UDC 004.891

AUTOMATION SYSTEM OF QUANTITATIVE ESTIMATION OF OPERATIVE RISK

I.A. Dragun, G.G. Ustinov, P.M. Zatsepin

Altay State University, Barnaul city
E-mail: dragun_igor@mail.ru, zpm@phys.asu.ru

Automation system of quantitative analysis of operative risk which is an expert object-module support system in making clinical decisions by means of analysis by neuro-network and probability-statistic methods of medical-biological data and obtaining knowledge from empirical information has been developed and applied.

Introduction

In a clinical practice, during realization of patients' scheduled treatment, based on own experience, doctor-clinical physician diagnoses, qualitatively estimates a gravity degree of patient's condition, estimates risk of medical actions realization. In situations with difficultly established or not clear diagnosis it is very important to estimate quantitatively gravity of disease while universal methods of quantitative estimation of gravity do not exist.

One of the most important problems in modern medicine is complexity of quantitative estimation of operational risk. The estimation of operational risk means an estimation of gravity degree of preoperative patient's condition at threat of fatal outcome or possible postoperative complications. Complexity of condition gravity degree gravity estimation is caused by a big number of symptoms, various for each disease and the big lability of clinical displays of disease. In practice the gravity degree of a patient's condition is defined qualitatively. Thus the small set of 5–10 symptoms or risk factors are used. Values of which most obviously reflect patient's disease while less significant symptoms are not considered. Therefore one of ways of estimation efficiency increase of operational risk is application of mathematical methods of medical and biologic data processing and development of problem-focused systems of information analysis. It will allow us to estimate more precisely gravity degree of preoperative condition at certain patient's disease and to conduct medical activities in due time. Thereby, the problem of development and realization of the automated system of quantitative estimation of operational risk is relevant.

Initial data

Researches of quantitative estimation of operational risk had been done on a statistical material bilious-stone disease patients (BSD), operated within last 12 years (1990–2002) in the Hospital № 5 of Barnaul city. The choice of biliously – stone disease is caused by its frequency and quite high postoperative lethality at complications. The base of initial data contains clinical-laboratory parameters of 1172 patients, 268 of them ill with BSD have died after surgery, and 904 patients had been operated on successfully. Each record in data base contains the information consisting of three logic blocks:

- information on gravity of the preoperative condition, characterized by a lethal outcome of a surgery or its absence;
- qualitative information on presence of the basic disease complications found in the patient, accompanying diseases, on deviations in functioning of this or that system of an organism vitality (22 parameters);
- quantitative data of laboratory and tool tests of BSI patient's preoperative condition (30 parameters).

Thus, the gravity of preoperative condition of BSD patients, characterized by postoperative lethality, has been estimated by preoperative clinical data. At such statement of a problem the estimation of operational risk can be interpreted as a problem of clinical forecasting of postoperative condition by preoperative data, as well as problem of patient's preoperative condition diagnostics. That is, to solve the problem of a quantitative estimation of an operational risk degree, mathematical methods of diagnostics and forecasting are applicable.

Mathematical methods of diagnostics and forecasting

The analysis of existing and traditionally applied in medicine for diagnostics and forecasting methods has shown that probabilistic-statistical and neuronetwork models allow specialists to represent and analyze the system of complex interrelations of disease symptoms most adequately. Probabilistic methods are more flexible and convenient for clinicolaboratory data processing rather than determined multipleparameter statistical models burdened by conditions of normality of values distributions, orthogonality of considered space, etc. [1]. Besides, the mathematical device of artificial neural networks (ANN) had been developed for biological objects behavior modeling [2], therefore, the given methods allow more precise estimation of patients' preoperative state in conditions of big lability of clinical displays.

Probability-statistical methods. The basic methods solving a problem of diagnostics in probabilistic space of symptoms are methods of images recognition, leaning on strategies of Bayess, Neumann – Pirson and Wald. Introduction of probabilistic measure in space of clinicolaboratory attributes is realized by calculation of frequencies values in the intervals formed by quantization of range of an attribute values. The most powerful criterion of distinction of distributions is the nonparametric information measure of Kulbak (1), allowing us to find borders of intervals so that to emphasize the most significant differentially-diagnostic information and to minimize influence of casual fluctuations [1]. For two discrete distributions A_1 and A_2 of an attribute x_i Kulbak's measure $J(x_i;A_1,A_2)$ is expressed by the formula:

$$J(x_i; A_1 : A_2) = \\ = \sum_s [P(x_{is} / A_2) - P(x_{is} / A_1)] \ln \frac{P(x_{is} / A_2)}{P(x_{is} / A_1)},$$

where $P(x_{is}/A_1)$ and $P(x_{is}/A_2)$ are the probabilities of an attribute x_i presence, which values x_i belong to an interval with the number *s*. At validity of a zero hypothesis (i. e. samples are indiscernible), the value *J* asymptotically has distribution χ^2 with s–1 degrees of freedom. Thus, in view of the significance value 0,05 accepted in medical-biologic researches, the attribute is considered informative with corresponding self-descriptiveness *J* and significance value P_J.

The intervals of one-dimensional symptom calculated by means of such quantization refer to as interval structure, an attribute generated by a pair of interval structures – binary structure. One-dimensional and two-dimensional attributes, significance value of which is less than 0,05, are ranged on their self-descriptiveness decrease (value of Kulbak's measure), thereby allocates a subset of the most informative, significant attributes.

Researches by probability-statistical methods are organized by the following three strategies of images recognition, each of which gives probability of surgery lethal outcome at certain patient and by that quantitatively estimates risk of operation [1]:

- Wald's strategy is used when decision-making and is based on calculation of the likelihood relation consistently for the attributes, ordered on *self-descriptiveness decrease;* at the same time the relations of likelihood in each step are compared to two thresholds, at excess of one of them decision on the most probable clinical situation is passed;
- Bayess' strategy is appened when the decision is accepted according to Bayess' formula. The probability of clinical situation A_j at value *x* an attribute *x* is defined through probabilities of this value for each of clinical situations and aprioristic probabilities of these clinical situations;
- Bayess' consecutive strategy is when attributes are ordered on self-descriptiveness decrease and in each step the Bayess' formula is used, where aprioristic probability is considered to be posterior probability, calculated in the previous step; the decision is accepted in last step in favor of that hypothesis for which posterior probability appears the greatest.

For optimum work of the given strategies we have developed and realized algorithms of search for a set of risk factors. This subset of information-valuable attributes, providing not less results of recognition than all informative attributes and being optimum concerning quantity and the order of attributes at use of consecutive strategies of recognition.

In table 1 the average results of learning and test data of cross testing for specified strategies of recognition of a dangerous condition (fatal outcome) are presented. According to the level accepted among clinical physicians, the forecast which probability exceeds 0,8 is considered to be true.

Table 1. *Results of cross testing at forecasting of a surgery's outcome by probability-statistical methods*

Conditions	Recognition strategy	Percentage of right forecasts	Percentage of wrong forecasts	Se, %	Sp, %
A1 (successful surgery)	Bayess'	81,53	18,47	81,53	85,14
	Bayess' (multistep)	77,71	22,29	77,71	81,10
	Wald's	88,32	11,68	88,32	88,32
	Average for all strategies	82,52	17,48	82,52	84,85
A2 (fatal outcome after surgery)	Bayess'	76,86	23,14	76,86	80,99
	Bayess' (multistep)	80,99	19,01	80,99	82,64
	Wald's	58,68	41,32	58,68	58,68
	Average for all strategies	72,18	27,82	72,18	74,10

In the last two columns of table 1 values of the parameters are shown, often used in medicine for an estimation of quality of a diagnostics method – sensitivity (Se) and specificity (Sp), being the analogues of mistakes of the first and second sort of the theory of decision-making of Neumann-Pirson [1]. As evident from table 1 probability-statistical methods displace risk estimation more to the side of specific tests, while the main requirement to the method of risk estimation should be considered its sensitivity and specificity. However, it is necessary to note some peculiarities of information analysis by probabilistic methods, such as: absence of necessity of preliminary data processing, opportunities revealing differentially-diagnostic borders of separate attribute values and allocation of risk factors – an optimum feature set for diagnostics.

Neuronetwork methods. The multilayered artificial neural networks of direct distribution, also called as multilayered perceptrone, are often used in medicine for forecasting and diagnostics. Owing to ability of multilayered perceptrone to generalization of the analyzed information, at training by implicit way interrelations between input and output data come to light. The result of work of each output neuron of a trained neuronetwork has indistinct character and gives quantitative estimation of confidence in conformity of input data to output. Thus, at recognition of a heavy, menacing condition on neuronetwork output we get a quantitative estimation of confidence of presence of this condition, thereby we quantitatively estimate risk. Besides for trained neuronetwork, it is possible to quantitatively estimate contribution of each input neuron at classification or recognition [2]. Therefore, it is possible to allocate a set of the most informative attributes at risk estimation – risk factors.

Since neuronetworks are capable of processing only quantitative information, the preliminary processing of initial data has been conducted, consisting of stages:

- restoration of the missing information, missed during gathering.
- preprocessing of quantitative information consists in transformation to dimensionless kind, by scaling to a single scale (from 0 up to 1);

- preprocessing of qualitative information by numerical values coding of presence (1) or absence (0) of the attribute.

For the analysis of applicability of multilayered perceptrone model at a quantitative estimation of operational risk, the numerical experiment on search of the most correct architecture of neuronetworks, learned on algorithm of mistake return distribution, has been conducted. In experiment, at cross training and testing, sensitivity and specificity of models of forecasting of a fatal outcome (1 output neuron) and classifications of a patient's preoperative condition (2 output neurons) were analyzed. In tables 2 and 3 characteristics of quality of work of three neuronetworks architectures are presented. Their average values of accuracy at cross testing have appeared the greatest. The forecast and the diagnosis, as well as for probabilistic methods, were considered true if a network output was more than 0,8.

Besides, the experiment on neuronetworks training on the most informative attributes, allocated by probability-statistical methods and comparison with the risk factors, allocated by neural networks has been conducted. As a result: neuronetworks cross training and testing on the informative attributes found by probabilistic methods, gave about the same parameters of sensitivity and specificity as all the data, with the only difference that the latent layer of networks contained smaller quantity of neurons. Comparison of sets of the risk factors allocated by neuronetworks and probabilistic methods has not revealed any special distinctions in sets, while both methods allowed reduction of quantity of attributes, necessary for risk estimation with good accuracy, at least twice.

Table 2. *Results of neuronetworks learning with one output neuron*

Conditions	Neuronetwork architecture				Percentage of right forecasts	Percentage of wrong forecasts	Se, %	Sp, %
	Layers	Neurons in a layer	Sigmoid's description	Training speed				
A1 (successful surgery)	2	5	2,5	0,6	19,29	80,71	19,3	20,3
	2	5	3	0,5	17,77	82,23	17,8	18,3
	2	10	2,5	0,3	6,59	93,41	6,6	7,11
A2 (fatal outcome after surgery)	2	5	2,5	0,6	97,34	2,66	97,3	97,3
	2	5	3	0,5	96	4	96	96
	2	10	2,5	0,3	94,67	5,33	94,7	96

Analyzing parameters of sensitivity and specificity of quantitative estimation of operational risk at BSI patients by means of neuronetwork models, it is necessary to note that the analysis of medical-biologic data by multilayered ANN is a high-sensitivity and high-specificity test. Meaning that an estimation of operational risk by means of neuronetworks, the quantity of wrong answers, as well the quantity of not distinguished conditions, will be admissible small. Considering available average percent of fatal outcomes after surgery at patients with bilious – stone illness in Hospital № 5 of Barnaul city, it is possible to assume that application of neuronetwork technologies for an estimation of surgeriy risk will help to lower the level of lethality to an admissible norm.

Table 3. *Results of neuronetworks learning with two output neurons*

Conditions	Neuronetwork architecture				Percentage of right forecasts	Percentage of wrong forecasts	Se, %	Sp, %
	Layers	Neurons in a layer	Sigmoid's description	Training speed				
A1 (successful surgery)	3	25	1	0,3	94,42	5,58	94,4	95,9
	2	5	2	0,1	94,42	5,58	94,4	95,9
	2	25	1,5	0,1	93,91	6,09	93,9	96,5
A2 (fatal outcome after surgery)	3	25	1	0,3	94,67	5,33	94,7	96
	2	5	2	0,1	93,34	6,66	93,3	94,7
	2	25	1,5	0,1	93,34	6,66	93,3	97,3

Considering characteristics, merits and demerits of probability-statistical and neuronetwork methods, during development of the automated system of quantitative estimation of operational risk, it is possible to recommend probability-statistical models for fast preliminary data processing at risk estimation, while neuronetwork methods for more exact data analysis.

Expert system

The automated system of quantitative estimation of operational risk (ASQEP), realized by us, is an expert objective-modular system of support of clinical decisions making by means of neuronetwork and probability-statistical methods analysis of clinical-laboratory data and extraction of knowledge from the empirical information. The result of data processing is the system main object – the computer image of a condition containing information structures, reflecting interrelations between an interesting clinical situation and the empirical information, figure.

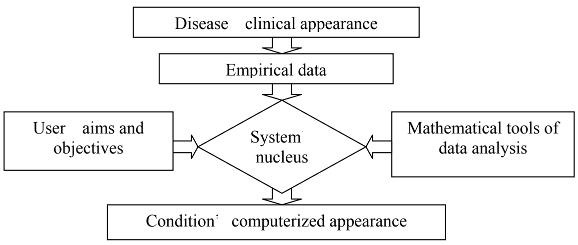

Figure. *Correlation between objects ASQEP*

System ASQEP includes:

- system of projects management;
- designer of a subject domain shell (generator of condition computer image);
- information system of forms, for convenient and fast input of the clinical, laboratory and tool information;
- research module – mathematical toolkit of the adequate analysis of empirical data;
- mechanism of knowledge base formation on the basis of results of the research module work;

- expert module, providing dialogue with the knowledge base and expert decisions-making.

The designer of a subject domain shell is the module creating a system of forms for gathering of medical-physiological data, clinical situations are also concretized here, recognition of which is done by the expert module and the processing of statistical material (the research module) is taking place.

The research module contains following functions:

- preliminary data processing: restoration of missed during gathering information, coding of qualitative and transformation of quantitative data;
- formation of interval and binary structures – the analysis of distributions of attribute values in probabilistic space;
- search of informative attributes multitude – risk factors;
- learning of neural networks and probabilistic recognition strategy;
- percentage calculation of system training.

At user's desire, based on results of the research module work, the entire report, where the process of data processing and testing results are reflected, can be generated.

For restoration of the missed information two methods are realized in ASQEP: filling blanks with the most probable values on extracts and offered by us development modification of Laboratory of nonequilibrium systems of Institute of computing modeling of the Siberian Branch of the Russian Academy of Science (Krasnoyarsk city). The given method represents iterative process of sequence construction of one-dimensional quasi-linear models, as a result of which the initial table of data with blanks is represented in the form of giving way of plausible restoration of the missed information and approximating known data [3].

Decision-making process in ASQEP is realized as follows. By means of the system of forms, cooperating with the knowledge base, the computer image of a patient's condition is formed. The expert module provides a quantitative estimation of conformity of a patient's computer image to a distinguished condition by neuronetwork and probability-statistical methods. The result of the expert module work is the report of the expertise in which the following factors are fixed:

- values of all medical-biologic parameters of a patient;
- results of quantitative estimation of risk condition (gravity) on the basis of which the decisions on prescription or change of scheduled clinical treatment can be made;
- the table of a certain patient's risk factors – features and values of which reflect pathology of those or other organs.

All system modules were developed by modern visual means of programming Windows – applications. The reports, formed by ASQEO, have formats of Microsoft Excel (the research module) and Microsoft Word (the expert module) files and can be printed directly by the system or specified programs.

Conclusions and results of researches

Carried out researches at quantitative estimation of operational risk at bilious-stone illness patient's allow specialists to judge about high efficiency of neuronetwork and probability-statistical methods application in the given area. Developed and realized by us system ASQEP realizes the newest methods of analysis and processing of medical-biologic data and can be applied for:

- quantitative estimation of surgerie risks;
- quantitative estimation of gravity of clinical situations;
- diagnostics and forecasting of one of the two alternative diseases
- support of clinical decisions-making.

The work of the ASQEP system has been tested on two projects of quantitative estimation of operational risk of patients with mechanical jaundice and bilious – stone illness and is introduced for practical application in Hospital № 5 of Barnaul city. Also, the official registration of the expert module ASQEP [4] has been obtained.

REFERNCES

1. Genkin A.A. New information technology of medical data analysis. – SP6.: Polytechnica, 1999. – 191 p.
2. Neuroinformatics / A.N. Gorban, V.L. Dunin-Barkovskiy, A.N. Kirdin, a. o. – Novosibirsk: Nauka, 1998. – 296 p.
3. Dragun I.A., Polyakov V.V., Ustinov G.G., Zatsepin P.M. Data processing for neuronetwork estimation of operational risk // Neuroinformatics and its appendices: Materials of the XIII Russian convention. – Krasnoyarsk, 2005. – P. 60–61.
4. Certificate of the computer program for PC № 2006611654 official registration. Forecasting of surgery's fatal outcome for patients with mechanical jaundice (Forecasting) / I.A. Dragun, G.G. Ustinov, V.V. Polyakov, P.M. Zatsepin. Registered in the Register of computer programs for PCs on 17/05/2006.

Arrived on 21.06.2006