Research and Methodological Approach to Diagnostic Prediction for Use in Body Area Sensory Networks for Medical Purpose

IGOR ILICH KOLTUNOV[1], ANTON VLADIMIROVICH PANFILOV[2], IVAN ALEKSANDROVICH POSELSKY[3]*, NIKOLAY NIKOLAEVICH CHUBUKOV[4] and IVAN VLADIMIROVICH KRECHETOV[5]

[1]Department of Information Systems and Technologies, Faculty of Information Technologies , Moscow Polytechnic University, Moscow, Russia.
[2]Tradition Group LTD / Moscow, Russia.
[3]NTC Automated Technical Systems, Moscow Polytechnic University, Moscow, Russia.
[4]Tradition Group LTD, Moscow, Russia
[5]Office of Scientific Research and Development, Moscow Polytechnic University, Moscow, Russia.
*Corresponding author E-mail: poselskiyi@bk.ru

http://dx.doi.org/10.13005/bpj/1323

(Received: October 26, 2017; accepted: November 07, 2017)

ABSTRACT

The content of the article is aimed at the formation of the scientific and technical concept of medical diagnostics and diagnostic prediction (D&DP) within the framework of the daily wearable non-invasive on-body biosensory system for monitoring physiological indicators of human health in complicated conditions of application determined by: noisiness of sensory reception, difficulties in dynamic modeling of physical activity of the observation object, specificity of a priori biomedical information consideration, and peculiarities of solving the problem of reducing compared disparate information streams: interactively updated digital data from the sensor network to a single semantic context, and reference medical information. A harmonious methodology for the design of intelligent sensor networks for medical purposes to carry out remote on-line diagnostics has not been developed in science up to the present time, which determines the scientific novelty and theoretical significance of ongoing research. The topicality of the project is determined by the urgent need for active implementation of achievements in the fields of informatics, microelectronics and communications in the equipping of medical institutions for improving the healthcare quality. Prospects of the research are expressed in the fact that the project goals are in line with the general trends in the telemedicine development, and as a result of applying sensor networks, automation and speed of solving the tasks of remote functional D&DP are expected to increase. The article reflects the priority of the social aspect of the project, as the state of health on the basis of the current analysis of the characteristic combinations of life indicators in nosologies and the results of long-term collection, processing and semantic classification of biomedical data is assessed during the public life of patients. The practical and theoretical importance of the work consists in the weighty contribution of the results obtained to the development of the medical sensory networks methodology in terms of developing a structural scheme that determines the interconnection of the functional modules of the hardware-software complex Sensory Body Network (HSC SBN). Also positive results were obtained in the issues: formalization of the subject area on the basis of the developed structural chart of D and DP; reducing the data array to the standard conditions for solving the statistical task of the current restoration of the diagnostic features of the human body; formation of a closed calibration loop of

Published by Oriental Scientific Publishing Company © 2017

the HSC SBN; the construction of the criterion for the information retrieval task of identifying the most reliable diagnosis; structural synthesis of the external diagnostic circuit

Keywords: Bioinformatics; biosensor; diagnostic value; diagnostic prognosis; non-invasive monitoring; cloud service; sensory body network; telemedicine.

INTRODUCTION

Problem description

The search for new comprehensive approaches to an objective evaluation of the health status of Russian citizens on the basis of modern technologies has a high social significance. Diagnostics is one of the key areas of medicine; being timely organized and conducted it determines the overall situation of the incidence in the country in many ways. Moreover, the early diagnostics has the greatest therapeutic effect.

This circumstance stimulates investigations to create systems capable of detecting hidden signs of diseases at the initial stages, prognosticating on the basis of multi-criterial choice of possible alternatives to the disease development trend[1], timely adjusting the treatment strategy and tactics, and minimizing the risks of failure in the event of complications. These procedural tasks should be solved within the modern regulatory framework, standardizing not only the issues of technical regulation[2], but also taking into account the current requirements for the organization of work of the diagnostic units of medical institutions[3].

A groundbreaking technology in the line of creating such systems is production of special clothes for everyday wear with an integrated network of small non-invasive biosensors for continuous monitoring of many patients' health. The achievements of science in the fields of bioinformatics and wireless communications, as well as the extensive range of biometric devices offered by the medical equipment market, allowed for the initiation of work on the creation of the HSC SBN, the design concept of which combines the methods of network intelligence, telemedicine and cloud technologies[4, 5].

Characteristics of the HSC SBN measuring component largely depend on the measurement transducers comprising it. Their composition can be flexibly configured depending on the specificity of the medical challenges addressed.

Theoretical framework and objectives of the research

Diagnostic prognosis provides for prediction of the disease course and its clinical outcome on the basis of the patients' individual peculiarities, ongoing assessment of their condition and the expected consequences of treatment. Diagnostic prognosis is connected with diagnostics, being its logical continuation. Diagnostic prediction is done using mathematical modelling methods to determine the resistance of a human body to the effects of various factors, and to clarify its liability to a particular disease on the basis of known patterns of pathological processes and disease courses, and if it has already emerged, to predict its future course and the outcome.

The main objective of the research is to develop a scientifically grounded mathematical apparatus for predicting variations of a disease progression by the multivariate analysis data and stratification of unfavorable predictors in patients. The essence of a DP making procedure implemented by means of HSC SBN consists in a comprehensive evaluation of a patient's functional status by identifying prognostically adverse symptoms, comparing them with the spectrum of clinical manifestations contained in the knowledge database, and further drawing up therapeutic and preventive recommendations.

The adaptability of the D and DP system to multitasking and variability of monitoring conditions is achieved by: controlling the vector of input parameters, ranking the risk factors, variability of the confidence limits of a diagnostic prognosis, optionality of controlling the body sensory network at the user's and the operator's levels.

MATERIALS AND METHODS

The wearable part of the complex is represented by a device for ongoing continuous monitoring of factors impacting the health status that ensures constant monitoring of physiological parameters critical for D and DP, taking into account individual (racial, gender-age) peculiarities of patients. The procedure of examining patients with the help of a sensory network, in addition to collecting anamnestic data and conducting clinical and laboratory studies, includes: a dynamic cycloergometric loading test, Holter ECG monitoring, pulse oximetry, motion activity monitoring, analysis of biological fluids (sweat, saliva), physiological parameters registration (temperature, pressure, skin moisture).

In the development of body area sensory networks, the following trends are observed:

- Increasing the functionality by expanding the range of sensors and improving the algorithms for data processing;
- Miniaturization due to a reduction of the mass-dimensional characteristics of the HSC SBN wearable part and integration of sensors into compact multi-channel measuring units;
- Improvement of the ergonomic and operational characteristics by using clothes with an elastic hypoallergenic fabric-backed frame resulting in unburdensome and convenient wearing;
- Increased automation and reduced human factor in the reception and decision-making processes;
- Primary processing of measurement data at the monitoring object level by an implantable device;
- Wireless data transmission;
- Enhancing the diagnostic data comprehension by an operator due to an improved data format, employing intuitive visualization tools (tables, diagrams, graphs, etc.);
- Using branching intelligent algorithms capable of making cross-analysis and prognosis, prioritizing tasks, implementing heuristic logic based on hypothesis generation and testing.

Designing HSC SBN functional chart

A human body is a complex biological object, its normal functioning being ensured by the joint operation of a number of systems whose parameters are closely interrelated. The mechanisms of self-regulation and homeostasis are explained by the fact that when one of the systems switches to a critical operation caused by a disease, its parameters 'alarm' the systems associated with it in terms of functions, for example, the immune system that 'swings' into the mode of hazard suppression.

This is determined by the presence of connections between the systems that are not obvious in every instance. Accordingly, 'taking' biomedical parameters is not sufficient for a reliable D and DP. If to establish and factor in the correlation of parallel measurement channels on the basis of ICD codes, the diagnostic and predictive power of the complex can be significantly increased.

Normalization of the factor scope of human body physiological status has made it possible to reduce the D and DP conditions to a conventional extremum problem of mathematical statistics. This statement allows the use of prediction methods for correlated time series that realize the extrapolation of a complex deterministic process that reflects the patient's current or expected health status.

According to the analysis report of statistically stable 'strong' correlations between factors, a structural-parametric synthesis of the diagnostic system has been performed (Figure 1). The set of developed criteria reflects the metrological quality of the D and DP findings in terms of quantitative indicators of accuracy, reproducibility, reliability.

Non-uniqueness and local conflictability of quality indicators determine the complexity of criterial design of the model and generate the need for stating a multi-extremum problem. The search for Pareto optimal solution is associated with striking a compromise between conflicting partial criteria[1].

The statistical value of D and DP is in registration of correlations; identification of patterns; singling out unique stable polymorphic combinations of risk factors; correlating them with the occurrence probabilities of respective events in a certain time perspective; comparison of semantically different information flows; risk assessment of a specific

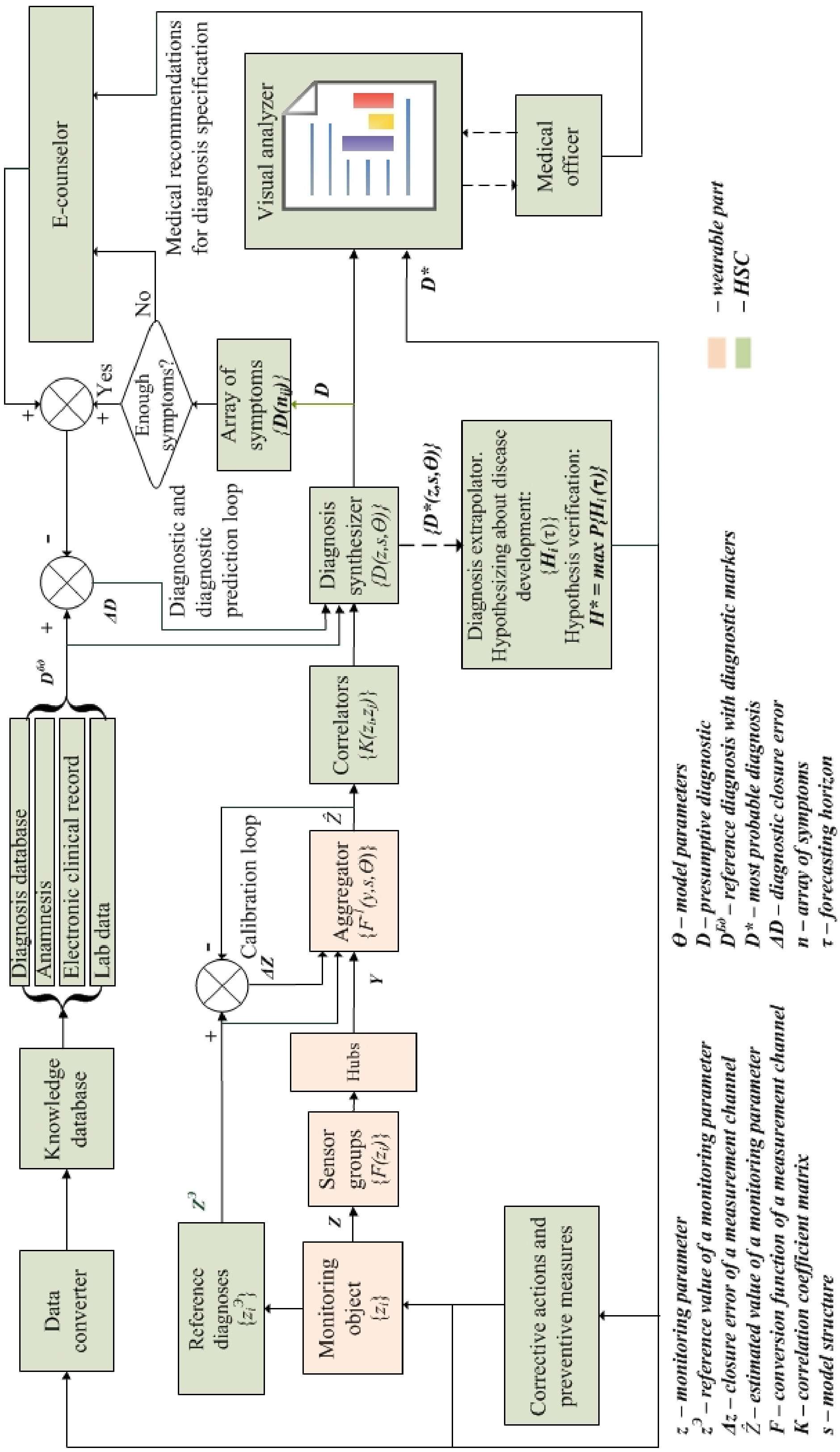

Fig. 1: HSC SBN general layout

disease incidence; hierarchization of possible diagnoses by degree of danger.

The D and DP medical aspect involves analysis of anamnestic data; testing; monitoring; recognition of characteristic clinical features; risk identification; identification, analysis, and ranking of high risk predictors; prediction of life-threatening complications; probability calculation of a disease exacerbation; calculation of a critical time horizon.

At the same time, the strict statement of the D and DP problems is hindered by:

- The need to take into account a variety of disparate factors, both individually and collectively in different combinations;
- Semantically heterogeneous data;
- The contradictory nature of the D and DP partial criteria;
- Interpretation ambiguity of findings.

Ethical and legal difficulties in implementing circuit solutions

The automation of the D and DP processes is not only hampered by the difficulties of an exhaustive and rigorous formal logical description of the subject area, but also by ethical and legal aspects. Medical professionals are an inextricable part of diagnostic process because a technical system, no matter how perfect it is, cannot be held responsible for possible errors.

Therefore, monitoring results should be specifically generalized, structured, and visualized in a human-readable fashion for experienced doctors' consideration and decision-making. When presenting monitoring results, geometric analogies are used to display data structures, graphs, and diagrams. In the prediction system, probabilistic statistical methods are used to estimate possible scenarios for the course of a disease.

Formation of an array of HSC SBN structural criteria

The contrariety of the local D and DP criteria requires techniques for finding compromise solutions. The quality of monitoring results is expressed by a set of indicators defined by metrological standards: reproducibility, accuracy, reliability [2]. Interpretation of these terms in the general case can vary significantly. Within the framework of this project, the following interpretation of these terms is proposed.

Reproducibility is a property of diagnostic monitoring expressed in the proximity of D and DP results in a series of repeated trials, under the same conditions, in the same patient, with one diagnosis. Reproducibility is determined by the statistical stability of monitoring results and is the most significant quality indicator of the D and DP.

Accuracy is determined by a minimum of diagnostic error norm. Accuracy is determined by a minimum of diagnostic error norm. In practice, the greatest accuracy of the results is achieved by the statistical convergence of the model value of D and DP to the real value of D or DD that is known beforehand. Accuracy depends on the class of mathematical operators of the diagnostic and prognosticating mathematical model, while accuracy criterion is a closure error of the output of D and DP system.

Reliability expresses the degree of confidence in D and DP results and is normalized through the probabilities of establishing an error-free diagnosis and prognosis.

The proposed interpretation of monitoring quality parameters meets the necessary conditions for development and normalization of the criterial basis of the D and DP model in structural operators according to[2].

Formalization of the domain

All the parameters used to compile the D and DP are mutually conditioned functions of time that collectively characterize the complex deterministic process of the vital activity of a human body. The degree of correspondence between this process and an object of monitoring is determined by the adequacy of an individual model of the physiological state of an observed patient, which links the following into a single integrated information structure:

- A priori information contained in the electronic patient record, anamnesis, and statistics of previous laboratory assessments;
- Biomedical data of the receptor level,

diagnostician's advice, results of the ongoing laboratory analysis;

- Expected scenarios for the development of identified trends;
- Recommendations for improving the health status.

The quality of the D and DP results is largely provided by the calibration loop (Figure 1). The type of measurements used in calibration is called codetermination, during which simultaneously registered reference values are compared, as well as the values of the calibrated channel homogeneous with them. Then, a calibration characteristic is definitely calculated as an analytical function.

The HSC SBN peculiarity is that its 'output' is rather a diagnosis – a hard-to-formalize verbal description of a certain physiological state of the biological object than a numerical representation of measurement result being conventional for classical metrology. This necessitates interpreting the term 'error' through uncertainty of establishing a diagnosis that can be quantitatively assessed by D and DP errors in probabilistic terms as follows.

In general, there are 4 possible diagnostic outcomes:

A 'healthy' patient is identified as 'healthy'An 'ill' patient is identified as 'ill'; A 'healthy' patient is identified as 'ill' – an error of the 1st kind;An 'ill' patient is identified as 'healthy' – an error of the 2nd kind

...(1)

A set of expressions (1) describes exhaustive events, the risks of which can be expressed through each other. For example, the probability of an accurate diagnosis *P*, when actually healthy one is defined as healthy, while actually ill one as ill, is:

$$P = 1 - (P_1 + P_2) \quad ...(2)$$

where,
P_1 is the probability of error of the 1st kind;
P_2 is the probability of error of the second kind.

Formulated for diagnostics, this expression is equally applicable to diagnostic prognosis. It is invariant to temporal shift: (2) can be replaced with an equivalent one written as a function of time: *Đ(t), Đ$_1$(t), P$_2$(t)*. Obviously, for any time shift, the identical equation *ô* is preserved:

$$P\,(t+\tau)= 1 - (P_1(t+\tau) + P_2(t+\tau)) \quad ...(3)$$

where
τ is the forecasting horizon. Only redistribution of shares between the probabilities P, P_1, P_2 will depend on the parameter τ.

However, this 'top-level' model for making predictive decisions is trivial. It involves establishing a diagnosis with accuracy of 'healthy'/'ill', while within the project the task should be solved with reasonable particularization of a diagnosis 'Ill with what' and a prognosis 'What disease can develop'.

Diagnostics is complicated by the fact that the decision-making field is ambiguous, multidimensional and non-linear and, therefore, is not susceptible of strict delimitation. In practice, there are multiple 'zones of uncertainty', 'overlapping zones', and 'fuzzy boundaries.' Thus, similar symptoms expressed by the same set of diagnostic markers and their relatively equal values can be induced by different causes and be peculiar to different diagnoses. As a result, collisions may occur when the system proposes several mutually exclusive measures to improve the health status.

For the foregoing reasons, D and DP cannot be implemented only on the basis of strict deductive logic. The diagnosis 'ill' 'disintegrates' into many unequally probable diseases, which generates a need for assessing the risks of a patient being in the state of each possible disease.

This picture is typical for standard medical diagnostic practice as well, where contradictions often emerge to be resolved in a consultative way with the participation of highly experienced doctors of different specialties. It is possible to reduce the influence of the uncertainty factor when using an intelligent technical system as a diagnostician only by accumulating the experience of erroneous diagnoses under similar conditions and adjusting the algorithm appropriately, which is the essence of training the decision support system with D and DP

[4]. This function is performed by a data converter included in the outer feedback loop (Figure 1).

Structure synthesis of HSC SBN

In more detail and mathematically more strictly, the problem can be formulated as follows. Let the database contain a set of *N* diagnoses, each of them being characterized by a combination of *M* signs.

$$D_i^{DB}(n_j);\ i=\{0,N\};\ j=\{0,M\} \qquad ...(4)$$

where

$D_i^{DB}(n_j)$ is a diagnosis pattern stored in the database, in which:

i is the index of a diagnosis number;

n_j is a fixed j-th combination of M diagnostic markers.

The HSC SBN database provisioning to form an array (4) is performed on the basis of reference medical literature, interviews, and questionnaires of experts. The database is a repository of reference diagnoses used to identify the personal diagnosis of a patient by correlating them.

The primary near-patient testing is performed at the aggregator level based on vitality tolerance limits. When control parameters are beyond the permissible values, an emergency signal is generated. These are such parameters as 'Cardiac Arrest', '45 bpm > Heart Rate> 130 bpm, 'Body Temperature <32°C', etc.

The next phase determines diagnostic analytics. During the registration, a procedure of generation and verification of hypotheses about the health status pertaining to i-th diagnosis is performed in a symptomatic evidence processing center. Preference is given to the hypothesis with maximum estimated probability:

$$H^* = max\ P\{Hk\ (\tau)\},\ \{k = 1,K\} \qquad ...(5)$$

where,

τ determines a prediction time interval;

K is the number of hypotheses made;

P is an estimate probability that the hypothesis $H_k(v)$ is true.

The remaining hypotheses are discarded as untenable.

Next, the correlates are established for an out-of-range parameter or for a parameter having a steady trend to over-ranging, and the current values of its relationships with adjacent parameters in different diagnoses are checked based on the previously formed correlation matrix that determines the structure of the interrelations of the set of diagnostic markers:

$$Kij = K(zi,zj),\ i \neq j \qquad ...(6)$$

If values of the correlates approach one of the reference sets of diagnostic markers from the database, or a critical dynamics of approaching one of such sets is registered, the grounds for prognosis generation by testing the hypothesis of respective diagnosis validity through the estimated period of lead time are considered sufficient.

It should be borne in mind that symptomatology coming from the sensory network is most likely to be incomplete. This circumstance requires the hypothesis about i-th disease to be supplemented in a certain way. The function of adding control parameters to complete the required composition of diagnostic markers ensuring an accurate diagnosis or diagnostic prognosis is performed by a synthesizer of diagnoses. This module generates queries:

- To the database to extract non-formalized information and feed it in the system;
- To the laboratory to compile required analyses;
- To the monitor unit of a diagnostic operator in order to use their knowledge to confirm/reject/correct the hypothesis of i-th diagnosis.

The composition of additional markers comes from the database and from the laboratory upon request, and is further defined and confirmed by a diagnostician. Based on this information, an e-counselor issues recommendations to a patient's cellphone to take actions that ensure the diagnosis specification: to attend a therapist appointment, to have certain tests, etc.

After a complete composition of the markers has been formed, the current status at the diagnosis synthesizer output is compared with the set of markers from the database, and the most probable diagnosis D_i^* is identified, such that minimizes the closure error:

$$\Delta Di^* = min\ \{Di^{\delta\delta}\ (nj) - Di(nj)\} \qquad ...(7)$$

Thus, the following challenges can be met:

1. Identification of the current diagnosis by coincidence of markers (5).
2. With partially matching markers by characteristic trends and calculated based on time-expanded measurements of the dynamics of markers approaching critical levels, a sample is extrapolated to the time horizon determined by the most reliable prediction of the moment the *i*-th diagnosis comes into effect. This is how a particular disease and the onset of certain stages of the disease are predicted.
3. If there are several hypotheses about a disease or its precursors, they are ranked according to the hazard rate and are checked for accuracy according to (5).
4. If a disease treatment modality is sufficiently formalized, the HSC SBN issues recommendations for choosing therapeutic or preventive measures to reduce the dynamics of the disease progression. Otherwise, treatment and prevention scheme arrangement remains at the discretion of the duty medical officer.

RESULTS

As part of the study, a structural chart of HSC SBN functioning and the D and DP mathematical apparatus in structural operators that allows for building special software to process data coming online from the sensory networks have been developed.

The quality requirements for medical monitoring findings in patient groups have been stated, and also interpretation of metrological criteria of reproducibility, accuracy and reliability is given with respect to the project objective, satisfying the norms of technical regulation and the requirements of medical standards [2, 3].

The proposed D and DP algorithm methodically predetermines automatic early recognition aimed at detection of disease symptoms in stages when the disease has not yet seriously damaged the body, and its treatment does not yet require high costs for expensive medication, surgery, or rehabilitation.

Algorithmization of the D and DP model on the basis of the developed HSC SBN structural chart makes it possible to obtain fairly accurate and reliable diagnosis estimates to ensure intelligent support of medical personnel, corresponding to modern concepts of artificial intelligence systems [4].

DISCUSSION

In the course of the research, the authors of this article have continued the studies on enhancing the technology advance in constructing circuit solutions and algorithmization of HSC SBN functionality on the basis of established regularities and the developed structural chart.

These findings are a significant contribution to the scientific and methodological groundwork for a successful development of the project and the methodology of medical sensory networks in general. In the medium term the research output taken to the level of a software product enables to solve a wide range of urgent problems of modern medicine on the basis of the statistical approach: simultaneously observing a large number of patients, grouping them by characteristic symptomatology and territorial grounds, assessing stages of a disease, taking emergency measures in life-threatening situations, monitoring and predicting spread of the nidus of contagious diseases, analyzing regional incidences based on statistics per totality of patients, choosing effective area-based preventive measures, planning effective operational procedures to control the epidemiological situation, etc.

The proposed data systematization during the HSC SBN development promotes taking consistent approaches to the assessment of exposure to diseases and making D and DP. It is a significant scientifically grounded methodological support for medical staff, not only in diagnostics, but also in choosing an optimal therapeutic strategy, deciding on the advisability of surgical treatment, creating the most rational and the most effective structured methodology of individual disease prevention based on modern technology. The effectiveness of health status remote monitoring by applying the HSC SBN is ensured by combining the design output with the experience of using telemedicine technologies, and the formation of client networks based on cloud services [5].

ACKNOWLEDGEMENT

This research is carried out with the financial support from the Ministry of Education and Science of the Russian Federation under the agreement No.14.577.21.0232 of September 29, 2016 (unique number RFMEFI57716X0232), the applied scientific research is conducted on the subject 'Research into scientific and technical solutions and development of an intelligent wearable biosensor platform for preventive monitoring and evaluation of vital signs of a human body 'Sensory body network' capable of correlating the data received from various sensors in a high-noise environment'.

REFERENCES

1. Nogin, V. D. Decision-making in a multicriteria environment: a quantitative approach. Fizmatlit, Moscow (2005).
2. GOST R 8.596-2002. GSI. Metrological support of measuring systems. Basic provisions. METROB.RU (2002).
3. Order of the Ministry of Health of the Russian Federation No. 283 of November 30, 1993 "On the improvement of the service of functional diagnostics in health care institutions of the Russian Federation".
4. Alpaydin, E. Machine Learning: the New AI. Tochka Publishing Group, Moscow (2017).
5. Shanahan, M. The Technological Singularity. Tochka Publishing Group, Moscow (2017).