REVIEW

Open Access

WSES project on decision support systems based on artificial neural networks in emergency surgery

Andrey Litvin[1]* , Sergey Korenev[1], Sophiya Rumovskaya[2], Massimo Sartelli[3], Gianluca Baiocchi[4], Walter L. Biffl[5], Federico Coccolini[6], Salomone Di Saverio[7], Michael Denis Kelly[8], Yoram Kluger[9], Ari Leppäniemi[10], Michael Sugrue[11] and Fausto Catena[12]

Abstract

The article is a scoping review of the literature on the use of decision support systems based on artificial neural networks in emergency surgery. The authors present modern literature data on the effectiveness of artificial neural networks for predicting, diagnosing and treating abdominal emergency conditions: acute appendicitis, acute pancreatitis, acute cholecystitis, perforated gastric or duodenal ulcer, acute intestinal obstruction, and strangulated hernia. The intelligent systems developed at present allow a surgeon in an emergency setting, not only to check his own diagnostic and prognostic assumptions, but also to use artificial intelligence in complex urgent clinical cases. The authors summarize the main limitations for the implementation of artificial neural networks in surgery and medicine in general. These limitations are the lack of transparency in the decision-making process; insufficient quality educational medical data; lack of qualified personnel; high cost of projects; and the complexity of secure storage of medical information data. The development and implementation of decision support systems based on artificial neural networks is a promising direction for improving the forecasting, diagnosis and treatment of emergency surgical diseases and their complications.

Keywords: Decision support system, Artificial neural networks, Emergency surgery, Acute appendicitis, Acute pancreatitis, Acute cholecystitis, Bowel obstruction, Perforated gastroduodenal ulcers, Peptic ulcer bleeding, Strangulated hernias

Introduction

Currently, computer decision support systems (DSS) based on algorithms developed with the use of the methods of intellectual analysis of medical data are receiving special development [1–3]. It is known [3] that when making medical decisions, there are such problems as lack of knowledge, limited time resources, inability to attract a large number of competent experts, incomplete information about the patient's condition, etc. There is an urgent need to use different methods of computer decision support in emergency surgery wider [4]. Decision support systems can enable a surgeon to check his own prognostic and diagnostic assumptions in emergency and to use artificial intelligence in complex urgent clinical cases [5, 6]. DSS frequently are developed based on artificial neural networks (ANNs) [7–9].

ANNs are mathematical models that allow, as a rule, the classification of objects somewhat better than probabilistic (stochastic) models [10]. ANNs are based on some principles of information processing in biological systems. ANN consists, in most cases, of relatively simple,

*Correspondence: alitvin@kantiana.ru
[1] Department of Surgical Disciplines, Immanuel Kant Baltic Federal University, Kaliningrad, Russia
Full list of author information is available at the end of the article

Litvin *et al. World J Emerg Surg* (2021) 16:50

Page 2 of 9

elements that imitate the function of brain neurons. Each artificial neuron is analogous with nerve cells in the brain, which can be fired or inhibited. The artificial neuron receives one or more inputs (representing excitatory postsynaptic potentials and inhibitory postsynaptic potentials at neural dendrites) and sums them to produce an output (or activation, representing a neuron's action potential which is transmitted along its axon) [11]. ANNs are characterized by the principle of parallel signal processing, which is achieved by combining a large number of neurons into so-called layers. Neurons of different layers are connected in a specific way. The strength of synaptic links is modified in the process of extracting knowledge from the training dataset (training mode), and then used when obtaining a result on new data (execution mode) [10]. Neural networks make decisions based on hidden patterns they reveal in multidimensional data.

Neural networks are self-learning in distinction to classical machine learning. There is a deep learning in neural networks, which is the most popular and high-demand because an amount of data and computing power are constantly growing [12]. It is considered that the complexity of the neural network should correspond to the complexity of the phenomenon under study. Moreover, in many cases, very simple neural networks with two layers are even more accurate than very deep networks with many layers. The reasons for this [13] may lie in the insufficient amount of data, in the excessive complexity of the network in comparison with the phenomenon under study, etc. In general, the trend is that neural networks are increasingly replacing classical machine learning, and deep learning prevails over simple neural networks with a small number of layers [12].

Neural networks have found applications in many fields of medicine for solving multiple applied problems [14–18]. Examples of the use of computer prediction based on ANN technology in surgical specialties are also quite numerous [19–21].

The aim of the article is to scoping review the literature on the effectiveness of ANN-based DSS for the diagnosis, prognosis and treatment of urgent surgical diseases.

Methods

A scoping review was conducted following the guidelines of PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [22]. We assessed electronic database PubMed. The following search criteria have been used: "artificial neural networks" AND "acute appendicitis," "artificial neural networks" AND "acute pancreatitis," "artificial neural networks" AND "acute cholecystitis," "artificial neural networks" AND "bowel obstruction," "artificial neural networks" AND "perforated gastroduodenal ulcers," "artificial neural networks" AND "ulcer bleeding," "artificial neural networks" AND "strangulated hernias."

Results were reviewed based on predefined inclusion and exclusion criteria. We reviewed all peer reviewed published papers considering the use of ANNs in emergency surgery (acute appendicitis, acute pancreatitis, acute cholecystitis, ileus/bowel obstruction, perforated gastroduodenal ulcers, nonvariceal upper gastrointestinal/peptic ulcer bleeding, incarcerated/strangulated hernias), except publications older than 1990-year, case reports, editorials and commentaries.

Results

The initial search identified 61 publications in the PubMed database. Nineteen articles met the inclusion and exclusion criteria. Characteristics of each are shown in Table 1.

Acute appendicitis

Yoldaş et al. [23] investigated the diagnostic capabilities of ANN in acute appendicitis (AA) in patients with pain in the right lower abdomen. Data were collected from 156 patients with suspected acute appendicitis prospectively. The sensitivity, specificity, and positive and negative predictive values of the artificial neural network were 100%, 97.2%, 96.0% and 100%, respectively. The authors concluded that ANNs can be an effective tool for accurate diagnosis of acute appendicitis and can reduce the number of "unnecessary" appendectomies [23].

Park and Kim [24] also used ANN to diagnose AA. The data of 801 patients were used to create artificial neural networks of three types: ANN with a radial basis function, multilayer ANN and ANN with a probabilistic structure. The Alvarado clinical scoring system was used for comparison. The accuracy of the three types of ANN and Alvarado was 99.8%, 99.4%, 97.8% and 72.2%, respectively; ROC AUC (Area under the ROC (receiver operating characteristic) curve)—0.998, 0.993, 0.985 and 0.633, respectively. The ANN models proposed by the authors for the diagnosis of acute appendicitis showed better results than the Alvarado clinical assessment system ($P < 0.001$).

Reismann et al. [25] examined the informativeness of artificial intelligence (AI) for the diagnosis of the AA in childhood and adolescence. This study presents a method for the automatic diagnosis of appendicitis, as well as for identifying the differences between complicated and uncomplicated AA. It is based on generally accepted clinical data and diagnostic methods. The medical records of a total of 590 patients were retrospectively analyzed using AI (473 patients with acute appendicitis according to histological findings and 117 patients with negative

Litvin *et al. World J Emerg Surg*   (2021) 16:50

Page 3 of 9

Table 1 Results of using ANNs for diagnostics and prognosis in acute surgical diseases

Authors	The training set size	Objectives of the research	Results
Acute appendicitis (AA)			
Yoldaş et al. [23]	156	AA diagnostics	Sensitivity—100%, specificity—97.2%
Park and Kim [24]	801	AA diagnostics	The ANN proved to be more accurate in diagnosing AA (the accuracy of the three types of ANN—99.8%, 99.4%, 97.8%) than the Alvarado clinical scoring system (72.2%)
Reismann et al. [25]	590	AA diagnostics, prediction of a complicated course of the disease in pediatrics	The ANN has allowed a significant improvement of the accuracy of diagnosis (sensitivity 93%, specificity 67%), and complicated course of AA (sensitivity 95%, specificity 33%)
Park et al. [26]	667	AA diagnostics based on CT of patients with abdominal pains	The ANN showed good and very good diagnostic indicators of AA (accuracy > 90%)
Acute pancreatitis (AP)			
Kazmierczak et al. [27]	254	Diagnosis of the AP by the level of pancreatic enzymes in the blood serum	Lipase level has the highest diagnostic accuracy (accuracy lipase—82%, serum amylase—76%, lipase + amylase—84%)
Pofahl et al. [28]	156	Predicting the hospitalization length	Sensitivity 75%, specificity 81% and accuracy 79%, but the ANN predictive capabilities do not differ from Ranson and APACHE II
Keogan et al. [29]	92	Predicting the hospitalization length based on CT and laboratory tests	The ANN showed the best predictive accuracy (AUC = 0.83 ± 0.05) compared to the Ranson (AUC = 0.68 ± 0.06; $P < 0.02$) and Balthazar (AUC = 0.62 ± 0.06; $P < 0.003$)
Halonen et al. [30]	234	Predicting the potential mortality	The ANN predictive capabilities (AUC = 0.847) differ from Ranson (AUC = 0.655), APACHE II (AUC = 0.817) and Glasgow (AUC = 0.536)
Mofidi et al. [31]	496	Identification of the AP severity and predicting lethal outcome	The ANN proved to be more accurate in diagnosing of the AP (ANN was more accurate than APACHE II and Glasgow in predicting: AP severity—$P < 0.05$ and $P < 0.01$ Multiple organ failure $P < 0.05$ and $P < 0.01$ Lethal outcome—$P < 0.05$ and $P < 0.05$)
Andersson et al. [32]	139	Predicting the AP severity	The ANN proved to be more accurate (AUC = 0.92) in diagnosing of the severe AP in comparison with the logistic regression (AUC = 0.84, $P = 0.03$) and APACHE II (AUC = 0.63, $P < 0.001$)
Hong et al. [33]	312	Predicting the persistent (more than 48 h) organ failure	The ANN proved to be more accurate in predicting of the persistent organ failure (AUC = 0.96 ± 0.02) in comparison with the logistic regression (AUC = 0.88 ± 0.03, $P < 0.001$) and APACHE II (AUC = 0.83 ± 0.03, $P < 0.001$)
Fei et al. [34]	152	Predicting the severe AP associated with acute lung injury	The ANN proved to be more accurate (AUC = 0.859 ± 0.048) in predicting of the acute lung injury accompanying the AP in comparison with the logistic regression (AUC = 0.701 + 0.041)
Acute cholecystitis (AC)			
Eldar et al. [35]	180	Predicting the conversion from laparoscopic to laparotomic access in AC	The ANN demonstrated a good predictive ability to predict the conversion from laparoscopic to laparotomic approach (100% of cases respectively, 67%—prospectively) and to determine the group of patients requiring laparotomic cholecystectomy

Table 1 (continued)

Authors	The training set size	Objectives of the research	Results
Vukicevic et al. [36]	303	Prediction of choledocholithiasis in patients with gallstone disease and AC	The ANN demonstrated a good predictive ability to predict the choledocholithiasis and revealed informative clinical, laboratory, and instrumental signs (sensitivity—82.3%, specificity—94.7%, accuracy—92.2%, and AUC—0.934 in the validation set)
Ulcerative bleeding			
Rotondano et al. [37]	2380	Predicting the fatal outcome in patients with bleeding from the upper gastrointestinal tract	The predictive ability of the ANN is better than the score Rockall's one [38] (sensitivity—83.8% versus 71.4%, specificity—97.5% versus 52.0%, accuracy—96.8% versus 52.9%, and AUC—0.95 versus 0.67)
Wong et al. [39]	22,854	Identification of patients with a high risk of recurrent bleeding requiring surgical treatment and with a high risk of death	The ANN demonstrated AUC = 0.78, and accuracy—84.3%
Perforated gastroduodenal ulcers			
Søreide et al. [44]	117	Predicting the fatal outcome and determination of factors of the fatal outcome	AUC = 0.90, 0.95% CI [0.85–0.95]
Ileus/bowel obstruction			
Cheng et al. [45]	13,935 X-ray pictures	Ileus diagnostics	Sensitivity—91.4%, specificity—91.9%
Strangulated hernia			
Chen et al. [46]	762	Predicting the need for bowel resection	ANN revealed eight factors that are significantly associated with the need for bowel resection

histological findings). The accuracy of the developed automatic method for the diagnosis of the AA was 90% (sensitivity 93%, specificity 67%), and the accuracy of correct determination of the complicated course of the disease was 51% (sensitivity 95%, specificity 33%).

Park et al. [26] investigated the possibility of a neural network algorithm for the diagnosis of the AA using CT for patients with acute abdominal pain. The diagnostic characteristics of the developed model based on CT were "very good" and amounted to more than 90%.

Acute pancreatitis

The first attempt to predict the severity of acute pancreatitis using an ANN was undertaken by Kazmierczak et al. [27] by analyzing the serum concentration of pancreatic enzymes in blood serum. Pancreatic lipase was the best predictor of the severity of AP, with an accuracy of 82% (95% confidence interval (CI): 77–87). In contrast, using the serum amylase value, the prediction accuracy turned out to be quite low—76% (71–81). Meanwhile, the combination of such indicators as lipase and amylase did not increase the accuracy of the created ANN significantly— the predictive accuracy was 84% (79–89) [27].

Pofahl et al. [28] used ANN to predict the hospital length of stay (LOS) of patients with AP. In their study, they compared ANNs with the Ranson and APACHE II scoring systems in terms of its sensitivity, specificity, positive predictive value and accuracy. The authors provide information on the high sensitivity (75%), specificity (81%) and accuracy (79%) of ANN in identification of the acute pancreatitis severity as measured by LOS [28].

In research by Keogan et al. [29], a neural network model was used to predict a prolonged treatment duration for a patient with AP, longer than the average of 8.4 days. The ANN showed the best predictive accuracy (AUC$=0.83\pm0.05$) compared to the Ranson (AUC$=0.68\pm0.06$; $P<0.02$) and Balthazar (AUC$=0.62\pm0.06$; $P<0.003$). However, there were no significant differences in predicting the duration of treatment for acute pancreatitis (AUC$=0.82\pm0.05$; $P=0.53$) in comparison with the linear discriminant function. The disadvantage of their ANN model and the entire research was that it developed then tested on the same dataset, so the results will need to be verified by other studies.

Halonen et al. [30] developed two prognostic models to predict potential mortality in pancreatic necrosis [30]. The basis of the research was the retrospective analysis of medical records of 234 patients with severe AP. In the first model, mortality was predicted by the method of logistic regression, in the second, by the ANN. The predictive accuracy of the various models was compared using ROC analysis. The highest predictive accuracy was shown by logistic regression

(AUC$=0.862$) and the ANN (AUC$=0.847$). The rest of the scales demonstrated the following results in predicting mortality in severe AP: Glasgow scale— AUC$=0.536$, Ranson—AUC$=0.655$, MODS— AUC$=0.781$ and APACHE II—AUC$=0.817$ [30].

Mofidi et al. [31] developed a neural network model for classifying the severity of AP, predicting organ failure and death. For this study, the authors conducted a retrospective analysis of 664 case histories of patients with AP acute pancreatitis, including 181 patients with severe AP acute pancreatitis (27.3%). As a result, the developed ANN was based on 10 clinical parameters (age, the presence of hypotension, two or more signs of high-resolution survey, the level of PaO_2, LDH, glucose, urea, calcium, hematocrit, and the number of blood leukocytes) measured initially and then after 48 h. This model showed significantly better results than the APACHE II and Glasgow systems: ANN was more accurate than APACHE II and Glasgow in predicting severe acute pancreatitis ($P<0.05$ and $P<0.01$ respectively), in predicting the development of multiple organ failure ($P<0.05$ and $P<0.01$) and in predicting a lethal outcome ($P<0.05$). This work differs from those discussed above by the inclusion of a large number of patients (n$=664$), as well as by the fact that development and validation were performed on different groups of patients. Furthermore, all ten input variables are available to the doctor on duty within the first 6 h after hospitalization.

Andersson et al. [32] conducted a study aimed at developing and testing the effectiveness of the ANN model for early prediction of the severity of AP. The authors conducted a retrospective analysis of the results of treatment of 208 patients with AP. The area under the ROC curve for the neural network model was 0.92 (95% CI 0.85–0.99), 0.84 (0.76–0.92) for the logistic regression ($P=0.030$, χ^2), and 0.63 (0.50–0.76) when assessing the severity of acute pancreatitis using APACHE II ($P<0.001$, χ^2). The ANN is based on data obtained upon admission of the patient to the hospital. It is sufficiently accurate to predict the severity of AP.

Hong et al. [33] developed the ANN-based DSS for predicting persistent (more than 48 h) organ failure in patients with AP. The sample included 312 patients. The sensitivity of the initial model was 81.3%, specificity—98.9% and accuracy—96.2%. The predictive accuracy of the created ANN (AUC$=0.96\pm0.02$) was statistically significantly better in comparison with the model based on logistic regression (AUC$=0.88\pm0.03$, $P<0.001$, χ^2) and APACHE II (AUC$=0.83\pm0.03$, $P<0.001$, χ^2). The authors conclude that the created ANN can be useful for predicting the development of persistent organ failure in patients with AP.

Litvin *et al. World J Emerg Surg*   (2021) 16:50

Page 6 of 9

Fei et al. [34] developed an ANN to predict the risk of patients with severe AP developing an acute lung injury. When tested, their ANN showed a sensitivity of 87.5%, specificity of 83.3%, and accuracy of 84.4%. When predicting an acute lung injury, the ANN showed greater accuracy $AUC = 0.859 \pm 0.048$ than logistic regression ($AUC = 0.701 \pm 0.041$). The authors also identified 13 independent variables for predicting an acute lung injury, the most informative among which were the degree of pancreatic necrosis according to CT data, the level of lactate dehydrogenase, and oxyhemoglobin saturation.

Acute cholecystitis

Eldar et al. [35] used ANN to determine the predictors of conversion from laparoscopic access to laparotomic approach in the acute cholecystitis. Predictive models were selected using conventional statistical methods and ANN methods on the basis of data from 225 patients, who underwent laparoscopic cholecystectomy for acute cholecystitis AC. Direct logistic regression, direct linear discriminant analysis and ANN made it possible to predict the conversion in 0%, 27% and 100% of cases; a negative prediction was given and confirmed in 80%, 85.5% and 97%, respectively. When tested prospectively, the model predicted conversion in 0%, 25% and 67% of cases and non-conversion in 82%, 88% and 94%. According to the authors, a high degree of reliability of prediction reveals the potential of the ANN for allowing a decision to proceed directly to open cholecystectomy without initial laparoscopy.

The ANN was used to predict choledocholithiasis in patients with cholelithiasis and acute cholecystitis. The ANN was developed on the basis of data from 303 patients who underwent surgery for gallstones. The ANN revealed the most informative signs of possible choledocholithiasis were the level of bilirubin, alanine aminotransferase, the diameter of the common bile duct, the number of stones in the gall bladder, the size of the smallest stone, history of biliary colic, history of acute cholecystitis or acute pancreatitis. The authors concluded that ANN is a reliable and user-friendly system that can be successfully used to predict choledocholithiasis [36].

Upper GI bleeding

Rotondano et al. [37] developed an ANN predicting the probability of death in patients with bleeding from the upper gastrointestinal tract. The ANN was developed and tested on 2380 patients. Their ANN was compared with the Rockall scale [38]. The Rockall scale includes indicators selected by logistic regression—patient age, presence of shock, severity of comorbidities, endoscopic signs of recent bleeding and rebleeding. The developed ANN showed good sensitivity (83.8% versus 71.4%), specificity (97.5% versus 52.0%), accuracy (96.8% versus 52.9%) and AUC (0.95 versus 0.67) [37].

ANN models have also been developed to identify patients with a risk of recurrent bleeding and requiring surgical treatment. These models use clinical, instrumental data and allow identifying such patients with an accuracy of more than 90% [39–43]. Wong et al. [39] developed ANN by retrospective analysis of 22,854 patients with peptic ulcer disease. This ANN was able to identify patients with recurrent ulcer bleeding based on their age, hemoglobin level, localization of the ulcer in the stomach, the presence of other diseases of the gastrointestinal tract, malignant neoplasms and infection. The model identified patients with recurrent ulcer bleeding with AUC of 0.78 and accuracy of 84.3%.

Perforated peptic ulcers

AI was used to create a model for predicting mortality in patients with perforated peptic ulcers of the stomach or duodenum [44]. Given the complex nature of this disease, which has many nonlinear associations with outcomes, the authors created an ANN to identify risk factors for death. The data of 168 patients were included in the neural network model. The data of 117 patients (70%) were used for the training set, and data of 51 (39%) were used for the test set. The ANN predicted mortality with $AUC = 0.90$ [95% CI 0.85–0.95], $P < 0.001$.

Ileus/bowel obstruction

ANN was used to diagnose acute small bowel obstruction based on X-rays. AUC for the developed neural network model was 0.803. AUC increased to 0.971 after additional training. The final ANN had a sensitivity of 91.4% and a specificity of 91.9%. The classification efficiency [45] increases with an increase in the size of the training sample, reaching a plateau for 200 positive training examples.

Strangulated hernia

Least often, artificial intelligence was involved in solving tactical issues in the treatment of a strangulated hernia. AI elements [46] were used for early recognition of the risk of bowel resection in patients with SH. Bowel resections were performed in 21.0% of 762 patients included in the study (160/762). With the help of ANN, eight factors were identified that are significantly associated with the risk of the need to perform bowel resection in case of its infringement: female gender, age, age > 65 years, femoral hernia, intestinal obstruction, duration of infringement (measured in hours), the number of leukocyte and neutrophilic leukocytes.

Litvin *et al. World J Emerg Surg*   (2021) 16:50

Page 7 of 9

Discussion

In general, a variety of approaches and mathematical algorithms for the construction of decision support systems in medicine have been accumulated and systematized to date. Recently, experts have concluded that the majority of modern and successful models are based on ANN technology, first of all, deep neural networks and deep learning [12]. However, no definitive application has been demonstrated in emergency general surgery. In fact, this field often deals with situations and diseases for which the data accrual and the mathematical model testing is difficult and limited by the circumstances. As a counterpart especially in emergency general surgery, a real need to overcome the limited diagnostic tools requires a definitive improvement in technology and even the introduction of AI tools.

ANN-based decision support systems have been used to diagnose and treat patients with acute surgical pathology over the past 25 years. The first ANN models were mainly developed and intended to support clinical decision making in emergency surgical conditions and differential diagnosis with other diseases that do not require urgent surgery [47]. Subsequently, the main direction of research was to predict the course of the disease at its onset and predict the occurrence of complications. The latest ANNs are aimed at automatic prompting in the diagnosis and prediction of emergency surgical disease without the introduction of data by the surgeon (automated diagnostic methods, for example, analysis of X-ray or MRI images for the automatic detection of pathology, microscopic analysis of biological material, etc.) [48]. The last one will allow the ANN to identify and pay attention to routine pathology independently, reduce the time and cost of examination, and introduce remote diagnostics of emergency surgical conditions [47].

At the same time, all existing ANN-based decision support systems for predicting and diagnosing emergency surgical diseases have certain drawbacks and limitations in their use. Firstly, data from most studies were evaluated retrospectively, which could lead to known biases in results. Secondly, the data were obtained in hospitals of various levels. The last casts doubt on the question of the reproducibility of the method based on data from other clinics. Third, the sample size in most studies was insufficient for developing and testing the ANN (not all studies have trained and tested the ANN on different patient samples). Finally, the solution to the problem of diagnosis and treatment of urgent surgical diseases is strongly associated with the emergence of various interfering factors (confounders), associated mainly with the multi-causality of these diseases, the expressed heterogeneity of the samples of patients with acute surgical diseases. It is usually extremely difficult to assess the degree of various confounding factors impact on the final result [48].

Currently, there are five main limitations of the implementation of decision support systems based on artificial neural networks in medicine and surgery.

First, the quality and reliability of medical information are not always known. The data accumulated in the patient's medical records may be incomplete, contain errors, inaccuracies and non-standard terms. Currently, there are no effective mechanisms for collecting accurate information. Attempts to improve the quality of these analyses often fail due to the complexity of the process. To eliminate this problem, methods for training ANNs on small amounts of reliable information are now proposed [3, 7].

The second significant limitation is the lack of transparency in the decision-making process by the intellectual core of the system. ANN works according to the "black box" principle. If there is an error in the algorithm, and the system made the wrong decision, then it will be extremely difficult to answer the question "why" [3]. Currently, research is underway toward the development of hybrids of the ANN-expert system, which are aimed at improving the understanding of doctors of the way the system makes decisions [7].

The next significant limitation is the selection and development of personnel capable of effectively using and maintaining intelligent systems [8].

The fourth limitation is a high cost of projects, which is associated with the need to configure the new system for the data accumulated in a particular medical institution and to form a qualified and motivated team [9].

One of the most important limitations is that data sets have to be taken outside a medical institution, and this threatens the security of storage. It is no coincidence that many projects of the introduction of AI were stopped because of the risks related specifically to information security [7].

Despite the existing problems, researchers see the further use of neural networks in software that will quickly and accurately process large amounts of data. This will ultimately lay the foundation for high-performance medicine, which will be based on big data and reduce dependence on human resources [5]. The need for quick decision-making in emergency surgery when there is a limited resource is especially important. It will require the development of simple and accurate decision support systems based on AI [6, 48, 49].

Currently, the term Surgery 3.0 [6] has appeared. It characterizes modern changes in surgery by analogy with the Internet. The Internet is evolving from a collection of passive readers (Web 1.0) to a modern, interactive, AI-powered audience (Web 3.0). Surgery will change

Litvin *et al. World J Emerg Surg*   (2021) 16:50

Page 8 of 9

radically with the development of computational science and AI toward surgical artificial intelligence (Surgery 3.0) in order to improve the results of treatment of surgical patients [6].

However, artificial intelligence is currently one of the most controversial issues in the world. Written AI algorithms can contain errors that can lead to unintended consequences and unfair results. AI researchers must consider the need to prevent ethical violations [50].

Conclusion
The development and implementation of a DSS based on ANNs is a promising direction for improving the diagnosis and treatment of emergency surgical condition and their complications. At the same time, further improvement of the ANN is necessary, taking into account the shortcomings of previous studies. Developed DSS in medicine in general and, especially, in emergency surgery should be simple, accurate and as close as possible to the doctor's workplace.

Abbreviations
DSS: Decision support systems; ANN: Artificial neural network; AI: Artificial intelligence; AA: Acute appendicitis; AP: Acute pancreatitis; APACHE II: Acute physiology and chronic health evaluation; SPS: Severity of physiological state; AC: Acute cholecystitis; AUC: Area under curve; ROC: Receiver operating characteristic; CI: Confidence interval; LOS: Length of stay; MODS: Multiple organ dysfunction score; LDH: Lactate dehydrogenase.

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Surgical Disciplines, Immanuel Kant Baltic Federal University, Kaliningrad, Russia. [2]Kaliningrad Branch of Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Kaliningrad, Russia. [3]Department of Surgery, Macerata Hospital, Macerata, Italy. [4]Surgical Clinic, Department of Experimental and Clinical Sciences, University of Brescia, Brescia, Italy. [5]Division of Trauma and Acute Care Surgery, Scripps Memorial Hospital La Jolla, La Jolla, CA, USA. [6]General, Emergency and Trauma Surgery Department, Pisa University Hospital, Pisa, Italy. [7]Department of Surgery, Cambridge University Hospital, NHS Foundation Trust, Cambridge, UK. [8]Department of General Surgery, Albury Hospital, Albury, Australia. [9]Department of General Surgery, Rambam Healthcare Campus, Haifa, Israel. [10]Department of Gastrointestinal Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [11]Donegal Clinical Research Academy, Letterkenny University Hospital, Donegal, Ireland. [12]Department of Emergency and Trauma Surgery of the University Hospital of Parma, Parma, Italy.

References
1. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017;69S:S36–40. https://doi.org/10.1016/j.metabol.2017.01.011.
2. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719–31. https://doi.org/10.1038/s41551-018-0305-z.
3. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. Minim Invasive Ther Allied Technol. 2019;28(2):73–81. https://doi.org/10.1080/13645706.2019.1575882.
4. Rimmer L, Howard C, Picca L, Bashir M. The automaton as a surgeon: the future of artificial intelligence in emergency and general surgery [published online ahead of print, 2020 Jul 26]. Eur J Trauma Emerg Surg. 2020. https://doi.org/10.1007/s00068-020-01444-8.
5. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg. 2018;268(1):70–6. https://doi.org/10.1097/SLA.0000000000002693.
6. Mirnezami R, Ahmed A. Surgery 3.0, artificial intelligence and the next-generation surgeon. Br J Surg. 2018;105(5):463–5. https://doi.org/10.1002/bjs.10860.
7. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. PLoS ONE. 2019;14(2): e0212356. https://doi.org/10.1371/journal.pone.0212356.
8. Cresswell K, Callaghan M, Khan S, Sheikh Z, Mozaffar H, Sheikh A. Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: a systematic review. Health Informatics J. 2020;26(3):2138–47. https://doi.org/10.1177/1460458219900452.
9. Wang F, Preininger A. AI in health: state of the art, challenges, and future directions. Yearb Med Inform. 2019;28(1):16–26. https://doi.org/10.1055/s-0039-1677908.
10. Drew PJ, Monson JR. Artificial neural networks. Surgery. 2000;127(1):3–11. https://doi.org/10.1067/msy.2000.102173.
11. Forsström JJ, Dalton KJ. Artificial neural networks for decision support in clinical medicine. Ann Med. 1995;27(5):509–17. https://doi.org/10.3109/07853899509002462.
12. Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? J Arthroplasty. 2018;33(8):2358–61. https://doi.org/10.1016/j.arth.2018.02.067.
13. Miller DD, Brown EW. How cognitive machines can augment medical imaging. AJR Am J Roentgenol. 2019;212(1):9–14. https://doi.org/10.2214/AJR.18.19914.
14. Durstewitz D, Koppe G, Meyer-Lindenberg A. Deep neural networks in psychiatry. Mol Psychiatry. 2019;24(11):1583–98. https://doi.org/10.1038/s41380-019-0365-9.
15. Hessler G, Baringhaus KH. Artificial intelligence in drug design. Molecules. 2018;23(10):2520. https://doi.org/10.3390/molecules23102520.
16. Dande P, Samant P. Acquaintance to artificial neural networks and use of artificial intelligence as a diagnostic tool for tuberculosis: a review. Tuberculosis (Edinb). 2018;108:1–9. https://doi.org/10.1016/j.tube.2017.09.006.
17. Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, Danese S, Peyrin-Biroulet L. Application of artificial intelligence to gastroenterology and hepatology. Gastroenterology. 2020;158(1):76-94.e2. https://doi.org/10.1053/j.gastro.2019.08.058.
18. Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. World J Gastroenterol. 2019;25(14):1666–83. https://doi.org/10.3748/wjg.v25.i14.1666.

19. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. Surv Ophthalmol. 2019;64(2):233–40. https://doi.org/10.1016/j.survophthal.2018.09.002.

20. Bashir M, Harky A. Artificial intelligence in aortic surgery: the rise of the machine. Semin Thorac Cardiovasc Surg. 2019;31(4):635–7. https://doi.org/10.1053/j.semtcvs.2019.05.040.

21. Panchmatia JR, Visenio MR, Panch T. The role of artificial intelligence in orthopaedic surgery. Br J Hosp Med (Lond). 2018;79(12):676–81. https://doi.org/10.12968/hmed.2018.79.12.676.

22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl EA, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garritty C, Lewin S, Godfrey CM, Macdonald MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467–73. https://doi.org/10.7326/M18-0850.

23. Yoldaş Ö, Tez M, Karaca T. Artificial neural networks in the diagnosis of acute appendicitis. Am J Emerg Med. 2012;30(7):1245–7. https://doi.org/10.1016/j.ajem.2011.06.019.

24. Park SY, Kim SM. Acute appendicitis diagnosis using artificial neural networks. Technol Health Care. 2015;23(Suppl 2):S559–65. https://doi.org/10.3233/THC-150994.

25. Reismann J, Romualdi A, Kiss N, Minderjahn MI, Kallarackal J, Schad M, Reismann M. Diagnosis and classification of pediatric acute appendicitis by artificial intelligence methods: an investigator-independent approach. PLoS ONE. 2019;14(9): e0222030. https://doi.org/10.1371/journal.pone.0222030.

26. Park JJ, Kim KA, Nam Y, Choi MH, Choi SY, Rhie J. Convolutional-neural-network-based diagnosis of appendicitis via CT scans in patients with acute abdominal pain presenting in the emergency department. Sci Rep. 2020;10(1):9556. https://doi.org/10.1038/s41598-020-66674-7.

27. Kazmierczak SC, Catrou PG, Van Lente F. Diagnostic accuracy of pancreatic enzymes evaluated by use of multivariate data analysis. Clin Chem. 1993;39(9):1960–5.

28. Pofahl WE, Walczak SM, Rhone E. Predicting length stay in acute pancreatitis with a neural network. Gastroenterology. 1998;114(S1):A491. https://doi.org/10.1016/S0016-5085(98)81989-6.

29. Keogan MT, Lo JY, Freed KS, Raptopoulos V, Blake S, Kamel IR, Weisinger K, Rosen MP, Nelson RC. Outcome analysis of patients with acute pancreatitis by using an artificial neural network. Acad Radiol. 2002;9(4):410–9. https://doi.org/10.1016/s1076-6332(03)80186-1.

30. Halonen KI, Leppäniemi AK, Lundin JE, Puolakkainen PA, Kemppainen EA, Haapiainen RK. Predicting fatal outcome in the early phase of severe acute pancreatitis by using novel prognostic models. Pancreatology. 2003;3(4):309–15. https://doi.org/10.1159/000071769.

31. Mofidi R, Duff MD, Madhavan KK, Garden OJ, Parks RW. Identification of severe acute pancreatitis using an artificial neural network. Surgery. 2007;141(1):59–66. https://doi.org/10.1016/j.surg.2006.07.022.

32. Andersson B, Andersson R, Ohlsson M, Nilsson J. Prediction of severe acute pancreatitis at admission to hospital using artificial neural networks. Pancreatology. 2011;11(3):328–35. https://doi.org/10.1159/000327903.

33. Hong WD, Chen XR, Jin SQ, Huang QK, Zhu QH Pan JY. Use of an artificial neural network to predict persistent organ failure in patients with acute pancreatitis. Clinics (Sao Paulo). 2013;68(1):27–31. https://doi.org/10.6061/clinics/2013(01)rc01.

34. Fei Y, Gao K, Li WQ. Artificial neural network algorithm model as powerful tool to predict acute lung injury following to severe acute pancreatitis. Pancreatology. 2018;18(8):892–9. https://doi.org/10.1016/j.pan.2018.09.007.

35. Eldar S, Siegelmann HT, Buzaglo D, Matter I, Cohen A, Sabo E, Abrahamson J. Conversion of laparoscopic cholecystectomy to open cholecystectomy in acute cholecystitis: artificial neural networks improve the prediction of conversion. World J Surg. 2002;26(1):79–85. https://doi.org/10.1007/s00268-001-0185-2.

36. Vukicevic AM, Stojadinovic M, Radovic M, Djordjevic M, Cirkovic BA, Pejovic T, et al. Automated development of artificial neural networks for clinical purposes: application for predicting the outcome of choledocholithiasis surgery. Comput Biol Med. 2016;75:80–9. https://doi.org/10.1016/j.compbiomed.2016.05.016.

37. Rotondano G, Cipolletta L, Grossi E, Koch M, Intraligi M, Buscema M, et al. Artificial neural networks accurately predict mortality in patients with nonvariceal upper GI bleeding. Gastrointest Endosc. 2011;73(2):218-226.e2262. https://doi.org/10.1016/j.gie.2010.10.006.

38. Rockall TA, Logan RF, Devlin HB, Northfield TC. Risk assessment after acute upper gastrointestinal haemorrhage. Gut. 1996;38(3):316–21. https://doi.org/10.1136/gut.38.3.316.

39. Wong GL, Ma AJ, Deng H, Ching JY, Wong VW, Tse YK, et al. Machine learning model to predict recurrent ulcer bleeding in patients with history of idiopathic gastroduodenal ulcer bleeding. Aliment Pharmacol Ther. 2019;49(7):912–8. https://doi.org/10.1111/apt.15145.

40. Das A, Ben-Menachem T, Cooper GS, Chak A, Sivak MV Jr, Gonet JA, Wong RC. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. Lancet. 2003;362(9392):1261–6. https://doi.org/10.1016/S0140-6736(03)14568-0.

41. Das A, Ben-Menachem T, Farooq FT, Cooper GS, Chak A, Sivak MV Jr, Wong RC. Artificial neural network as a predictive instrument in patients with acute nonvariceal upper gastrointestinal hemorrhage. Gastroenterology. 2008;134(1):65–74. https://doi.org/10.1053/j.gastro.2007.10.037.

42. Ayaru L, Ypsilantis PP, Nanapragasam A, Choi RC, Thillanathan A, Min-Ho L, Montana G. Prediction of outcome in acute lower gastrointestinal bleeding using gradient boosting. PLoS ONE. 2015;10(7): e0132485. https://doi.org/10.1371/journal.pone.0132485.

43. Sengupta N, Tapper EB. Derivation and internal validation of a clinical prediction tool for 30-day mortality in lower gastrointestinal bleeding. Am J Med. 2017;130(5):601.e1-601.e8. https://doi.org/10.1016/j.amjmed.2016.12.009.

44. Søreide K, Thorsen K, Søreide JA. Predicting outcomes in patients with perforated gastroduodenal ulcers: artificial neural network modeling indicates a highly complex disease. Eur J Trauma Emerg Surg. 2015;41(1):91–8. https://doi.org/10.1007/s00068-014-0417-4.

45. Cheng PM, Tran KN, Whang G, Tejura TK. Refining convolutional neural network detection of small-bowel obstruction in conventional radiography. AJR Am J Roentgenol. 2019;212(2):342–50. https://doi.org/10.2214/AJR.18.20362.

46. Chen P, Huang L, Yang W, He D, Liu X, Wang Y, et al. Risk factors for bowel resection among patients with incarcerated groin hernias: a meta-analysis. Am J Emerg Med. 2020;38(2):376–83. https://doi.org/10.1016/j.ajem.2019.09.023.

47. Loftus TJ, Upchurch GR Jr, Bihorac A. Use of artificial intelligence to represent emergent systems and augment surgical decision-making. JAMA Surg. 2019;154(9):791–2. https://doi.org/10.1001/jamasurg.2019.1510.

48. Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, et al. Artificial intelligence and surgical decision-making. JAMA Surg. 2019. https://doi.org/10.1001/jamasurg.2019.4917.

49. Bronsert M, Singh AB, Henderson WG, Hammermeister K, Meguid RA, Colborn KL. Identification of postoperative complications using electronic health record data and machine learning. Am J Surg. 2020;220(1):114–9. https://doi.org/10.1016/j.amjsurg.2019.10.009.

50. Keskinbora KH. Medical ethics considerations on artificial intelligence. J Clin Neurosci. 2019;64:277–82. https://doi.org/10.1016/j.jocn.2019.03.001.

Publisher's Note