Statistically Based Criteria for Complex ECG Signal Assessment at Preventive Cardio Control

Aneliya Manukova
Department of Electronics
Ruse
Ruse, Bulgaria
amanukova@uni-ruse.bg

Ivan Georgiev
Department of Applied Mathematics and Statistics
University of Ruse
Ruse, Bulgaria
irgeorgiev@uni-ruse.bg

Matey Marinov
Department of Computer Science
University of Liverpool
Liverpool, United Kingdom
matey.marinov@gmail.com

ABSTRACT

The article presents the main studies, statistical results and justification of the criteria for complex evaluation in the analysis of electrocardiographic signals of clinically healthy people in preventive control to determine the current cardiac abnormalities. The presented study and statistically valid criteria are part of a software system for the evaluation of the individual medical information from the real life of healthy people and a specific method of analysis of the results.

CCS CONCEPTS

• Software and its engineering – Software creation and management – Designing software – Software design engineering, High Relevance

• Theory of computation – Models of computation, Medium Relevance

• Mathematics of computing – Probability and statistics – Statistical paradigms, Medium Relevance

• Applied computing – Life and medical sciences – Health care information systems, High Relevance

KEYWORDS

ECG, statistics, evaluation, analysis, preventive control.

ACM Reference format:

Aneliya Manukova, Ivan Georgiev and Matey Marinov. 2020. Statistically Based Criteria for Complex ECG Signal Assessment at Preventive Cardio Control. In Proceedings of the 21st International Conference on Computer Systems and Technologies (CompSysTech '20), June 19-20, 2020, Ruse, Bulgaria. *ACM, New York, NY, USA, 6 pages.*
https://doi.org/10.1145/3407982.3407985

ACM ISBN 978-1-4503-7768-3/20/06...$15.00

https://doi.org/10.1145/3407982.3407985

1. Introduction

In medical science, statistical probability and stochastic methods and approaches for processing, evaluation and analysis of accumulated information to achieve effectiveness in doctors diagnostic practice are increasingly being used [1, 2, 4, 11, 13]. Modern software programs for statistical processing (STATISTIKA, STATGRAPHICS, SAS, SYSTAT, SPSS, etc.) have proven their functionality in various fields of science – engineering, social, medical and marketing [5].

The use of a new approach in the development of an electronic system for assessing individual medical information of healthy people and a specific method of analysis of the results are the basis of the proposed statistical criteria for the complex evaluation of the momentary ECG signal abnormalities in preventive control through early and timely assessment of their momentary state. Estimating quality parameters in the treatment of patients using digital quantitative indicators is an effective method of equitable diagnosis and an excluded subjective factor [9]. This also determines the need to develop criteria for a comprehensive assessment of the current ECG signal disturbances in preventive cardio control based on statistical analysis of the cross-sectional data of the investigated people.

The aim of the article is to define statistically valid criteria for the complex evaluation of the current disturbances of electrocardiographic signals of healthy people in preventive control after emotional or physical exercise, based on presented results from the analysis of the processed ECG parameters.

2. Methodology

In the present study, statistical analyzes were carried out using one of the most widely used specialized software products for statistical processing – SPSS (Statistical Package for the Social Sciences). The advantages of choosing the product SPSS [4, 5, 7], are:

- obtaining aggregated distributions of the values of a variable, along with most of the statistical indicators characterizing the distribution;

- quickly modify and restructure the data array;
- sectional summation of each of the variables in the dataset;
- opportunities for applying a variety of complex methods of statistical analysis – variation, correlation, regression, dispersion, discriminative, factorial, cluster, etc.

To verify the effectiveness of the experimental methodology, links are examined and statistical hypotheses are examined for the sample subunits studied. The SPSS is structured with an array of survey data to maximize product integration and statistical data analysis. For the evaluation of the relations, different types of correlation coefficients are defined depending on the type of variables – quantitative, regular and qualitative. All elevated statistical hypotheses were made at a level of significance $\alpha=0.05$. For all tests for significance of the correlation coefficients in which values of p-value is less than α should be interpreted as statistically significant [4, 5].

To analyze carefully connections between quantitative traits in the survey is used, a correlation coefficient of Pearson (1), was evaluated its statistical significance.

$$\rho_{Pearson} = \frac{\sum \frac{XY}{n-1} - \frac{\sum X \sum Y}{n.(n-1)}}{\sigma_X . \sigma_Y}, \tag{1}$$

where: σ_X is a standard deviation of X; σ_Y is a standard deviation of Y; и n – volume of the sample.

The Pearson Correlation Test uses the Student criterion (T-criterion) by comparing the empirical (2) and theoretical (3) characteristics. With a less empirical versus theoretical characteristic, the Pearson correlation coefficient is considered statistically insignificant [5, 7].

$$T_{emp} = \frac{\rho_{Pearson} . \sqrt{n-2}}{\sqrt{1-\rho^2_{Pearson}}} \tag{2}$$

$$T_{th} = T_{\frac{1-\alpha}{2};n-1} \tag{3}$$

The formula (3) is empirical characterization characteristic – a quantum of Students distribution at a specified level of significance α, with (n-1) degrees of freedom.

In the study, Spearmen correlation coefficient (4) and Kendalls tau -b correlation coefficient (5) were used to determine relationships between rancid signs. For both coefficients, their statistical significance was assessed.

$$\rho_{Spearman} = 1 - \frac{6.\sum d^2}{n.(n^2-1)}, \tag{4}$$

where: d is the difference in the ranks and n is the volume of the sample.

$$\rho_{Kendall'tau-b} = 1 - \frac{n_c - n_d}{\sqrt{(n_0 - n_1).(n_0 - n_2)}}, \tag{5}$$

where:

$$n_0 = \frac{n.(n-1)}{2}; \quad n_1 = \sum_i \frac{t_i.(t_i-1)}{2}; \quad n_2 = \sum_j \frac{u_j.(u_j-1)}{2};$$

n_c is the number of concordant pairs; n_d is the number of discordant pairs;

t_i is number of values in the i^{th} group of ties for the first quantity;

u_j is number of values in the j^{th} group of ties for the second quantity [7].

For the connections developed between qualitative features, the Phi and Cramers V (6) coefficient was used in the study and its statistical significance.

$$\rho_{Cramer's\,V} = \sqrt{\frac{\chi^2}{n.(k-1)}}, \tag{6}$$

where: χ^2 is the Pearson chi-square statistic from the aforementioned test; k is the lesser number of categories of either variable and n – volume of the sample.

Regarding the normal distribution of the samples, a classic normality test was performed according to the criterion of Kolmogorov–Smirnov.

Parametric and nonparametric criteria for verifying the relevance of the links between the signs are used in the course of the study. The mean values of the groups under consideration, the values of the coefficients obtained, as well as the respective levels of significance in the two-sided critical area (2-Tail Sig.), are obtained by SPSS using the indicated methods.

3. Analysis of experimental statistical data

In the series of experimental studies in laboratory and field conditions on the territory of the University of Ruse from 2016 to 2018, statistically significantly enough respondents voluntarily participated, whether they agreed to study and declared that they did not use drugs for cardiac control.

Each of them has been tested through the developed AMEG software system using the standard method of analyzing ECG signal by medical criteria [14, 15] and by the developed criteria for preventive control [3, 8, 10]:

- FOM, % – The myocardial rest phase index gives the ratio of the myocardial rest phase to the entire heart cycle;
- Fcc, % – The index expresses the physiology of the heart systole by signs and defines AV connection violations – accelerating or slowing AV conduction;
- PQc, s – The adjusted interval estimates the physiological time distribution. The whole heart systole was distributed in the following proportions: the ventricle systole occupies 70%, the interval PQ occupies 30%.
- PQs, % – The index is a criterion for assessing the accuracy of the AV connection and objectivizing the risk of cardiac arrhythmias in tachycardia;
- QTc, s – Appropriate heart rate QT interval using a constant or actual QT interval;

- Max FHS, bpm – the maximum possible rhythm for the given cardiogram determined on the basis of its actual resting RV interval;
- SI, % – The index assesses the state and dynamics of the chamber systole;
- FP, % – The index measures the condition of the myocardial ablation function;
- PQd, % – Relative deviation of the PQ interval.

The Kolmogorov–Smirnov test for normality found that the distribution of patients by age was normal and respected the Gaussian curve. The sample complies with all requirements for statistical processing and analysis: (i) The volume of the sample is large enough; (ii) The selection of the study groups is done in accordance with the principle of randomness and observing the homogeneity requirement of the samples at the beginning of the experiment. Accordingly, the sample in question is representative and faithfully reflects the general population.

The variation analysis of experimental results on the tested evidence of established database in SPSS determine basic numerical characteristics of indicators examined two states – stress (ST) and calm (CA).

Examine the average of prevention indicators

To achieve the goal and solve the tasks of the study, a partial variance analysis of the variables to determine the mean value as a basic numerical characteristic of the variables was performed and the trend of its variation for each of the experimental groups was determined: (AU) driving in extreme conditions, (EX) emotional stress during the examination, (FIT) physical exercise in the gym, (DAN) physical exercise during the play of Bulgarian folk dances, (TIN) physical exercise during swimming.

Researches of variance analysis for different groups and states are represented by their numerical characteristics in various sections by selected attributes in Figure 1 for one of the closely related and significant indexes – FOM. The trend of their change is determined.

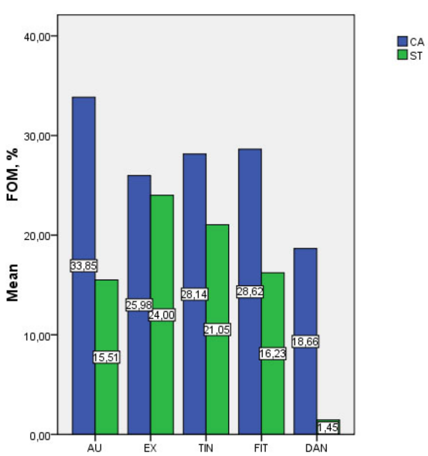


Figure 1: Variation analysis of Index FOM by groups and condition, in percentages.

Study of correlation dependencies of quantitative and qualitative signs

To achieve the goal, we have determined the coefficients linking the investigated indices and the ECG signal parameters determined by the criteria of conviction and the medical criteria. A statistical analysis has been carried out to prove the strength and significance of the dependencies between the individual quantitative and qualitative features. The indications of the indexes used and the coefficients of significance of the coefficient and its strength are presented in Table 1.

The introduced symbol from Table 1 presents the results of the verification of the statistical hypothesis on the significance of the Pearson, Kendal and Spearmen correlation coefficients, Tables 2 and 3. A color scale based on the obtained coefficients was used for the visual perception of the different relationships between the rating indices.

Table 1. Denominations of coefficients of significance of the coefficient and its strength

Mark	**Distribution of link strength**
A	over 0.9 very strong, almost functional
B	0.7 ... 0.9 very strong connection
C	0.5 ... 0.7 strong connection
D	0.3 ... 0.5 moderate
E	0.1 ... 0.3 weak connection
F	under 0.1 very weak link or missing one
Grade	**Significance factor**
0	Statistically insignificant coefficient Very weak link, almost absent
1	Statistically significant, at significance level α = 5% Low connection
2	Statistically significant, at significance level α = 1% Moderate, strong to very strong relationship

Table 2 presents the results of the study at a standard level of statistical significance α=0.05 and α=0.01. On the basis of the results obtained, the study indicators for the complex evaluation were selected.

Table 2

Correlational dependence (between quantitative variables) on the prevention indicators, estimated by the strength and statistical significance of the Pearson coefficient

Index	Fcc	PQd	PQs	SI	FOM	FP	QTc	PQc	FHS
Fcc	–	B2	F0	E1	D2	E1	F0	E2	E2
PQd	B2	–	E0	D2	F0	E1	E0	D2	E1
PQs	F0	E0	–	F0	F0	F0	F0	E1	E1
SI	E1	D2	F0	–	B2	F0	D2	C2	D2
FOM	D2	F0	F0	B2	–	F0	D2	D2	D2
FP	E1	E1	F0	F0	F0	–	F0	F0	F0
QTc	F0	E0	F0	D2	D2	F0	–	D2	D2
PQc	E2	D2	E1	C2	D2	F0	D2	–	C2
FHS	E2	E1	E1	D2	D2	F0	D2	C2	–

Table 3
Correlational dependence (between qualitative variables) of indicators on medical criteria, estimated by the strength and statistical significance of the Spearmen and Kendal coefficient

State	1	2	3	4	5	6	7	8	9	10	11	12
1	–	D2	E2	F0	E1	E1	D2	F0	D2	E1	F0	E0
2	D2	–	E2	F0	F0	F0	E1	F0	E2	F0	E0	E1
3	E2	E2	–	F0	F0	F0	F0	F0	E2	F0	F0	F0
4	F0	F0	F0	–	F0	F0	F0	F0	F0	D2	F0	E1
5	E1	F0	F0	F0	–	F0	E2	F0	E1	F0	E2	F0
6	E1	F0	F0	F0	F0	–	E1	D2	F0	F0	E0	E1
7	D2	E1	F0	F0	E2	E1	–	D2	D2	E2	F0	F0
8	FO	F0	F0	F0	F0	D2	D2	–	E0	F0	F0	F0
9	D2	E2	E2	F0	E1	F0	D2	E0	–	E2	E0	F0
10	E1	F0	F0	D2	F0	F0	E2	F0	E2	–	E0	E1
11	F0	E0	F0	F0	E2	E0	F0	F0	E0	E0	–	E2
12	E0	E1	F0	E1	F0	E1	F0	F0	F0	E1	E2	–

Legend:
1 Risk of accelerating/delaying AV joints (Fcc index rating)
2 Risk of AV Disorders (PQd Assessment)
3 Risk of disturbed rhythm in AV compounds (Evaluation of PQs rhythm)
4 Risk of disturbed PQ interval (Evaluation of PQs-PQ)
5 Risk to the structural and functional pathologies (Evaluation by FOM Index)
6 Risk of abnormal heart function impairment (FP Rating Index))
7 Overloaded P wave (medical evaluation)
8 Prerequisite for AV block (medical evaluation)
9 A prerequisite for WPW syndrome (medical evaluation)
10 Prerequisite for LGL Syndrome (medical evaluation)
11 Abnormal Q cranium (medical evaluation)
12 Abnormal T wave (medical evaluation)

According to the correlation analysis, the indicator Index Fcc has a strong direct dependence on Deviation PQd (0.835) and in moderate inverse dependence on Index FOM (-0.306). The level of significance is low at Index SI (-0.140) and Index FP (0.169). There is significantly strong relationship between Deviation PQd with Index Fcc, Index SI and Interval PQc. Low significance is obtained between Index FP и Index Max FHS. The Index PQs and the Index FP lack a strong significant relationship between the individual indicators. Index FOM is in strong to moderate correlation with the following indicators Index Fcc, Index SI (-0.883), Interval QTc (-0.434), Interval PQc (-0.465) and Index Max FHS.

The sign of the coefficients obtained is related to the direction of the link. The positive sign determines the right dependence – the increase of one parameter corresponds to the increase of the other. In case of a negative sign the connection is inverse - the increase of one parameter corresponds with the reduction of the other.

For each indicator in Tables 2 and 3, significant relationships are established that lead to accurate ECG curve analysis. The established dependencies correspond to the medical indicators and prove the significance of the proposed evaluations in the preventive control.

4. Results and discussion

Evaluating the Impact of the Indicators

The ST interval is part of the QT interval and provides the information needed to analyze cardiac activity. The area under consideration is related to the main phase of repolarization of the cardio myocyte action potential and the plateau phase that provides the strength and quality of the myocardial abbreviation [6,12]. The phase-to-plate ratio of the electrical systolic is more than 50%, and the FP lowering of less than 45% may indicate the occurrence of myocardial contracture irregularities. The myocardial ablation function should be monitored not only at rest but also at exercise.

When analyzing the range of chart values, including the AU, EX, TIN, and FIT groups at CA, the average index value increases and the STF decreases at load (ST). In the DAN group, the load is concentrated and the determination of the ratio of the phase length of the plateau to the duration of the overall repolarization of the action potential is different. After correcting the cellular metabolism of ECG cardio myocytes, the plateau phase is normalized to normal ranges.

The FOM in a quiet mode for all study groups expresses the clinically significant chamber systolic pathology associated with an increase in systolic duration, expressed in the QRS extension and the QT interval above a threshold. The influence of the FOM index in a stress state expresses the phase of myocardial rest and is less at load.

The SI index is used to diagnose prolonged and shortened QT interval. In a quiet mode, a lower level is indicated, and in the case of a stress state in all groups the assessment of the state and dynamics of the electric systole of the cameras is with a higher percentage. An exception is an EX group where the stress regime is not physical but emotional.

Criteria for complex cross-sectional evaluation of the ECG signal generated on the basis of statistical evaluation.

The conducted statistical analysis provides a quantitative assessment of the qualitative ranged indicators for the ECG signal assessment and proves the correlation between the nine established assessment indicators and the medical norms, with the overall ECG signal assessment being based on the current state of prevention.

Quantification estimates based on significant proven relationships and proposed analysis algorithms [8] are summarized and defined in two new weighting ratios: K_{PreC} – for qualitative assessments on preventive control criteria (7), and K_{MedC} – for qualitative assessments of medical criteria (8).

$$K_{PreC} = \frac{1}{6}\sum_{j=1}^{6} a_j \,, \tag{7}$$

$$K_{MedC} = \frac{1}{6}\sum_{j=1}^{6} b_j \,, \tag{8}$$

Any change in the status of the test subject is further assessed by quantitative mass evaluation a_j and b_j, by assigning 1 in the absence of a physiological change, and 2 in the presence of such an alteration, evaluated according to the criteria for preventive control and the medical criteria in Tables 2 and 3.

The resulting aggregated weighting factors K_{PreC} and K_{MedC} are changed in the range from 1 to 2. In the absence of momentary changes, they are of minimal significance 1 and in the case of all changes – by a maximum of 2. The obtained evaluation interval is divided into three zones, Table 4, and they describe unambiguously the state of people in prevention.

The coefficients 1.3 and 1.7 in table 4 are defined from the statistical researches based on the studied group of people, their physical and emotional conditions and the area they live in.

Table 4

Evaluation intervals of aggregate weighting factors K_{PreC} and K_{MedC}.

Interval	Emoticon	Description
$1 \le K_{PreC} \le 1{,}3$ $1 \le K_{MedC} \le 1{,}3$		Good general condition
$1{,}3 < K_{PreC} < 1{,}7$ $1{,}3 < K_{MedC} < 1{,}7$		Need to rest
$1{,}7 \le K_{PreC} \le 2$ $1{,}7 \le K_{MedC} \le 2$		Recommended consultation with a doctor

The two weighting factors K_{PreC} and K_{MedC} have been tested for statistical dependence with level of significance $\alpha=0.05$. A strong positive relationship has been demonstrated between them in both cases: Evaluation in preventive control as a function of the medical criteria assessment – $K_{PreC} = f(K_{MedC})$, and Medical Criteria Assessment as a function of the Preventive Control Assessment – $K_{MedC} = f(K_{PreC})$. The resulting dependencies determine a positive relationship between the prevention criteria and the medical criteria.

All coefficients in the deduced two linear regression models (right and inverse dependence) – K_{PreC} (9) and K_{MedC} (10), are statistically significant (Sig.<α). For both models the coefficient of determination R^2 is statistically significant (Sig.<α), which states that the models are adequate.

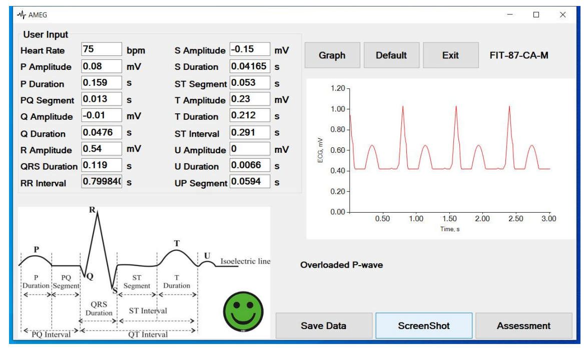
$$K_{PreC} = 0.215 \times K_{MedC} + 1.359, \quad (9)$$

$$K_{MedC} = 0.339 \times K_{PreC} + 0.656, \quad (10)$$

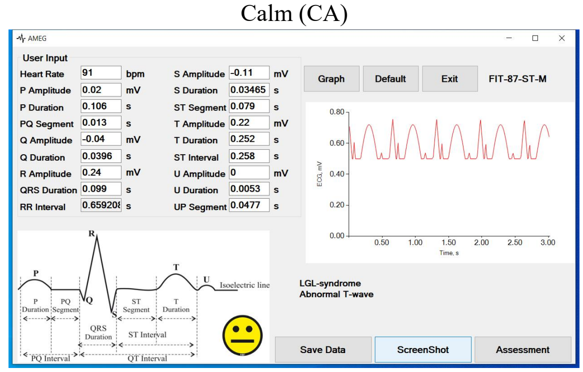
Some interest is the aggregate coefficient $K_{MedC} = f(K_{PreC})$ (10), by which a medical assessment is made on the criteria for preventive control, i.e. feedback is provided to the groups of people surveyed. The deduced dependencies prove that the proposed methodology for analyzing ECG signals for preventive control and the evaluation obtained are significant and should be used for various emotional or physical exertions.

Fig. 2 presents the result of the software analysis system with the visual evaluation of ECG signals after the statistical evaluation of the introduced coefficients of weighting K_{PreC} and K_{MedC}. The visual emoticons give a quick and clear analysis of the current state or current disruption of the heart cycle.

When comparing the two criteria – medical and prevention – it was found that the analysis by the prevention criteria was significantly more extensive in the case of automated data processing than the medical criteria. The reconciliation of the two methods of investigation gives an extremely clear and accurate assessment of the current status of those investigated for prevention.

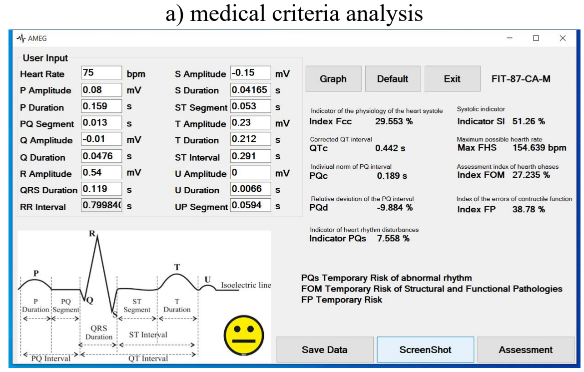


Calm (CA)

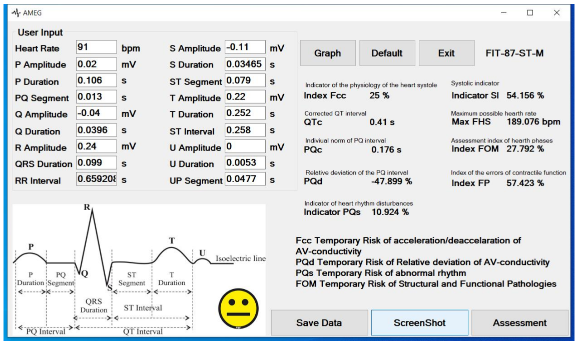


Stress (ST)

a) medical criteria analysis



Calm (CA)



Stress (ST)

b) analysis on preventive control criteria

Figure 2: Experimental ECG data in a relaxed and stressful mode for the study group FIT.

5. Conclusion

There have been conducted statistic studies to assess the current status of the ECG signal by quantitative, qualitative and qualitative indicators and a correlation between the introduced preventive control indices and the preventive medical standards.

Through the visual assessment for the categorization of the current disorders and the state of the heart cycle, the two methods of study were reconciled and the prevention assessment was made more precise.

Statistical criteria for complex assessment of the current disturbances of electrocardiographic signals of healthy people in preventive control after emotional or physical load are justified and defined on the basis of presented results of the analysis of the processed ECG parameters.

ACKNOWLEDGMENTS

The article reflects the results of the work on the project No. 2020–FOZZG–02 "Development of algorithms for diagnosis and treatment of problematic skin wounds by the method of enriched platelet rich plasma", funded by the Research Fund of the University of Ruse.

REFERENCES

[1] Anouncia, M. S., Cl. Madonna L. J., Jeevitha P., Nandhini R. T. Design of a Diabetic Diagnosis System Using Rough Sets. – Cybernetics and Information Technologies, Vol. 13, 2013, No 3, pp. 124-139, DOI: 10.2478/cait-2013-0030.

[2] Boveda, S., et al. Prognostic Value of Heart Rate Variability in Time Domain Analysis in Congestive Heart Failure. – Journal of Interventional Cardiac Electrophysiology, Vol. 5, 2001, pp. 181-187.

[3] Dotsinsky, I. Software Real Time QRS Detection. – In: Proc. of World Congress on Medical Physics and Biomedical Engineering, 7-12 July 1991, Kyoto, Japan. Medical & Biological Engineering & Computing (supplement), 29, p. 337.

[4] Field, A. Discovering Statistics Using IBM Statistics, 4th ed, SAGE, 2013, ISBN 978-1-4462-4917-8, p. 915.

[5] Ganeva, Z. Discovering Statistics Using IBM SPSS Statistics, Publ. ELESTRA, 2016, ISBN 978-619-7292-01-5, p. 713.

[6] Himanshu, G., S. Kedawat, R. Kumar. Cardiac Arrhythmias Detection in an ECG Beat Signal Using Fast Fourier Transform and Artificial Neural Network. – Journal of Biomedical Science and Engineering, Vol. 4, 2011, pp. 289-296.

[7] Landau S., B. Everitt, A Handbook of Statistical Analyses using SPSS. CHAPMAN & HALL/CRC, 2004, 339 p, ISBN 1-58488-369-3.

[8] Manukova, A., M. Marinov, M. Grozeva. An Approach to Evaluation of Clinically Healthy People by Preventive Cardio Control. – Cybernetics and Information Technologies, Vol.19, 2019, No 2, pp. 133-145, ISSN 1311-9702, DOI: 10.2478/cait-2019-0020.

[9] Manukova-Marinova A., Ts. Sokolov, M. Marinov. Algorithm and Software System for Treatment Application of Platelet-Rich Plasma on Problematic Skin Wounds. Cybernetics and Information Technology, BAS, 2020, Volume 20, No 1, pp. 129-137, DOI: 10.2478/cait-2020-0009, ISSN Print ISSN:1311-9. (SJR rank: 0.22 /2018, https://www.scimagoir.com/)

[10] Manukova, A., M. Grozeva, M. Marinov. Criteria for Evaluation, Methodology and Analysis of Electrocardiographic Signal Results from Experimental Studies in Preventive Control of Clinically Healthy People. – Electronics – Design, Technology, Applications, Poland, Vol. 1, 2018, pp. 40-44. DOI: 10.15199/13.2018.1.9.

[11] Prutchi, D., Norris M. Design and Development of Medical Electronic Instrumentation: A Practical Perspective of the Design, Construction and Test of Medical Devices. Wiley Interscience, 2004.

[12] Rudenko, M., V. A. Zernov, O. K. Voronova. Study of Hemodynamic Parameters Using Phase Analysis of the Cardiac Cycle. – Biomedical Engineering, Springer, New York, Vol. 43, 2009, No 4, pp. 151-155.

[13] A. T. Reisner, G. D. Clifford, R. G. Mark, Eds. Advanced Methods & Tools for ECG Data Analysis. Chapter 1: The Physiological Basis of the Electrocardiogram. Accessed on 16.08.2017. http://www.mit.edu/~gari/ecgbook/ch1.pdf.

[14] Vorobiov, L.V. ECG Analysis of Cardiac Activity of a Healthy Person. – International Journal of Applied and Fundamental Research, Vol. 10, 2016, No 4, pp. 549-553 (in Russian). ISSN 1996-3955.

[15] Marchev, S. How to Read an Electrocardiogram. LaxBook, Plovdiv, 2014, ISBN 978-619-189-002-6, 159 p. (in Bulgarian).