were trained in all cases from the database using the algorithm of backpropagation. The authors concluded that the predictive ability of ANN (AUC=0.83±0.04) in the early identification of groups of patients, threatened by the development of heavy AP, significantly (P<0.001, χ²) exceeds the capabilities of evaluation systems of Ranson (AUC=0.55±0.06), Glasgow (AUC=0.58±0.06), TPS (AUC=0.53±0.06), APACHE II (AUC=0.58±0.06) and USD/CT criteria (AUC=0.68±0.06). When predicting the risk of death, the authors showed that the area under the ROC-curve for ANN was equal to 0.83±0.04, scales of Ranson — 0.55±0.06, Glasgow — 0.58±0.06, TFS — 0.53±0.06, APACHE II — 0.58±0.06, SOFA — 0.72±0.05 and USD/CT criteria — 0.68±0.06. ANN predictive value in determining the risk of death was higher than all other systems (P<0.001, for the scales of Ranson, Glasgow, TFS, APACHE II), (P=0.033 for USD/CT criteria), (P=0.046 for the scale SOFA) [10].

A. A. Litvin et al. (2012) developed a system of prediction of infected pancreatic necrosis on the basis of survey data obtained in the first day after admission of patients with AP to the hospital [8]. We used retrospective clinical, laboratory and instrumental data of 398 patients who were treated from 1995 to 2005.
The sample of 398 objects was divided randomly into two groups: the first group of objects was used to train ANN (298 cases); objects of the second group were used for testing the trained ANN (100 examples). In order to determine the validity of the developed system of forecasting, we performed an analysis of outcomes in the examination sample of 128 patients with severe AP observed in 2006-2008. With the help of genetic algorithm for feature selection, the authors highlighted 12 most informative features of possible development of infectious complications of pancreatic necrosis:

1) time from onset of disease to hospitalization;
2) “early” operations in the history of the patients;
3) severe pain syndrome (relieved by narcotic analgesics);
4) body mass index;
5) heart rate;
6) respiratory rate;
7) sharp fluid accumulations (by USD and CT) or palpable infiltration;
8) bloating (on clinical data);
9) number of stab forms of white blood cells;
10) glucose;
11) urea levels of serum;
12) effect of a complex intensive therapy within 24 hours of hospitalization.

The developed computer program showed relatively high predictive ability with respect to infected pancreatic necrosis in a sample of patients with severe AP: sensitivity — 85.5% (95% CI: 73.3-93.5), specificity — 91.8% (83.0-96.9). This system was more accurate than comparable scale determining the risk of infectious complications of severe AP: differences with M-APACHE II and scale of D. A. Taganovich were found with P=0.005, Z-test, with scale of S. I. Tretiak et al. — P=0.003, Z-test.

The same authors also developed a system of diagnostics of infected pancreatic necrosis, used in the process of follow-up of patients with severe AP to identify transition of sterile pancreatic necrosis in the infected one. With ANN, 14 features
were identified that, when used as the ANN input parameters, allow with a high probability to diagnose infectious pancreatic necrosis complications. Some of informative indicators included traditional clinical research ("time of onset", "early operation in history", "AP-patient treatment history", "body mass index", "body temperature", "heart rate", "frequency of breath", "bloating", "effect of medical treatment within 24 hours"); laboratory parameters ("white blood cells", "ESR", "number of stab forms of white blood cells," "blood glucose") and instrumental data ("presence of infiltration or liquid clusters"). ANN showed very good quality of diagnostic model in recognition of infected pancreatic necrosis — the area under the ROC-curve was 0.854 (95% CI: 0.791-0.917). ANN in the study of sample of patients demonstrated a sensitivity of 81.8% (75.3-88.3), specificity — 89.0% (83.5-94.5). In a comparative study of the diagnostic capabilities of ANN, M-APACHE II, Infection Probability Score and SIRS 3-4 statistically significant differences were revealed — P=0.03, P=0.001 and P=0.005 (Z-test), respectively [8].

Hong Wan-dong et al. (2013) developed DSS based on ANN for predicting persistent (more than 48 hours) organ insufficiency in patients with AP [26]. The sample included 312 patients with AP, as a result of the factor analysis, 13 of the most informative were selected in the first days of the date of admission. The final three-layer ANN (5-5-1) with five input parameters: age, hematocrit, blood glucose, urea, calcium, was trained on 312 examples. The sensitivity of the trained network was 81.3%, specificity — 98.9%, accuracy — 96.2%. The predictive accuracy of established ANN (AUC=0.96±0.02) was significantly better than the model constructed on the basis of logistic regression (AUC=0.88±0.03, P<0.001, $\chi^2$) and APACHE II (AUC=0.83±0.03, P<0.001, $\chi^2$). The authors conclude that the established ANN can be useful for predicting the development of persistent organ insufficiency in patients with AP [26].

The following table summarizes the results of the development and use of ANN in AP [8, 10, 19, 20, 21, 22, 23, 25, 26].

Table 1

**Results of using the ANN with the aim of diagnostics and prognosis in AP**
<table>
<thead>
<tr>
<th>Authors</th>
<th>The number of observations in the training of ANN/number of observations during testing the ANN</th>
<th>Objectives of the study</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazmierczak et al. (1993) [20]</td>
<td>254/254</td>
<td>AP diagnostics by pancreatic enzymes in serum</td>
<td>The level of lipase has the highest diagnostic accuracy</td>
</tr>
<tr>
<td>Pofahl et al. (1998) [25]</td>
<td>156/39</td>
<td>Prediction of duration of hospital stay</td>
<td>The prognostic value of ANN did not differ from Ranson, APACHE II</td>
</tr>
<tr>
<td>Keogan et al. (2002) [21]</td>
<td>92/92</td>
<td>Prediction of length of stay in hospital by CT and laboratory studies</td>
<td>The prognostic value of ANN did not differ from Ranson, APACHE II</td>
</tr>
<tr>
<td>Halonen et al. (2003) [22]</td>
<td>234/60</td>
<td>Prediction of death</td>
<td>The prognostic value of ANN did not differ from Ranson, APACHE II, Glasgow</td>
</tr>
<tr>
<td>Mofidi et al. (2007) [19]</td>
<td>496/166</td>
<td>Identification and prediction of heavy AP and death</td>
<td>ANN proved to be more accurate in predicting the severity of AP</td>
</tr>
<tr>
<td>Andersson et al. (2011) [23]</td>
<td>139/69</td>
<td>Prediction of severe AP</td>
<td>ANN proved to be more accurate in predicting heavy AP as compared to logistic regression, APACHE II</td>
</tr>
<tr>
<td>Mironov et al. (2011) [10]</td>
<td>100/100</td>
<td>Identification of heavy AP and prediction of its outcomes</td>
<td>Predictive ability of ANN in early identification of severe AP and risk of death exceeds the system Ranson, Glasgow, TFS, APACHE II, Balthazar</td>
</tr>
<tr>
<td>Litvin et al. (2012) [8]</td>
<td>298/100</td>
<td>Prediction of infected pancreatic necrosis, identification of AP infectious complications</td>
<td>ANN proved to be more accurate in the prediction and diagnostics of infected pancreatic necrosis as compared with other multiparameter scales</td>
</tr>
<tr>
<td>Wan-dong et al. (2013) [26]</td>
<td>312/312</td>
<td>Prediction of persistent (more than 48 hours) organ insufficiency</td>
<td>ANN proved to be more accurate in the prediction of persistent organ insufficiency as compared to logistic regression, APACHE II</td>
</tr>
</tbody>
</table>

Thus, as it’s shown in the table, ANN has been used for diagnosing and treating patients with AP during the last 20 years. First ANN was developed to determine the information content of laboratory parameters in the diagnostics of severe AP [20]. Subsequently, the main area of research was the prediction of disease severity at the beginning of AP with noting the patients who subsequently will have mild or severe AP. In this case, the definition of AP severity was first held indirectly through the length of stay of patients in hospital [21, 25], the probability of death [22]. Then the researchers used the AP classification "Atlanta 1992" with predicting primarily severe AP [10, 19, 23]. ANN was also used to predict the possible development of AP infectious complications for the purpose of timely prevention and
for the diagnosis of transfer of the sterile pancreatic necrosis in the infected one with the aim of correct treatment [8].

In recent years there has been an improvement in ANN prediction accuracy upon AP as compared to conventional scales determining AP severity (Ranson, Glasgow), severity of the physiological state of patients with AP (APACHE II). If the first publications on the use of ANN in predicting severe AP artificial neural networks did not show significant differences with traditional scales [20, 21, 22, 25], in all subsequent studies authors noted prognostic superiority of ANN [8, 10, 19, 23, 26].

However, all of the studies on the use of ANN for predicting AP have certain disadvantages and limitations. First, the data from all studies were evaluated retrospectively, which could lead to a known offset in the results. Second, the data were obtained in hospitals of different levels, which casted doubt on the reproducibility of methods based on data from other clinics. Third, the sample size in most studies was insufficient for training and testing of ANN, not in all the studies ANN training and testing was carried out on different samples of patients with AP. Finally, the development of the problem of AP strongly interfaced with the advent of various confounding factors (confounders), mainly related to the multicausality of severe AP development, expressed heterogeneity of samples of patients with necrotizing pancreatitis. It is extremely difficult to assess the impact of these confounding factors on the final result.

Thus, the development and implementation of DSS based on the use of ANN is a promising direction of improving the prediction, diagnostics of severe AP and its complications. However, we need further improvement of ANN with the shortcomings of previous studies, the development of a simple, accurate and close to the surgeon’s workplace decision support systems in the problem of AP.
References


Decision support systems based on artificial neural networks in the diagnostics and treatment of acute pancreatitis: literature review

A. A. Litvin¹, O. Y. Rebrova²

¹Gomel Regional Hospital, Gomel State Medical University, Gomel, Belarus
²Russian National Research Medical University n. a. N. I. Pirogov, Moscow, Russia

Key words: decision support system, artificial neural networks, diagnostics and treatment, severe acute pancreatitis, infected pancreatic necrosis

This paper presents a review of the literature on the use of artificial neural networks in the diagnostics and treatment of acute pancreatitis. The authors provide modern literature data on the efficacy of decision support systems based on artificial neural networks to determine the severity of acute pancreatitis outcome, prognosis and diagnostics of infected pancreatic necrosis.