Neural network classifiers with descriptors obtained on the basis of analysis of the system rhythms in intellectual prediction systems for non-hospital pneumonia

M B Myasnyankin¹, A A Kuzmin¹ and S A Filist¹

¹ Southwest State University, Russia, 94, 50 years of October st., Kursk, 305040, Russia

E-mail: SFilist@gmail.com

Abstract. The description of the synthesis of the risk classifier of non-hospital (community-acquired) pneumonia based on predictors is presented in the article. Predictors are compiled by analyzing the synchronicity of the systemic rhythms of the cardiovascular system and the respiratory system. Spectr of cardiac signals and wavelet coefficients of cardiac signals are used as sources of systemic rhythms. An algorithm for classifying the risk of community-acquired pneumonia in terms of synchronicity of system rhythms is proposed in the article.

1. Introduction
A systematic approach is necessary when assessing the functioning of the human body [2]. The technology for assessing the cumulative risk of major socially significant diseases (called "ORISCON") has been developed recently. This technology is the original expert system. It assesses the risk of developing major chronic diseases that have common risk factors and common conceptual frameworks for prevention [3]. The main socially significant diseases include major cardiovascular diseases (coronary artery disease, cerebrovascular disease and other diseases), cancer, lung disease, endocrine diseases, diseases of the digestive system, as well as a number of loosely established causes.

The cardiorespiratory system, consisting of the cardiovascular system and the respiratory system, is the most sensitive indicator of the physiological state of the body. The structures of the brain that are responsible for the rhythmogenesis of respiration and heart are connected with each other. The coordination of the rhythmogenesis of respiration and heart rate is so significant that these structures form a functional unity. In the early stages of the disease or in a state of pre-disease (in a premorbid state), this unity is disrupted, which is primarily reflected in the change in the spectral characteristics of systemic rhythms and their correlation indicators. Heart rate analysis is used to monitor human adaptive potentials, which can be used as a marker of human immunity. Bronchial obstruction affects the variability of heart rate and respiration. Breathing rhythm analysis is less covered in the scientific literature. Nevertheless, a number of experimental works are known, which show that the electromyogram of the respiratory muscles is an indicator of the state of the respiratory system. In these works, it was shown that the electrical activity of the respiratory muscles is an informative parameter in the diagnosis of violations of the functional state of the respiratory system [4, 5, 6].
2. Methods

The technical solutions proposed in [7, 8, 9] are taken as the basis for building a model for predicting community-acquired pneumonia. If we consider the pulse wave as a carrier signal $s_1(t)$, which is modulated by one of the system rhythms $s_2(t)$, then to obtain $s_2(t)$ we can use a correlation detector that performs a mathematical operation

$$R_{s s_1}(\tau) = \frac{1}{T} \int_0^T s(t) s_1(t + \tau) dt,$$  \hspace{1cm} (1)

where $T$ is the observation interval of the pulse wave signal $s(t)$, $\tau$ is the time shift.

The reference signal $s_1(t)$, which is not available for measurement and analysis, is necessary for us to use equation (1) in order to form descriptors. Since the signal $s_1(t)$ is the signal of the cardiocycle, which occupies a certain frequency band. This frequency band varies within a fairly wide range from patient to patient and at different points in the observation of the same patient. With this in mind, we use the wavelet function as the reference signal $s_1(t)$ in expression (1). The frequency band of this wavelet function corresponds to the frequency band of the system rhythm selected as an indicator of the synchronicity of the system rhythms. Therefore, expression (1) is calculated for $s_1(t) = \Psi(t / a^*)$, where $\Psi(t)$ is the mother wavelet, and $a^*$ is the wavelet scale with the frequency range corresponding to the selected systemic rhythm. In this case, we call this expression (1) as the resonant wavelet slice signal, and we call a wavelet with a scale $a^*$ as resonant wavelet. The resonant wavelet was determined empirically. To construct the feature space, we do not use resonant wavelet slices directly, since they have a large dimension, which exceeds the dimension of the original space. In this case, we select only with a narrow-band signal characterizing the respiratory component. The mutual spectrum of signals can serve as a tool that allows the transition to a narrow-band informative signal zone.

The mutual spectrum $X_{12}(\omega)$ of two signals, $x_1(t)$ with the spectrum $X_1(\omega)$ and $x_2(t)$ with the spectrum $X_2(\omega)$, is a complex number. If we do not consider the phase characteristics of the signals, then it can be defined as

$$X_{12}(\omega) = X_1(\omega) \cdot X_2(\omega).$$  \hspace{1cm} (2)

Analyzing formula (2) and wavelet transforms of cardiac signals, we come to the conclusion that the mutual spectrum of the cardio signal and the resonant wavelet will be nonzero in a small spectral range that the resonant wavelet occupies. This spectral range does not exceed 0.4 Hz. At a frequency resolution of 0.033 Hz (in the analysis window with a duration of 30 seconds), the mutual spectrum has no more than several tens of counts other than zero at 30,000 initial samples of the cardiac signal.

The cut of the wavelet plane of the cardiac signal in the respiration frequency band and the corresponding mutual spectrum are shown in Fig. 1. The range of significant samples of the mutual spectrum of the two received signals does not exceed fifty.

Figure 1. Wavelet slice ($a^*$ corresponds to the 0.396 Hz band) and the mutual spectrum (right) of a 57-year-old man with a high risk of community-acquired pneumonia.
Since the respiration rate is unstable, the empirical choice of a resonant wavelet is difficult, therefore, it is advisable to use a set of wavelets that lie in the frequency domain of chosen systemic rhythm.

In connection with the above, we select the number of lines in the wavelet plane and the step of changing the scale in such a way that with the minimum number of analyzed wavelets, all harmonics of the selected system rhythms are represented on the wavelet plane of the cardiac signal.

To obtain the mutual spectra of the cardiac signal with test signals, we use the Wiener-Khinchin equation

$$X_{12}(\omega) = \int_{0}^{T} R_{x_1 x_2}(\tau) \cdot e^{-j\omega \tau} d\tau,$$

where $R_{x_1 x_2}(\tau) = \frac{1}{T} \int_{0}^{T} x_1(t) x_2(t + \tau) dt$; $T$ is the interval of observation of the cardiac signal $x_1(t)$; $x_2(t)$ is the test signal.

Equation (3) we apply to each line of the wavelet plane, therefore, the mutual spectrum of the cardio signal and the set of test signals will also be a plane. In contrast to the wavelet plane, the plane of the mutual spectra will be as sparse matrix, since the wavelet function is a narrow-band signal, in contrast to the cardio signal.

An example of a line-by-line scan of the matrix of mutual spectra for the wavelet planes of one of the patients obtained before and after pneumonia is shown in Fig. 2.

Figure 2. The mutual spectra for twenty wavelet lines - the plane of the cardio signal of a 57-year-old man before (left) and after (right) pneumonia.

Figure 3 shows mutual spectra similar to those shown in Fig. 2, but only for one to fourteenth row of the mutual spectra matrix.
We carried out statistical studies of the matrices of mutual spectra similar to those shown in Fig. 3 patients and volunteers with different pneumonia risks. We used expert judgment and instrumental research. As a result, we came to the conclusion that the relative power of those lines in the mutual spectrum, which are associated with the selected systemic rhythm (respiratory), increases after pneumonia. This can be explained on the basis of the fact that with a disturbing effect on the system, intrasystemic connections weaken and, therefore, the system is more susceptible to external influences. That is, the system is more susceptible to modulation by systemic rhythms, which will be external disturbances in relation to the vascular system.

The statistical analysis of such mutual spectra allowed us to conclude that it is advisable to use the ratio of the mutual power of the cardiac signal and reference narrow-band signals as a risk factor. The frequency range of these signals should cover the frequency range corresponding to the frequency range of the breathing cycle. The range of significant readings of the mutual spectrum does not exceed twenty, and the number of lines per breathing rhythm does not exceed ten.

We define the mutual power $P_i$ of the $i$-th row of the matrix of mutual spectra as

$$P_i = \frac{1}{N} \sum_{j=0}^{N} X_{ij}(\omega_j)^2,$$  \hspace{1cm} (4)

where $N$ is the number of columns in the wavelet image matrices - planes in Fig. 1.

The risk factor for community-acquired pneumonia $FR_i$ is defined for one row of the cross-spectrum matrix as

$$FR_i = P_i.$$  \hspace{1cm} (5)

The set of $FR$ is used to construct the space of informative signs in conjunction with other markers of community-acquired pneumonia. Particular decision rules with subsequent aggregation of decisions are based on this data.

3. Results

We selected eleven signs $X0 \ldots X10$ from human vital signs indirectly associated with the risk of pneumonia. The first sign $X0$ is a vector quantity and is determined according to formula (5). $X1$ is an increase in body temperature up to 38 - 39.5 °C; $X2$ there is a cough with profuse expectoration; $X3$ is discomfort in the chest; $X4$ there is shortness of breath during physical exertion (sometimes even at rest); $X5$ is smoking; $X6$ is stress and psycho-emotional factors; $X7$ is hypodynamia; $X8$ is sex; $X9$ is age; $X10$ is chronic alcohol use.

The decision-making model includes the core of a hierarchical fuzzy neural network that predicts community-acquired pneumonia. The forecast is formed on the basis of the mutual spectra of the cardiogram and test signals occupying a frequency band correlated with systemic rhythms. Also, the forecast takes into account many particular decision rules. Decision rules are synthesized based on the
above eleven risk factors. In this case, a number of factors, and, consequently, the corresponding decision rules may be absent.

The fuzzy hierarchical structure is used as a base model for the analysis and aggregation of selected risk factors. Decision rule models are formed at the lower hierarchical level of this structure. Decision rules use the corresponding subsets of informative features. Membership functions for pneumonia risk classes are determined on the basis of expert assessments and statistical studies of the influence of a particular sign on the risk of this disease.

The model has a core, which is a neural network. The vector of informative features \( \tilde{X}_O \) is fed to the input of this network. In the model, aggregators combine only two confidence factors according to an algorithm based on an enumeration method (genetic algorithm).

The fuzzy operation that one of the ten aggregators can implement is a function of two arguments, so the number of possible fuzzy operations in each aggregator is defined as

\[
L = n!/(2(n - 2)!),
\]

where \( n \) is the number of elements in the selected tuple of fuzzy operations.

The length of a tuple \( n \), as a rule, does not exceed 8. A fuzzy operation in the \( i \)-th aggregator is described by the expression

\[
KV_i = f_i(KV_{i-1}, \mu_{FR}(X_i)),
\]

Where \( i = 1,13 \), \( KV_i \) is the fuzzy number at the output of the \( i \)-th aggregator, \( \mu_{FR}(X_i) \) is the membership function describing the risk of pneumonia by the carrier \( X_i \), which indicates the degree of belonging of the object with the parameter \( X_i \) to the fuzzy set FR.

A training sample with objects for which the risk values FR and informative signs \( X_1,...,X_{10} \) are known is needed to customize the classifier model. In addition, we must specify a tuple of fuzzy operations used in aggregators. Further, the value \( KV_0(j) \) of the risk of pneumonia by the vector \( \tilde{X}_O \) must be obtained for each \( j \)-th object of the training sample. This is done by a neural network.

The gatool module built into Matlab was used to configure the model aggregators. This module minimizes the number of errors for a given objective function. Three grades of risk: "Absent", "Available", "High" were used when setting up the classifier model.

**Table 1.** Experimental data on predicting community-acquired pneumonia on a training sample.

| Surveyed | Trigger results of the fuzzy inference module | DS, % | DSp, % |
|----------|-----------------------------------------------|-------|-------|
|          | «High» | «Present» | «Absent» |       |       |
| \( n_{e1} = 400 \) | 320     | 68        | 12     | 80    | 100   |
| \( n_{e2} = 80 \)  | 0       | 62        | 18     | 78    | 78    |
| \( n_{e3} = 120 \) | 0       | 25        | 95     | 79    | 80    |
| Total    | 320     | 155       | 125    | DE = 80%  |

Patients undergoing inpatient treatment at the regional hospital in Kursk were selected to test a fuzzy neural network model of pneumonia risk. Men and women of all ages with a relatively stable (without exacerbation) health were included in the study. All of them were on treatment and all had an established diagnosis of the underlying disease based on the results of inpatient examination prior to the period of technology testing. Patients with an unknown disease and in a state of exacerbation (instability) of the disease were not included in the study.
The results of the study of the neural network model on the control samples when assessing the risk of community-acquired pneumonia in terms of diagnostic sensitivity (DS), diagnostic specificity (DSp) and diagnostic efficiency (DE) are shown in Table 1. The analysis of the obtained results showed that statistical tests on control samples make it possible to recommend the obtained decision rules and algorithms for practical use.

4. Conclusion
Thus, at the first stage, we chose the carrier of systemic rhythms for the classification of non-hospital (community-acquired) pneumonia. Then we selected the systemic rhythms most correlated with the adaptive status of the body systems through iterative procedures. Next we built a feature space based on the frequency decomposition of systemic rhythms. Then we developed a risk classifier for community-acquired pneumonia. Hybrid neural network structures are used for this purpose. They make it possible to use both the intuition and experience of the researcher himself, and training samples for constructing classifying systems. At the last stage, we synthesized a decisive module that allows us to determine the risk of community-acquired pneumonia based on the adaptive statuses of individual body systems.

5. Acknowledgments
The reported study was funded by RFBR, project number 20-38-90058.

References
[1] Pavlov S E and Kuznetsova T N 1999 Adaptation and stress in sports Actual issues of medical rehabilitation in modern conditions 307-312
[2] Ministry of Health and Social Development of the Russian Federation. The incidence of the population of Russia in 2007: Statistical materials. M., 2008.
[3] Shalnova S A, Kalinina A M, Deev A D and Pustelenin AV 2013 Russian expert system for risk assessment of major noncommunicable diseases (ORISCON) Cardiovascular therapy and prevention 4(12) p 55
[4] Ershov S P and Perelman Yu M Electrophysiological characteristics of respiratory muscles in patients with chronic bronchitis Bulletin of physiology and pathology of respiration
[5] Tetenev F F et al. 2005 Peak expiratory flow rate and bronchial resistance in patients with community-acquired pneumonia Siberian medical journal 58(8) 43-45
[6] Chuchalin A G and Aisanov Z R 1988 Dysfunction of the respiratory muscles in chronic obstructive pulmonary diseases Therapeutic archive 60(7) 126-131
[7] Filist S A, Shatalova O V and Efremov M A 2014 Hybrid neural network with macro layers for medical applications Neurocomputers. Development and application 6 35-39
[8] Filist S A, Tomakova R A and Yaa Z D 2012 Universal network models for classification problems of biomedical data Bulletin of the South-West State University 4(43) 44-5
[9] Korenevsky N A, Filist S A and Kraskovsky A B et al. 2009 Theory of designing fuzzy network expert systems for the management of medical and environmental safety Scientific Bulletin of Belgorod State University. Series: History. Political science. Economy. Informatics 9-1(64) 146-151
[10] Petrova T V, Filist S A, Degtyarev S V, Kiselev AV and Shatalova O V 2018 Predictors of synchronicity of systemic rhythms of living systems for classifiers of their functional states System analysis and control in biomedical systems 17(3) 693-700