Представлено метод прогнозування рецидивуючого інфаркту міокарда (РІМ) на основі розробленого інтерполяційного діагностичного полінома для визначення можливості виникнення і результату РІМ. Розроблений метод базується на методології вербального аналізу рішень і дозволяє, з урахуванням сукупності ознак захворювання, їх комбінації і взаємовпливу, підвищити точність прогнозу на 2,7 % (у порівнянні з методом-прототипом), що надає можливість попередити рецидив і раптову коронарну смерть. Запропонований метод становить практичний інтерес і може бути використаний для діагностики та прогнозування розвитку інших захворювань серцево-судинної системи людини

Ключові слова: метод прогнозування, рецидивуючий інфаркт міокарда, ординальна класифікація станів, діагностичний інтерполяційний поліном

Представлен метод прогнозирования рецидивирующего инфаркта миокарда (РИМ) на основе разработанного интерполяционного диагностического полинома для определения возможности возникновения и исхода РИМ. Разработанный метод базируется на методологии вербального анализа решений и позволяет, с учетом совокупности признаков заболевания, их комбинации и взаимовлияния, повысить точность прогноза на 2,7 % (по сравнению с методом-прототипом), что дает возможность предупредить рецидив и внезапную коронарную смерть. Предложенный метод представляет практический интерес и может быть использован для диагностики и прогнозирования развития других заболеваний сердечно-сосудистой системы человека

Ключевые слова: метод прогнозирования, рецидивирующий инфаркт миокарда, ординальная классификация состояний, диагностический интерполяционный полином

UDC 004.02-89:616.127-005.8:61
DOI: 10.15587/1729-4061.2016.81004

DEVELOPING A METHOD FOR PREDICTION OF RELAPSING MYOCARDIAL INFARCTION BASED ON INTERPOLATION DIAGNOSTIC POLYNOMIAL

S. Yakubovska
Assistant*
E-mail: sofiya.yakubovska@nure.ua
O. Vysotska
Doctor of Technical Sciences, Professor**
E-mail: olena.vysotska@nure.ua
A. Porvan
PhD, Assistant Professor**
E-mail: porvan_a_p@mail.ua
D. Yelchaninov
PhD, Associate Professor
Informatics and Intellectual Property Department
National Technical University
«Kharkiv Polytechnic Institute»
Bagaliya str., 21, Kharkiv, Ukraine, 61002
E-mail: elchaninov@mipk.khakiv.edu
E. Linnyk
PhD, Associate Professor**
E-mail: elena.linnyk@nure.ua
*Department of Economic Cybernetics and Management of Economic Security***
Department of Biomedical Engineering*
***Kharkiv National University of Radio Electronics
Nauky ave., 14, Kharkiv, Ukraine, 61166

1. Introduction

At present, there are certain achievements gained in the field of diagnosing cardiovascular diseases. However, myocardial infarction (MI) is still the most common cause of disability and mortality among the social-significant group of population group in the world. According to data of the World Health Organization, every year 32,4 million of cases of MI and strokes occurrence are recorded worldwide. Patients with MI belong to the increased risk group for the occurrence of further coronary events. The patients who survived MI have the increased risk of relapses and repeated infarctions. This fact increases mortality by 5 % [1]. The State Department of Statistics in Ukraine registered 1 879 963 people with the circulatory system diseases in 2014. More to the point, the contribution of MI to the structure of mortality from the circulatory system diseases over the past 4 years has increased by 14,3 %. There were revealed 5779 cases of disease registered for the first time for 100000 of the population in the Kharkov Region (Ukraine) alone. The index of mortality from the acute MI in the structure of total mortality of the population of Ukraine reached 18,7 %.

An increase in the stay of patients in the hospital and the lethal outcome is caused by the relapsing MI (RIM) that is a variant of the disease, at which new sections of the necrosis of myocardium are developed during 28 days from the beginning of previous MI [2]. Patients with MI need adequate therapeutic measures, directed toward survival and averting threatening complications. That is why the prediction of possible development and outcome of RIM must be timely, taking into account possible changes in the state of patient during the period of stay in the clinic.

Thus, the relevance of the problem is defined by the necessity of increasing the accuracy in predicting both the occurrence of RIM and its outcome in the patients with MI for the purpose of averting threatening states and reduction in the hospital lethality.

2. Literature review and problem statement

For solving the problems of diagnosis and prediction in medicine, in particular cardiology, different mathematical methods and approaches are actively applied [3, 4].

The problem of prediction of the course of disease in patients with MI is extremely complex. A number of known methods can be used for its solution. Papers [5, 6] propose the method of predicting the relapses of MI in patients, who are in the rehabilitative period. This method is based on the mathematical apparatus of fuzzy logic. The main disadvantages include the need for a large volume of the training sample, which is not always acceptable under conditions of decision making in emergency.

There is a known method of predicting the possibility of occurrence of RIM, which is based on the mathematical apparatus of artificial neural networks and standard algorithm of reverse propagation [7]. This method is labor-consuming and requires the redundancy of information about the patient (a doctor has to enter minimum 25 factors to the map of patient).

A large number of existing methods and approaches to the prediction of possibilities of occurrence of different complications after previous MI are based on the methods of mathematical statistics. Thus, papers [8, 9] describe analysis of survival of the patients who had MI with the use of regression analysis, namely the multifactor regression model of Cox. The authors on the basis of the methods of mathematical statistics also determined significant predicting factors, which influence the risk of repeated and relapsing MI.

Authors of article [10] proposed a mathematical model of proportional intensities for the prediction of MI and repeated cardiovascular catastrophes. All predictors of the model synthesized by the authors were determined with the use of the long-run test and Bayesian approach. Authors of paper [11], on the basis of dispersion data analysis, determined the influence of indices on the possibility of occurrence of relapsing MI and proposed a multiparametric mathematical model for predicting lethal outcome from this complication.

The shortcomongs in the application of the considered methods are:

– mandatory existence of a large number of tests;

– slow convergence of computational procedures with a large number of factors, which complicates the process of obtaining reliable solution;

– the margins of error in the calculations in the synthesis of mathematical models are not accurately determined and have certain randomness, which is caused by possible incompleteness of the initial data.

The application of statistical methods also does not warrant the construction of complete and precise classification of the states of patient for the prediction of RIM under conditions of uncertainty of the initial information, which may subsequently affect the accuracy of prediction.

Thus, it is necessary to design a method for predicting RIM taking into account the enumerated shortcomings.

For the solution of the set problem, it is expedient to use the methods of verbal decision analysis (VDA), developed in the Institute of Systems Analysis of the Russian Academy of Sciences. These methods are divided into three classes: distribution of alternatives according to the classes of solutions, ordering of alternatives, and detection of the best alternative. The RIM prediction problem may be attributed to the class "distribution of alternatives according to the classes of solutions".

There are three methods, with the aid of which they perform the distribution of alternatives according to the classes of solutions: ORCLASS (ordinal classification of states), CYCLE, CLARA.

The CLARA method categorizes not all possible vectors of state, comprised of the criteria of evaluation of the states of patient for the prediction of RIM, but only some of them. In the CYCLE method, for constructing full classification during the formulation of diagnosis, the doctor deals with two vectors y' and y", those describing the state of one patient [12]. In this case, the method of constructing full classification, although is the fastest, however does not describe all possible processes, necessary for the prediction of RIM, in contrast to the ORCLASS method, by which, for constructing full classification, a doctor-cardiologist is presented with one vector of state of the patient. The application of the ORCLASS method also provides the possibility of comprehensive consideration of the factors, which influence the occurrence and outcome of RIM, their combinations and mutual influence, which is extremely important.

3. The purpose and objectives of the study

The purpose of present study is to develop a method of prediction of RIM using the method of verbal decision analysis ORCLASS.

To achieve the set goal, the following tasks are to be solved:

– to determine the set of criteria for evaluation of the state of patient for the prediction of RIM;

– to build an ordinal classification of the states of patients for predicting the development of RIM;

– to design interpolation diagnostic polynomial for determining the measure of proximity between the classes and the group of attributes of RIM.

4. Development of method for predicting relapsing myocardial infarction

At the first stage, for the prediction of RIM, based on the method for determining significance of opinions of experts during formation of an expert group [13–16], we selected a group of experts, which includes the best 5 specialists, who have been working in the field of cardiology for not less than 10 years.

At the second stage, experts define a set of criteria for evaluation of the state of a patient: $K=|K_i|,\ i=\overline{1,z}$ where i is the ordinal number of the criterion for evaluation of the state, z is the quantity of criteria. For the problem of prediction of RIM, the experts determined the following 5 criteria (Table 1).

Table 1

Criteria for evaluation of the state of patient for the prediction of RIM

Denotation	Name
K_1	type infarction
K_2	Killip heart failure, determined by the Killip scale
K_3	hypertonia based on gender and age grading
K_4	associated diseases
K_5	pain syndrome (pain intensity)

Let us denote through k_{ij} the j-th estimation of the i-th criterion of RIM specificity, which is regulated descendingly $j=1,n(i)$, n(i) is the number of values of estimations on the scale of the i-th criterion (different for each criterion; therefore, depends on i). Possible values of criteria for the prediction of RIM are represented in Table 2.

Table 2

Values of estimations of criteria of the state of patient for the prediction of RIM

i	j	Significance, k_{ij}
1	1	primary microfocal MI
	2	primary macrofocal MI
	3	repeated microfocal MI
	4	repeated macrofocal MI
2	1	MI without signs of circulatory failure
	2	MI with signs of moderate heart failure
	3	MI with acute left ventricular failure (pulmonary edema)
	4	MI with cardiogenic shock
3	1	male or female to 40
	2	male or female from 41 to 50
	3	female from 51 to 60
	4	male from 51 to 60
	5	male or female from 61 to 70
	6	male or female older than 71
4	1	no associated diseases
	2	existence of the associated diseases that are more frequently do not lead to the lethal outcome, such as the kidney deficiency, anemia (not linked to the oncologic diseases), chronic obstructive disease of lungs (CODL)
	3	the presence of the associated diseases, is more frequent than leading to the lethal outcome, for example, diabetes mellitus
5	1	no pain
	2	weak pain, aching pain without irrodation (spontaneous)
	3	pain with irrodiation to left shoulder (lasting longer)
	4	constant pain, periodically intensifies in the heart area, radiating to the left shoulder, shoulder blade, arm, left part of the head, not stopped with medication

At the third stage all values of the i-th criterion are represented in the form of the rank qualitative scale in the ascending order of the attribute of specificity of the degree of severity of RIM (from the best to the worst) (Fig. 1):

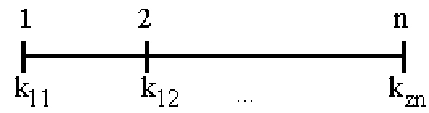

Fig. 1. Rank qualitative scale of the criteria of evaluation of the state of a patient

At the fourth stage we determine the set of all hypothetically possible states of patient, which is the Cartesian product of the sets of values of the criteria K, of the form:

$$A = K_1 \times K_2 \times K_3 \times K_4 \times K_5,$$

cardinality of the set A:

$$|A| = |K_1| \times |K_2| \times |K_3| \times |K_4| \times |K_5|. \quad (1)$$

Set A determines the space of states of patient, subject to the classification:

$$A = \{a_i\}, \quad (2)$$

where $a_i=(a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5})$ is the vector estimation of the i-th state of patient, $a_{ij} \in K_i$.

At the fifth stage we determine the class for the prediction of RIM, to which the patient may be assigned with his vector of state.

The ratio of strict domination is introduced for this purpose: vector of state a_i of the patient is better than the vector of state a_j of the patient if, by all the criteria, vector of state a_i has estimations that are not worse than vector of state a_j, and at least by one criterion has a bettter estimation.

All a_{ij} states of patients are coded, using only ordinal numbers of the values of the corresponding criteria and they are ordered in the lexicographical order (from the best to the worst).

Then the set A is divided into nonintersecting classes of the state of patient, regulated by the growth of manifestation of RIM:

$$A = \cup_{i=1}^{l} C_i, \quad (3)$$

where C_i is the i-th class of the state of patient, i=1...l is the number of the classes of states of patient ordered from the best to the worst.

Classes C_i are meaningfully the diagnoses, regulated from the best to the worst. In this case, the principle is used: for the best states, the worst classes are forbidden (and vice versa – for the worst states, forbidden are the best classes). Formally this means that if:

– $a_i \in C_1$ then the best vector of state $a_j \notin C_2 \cup C_3$;

– $a_i \in C_2$ then the best vector of state $a_j \notin C_3$.

Belonging a_i of the vector estimation of the state of patient to one or another class is determined by the survey of experts (straight classification), for the best/worst states, forbidden are the worst/best classes (indirect classification).

Since the belonging of states to several classes C_i is the sign of incompletness of the process of classification of the states of patients, then it is necessary to determine the set of numbers of the classes G_i, permissible for the vector of state a_i.

The quantity of indirectly classified vectors of state a_i depends on what point of a multidimensional space, formed by the Cartesian product of the scales of criteria of evaluation of the state of patient is presented to an expert, and on the class of the state C_i to which he/she will assign the state in question. Taking into account that the probability p_{il} of belonging a_i of the vector of state to the C_i class depends on the proximity of this state to this class, then for each a_i vector of state it is possible to determine the evaluation of the obtained information for possible answers of experts and the estimation of proximity of the state of patient to each of the possible classes C_i.

The process of classification consists in the determination of the most informative state, its direct classification (presentation to a decision making person (DMP) and assignment to a particular class) and indirect classification of the best states (forbidding the worst classes) and the worst states (forbidding the best classes).

As a result of classification, each state of patient must belong only in one class, that is, a definite diagnosis is formulated for the patient in the given state. Thus, it is necessary to determine a number of classes C_i, forbidden for the best and worst states, when assigning the vector of state a_i to one or another class.

At the sixth stage, for minimization of the number of variants presented to experts during the division of set A into C_i classes, in each iteration of a multi-iterative process, we determine a_i vector of state, which provides with any answer of experts for a maximum amount of expected information in relation to other states. Let us describe this procedure.

The measure of proximity of vector of state $a_i \in A$ to a certain class C_i is introduced, which will characterize the probability that a_i will be assigned by experts to class C_i. It is necessary to determine a maximally informative state of the patient, which will be presentred to DMP, for assigning to the class. For this purpose, we calculate the center s_l of non-empty class C_l in the following way:

$$s_l=(s_{l1},s_{l2},...,s_{lj}),$$

where

$$s_{lj}=\left(\sum_{a_i \in C_l} a_{ij}\right)/|C_l|. \tag{4}$$

Then we determine distance d_{il} from the state a_i to the center s_l of class C_l by the following expression:

$$d_{i,l}=\sum_{j=1}^{n(i)}|a_{ij}-s_{lj}|. \tag{5}$$

After this, we determine a maximally possible distance d_{max} between two vectors of state, which belong to the set A (the best and the worst state):

$$d_{max}=\sum_{j=1}^{n(i)}(n(i)-1). \tag{6}$$

Then we determine the probability p_{il} of the fact that DMP will assign state a_i to the valid class C_l:

$$p_{i,l}=(d_{max}-d_{i,l})/\sum_{w\in G_i}(d_{max}-d_{i,w}). \tag{7}$$

where

$$w=\prod_{j=1}^{i} n(i)$$

is the total number of possible combinations of estimations according to the i-th criteria with n(i) number of values of estimations by the scale. The probability of p_{il} is the larger the less is the distance between the vector of estimations a_i and the center of class C_l. It is possible to consider that in this case the possibility of assigning vector of state a_i to the class C_l will be higher.

Let us denote through $g_{j>i,l}$ the number of classes (amount of additional information), forbidden for the best vector of state a_i if it is assigned to class C_l. Its value is calculated by formula:

$$g_{j>i,l}=\left|G_j \cap \{l+1,\dots,l\}\right|. \tag{8}$$

Let us denote through $q_{j>i,l}$ the number of classes, forbidden for all best states with the vector of state a_i assigned to class C_l. Its value is calculated by formula:

$$q_{j>i,l}=\sum_{a_j>a_i} q_{j>i,l}, \tag{9}$$

where the summing is conducted by all vectors of state a_j that are better than a_i.

The best classes are forbidden for the worst states. Formally this means that if: $a_i \in C_l$, then the classes C_l,..., $C_{l-1}$are forbidden.

Let us denote through $q_{j<i,l}$ the number of classes, forbidden for all the worst states with the vector of state a_i assigned to class C_i. Its value is calculated by formula:

$$q_{j<i,l}=\left|G_j \cap \{1,\dots,l-1\}\right|. \tag{10}$$

Let us denote through $q_{j<i,l}$ the number of classes, forbidden for all the worst states with the vector of state a_i assigned to class C_l. Its value is calculated by formula:

$$q_{j<i,l}=\sum_{a_j<a_i} q_{j<i,l}, \tag{11}$$

where the summing is conducted by all vectors of state a_j that are worse than a_j.

Let us denote through $q_{i,l}$ the number of additional information in relation to all classes of the states of patient, forbidden for all other states at $a_i \in C_l$ Its value is calculated as follows:

$$q_{i,l}=q_{>i,l}+q_{<i,l}. \tag{12}$$

Then for each vector of state a_i we determine evaluation of its informativeness I_{il} relative to class C_l:

$$I_{il}=p_{il}\times q_{il} \tag{13}$$

and we determine the uniform quantitative index of informativeness I^{max} of each, not yet classified, vector of state $a_i \in A$ [12]:

$$I^{max}=\sum_{l\in G_i} I_{il}. \tag{14}$$

At the seventh stage, as a result of ordinal classification of the states of patients, the decisive rule is formulated for the doctor with all a_i vectors of state, distributed by the Ci classes. The classes of possible states, to which a patient may be assigned, are presented in Table 3.

Table 3

Classes of possible states of patients

Class	Value
C_1	without relapse
C_2	relapse is possible without lethal outcome
C_3	relapse is possible with lethal outcome

At the eighth stage, based on the data obtained as a result of classification, we synthesize interpolation diagnostic polynomial. This polynomial will make it possible to determine the measure of proximity of the vector of state a_i to class C_i and thus to predict possible state of a patient.

Interpolation diagnostic polynomial is presented in the form of function whose arguments are the vectors with the code of the state of a patient, formed by the ordinal numbers of values of the corresponding criteria, and the values are the numbers of classes, to which the states of the patient belong:

$$f: K_1 \times \ldots \times K_i \to \{1 \ldots l\},$$

where l is the quantity of classes of the states of patients. Then the interpolation diagnostic polynomial may take the following form:

$$f\left(k_{i1} \ldots k_{ij} \ldots k_{zn}\right) = \sum_{i_1=1}^{|K_1|} \ldots \sum_{i_j=1}^{|K_j|} \ldots \sum_{i_n=1}^{|K_n|} b_{i_1 \ldots i_j \ldots i_n} k_{i1}^{|K_1|-i_1} \ldots \times$$
$$\times k_{ij}^{|K_j|-i_j} \ldots \times k_{zn}^{|K_n|-i_n}, \quad (15)$$

where $(k_{i1} \ldots k_{ij} \ldots k_{zn})$ is the vector of state of patient; $b_{i1 \ldots ij \ldots in}$ are the coefficients at the appropriate term.

It is known that two points uniquely determine a straight line on the plane, three points – parabola, four – cubic parabola, etc. That is why the corresponding interpolation polynomial has a degree one unity less than the number of points, in which the value of function is known.

At the ninth stage a doctor-cardiologist collects anamnesis of the patient, which include: results of inspection, of conducted clinical-instrument (electro- and echocardiography, angiography, etc.) and clinical laboratory examinations (clinical and biochemical analyses of blood and urine, coalography, etc.).

At the tenth stage, based on the experience and by the results of analysis of information, received about the patient, a doctor selects the values of each attribute of the degree of manifestation of RIM k_{ij} for the i-th criterion from the set K, proposed by the experts.

At the eleventh stage a doctor forms the vector of state of patient a_i and, with the use of polynomial (15), determines the values of measure of proximity to each class C_i. After this, based on the obtained values, a possible class of the state is determined. Thus, a patient will be assigned to class C_1 if

$$(K_1 \in 1..4) \wedge (K_2 \in 1..4) \wedge (K_3 \in 1..5) \wedge$$
$$\wedge (K_4 \in 1..3) \wedge (K_5 \in 1..4),$$

C_2 – if

$$(K_1 \in 1..4) \wedge (K_2 \in 1..4) \wedge (K_3 \in 1..6) \wedge$$
$$\wedge (K_4 \in 1..3) \wedge (K_5 \in 1..4)$$

and C_3 – if

$$(K_1 \in 1..4) \wedge (K_2 \in 1..4) \wedge (K_3 \in 2..6) \wedge$$
$$\wedge (K_4 \in 1..3) \wedge (K_5 \in 1..4).$$

At the twelfth stage, a diagnostic conclusion is formed for making a final decision by the cardiologist.

5. Ordinal classification of the states of a patient for predicting relapsing myocardial infarction

The process of ordinal classification of the states of a patient based on the proposed method is the one-time multi-iterative procedure, which can be presented in the following form.

Let us determine cardinality of the set |A| of hypothetically possible states of patient by (1):

$$|A| = 4 \times 4 \times 6 \times 3 \times 4 = 1152.$$

The space of states of a patient, due to be classified, takes the form by (2):

$$A = \{a_i\}_{i=1}^{1152}.$$

Initially (in the first iteration), class C_1 includes the vector of state a_1=(1, 1, 1, 1, 1). This class is not empty, it consists of one element, which determines the center of the class: s_1=(1, 1, 1, 1, 1). By analogy, for class C_3 – s3=(4; 4; 6; 3; s_3=(4; 4; 6; 3; 4).

The center of class C_2 is defined as the arithmetic mean of values of the corresponding components of the centers of classes s_1 and s_3 by (4):

$$s_2 = \left(\frac{1+4}{2}; \frac{1+4}{2}; \frac{1+6}{2}; \frac{1+3}{2}; \frac{1+4}{2}\right) = (2{,}5; 2{,}5; 3{,}5; 2; 2{,}5).$$

Let us calculate the informativeness I_i of vector of state (1, 1, 1, 1, 2).

Let us first determine distances d_1, d_2 and d_3 from this state to the centers of classes C_1, C_2 and C_3 according to (5):

$$d_1 = |1-1| + |1-1| + |1-1| + |1-1| + |1-2| = 1;$$

$$d_2 = |2{,}5-1| + |2{,}5-1| + |3{,}5-1| + |2-1| + |2{,}5-2| = 7;$$

$$d_3 = |4-1| + |4-1| + |6-1| + |3-1| + |4-2| = 15.$$

Maximum distance (distance between the best and the worst state) by (6) can be determined as

$$d_{max} = |1-4| + |1-4| + |1-6| + |1-3| + |1-4| = 16.$$

Then by (7) we will determine probabilities p_1, p_2 and p_3 that DMP will assign state (1, 1, 1, 1, 2) to classes C_1, C_2 and C_3, respectively:

$$p_1 = \frac{d_{max} - d_1}{d_{max} - d_1 + d_{max} - d_2 + d_{max} - d_3} = \frac{15}{15+9+1} = \frac{15}{25};$$

$$p_2 = \frac{d_{max} - d_2}{d_{max} - d_1 + d_{max} - d_2 + d_{max} - d_3} = \frac{9}{25};$$

$$p_3 = \frac{d_{max} - d_3}{d_{max} - d_1 + d_{max} - d_2 + d_{max} - d_3} = \frac{1}{25}.$$

If DMP assignes vector of state (1, 1, 1, 1, 2) to class C_1, then it will give no additional information about the belonging of other states to one or another class. Actually, based on the fact that the best states cannot belong to the worst classes, for the vector of state a_1 it is necessary to forbid classes C_2 and C_3. From the other hand, based on the fact that the worst states cannot belong to the best classes, then for these states it is necessary to forbid the classes, which are better than C_1. But there are no such classes because by condition C_1 is the best class, that is, amount q_1 of additional information with DMP's assigning of vector of state (1, 1, 1, 1, 2) to class C_1 is equal to zero (q_1=0).

If DMP assigns vector of state (1, 1, 1, 1, 2) to class C_2, then for the best state a_1, it is necessary to forbid the worst class C_3. And for the worst states, it is necessary to forbid the best class C_1. Their quantity is determined based on that the values of all components of all the worst states are not better than (1, 1, 1, 1, 2):

– the first component changes from 1 to 4 (4 variants are possible);

– the second component changes from 1 to 4 (4 variants are possible);

– the third component changes from 1 to 6 (6 variants are possible);

– the fourth component changes from 1 to 3 (3 variants are possible);

– the fifth component changes from 2 to 4 (3 variants are possible);

Thus, there are 4×4×6×3×3=864 variants of combinations of the values of components. These combinations include vector of state (1, 1, 1, 1, 2) and the worst vector of state (4, 4, 6, 3, 4), it does not belong to class C_1 because originally $a_{1152} \in C_3$. Therefore, these states must be excluded from the total quantity of variants.

Consequently, for each of 862 worst variants, the best class C_1 will be forbidden. In other words, amount q_2 of additional information with DMP's assigning of vector of state (1, 1, 1, 1, 2) to class C_2 is equal to 862: $q_2 = 1 \times 862 = 862$.

If DMP assigns vectors of state (1, 1, 1, 1, 2) to class C_3, then for the best state a_1, it is necessary to forbid the worst class, but there are no such classes because by tcondition C_3 is the worst class. But for the worst states, it is necessary to forbid the best classes C_1 and C_2. In other words, amount q_3 of additional information with DMP's assigning vector of state (1, 1, 1, 1, 2) to class C_3 is equal to: $q_3 = 2 \times 862 = 1724$.

Next, in accordance with (13), let us determine the estimation of informativeness I_1, I_2 and I_3 of vector of state (1, 1, 1, 1, 2) relative to classes C_1, C_2 и C_3, respectively:

$$I_1 = p_1 \times q_1 = \frac{15}{25} \times 0 = 0;$$

$$I_2 = p_2 \times q_2 = \frac{9}{25} \times 862 = 310,32;$$

$$I_3 = p_3 \times q_3 = \frac{1}{25} \times 1724 = 68,96.$$

Next by (14) we will calculate the uniform quantitative index I^{max} of vector of state (1, 1, 1, 1, 2).

$$I^{max} = I_1 + I_2 + I_3 = 0 + 310,32 + 68,96 = 379,28.$$

Thus, after the first iteration, for the selected vector of state we obtained the mean expected amount of additional information, with its presentation to DMP, equal to 379,28. All subsequent iterations are carried out analogously. The process of classification is completed as soon as each vector of state of a patient belongs only to one class of state.

The known values of function in the points, assigned by the coding vectors, are used for finding the values of coefficients of diagnostic polynomial. In this case, the problem comes down to solving the system of 1152 linear equations with 1152 unknowns of the form:

$$\begin{cases} 1b_{11111} + 1b_{11112} + \ldots + 1b_{44634} = 1; \\ 5668704b_{11111} + 1889568b_{11112} + \ldots + 1b_{44634} = 2; \\ 18345885696b_{11111} + 4586471424b_{11112} + \ldots + 1b_{44634} = 3. \end{cases}$$

The matrix of coefficients of this system has dimensions of 1152 by 1152 and consists of 1 327 104 elements and its solution can be found based on the Gauss method.

Fragments of the matrix of coefficients of the system of equations and the inverse matrix are given in Fig. 2, 3, respectively.

Fig. 4 demonstrates fragments of the column with values of the obtained coefficients.

	F	G	H	I	J	K	ARJ	ARK	ARL	ARM
6	1	1	1	1	1	1	1	1	1	1
7	8	4	2	1	8	4	8	4	2	1
8	27	9	3	1	27	9	27	9	3	1
9	64	16	4	1	64	16	64	16	4	1
10	4	4	4	4	2	2	1	1	1	1
11	32	16	8	4	16	8	8	4	2	1
12	108	36	12	4	54	18	27	9	3	1
13	256	64	16	4	128	32	64	16	4	1
14	9	9	9	9	3	3	1	1	1	1
15	72	36	18	9	24	12	8	4	2	1
1148	859963392	286654464	95551488	31850496	859963392	286654464	27	9	3	1
1149	2038431744	509607936	127401984	31850496	2038431744	509607936	64	16	4	1
1150	127401984	127401984	127401984	127401984	63700992	63700992	1	1	1	1
1151	1019215872	509607936	254803968	127401984	509607936	254803968	8	4	2	1
1152	3439853568	1146617856	382205952	127401984	1719926784	573308928	27	9	3	1
1153	8153726976	2038431744	509607936	127401984	4076863488	1019215872	64	16	4	1
1154	286654464	286654464	286654464	286654464	95551488	95551488	1	1	1	1
1155	2293235712	1146617856	573308928	286654464	764411904	382205952	8	4	2	1
1156	7739670528	2579890176	859963392	286654464	2579890176	859963392	27	9	3	1
1157	18345885696	4586471424	1146617856	286654464	6115295232	1528823808	64	16	4	1

Fig. 2. Fragment of the matrix of coefficients of the system of equations (parts of the first and last ten lines, and the first six and last four columns are shown)

The corresponding fragment of the interpolation diagnostic polynomial (the first and last ten members) by (15) takes the following form:

$$
\begin{aligned}
&f\left(k_{i1},k_{i2},k_{i3},k_{i4},k_{i5}\right)=-0{,}000000012k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{2}k_{i5}^{3}+\\
&+0{,}036863517k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{2}k_{i5}^{2}-0{,}184317327k_{1}^{3}k_{2}^{3}k_{3}^{5}k_{4}^{2}k_{5}^{1}+\\
&+0{,}203356591k_{i1}^{3}k_{2}^{3}k_{i3}^{5}k_{i4}^{2}k_{i5}^{0}+0{,}024575657k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{1}k_{i5}^{3}-\\
&-0{,}331771127k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{1}k_{i5}^{2}+1{,}156269917k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{1}k_{i5}^{1}-\\
&-1{,}092940152k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{1}k_{i5}^{0}-0{,}049151266k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{0}k_{i5}^{3}+\\
&+0{,}49826412k_{i1}^{3}k_{i2}^{3}k_{i3}^{5}k_{i4}^{0}k_{i5}^{2}+\ldots+35570{,}50611k_{i1}^{0}x_{i2}^{0}k_{i3}^{0}k_{i4}^{2}k_{i5}^{1}-\\
&-25819{,}50353k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{2}k_{i5}^{0}-7818{,}834735k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{1}k_{i5}^{3}+\\
&+63008{,}0101k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{1}k_{i5}^{2}-154453{,}1878k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{1}k_{i5}^{1}+\\
&+111034{,}512k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{1}k_{i5}^{0}+7652{,}334455k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{0}k_{i5}^{3}-\\
&-61110{,}50809k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{0}k_{i5}^{2}+148297{,}1837k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{0}k_{i5}^{1}-\\
&-105389{,}0096k_{i1}^{0}k_{i2}^{0}k_{i3}^{0}k_{i4}^{0}k_{i5}^{0}.
\end{aligned}
$$

	F	G	H	I	J	ARM
1159	1,92902E-05	-5,78705E-05	5,78704E-05	-1,92901E-05	-3,85803E-05	1,92901E-05
1160	-0,000173611	0,000462964	-0,000405093	0,000115741	0,000347223	-0,000115741
1161	0,000501544	-0,001099539	0,000810187	-0,000212192	-0,001003087	0,000212191
1162	-0,000462963	0,000694445	-0,000462964	0,000115741	0,000925926	-0,000115741
1163	-9,64507E-05	0,000289352	-0,000289352	9,64507E-05	0,000154321	-5,78704E-05
1164	0,000868056	-0,002314817	0,002025465	-0,000578704	-0,00138889	0,000347222
1165	-0,002507718	0,00549769	-0,00405093	0,001060958	0,004012348	-0,000636574
1166	0,002314816	-0,003472225	0,002314817	-0,000578704	-0,003703705	0,000347222
1167	0,000115741	-0,000347222	0,000347222	-0,000115741	-0,000115741	3,85802E-05
1168	-0,001041667	0,00277778	-0,002430557	0,000694445	0,001041668	-0,000231481
2301	-208,0000187	456,0000453	-336,0000369	88,00001008	416,0000299	-0,916667081
2302	192,000011	-288,0000268	192,0000219	-48,00000598	-384,0000174	0,500000243
2303	40,0000043	-120,0000102	120,0000083	-40,00000224	-64,00000669	0,250000087
2304	-360,000031	960,0000743	-840,0000605	240,0000165	576,0000485	-1,50000065
2305	1040,000065	-2280,000158	1680,00013	-440,0000354	-1664,000103	2,750001436
2306	-960,0000374	1440,000092	-960,0000757	240,0000208	1536,000059	-1,500000834
2307	-48,00000337	144,000008	-144,0000066	48,00000179	48,00000519	-0,166666733
2308	432,0000243	-1152,000059	1008,000048	-288,0000131	-432,0000375	1,000000501
2309	-1248,000051	2736,000125	-2016,000103	528,0000282	1248,000079	-1,833334448
2310	1152,000029	-1728,000072	1152,00006	-288,0000165	-1152,000045	1,00000653

Fig. 3. Fragment of the inverse matrix (parts of the first and last ten lines, and the first five and last of columns are shown)

1159	-1,21405E-08	2301	35570,50611
1160	0,036863517	2302	-25819,50353
1161	-0,184317327	2303	-7818,834735
1162	0,203356591	2304	63008,0101
1163	0,024575657	2305	-154453,1878
1164	-0,331771127	2306	111034,512
1165	1,156269917	2307	7652,334455
1166	-1,092940152	2308	-61110,50809
1167	-0,049151266	2309	148297,1837
1168	0,49826412	2310	-105389,0096

Fig. 4. Fragments of the column with values of the coefficients

In the process of predicting relapsing myocardial infarction, a doctor uses the obtained polynomial. Depending on the obtained value of function, in accordance with the decisive rules, a possible class of the state is determined, which a patient can experience.

6. Discussion of results of predicting relapsing myocardial infarction using the designed method

To verify the developed method, we examined the histories of diseases from186 people who had acute MI in the cardiological department of the 4th municipal clinical hospital of ambulance and emergency medical aid named after Professor Meshchaninov (Kharkov, Ukraine). All examinations were divided into three groups according to the proposed classification of possible states C_i: group 1 – patients without relapse (126 cases), group 2 – RIM without the lethal outcome (26 cases), group 3 – RIM with the lethal outcome (34 cases).

All observations were divided into a training sample (70 % of examinations, 87 cases of group 1, 18 cases of group 2 and 23 cases of group 3) and the test sample (30 % of examinations, 39 examinations from group 1, 8 examinations from group 2 and 11 examinations from group 3) (Table 4).

Based on the data on the patients, we performed estimation of deviation in values of the interpolation diagnostic polynomial from the assigned tabular values. Minimum deviation is approximately equal to zero and it is observed in point (1, 1, 1, 1, 1), the largest deviation, approximately equal to 0,0001, is in point (4, 4, 6, 3, 4).

The results of prediction for the training sample, obtained with the aid of the developed method, demonstrated that:

– 2 people from group 1 were mistakenly assigned to group 2; 85 – determined correctly;

– 1 person from group 2 was mistakenly assigned to group 1; 17 people – determined correctly;

– for group 3 all cases were determined without mistakes.

Thus, in the training sample, the prediction of RIM is determined correctly for 97,7 % of the object, which is by 2,7 % higher in comparison to the method–prototype [7].

Based on the comparison of predicted and actual states, in the test sample, the prediction of RIM is determined correctly for 98,3 % of the objects.

Table 4

Results of prediction of RIM with the aid of the proposed method and the method-prototype

Group	N	Designed method				Method-prototype	
		Training sample		Test sample			
		n	n’	n	n’	n	n’
1	126	85	2	39	0	121	5
2	26	17	1	8	0	23	3
3	34	23	0	10	1	32	2
Total	186	125	3	57	1	176	10

General results of prediction, obtained with the aid of the designed method, demonstrated that:

– 2 people from group 1 were mistakenly assigned to group 2; 125 – determined correctly;

– 1 person from group 3 was mistakenly assigned to group 2;

– 25 people were determined correctly;

– 1 person from group 2 was mistakenly assigned to group 1;

– 35 people were determined correctly.

The estimation of the predictive properties of the developed method was conducted using the ROC-analysis. The value of area under the ROC-curve, which makes it possible to estimate the diagnostic (predictive) value of the developed diagnostic polynomial, comprised 0,987 (0.981, 1.000) that indicates excellent quality of the model (Fig. 5). Sensitivity of the designed method reached 0,984, specificity – 0.967.

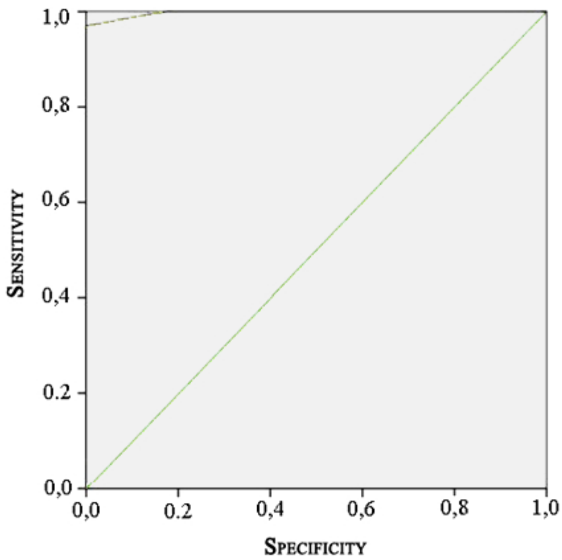

Fig. 5. ROC-curve

The developed method of predicting RIM on the basis of a mathematical model, represented in the form of interpolation diagnostic polynomial, is of practical interest for a doctor-cardiologist. This interest is in the possibility of predicting the relapse of disease and sudden coronary death by the qualitative indices that do not require considerable labor costs. The designed method also considers the totality of the attributes of disease, their combination and mutual effect, which is especially important when predicting such a complication.

The developed method may find further application when predicting not only RIM but also other human cardiovascular system diseases when it is also important to take the opinions of experts into account.

7. Conclusions

1. As a result of the study, based on expert estimations, we determined a set of criteria and classes of estimation of the states of patients, which makes it possible to conduct ordinal classification of vectors of these states for the prediction of RIM. The following criteria were selected: the type of infarction; heart failure, determined by the Killip scale; hypertonia with regard to gender-age gradation; pain syndrome (pain intensity); the associated diseases (kidney deficiency, chronic obstructive disease of lungs, diabetes mellitus, etc.).

2. A decisive rule is designed of the classification of possible states of a patient, based on the formalization of the experts' knowledge, which includes all possible vectors of states according to the assigned set of criteria and making it possible to determine one of the three possible classes for the prediction of RIM: without the relapse, the relapse is possible without the lethal outcome or relapse is possible with the lethal outcome.

3. The synthesized interpolation diagnostic polynomial, based on the use of full verbal classification, makes it possible to determine belonging of the state of a patient to the class of RIM by calulating the value of polynomial for the vector of state of a patient.

4. We developed a method for predicting RIM, based on the interpretation of the expert knowledge and the use of diagnostic interpolation polynomial, which allows, by quality indices that do not require considerable labor costs, determination of the possibility of relapse of the disease and sudden coronary death.

Acknowledgment

Authors of the article express appreciation to the Head of Department of Urgent States Medicine, Medicine of Catastrophes and Military Medicine of the Kharkov Medical Academy of Postgraduate Education, the Head of Center for "Infectious Endocarditis" (Ukraine), Academician of International Energy-Information Academy, Doctor of Medical Sciences, Professor Vadim Vladimirovich Nikonov for the provided expert information and experimental data, used in the course of verification of the developed method. While data processing, we observed all rules and regulations for confidentiality of the given information.

References

1. Prevention of Recurrences of Myocardial Infarction and Stroke Study. The PREMISE programme: country projects [Electronic resource]. – WHO, 2016. – Available at: http://www.who.int/cardiovascular_diseases/priorities/secondary_prevention/country/en/index1.html
2. Mendis, S. World Health Organization definition of myocardial infarction: 2008-09 revision [Text] / S. Mendis, K. Thygesen, K. Kuulasmaa, S. Giampaoli, Mahonen, M., Ngu Blackett, K., & Lisheng, L. // International Journal of Epidemiology. – 2010. – Vol. 40, Issue 1. – P. 139–146. doi: 10.1093/ije/dyq165
3. Vysotskaya, E. V. A method for prognosis of primary open-angle glaucoma [Text] / E.V. Vysotskaya, A. N. Strashnenko, Y. A. Demin, I. V. Prasol, C. A. Sinenko // International Review on Computers and Software. – 2013. – Vol. 8, Issue 8. – P. 1943–1949.
4. Ryabkova, E. B. Metod sinteza nechetkih reshayushchih pravil po informacii o geometricheskoj strukture mnogomernyh dannyh [Text] / E. B. Ryabkova, H. A. Korenevskij // Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta. – 2011. – Vol. 7, Issue 8. – P. 128–137.
5. Vorob'eva, O. M. Sintez reshayushchih pravil dlya prognozirovaniya infarkta miokarda po pokazatelyam perekisnogo okisleniya lipidov i antiokislitel'noj aktivnosti [Text] / O. M. Vorob'eva, V. N. Mishustin, I. V. Chernova // Izvestiya YUZGU. – 2012. – Vol. 2, Part 2. – P. 249–252.
6. Vorob'eva, O. M. Analiz rasprostranennosti i prognozirovanie infarkta miokarda, gipertonicheskoj bolezni v g. Kurske [Text] / O. M. Vorob'eva, I. V. Kolomiec // Vestnik novyh medicinskih tekhnologij. – 2009. – Vol. 2. – P. 16–17.

7. Pat. № 2197889 RF, MPK A61V 5/00. Sposob prognozirovaniya vozmozhnosti vozniknoveniya recidiviruyushchego infarkta miokarda [Text] / Sazonova N. S., Susnina A. V., Yakushin S. S. – zayavitel' i patentoobladatel' Ryazanskij gosudarstvennyj medicinskij universitet im. akad. I. P. Pavlova; declarated: 12.05.1999; published: 10.02.2003.
8. Nakatani, D. Incidence, Predictors, and Subsequent Mortality Risk of Recurrent Myocardial Infarction in Patients Following Discharge for Acute Myocardial Infarction [Text] / D. Nakatani, Y. Sakata, S. Suna, M. Usami, S. Matsumoto, M. Shimizu // Circulation Journal. – 2013. – Vol. 77, Issue 2, – P. 439–446. doi: 10.1253/circj.cj-11-1059
9. Leander, K. Primary risk factors influence risk of recurrent myocardial infarction/death from coronary heart disease: results from the Stockholm Heart Epidemiology Program (SHEEP) [Text] / K. Leander, B. Wiman, J. Hallqvist, T. Andersson, A. Ahlbom, U. de Faire // European Journal of Cardiovascular Prevention & Rehabilitation. – 2007. – Vol. 14, Issue 4. – P. 532–537. doi: 10.1097/hjr.0b013e328012e3cc
10. Chaudhry, S. I. National Trends in Recurrent AMI Hospitalizations 1 Year After Acute Myocardial Infarction in Medicare Beneficiaries: 1999-2010 [Text] / S. I. Chaudhry, R. F. Khan, J. Chen, K. Dharmarajan, J. A. Dodson, F. A. Masoudi // Journal of the American Heart Association. – 2014. – Vol. 3, Issue 5. – P. e001197–e001197. doi: 10.1161/jaha.114.001197
11. Cui, J. Semi-parametric risk prediction models for recurrent cardiovascular events in the LIPID study [Text] / J. Cui, A. Forbes, A. Kirby, I. Marschner, J. Simes, D. Hunt et. al. // BMC Medical Research Methodology. – 2010. – Vol. 10, Issue 1. doi: 10.1186/1471-2288-10-27
12. Larichev, O. I. Verbal'nyj analiz reshenij [Text] / O. I. Larichev. – Moscow: Nauka, 2006. – 181 p.
13. Vysockaya, E. V. Metod opredeleniya znachimosti mnenij ehkspertov pri formirovanii ehkspertnoj gruppy dlya prinyatiya medicinskih reshenij [Text] / E. V. Vysockaya // Sistemi obrobki informacii. – 2014. – Vol. 2, Issue 118. – P. 216–221.
14. Vysockaya, E. V. Koncepciya povysheniya kachestva prinyatiya reshenij vracha [Text]: tezi dop. XXIII mizhnar. nauch.-prakt. konf. / E. V. Vysockaya // Informacijni tekhnologii: nauka, tekhnika, tekhnologiya, osvita, zdorov'ya. – 2015. – Part III. – P. 32.
15. Vysockaya, E. V. Medical decisions under uncertainty [Text] / E. V. Vysockaya, A. I. Pecherskaya, A. I. Dovnar' // Eastern-European Journal of Enterprise Technologies. – 2014. – Vol. 3, Issue 4 (69). – P. 16–21. doi: 10.15587/1729-4061.2014.23935
16. Vysockaya, E. V. Razrabotka strategii povysheniya kachestva podderzhki prinyatiya reshenij vracha obshchej praktiki [Text] / E. V. Vysockaya // Sistemi obrobki informacii. – 2014. – Vol. 5, Issue 121. – P. 142–145.