

Big Data Clustering in Cardiology Based on Modeling of Electrical Dynamics of the Heart in the form of Fermi-Pasta-Ulam Auto-Recurrence as a New Tool for the Study of Cardiac Activity.

Novopashin MA ^{1*}, Shmid AV ¹, Zimina E.Y ¹, Berezin AA ¹

¹ EC-leasing.Tikhonov Moscow Institute of Electronics and Mathematics, Higher School of Economics, Russia

***Corresponding Author:** Novopashin MA, EC-leasing.Tikhonov Moscow Institute of Electronics and Mathematics, Higher School of Economics, Russia. Email: <mailto:mnovopashin@ec-leasing.ru>

Received date: July 20, 2018; Accepted date: July 27, 2018; Published date: August 13, 2018. Citation

this Article: Novopashin MA, Shmid AV, Zimina E.Y, Berezin AA. (2018) Big Data Clustering in Cardiology Based on Modeling of Electrical Dynamics of the Heart in the form of Fermi-Pasta-Ulam Auto-Recurrence as a New Tool for the Study of Cardiac Activity. J Clinical Cardiology and Cardiovascular Interventions. **Doi:**10.31579/2641-0419/008.

Copyright: ©2018. Novopashin MA, This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

The mass application of mobile cardiographs already leads to both explosive quantitative growth of the number of patients available for ECG study, registered daily outside the hospital (Big DATA in cardiology), and to the emergence of new qualitative opportunities for the study of long-term oscillatory processes (weeks, months, years) of the dynamics of the individual state of the Cardiovascular system of any patient.

The article demonstrates that new opportunities of long-term continuous monitoring of the Cardiovascular system state of patients' mass allow to reveal the regularities (DATA MINING) of Cardiovascular system dynamics, leading to the hypothesis of the existence of an adequate Cardiovascular system model as a distributed nonlinear self-oscillating system of the FPU recurrence model class [1]. The presence of a meaningful mathematical model of Cardiovascular system within the framework of the FPU auto-recurrence [2], as a refinement of the traditional model of studying black box, further allows us to offer new computational methods for ECG analysis and prediction of Cardiovascular system dynamics for a refined diagnosis and evaluation of the effectiveness of the treatment.

Keywords: electrocardiogram; veloergometer; pacemaker; dechaotization; cardiobarometer; general clustering

Introduction and put forward a Hypothesis

In our model of the heart based on the of the FPU recurrence phenomenon, the basic concepts are traditionally the "states" of the heart as a registered oscillatory process. However, since the heart is a self-oscillating system (in contrast to the previously studied theoretical and experimental models of the FPU recurrence based on passive nonlinear systems (plasma, nonlinear electrical lines, deep and shallow water waves dynamics [3,4,5,6] the perturbation energy is not required to start the oscillations, and in the study of the heart electrical dynamics, we can introduce the concept of the auto FPU recurrence.

The ideal ECG in the hypothetical case of the absence of any external influences and represents a periodically reappearing FPU auto-recurrence, characterizing the patient's condition.

In real life, various physical and emotional loads, therapy, etc. are external to the heart model "disturbances" and will be periodically displayed in the picture of the FPU auto-recurrence. Their influence on auto-recurrence picture can theoretically be revealed, allowing to establish biologically significant parameters of the equations of the FPU recurrence model, diagnosing the actual biological condition of the heart (for example, the area of the myocardium) predicting the future picture of the dynamics of cardiac activity.

Theoretically, the model of the FPU recurrence may contain an infinite number of states of the heart activity. However, in case of the ECG, for reasons of common sense for a practical cardiologist all the infinite number of pictures of auto-recurrences of heart conditions should fit into a finite number of possible diagnoses.

Therefore, the implementation of the clustering process is constructive: the integration of similar self-oscillating processes of the heart into subsets (clusters) based on belonging to the diagnosis.

In this case, the set of clusters becomes finite which seems obvious.

However, from the FPU auto-recurrence model [2] it also follows that the real picture of the heart's behavior can be much more complicated: in the ECG patterns may be such variants of heart behavior that "do not fit" any of the known diagnoses, and reflect the processes of the heart transition from state to state, including in completely healthy patients.

Theoretical (before the experiment) prediction of the possibility of such "strange" ECGs and their subsequent experimental detection can serve as a proof of the adequacy of the FPU model with automatic recovery of the physical picture of the heart.

Proof of the Hypothesis

The FPU model with the introduction of the concept of "auto-recurrence" [2] can adequately describe many processes taking place in the heart, at a level practically acceptable for solving the problem of clustering of ECG diagnosis.

The solution of the problem of content-based ECG clustering, in turn, opens up opportunities for a better solution to the problem of diagnosis and prediction of scenarios of heart disease development.

Now we introduce the concept of FPU auto-recurrence at the mathematical level.

The phenomenon of the canonical FPU recurrence in passive systems was first described as a result of numerical study of solutions of

differential – difference equations describing a chain of nonlinear coupled vibrators [1]. In these model chains, the phenomenon of dissipation has not been included.

To describe real dynamic processes, such as cardiac activity, a more adequate model was required, capable of describing the recurrences in autonomously functioning or self-oscillating systems.

A significant contribution in this direction was the work of American researchers Zabusky and Kruskal, who proposed to describe the FPU recurrence within the framework of the Korteweg de Vries (KDV) equation with periodically changing boundary and initial conditions [6].

Using the results of this work, it is possible to simplify the solution of the KDV equation in the form of knoidal waves, replacing them with the solutions of the Van der Pol equation, close to harmonic, periodic and using the low-frequency changes in boundary and initial conditions by the relaxation solutions of the Van der Pol equation.

At the same time, if we apply the theorem on the possibility of replacing the wave links by the delayed ones [6], then we can present a mathematical model of the electrical activity of the heart (ECG) in the form of the FPU auto-recurrence, described in the framework of coupled equations of Van der Pol with a time lag [7].

The computer study of the FPU auto-recurrence model [2] shows that in case of intact myocardium and under the absence of external influence at main heart 1 Hz frequency, the model reproduces the oscillations similar to a regular ECG (Fig. 1) with its Fourier spectrum given in Fig.2.

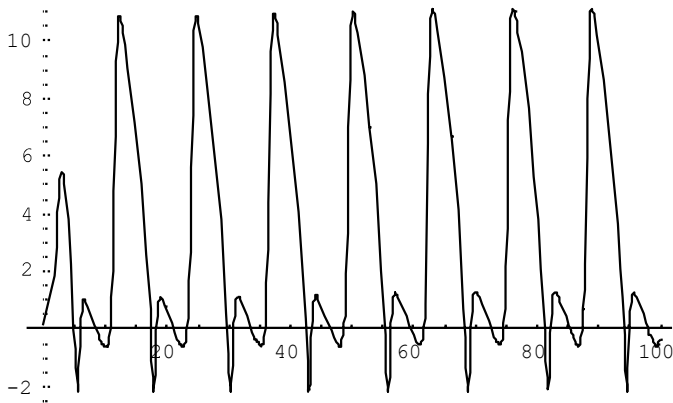


Figure 1. Model of normal ECG generated by the system (1) under the absence of external influence at 1 Hz frequency. Horiz. Axis – time, Vert.axis – voltage. Units conditional.

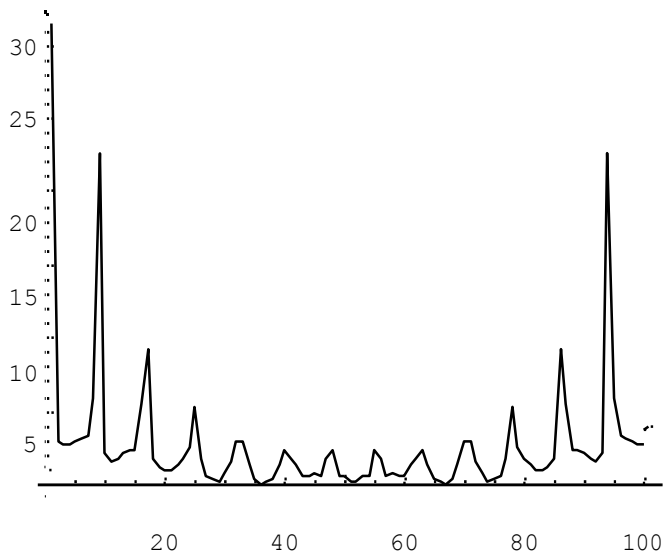


Figure 2. Fourier spectrum of the model ECG shown in Fig.1. Horiz. Axis – frequency, Vert.axis – amplitude. Units conditional .

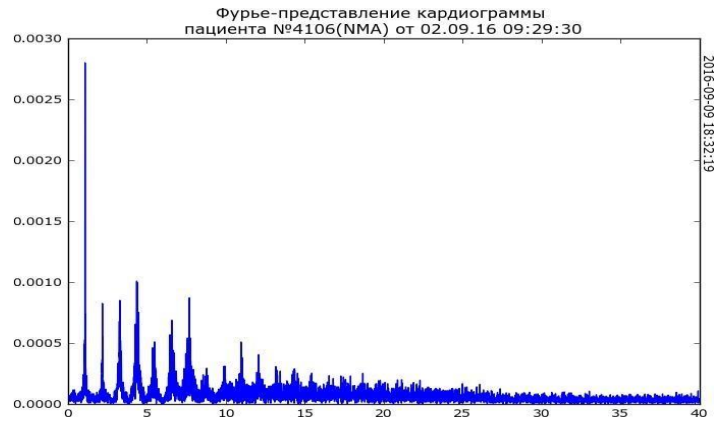


Figure 3. Fourier spectrum of normal electrocardiogram of a healthy person at the age of 28. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

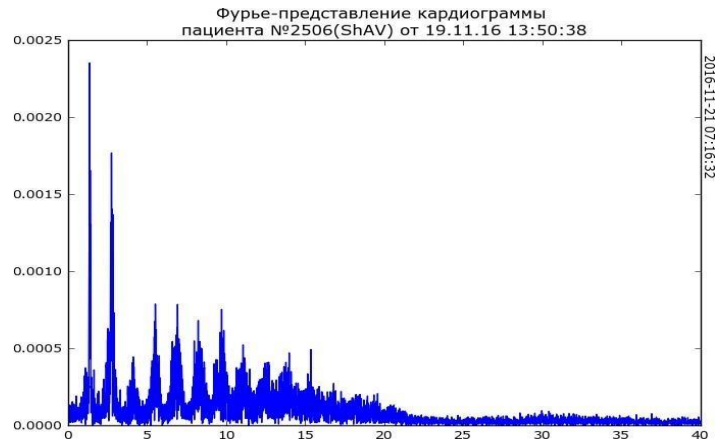


Figure 4. Fourier spectrum of normal electrocardiogram of a healthy person at the age of 70 years. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

The spectra shown in Figs. 3 and 4 correspond qualitatively to the model spectrum of the FPU auto-recurrence of the normal ECG obtained from the study of the FPU auto-recurrence model [2].

The obtained data result in a prediction of consequences of resonant external influence at 1 Hz frequency following from the hypothesis. In particular, an enough intense external resonant influence can discompose the regular model oscillations. The computer study shows that increasing the amplitude of the external perturbation by 50% brings the breakage of the rhythm of model oscillations corresponding to the state of the heart fibrillation (Fig.5.) This can be called a model infarct. At the same time, the hypothesis predicted the threshold of the external perturbation amplitude that would stop the discomposing of the model.

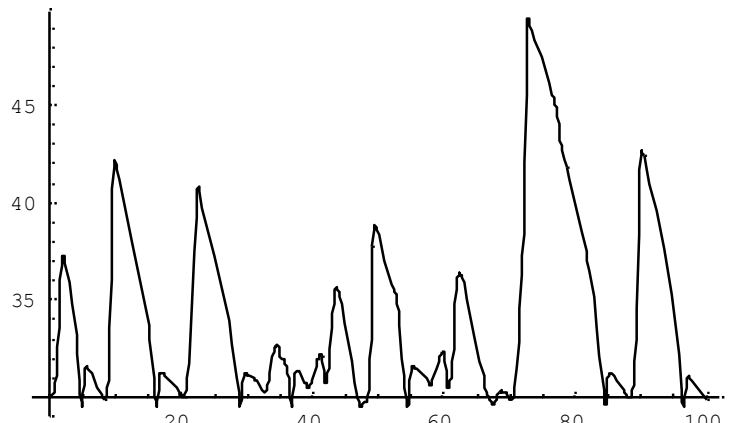


Figure5. Dynamics of the model ECGs corresponding to the state of the heart

fibrillation under the increase of the amplitude of external perturbation. Horiz. Axis – time, Vert.axis – voltage. Units conditional.

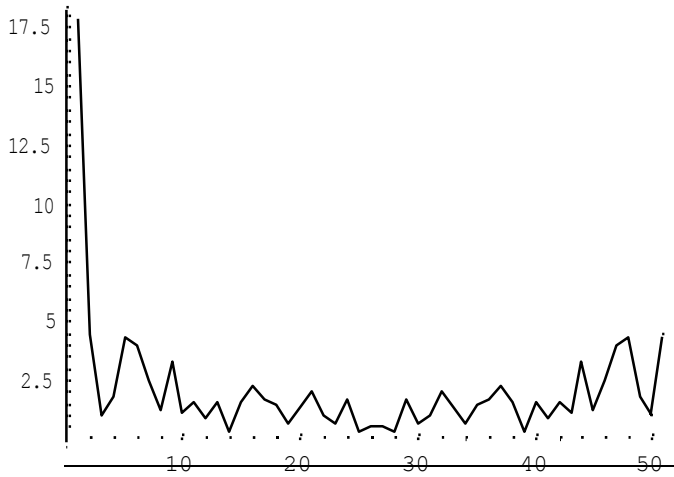


Figure 6. Fourier spectrum of the model fibrillation ECGs (Fig.5.) under the increase of the amplitude of external perturbation. Horiz. Axis – frequency, Vert.axis – amplitude. Units conditional.

The further increase of the external perturbation amplitude resulted in total synchronization of oscillations, reminding the dynamics of heart stimulated by a cardio stimulator (Fig.7).

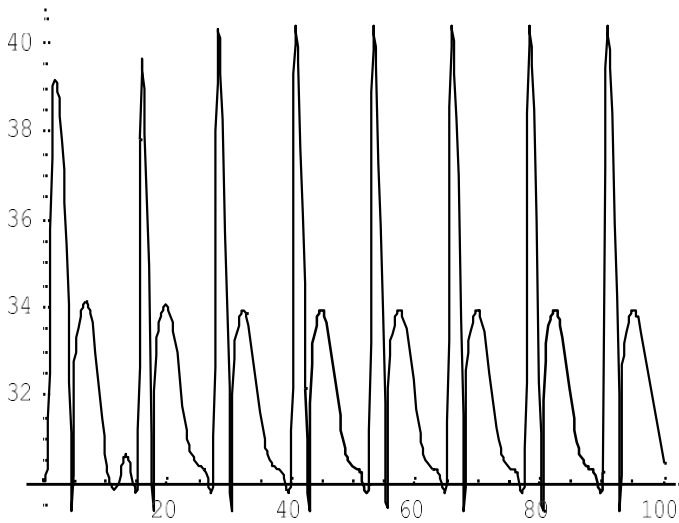


Figure 7. Model ECG in the model mode of “cardio stimulator”. Horiz. Axis – time, Vert.axis – voltage. Units conditional.

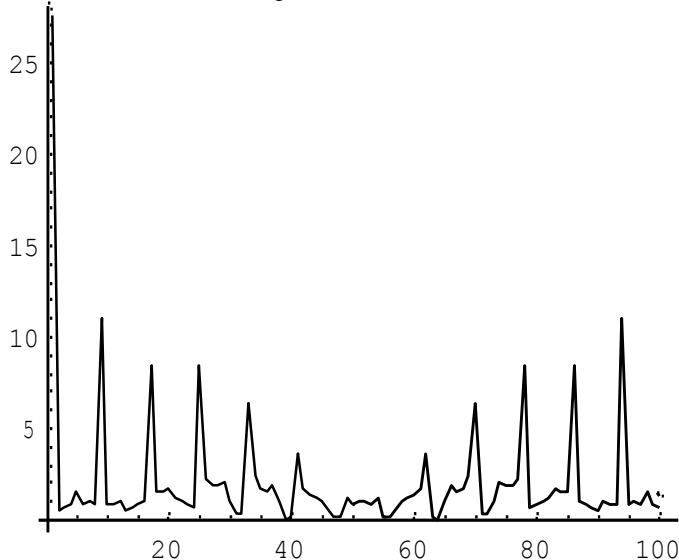


Figure 8. Fourier spectrum of the model ECG (Fig.7.) in the model mode of “cardio stimulator”. Horiz. Axis – frequency, Vert.axis – amplitude. Units conditional.

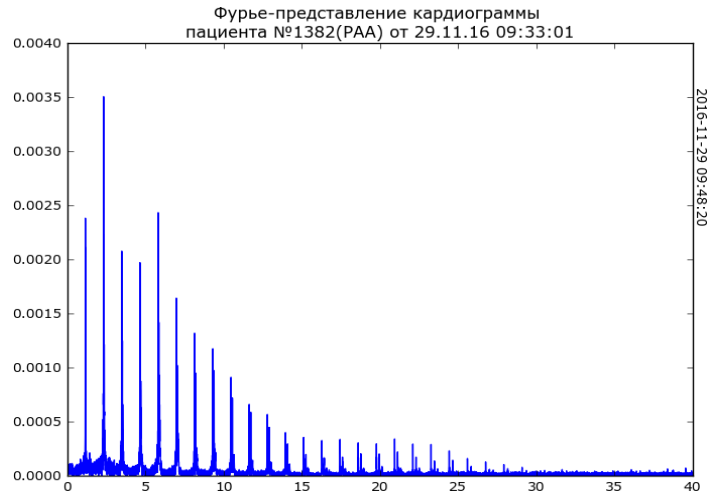


Figure 9. Fourier spectrum of electrocardiogram of a person with a cardio stimulator. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

Hypothesis

Considering the heart as a distributed self-oscillating system, which theoretically has an infinite number of auto-recurrences, we introduce the concept of a cluster obtained in the course of a series of comparisons of the ECG spectra with each other and the identification of repetitive ECG spectra. This allows us to identify similar clusters in different patients both in healthy and in patients with different cardio pathology. The hypothesis is that the proposed clusters can be used for the analysis of various pathologies in patients, and to assess the cardio logical status of healthy people. In addition, it is assumed that the above given approach will allow to analyze different dynamics of the Fourier ECG spectra in patients with pacemakers using clusters, since the dominant spectrum of the electronic generator of the pacemaker does not make it possible to evaluate its changes that directly characterize the parameters of the electrical dynamics of the heart itself.

In addition, the hypothesis of the possibility not only to diagnose coronary artery disease in patients, but also to assess their degree of development of the ischemic process.

Results of Experimental Studies Supporting the Hypotheses Put Forward

In order to test the validity of the hypotheses put forward, the following studies were carried out.

Material and methods

14500 cardiograms of patients aged 42-60 years with angina pectoris of different functional class were examined, of whom 36 people with typical pain syndrome and episodes of painless myocardial ischemia (men — 25, women - 11) and 30 people (men — 18, women — 12) with a typical pain syndrome, without episodes of painless myocardial ischemia.

Criteria for exclusion of patients from the study: permanent form of atrial fibrillation; complete atrioventricular (a-b) blockade, complete blockage of the left leg of a bundle of HIS, technical reasons for false-negative or false-positive detection in the identification of myocardial ischemia (recommendations for halter), life-threatening arrhythmias, acute cerebrovascular disease, tumor, severe anemia, thyroid disease in the stage of decompensation, lack of informed consent of the patient, the inability to carry out stress tests, doubtful results of stress tests and XM ECG.

The diagnosis was put taking into account clinical, anamnestic, instrumental methods of research. The diagnosis was verified by physical exercise tests (veloergometry (VEM) or stress echocardiography (stress Echocardiography). All patients underwent XM ECG.

Siemens-Elma veloergometer was used for VEM and stress-Echocardiography. Stress-Echocardiography was carried out using the program module Stress-12-Cardio automated system for ECG registration in 12 conventional leads (the manufacturer of JSC "Diamant", St. Petersburg). The interpretation of ST-segment changes was based on the generally accepted criteria of ischemic response to stress.

The Protocol of threshold load power with its step increase by 25 W every 3 minutes was used. to achieve ECG and/or Echo-criteria of termination of a load or submaximal heart rate (HR).

Principles of cluster construction

The form of the spectrum

Comparison of the Fourier spectra of ECG was made on the basis of the phenomenon of the FPU recurrence, on the basis of visual perception of a spectrum. For this purpose, the concept of the shape of the spectrum was introduced. The spectrum shape is a smooth curve based on the median of the upper 30% of the values in the 2 Hz window, taken in the range from 0 to 40 Hz. Window width 2 Hz is taken specially to eliminate the influence of pronounced harmonics for ECG with a heart rate less than 120 beats/min. The form of the spectrum is used further in the algorithm for determining regions of ischemia and clustering of ECG in the form of the spectrum.

For clustering, the form is approximated in increments of 0.1 Hz to 400 points and normalized by the arithmetic mean for the universality of comparison.

Clustering Algorithm

ECG clustering in the form of the spectrum is performed by the method of allocation of connected components. The distance between the spectrum forms of the compared ECG is taken Euclidean. In the process of clustering, the original fully connected graph is scattered into a group of subgraphs when dropping links of a longer length than the pre-defined value, which form the desired clusters.

To simplify the algorithm, clusters are formed automatically based on the classification of ECG as they arrive (the so-called online method). In this case, operations of adding a cluster with a new form of ECG, creating a new cluster and combining clusters based on the distance between them are implied.

Cluster

A cluster is a set of one or more ECGs and the associated set, the arithmetic mean of the spectrum, calculated automatically as a result of the process of ordered ECG classification.

Personal and General Clustering

For the analysis of individual patient conditions, clustering is performed only on the basis of the patient's ECG-personal clustering.

To identify common typical states, clustering is based on all available ECG – General clustering.

The Classification of the Clusters of the Patient

The resulting patient clusters differ not only on the basis of the averaged form, but the number of ECGs in the cluster is also important. It is practically obtained that there is always a cluster with a predominant number of ECG, caught in it. Presumably, the ECG of this cluster characterizes the normal state of the patient. ECG, trapped in other clusters, characterize to some extent the abnormal state of the patient, caused by external stress, and pathologies. Often there are clusters consisting of one ECG, - their appearance allows us to judge the intermittency of complex FPU auto-recurrences in the patient (when between repetitions of a certain state there is a change of several states, different from the repeated ones).

The cluster with the largest set of ECG is called the base cluster. For all other clusters, a measure of the cluster deviation from the base in the form of Euclidean distance between the corresponding averaged forms is introduced.

Clusters containing more than one ECG are called typical clusters.

Clusters consisting of one ECG are called "clusters in a chaotic field". Often, with the gradual completion of the patient's ECG set, some single clusters are supplemented with new ECG, moving into the category of typical clusters, or combined with other clusters.

Cluster grouping by pathology

On the basis of the characteristic features of the individual clusters obtained (both personal and general) and the data on the pathologies of the patients, the clusters similar in characteristics to the shape of the spectrum are distinguished and displayed in separate groups by pathologies. It is assumed that the selected clusters give a certain reference point in the 400-dimensional space of all forms of spectra, the distance to which characterizes the proximity of the shape of the patient's ECG spectrum to a particular pathology. This will further allow to judge in which direction the patient's therapy, training or rehabilitation is moving.

Several clusters of healthy people, patients with ischemia, and patients with a pacemaker were selected for such reference points. Examples of healthy people clusters are shown in Fig (11,12,13.) Fig.10 shows a typical Fourier spectrum of ECG of a healthy person.

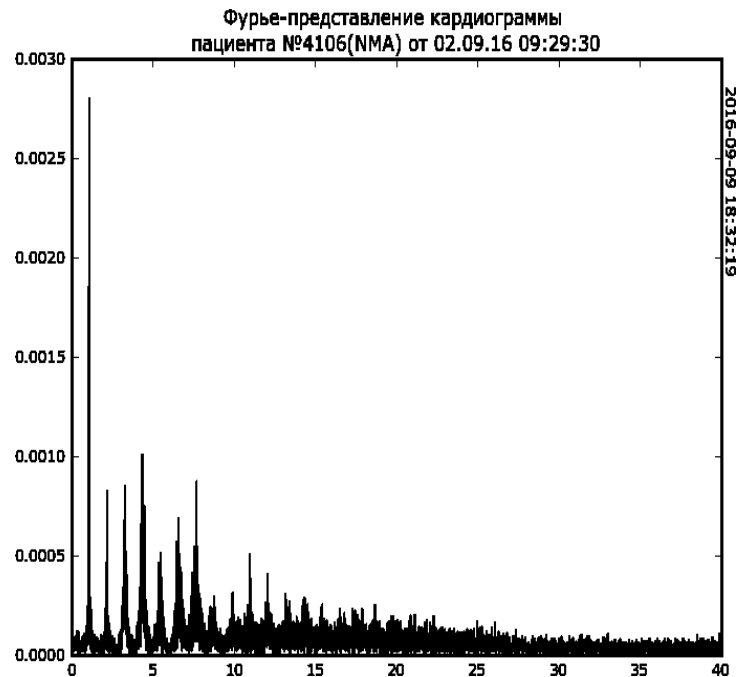


Figure 10. Typical Fourier spectrum of ECG of a healthy person. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

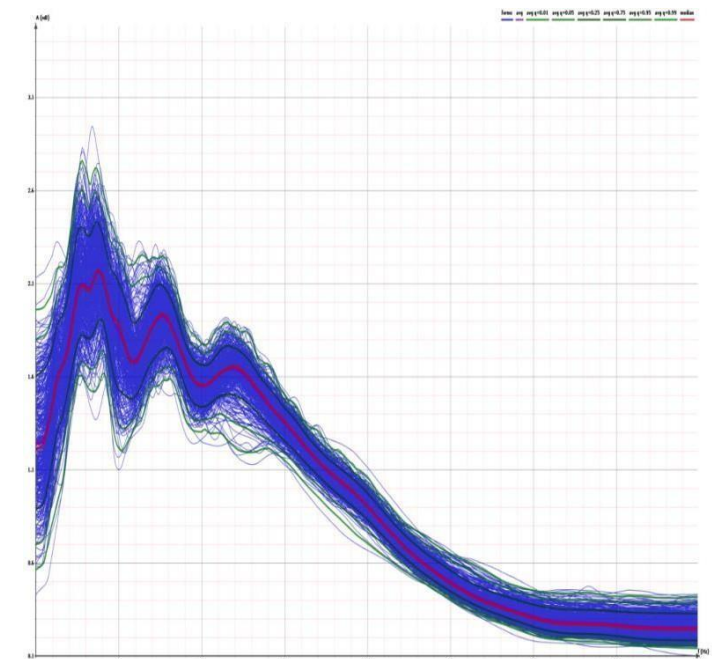


Figure.11. Cluster of a healthy person.

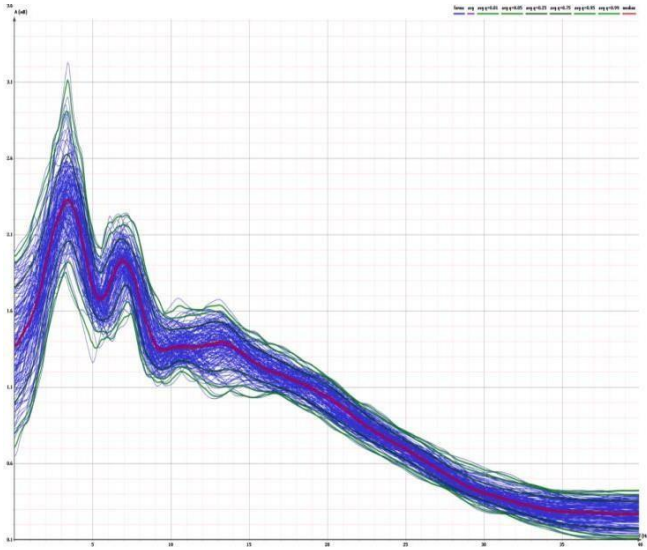


Figure.12 Cluster of a healthy person.

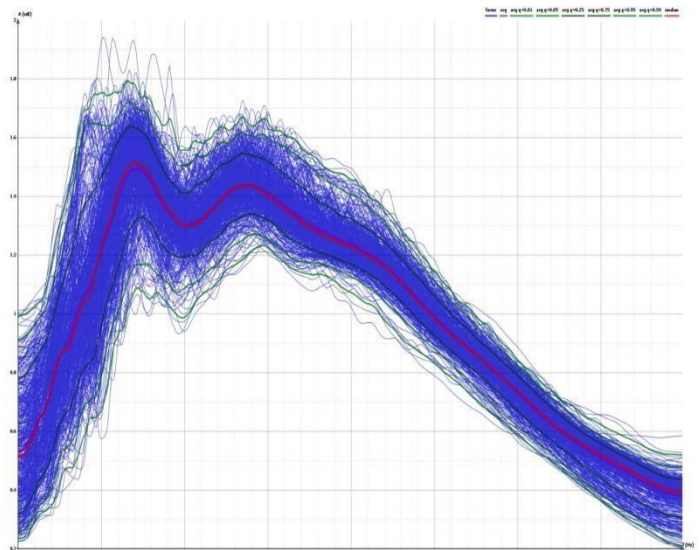


Figure15 .Cluster of a patient with diagnosed ischemia.

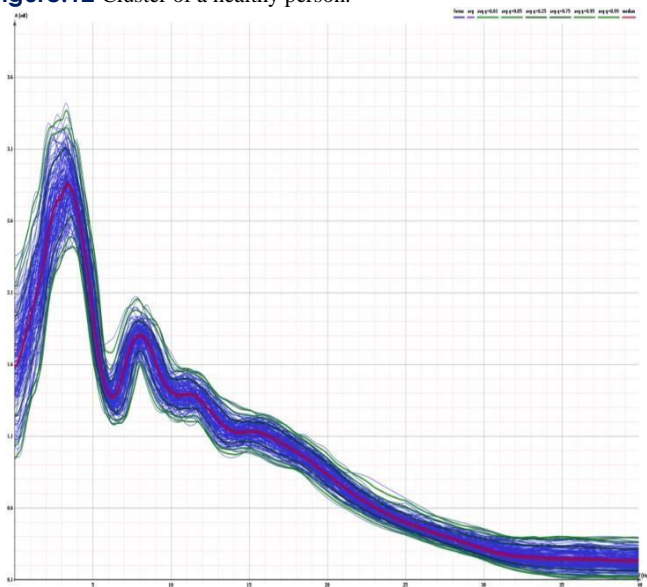


Figure 13. Cluster of a healthy person.

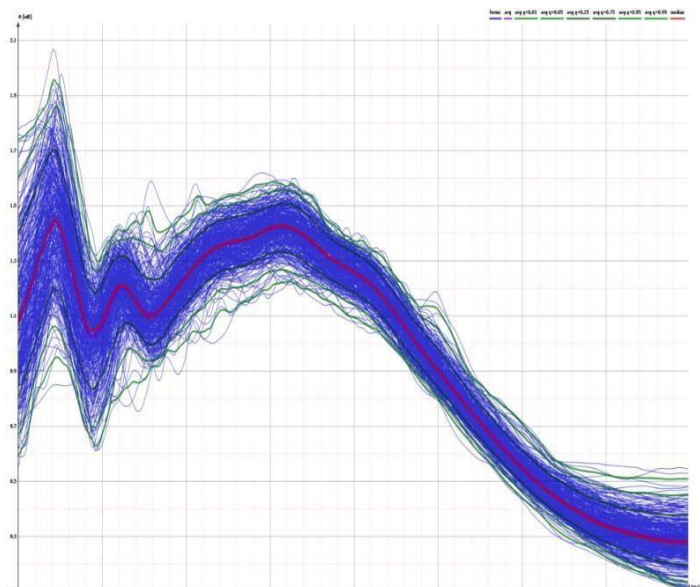


Figure 16. Cluster of a patient with diagnosed ischemia.

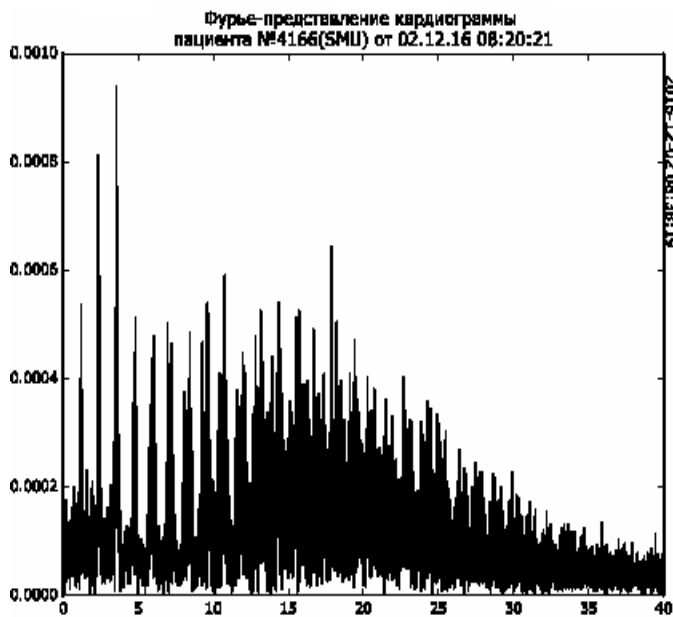


Figure 14. Typical Fourier spectrum of the patient's ECG with a developed ischemic process.

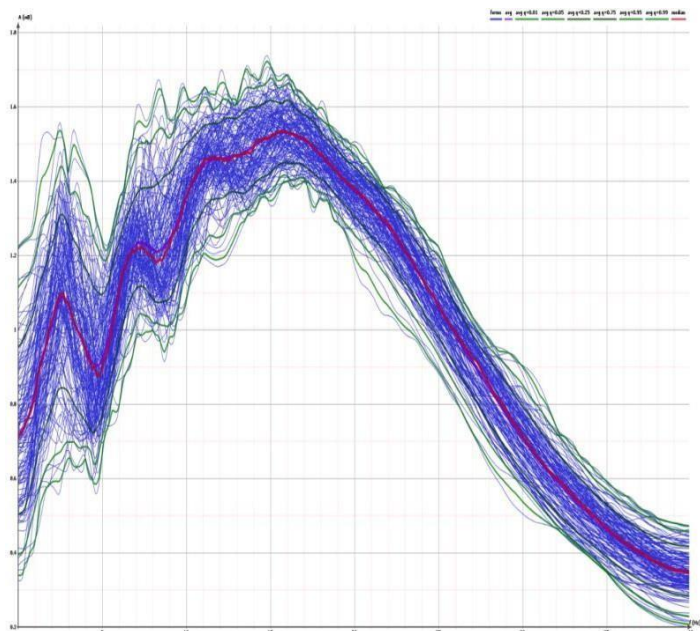


Figure.17 Cluster of a patient with diagnosed ischemia.

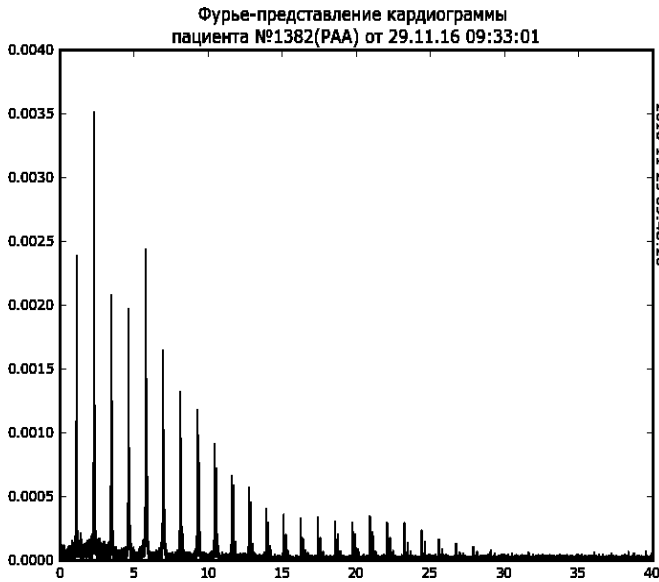


Figure.18 Shows a typical Fourier spectrum of an ECG of a patient with a pacemaker. Typical Fourier spectrum of an ECG of a patient with a pacemaker Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

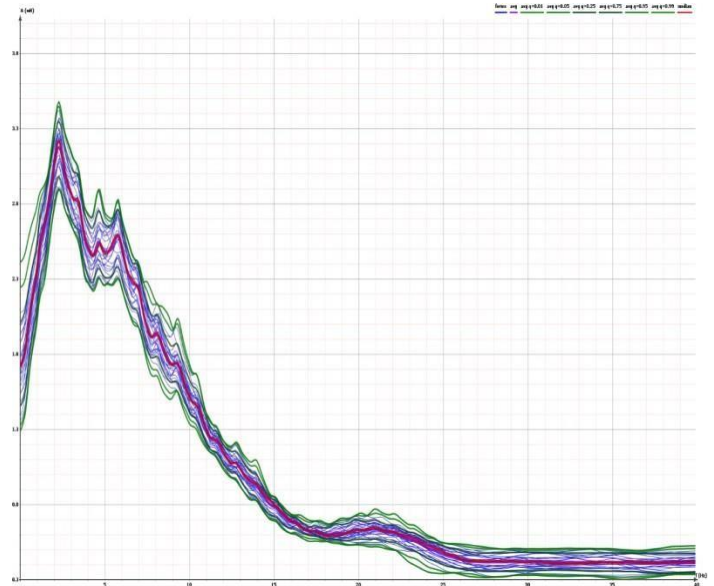


Figure 21. Typical cluster of a patient with a pacemaker.

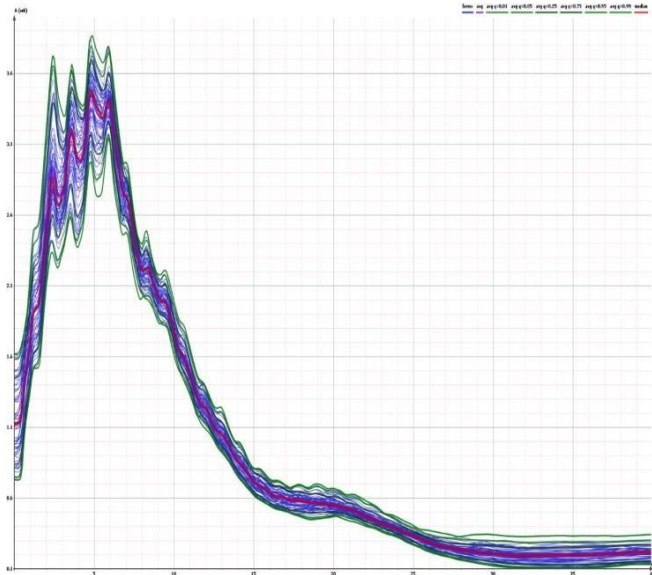


Figure 19. Typical cluster of a patient with apacemaker.

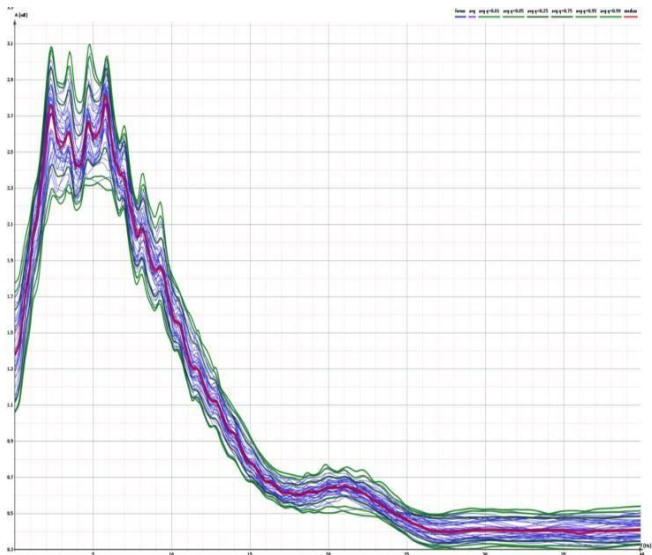


Figure.20 Typical cluster of a patient with apacemaker.

The developed ischemic process in accordance with the hypothesis is the formation of an autonomous dynamic system in the myocardium, which has oscillatory parameters different from the parameters of oscillations in normal myocardial tissue. In particular, the main frequencies of the ECG spectrum of a normally functioning heart lie in the range of 1-12 Hz, while the spectrum of the ischemic process is in the range of 14-20 Hz and can dynamically change. The dynamics of the ischemic process is similar to the dynamics of the Lorentz system [9], which is characterized by different types of attractors and can be reduced to the Van der Pol equation with time lag [8]. Fourier spectrum of ECG in a patient with developed ischemic process is shown in Fig 14. The peculiarity of clusters in patients with ischemia is the presence of a bulge due to the existence of an autonomous ischemic process (Fig.15,16,17). This convexity is absent in the picture of clusters of healthy people (Fig.11,12,13) and patients with pacemakers (Fig.19,20,21). Thus, the clustering method can be used for the diagnosis of coronary artery disease, determining the cardiological status of healthy people including aging parameters, as well as the analysis of the heart condition in patients with pacemakers.

Conclusion and Discussion

The results of experimental studies have shown that the hypothesis that the proposed cluster method can be used both for the analysis of various pathologies in patients as well as for the assessment of the cardiological status of healthy people has been confirmed. In particular, the use of clusters allowed not only to diagnose coronary artery disease in patients with different degrees of development of the ischemic process, but also to assess their level of heart failure. In addition, the cluster approach allowed to analyze the different dynamics of the Fourier spectra of the ECG in patients with pacemakers, despite the dominant spectrum of the oscillations of the electronic generator of the pacemaker, which, as a rule, does not allow to evaluate the spectrum parameters characterizing the condition of the electrical dynamics of the heart.

The introduction of the procedure of dechaotization has improved the analysis of the parameters of the spectra of the ECG in the event of a significant randomization in patients with coronary artery disease and also to differentiate the changes of the spectra in patients with Dystonia and with the symptoms of Ischemia.

Thus, the use of the cluster approach can be useful for the diagnosis of various degrees of ischemic heart disease, determining the cardiological status of healthy people with an assessment of its age-related changes, as well as for the analysis of the level of myocardial status in patients with pacemakers. In addition, the proposed cluster method will allow to evaluate the effect of the therapy in dynamics, and to change it in case of inefficiency.



There are given some essential examples of using the cluster method for diagnostics and age rating.

The process of teaching computer Analytics and ECG/PWdiagnostics

Clusters by category (building standards)

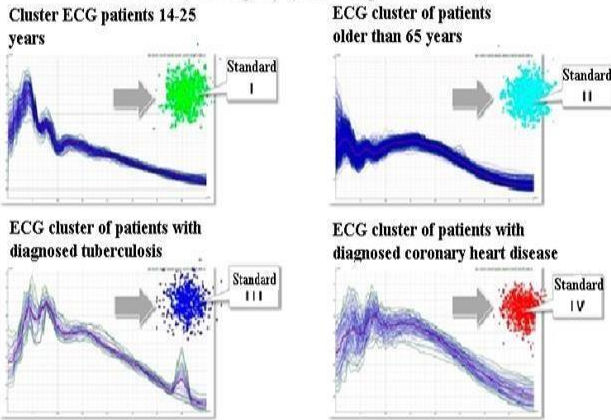


Figure 22.

Geometric interpretation of the patient's condition by the "distance" of his ECG to the standards of clusters (diagnoses)

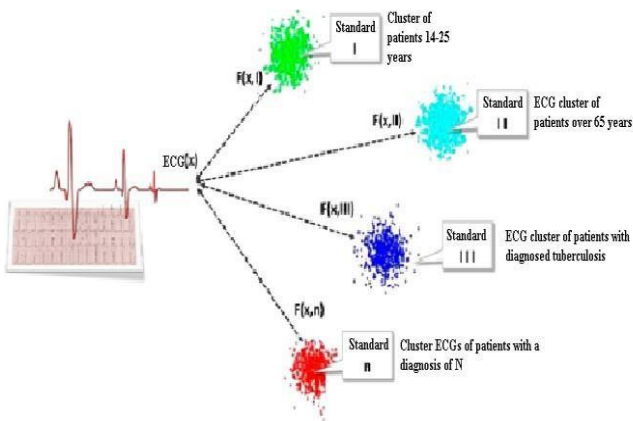


Figure 23.

Analysis of the dynamics of the patient's condition (real-time analysis of "biological" age)

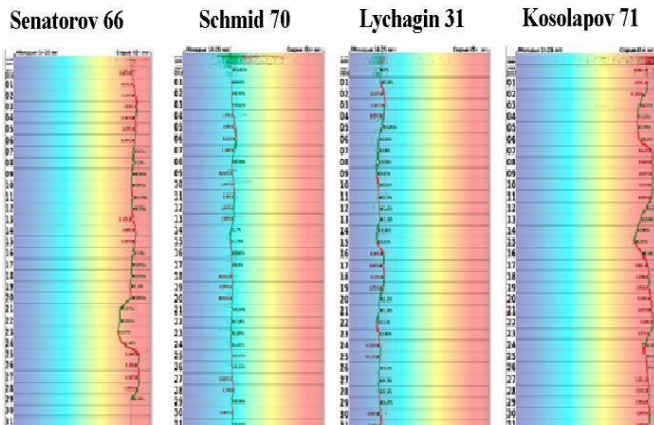


Figure 24.

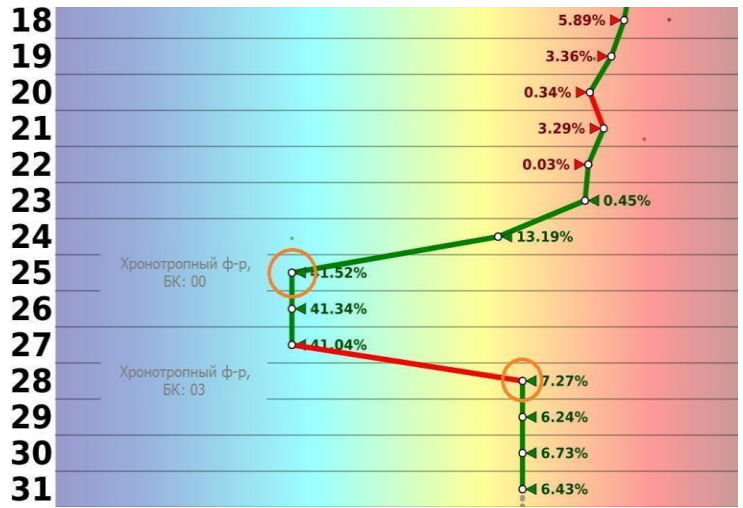


Figure 25. Transitional points indicating sharp changing in the organism functioning.

Statistics of deaths on the background of hypertension no other comorbidities

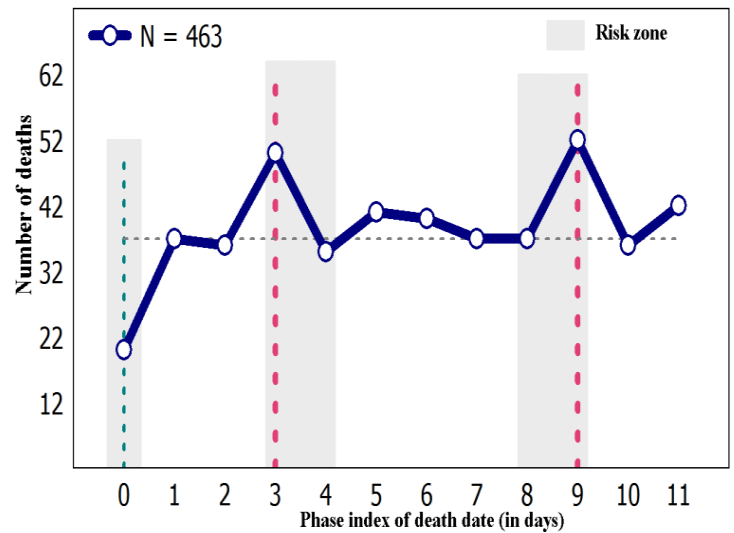


Figure 26 Death rate increase in the transitional points of bifurcation on some special days in a group of cardio patients.

An example of clustering application to identify the energy spectra of the ECG signs of diabetes (processed ~ 250 ECG)

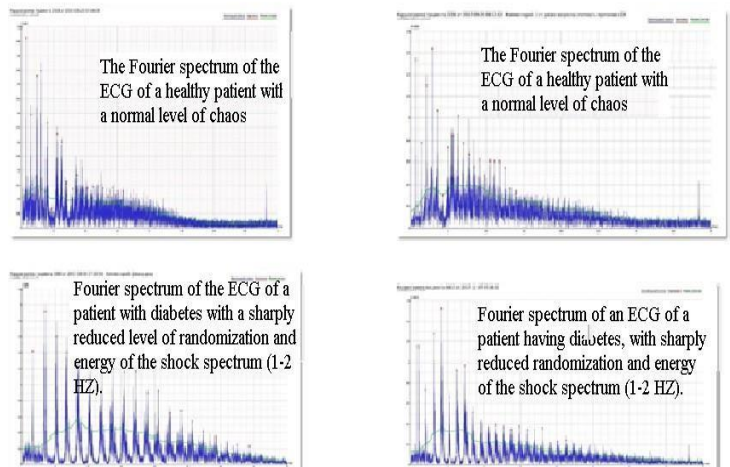


Figure 27.

Frequency ranges of classification of patient conditions

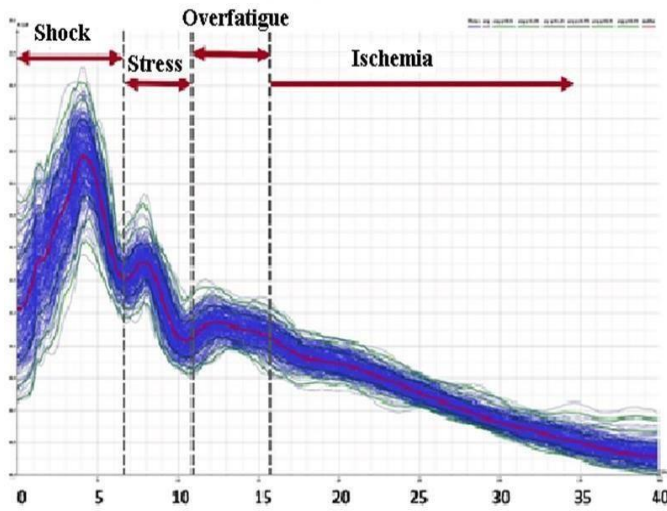


Figure 30.

A comparison of graphical representations of the clusters of EKG healthy patient and a patient with coronary artery disease

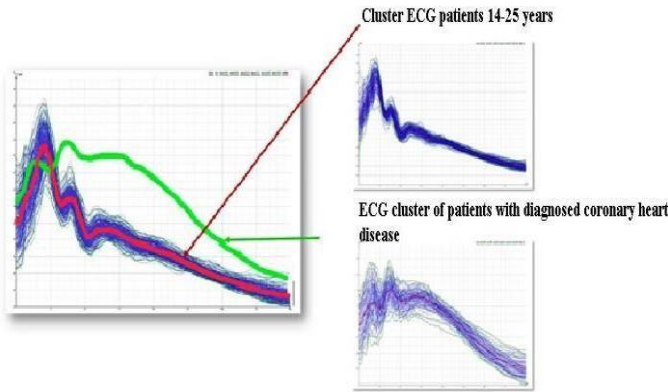
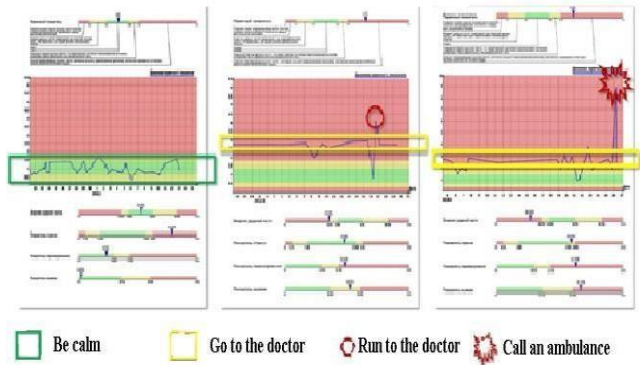


Figure 31.

Digital energy model of the heart (return of Fermi-Ulam Paste)



The analysis is not in the form of ECG, but in the ratio of the energy characteristics of the two main mechanisms of blood transmission: the main (1-6 HZ) and the reserve (18-22 HZ)

Figure 32. Stages of medical anxiety for a patient according to the readings of cardio barometer .

References

1. Fermi E., Pasta J, and Ulam S. In: Collected Papers of E. Fermi. (1955)2, 978.
2. Novopashin, M. A, Shmid, A. V, Berezin A. A. Fermi-Pasta- Ulam auto recurrence in the description of the electrical activity of the heart.//J. Medical Hypothesis.(2017) 102:12-16.
3. Ikezi.H. Solitons in plasmas. Solitons in Action. Academic Press. (1978)163-183.
4. Longren.K. Solitons in nonlinear electrical lines. Academic Press. 1978, p 139-162.
5. Yuen H.C. and Lake B.M. Non-linear dynamics of deep-water gravity waves. Advances in applied mechanics. (1982) 22:69- 229.
6. Zabusky N and Kruskal M. Interaction of solitons in a collisionless plasma and recurrence of initial states. Physical Review Letters (1965) 15(6-9).
7. Rubanik V.P. Oscillations of Quasi Linear Systems with a Time Lag. Ed.“Nauka”, Moscow (1969) 288. (In Russian).
8. Philip R and Scott. jr. Symmetrical Oscillator Having Unsymmetrical Solutions. Proceedings of IEEE. (1976) 64 (5)
9. Lorenz.E.// J. Atmos. Sci.30 (1963) 448.