Conference Poster

Data Analysis of Random Blood Measurements for Abnormal Condition Detection

Sokol, Y., Shchapov, P., Tomashevskyi, R., Veligorskyi, O., Picking, R and Chakirov, R

This is a paper presented at the 7th IEEE Int. Conference on Internet Technologies and Applications ITA-17, Wrexham, UK, 12-15 September 2017.

Data Analysis of Random Blood Measurements for Abnormal Condition Detection

Yevgen Sokol, Pavel Shchapov, Roman Tomashevskyi
Department of Industrial and Biomedical Electronics,
National Technical Uni. "Kharkiv Politechnical Institute",
Kharkiv, Ukraine

Oleksandr Veligorskyi
Depart. of Biomedical Radioelectronic Apparatus and Systems,
Chernihiv National University of Technology,
Chernihiv, Ukraine

Richard Picking
School of Applied Science, Computing and Engineering,
Glyndwr University, Plas Coch, Mold Road,
Wrexham, LL11 2AW, UK

Roustiam Chakirov
Department of Electrical and Mechanical Engineering,
Bonn-Rhein-Sieg University of Applied Science,
20 Grantham-Allee, Sankt Augustin, D-53757, Germany

***Abstract*—This paper discusses an approach of the abnormal condition detection of whole blood using piezo-synthetic effects in blood under dynamic external pressure. Three groups of samples having verified chemical and biological conditions were analysed to prove reliable detection: saline, whole blood and whole blood with colorectal cancer as an example of abnormal conditions. The procedure of a discrete differentiation process for obtained experimental data has been proposed as preliminary processing. Three information parameters have been selected to describe experimental data. Fischer F-statistics were used to determine the information content of the proposed information parameters. It has been proved that the proposed information parameters react on changing state of object under test and therefore can be effectively used for the abnormal condition detection.**

Keywords—biological object, factor load, information parameter, dynamic experiment, cancer cells detection

I. Introduction

There are many techniques to diagnose cancer. Some classical techniques such as colonoscopy [1], [2], mammography [3] or ultrasound [4] are applied to detect special types of cancer and can be used both standalone and as an additional procedure for cancer diagnosis. However, sometimes these methods can miss many cancer diseases, especially for mammography in case of dense-breasted cancer. Colonoscopy and biopsy [5] are invasive techniques, which can cause patient discomfort during the procedure. Other techniques, such as Magnetic Resonance Imaging (MRI) [6] or Positron Emission Tomography-Computerized Tomography (PET-CT) cause radiation exposure (in the case of PET-CT) and require expensive equipment.

Beside the classic methods some new techniques that used a patient's blood as an object of the analysis were proposed recently. Blood and cancer cells have different chemical and physical characteristics. Plasma membrane capacitance of cancer cells is two times more than for lymphocytes. That allows using special techniques like dielectrophoretic separation of cancer cells from blood [7] for the detection of cancer cells. For example, a novel approach for cancer detection based on biochemical analysis of peripheral blood plasma by using Fourier transform infrared spectroscopy was proposed [8]. That method provides minimal-invasive technique to detect any form of cancer, but requires special blood sample processing, like centrifugation, drying samples, etc.

In this paper we propose a new approach to detect the cancer cells in blood using pressure variable factor load, electrical potential registration and information analysis of measured signals. The paper presents an experimental investigation of different liquid biological material samples (whole blood), processing of experimental data and information parameter analysis, as well as proving difference in parameters for different samples that allow detecting blood with cancer cells.

II. Mechanic Factor Influences on Blood

Electrical characteristics of blood can be described using a classical well-known three-element model [9] shown in Fig. 1.

Resistance R_p of the model corresponds to the electrical resistance of plasma, the cell membrane capacitance of the erythrocytes in the blood is represented by the capacitor C_m. Erythrocytes interior cell resistance is described as resistance values for R_i. R_p, R_i and capacitor C_m. The values of the resistances and capacitor depend on different factors, for example, flow speed changes the electrical resistance R_p [10] whereas R_p and C_m depend on the velocity of blood.

Pressure is another external parameter that changes electric characteristics of whole blood. Blood cells became more polarized under mechanical loads that significantly decrease

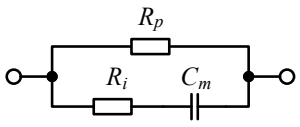

Fig. 1. Three-element electrical model of whole blood.

978-1-5090-4815-1/17/$31.00 ©2017 IEEE

the dielectric constant of membrane components of blood cells [11], [12]. This effect is called the "piezo-synthetic effect in biological tissues". As it was mentioned before, cancer cells have plasma membrane capacitance twice that of lymphocytes [8]. It allows detection of cancer cells in the blood under mechanical loads.

III. Dynamic Experiment

The experimental part of the paper describes the investigation of piezo-synthetic effects in different biological objects with and without pathological states.

Three groups of samples having verified chemical and biological states were chosen for the investigation of electro-physical parameters of liquid biological materials:

- sample S_0 – saline;
- sample S_1 – non-pathology whole blood;
- sample S_2 – whole blood with pathology – colorectal cancer.

Each sample has a volume of 1 cm^3. Variable factor load for biological samples $P(t)$ is a repeatable changing of pressure (up to 1 atmosphere). Different random and uncontrollable factors (perturbations) take place during the experiment, such as:

- non-stationarity on the mathematical expectation of electrochemical drift in the processes of biopotentials' formation at the electrodes of the primary measuring transducer, that causes additional additive conversion error;
- ambiguity in providing conditions for repeatability of the atmospheric pressure values and its changing over time, that causes the appearance of multiplicative errors in the output variable.

A. Experimental Setup

The experimental setup for a dynamic active biophysical experiment is shown in Fig. 2. Two gold-plated electrodes (E_1 and E_2) with different active surface areas have been used for measuring the electrical potential of liquid samples. Electrodes are placed inside the cylindrical cuvette with a piston changing the pressure $P(t)$.

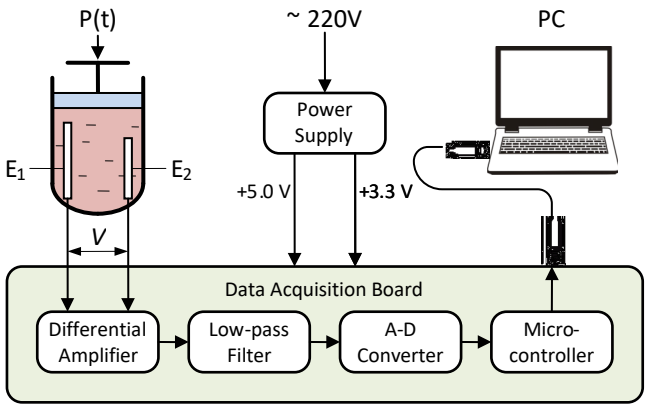

Fig. 2. Experimental setup for dynamic active biophisycal experiment with different blood samples.

This signal to the data acquisition board comes from the electrodes and goes through the differential amplifier, the low-pass filter, the analogue to digital converter (ADC) to the microcontroller. The differential amplifier with adjustable gain from 100 to 1000 provides preliminary amplification and filtering of raw measured difference of electrical potentials, that appears on the electrodes E_1 and E_2, caused by variable factor load $P(t)$.

The low-pass filter selects the measured signal at a frequency up to 100 Hz and sends it to the ADC to convert the signal into the 12-bit digital sequence. Then the microcontroller converts this data into the universal serial bus packages to transmit it to a PC for further processing, visualisation and storing.

Typical duration of cycle of variable factor load for biological samples $P(t)$ is around 4-5 seconds. An example of bioelectric potential response for two sequential cycle of pressure changing is shown on Fig. 3.

B. Preliminary Processing of Experimental Data

Dynamic behavior of factor load determines the selection of the special information parameters related to the dynamics of the measured bio-electric potentials, taking into account variable electrochemical inertia of ionic conductivity processes of samples with increasing and decreasing load. To demonstrate the difference, discrete (digital) differentiation of the measured signal $V(t)$ has been used.

Transformation of centralisation of measured signal has been carried out to eliminate drift additive error in the preliminary transducer (electrodes and differential amplifier):

$$\dot{V}(t) = V(t) - \tilde{V}(t) \tag{1}$$

where $\tilde{V}(t)$ is a parabolic regression, caused by drift of the settled value of the preliminary transducer's measured data, $\dot{V}(t)$ is a centralised value of the measured data.

Normalisation of the centralised data for amplitude level V_m has been used to decrease multiplicative errors caused by non-uniform parameters of the factor load, such as variation of pressure etc. Amplitude level of measured bio-electric potentials V_m has been used for all cycles of factor load $P(t)$.

$$\tilde{V}(t) = \frac{\dot{V}(t)}{V_m} \tag{2}$$

In addition, it was proposed to consider the parameters that have information about samples S_0, S_1 and S_2, as complex parameters, that are based on characteristics of $\tilde{V}(t)$ dynamics (separately for load $P(t)$ increasing and decreasing).

To extract these parameters every process $\tilde{V}(t)$ was differentiated, that allows this to be used for the analysis rate of biopotential's changing $\Delta V(t)$:

$$\Delta V(t) = \frac{d\tilde{V}(t)}{dt} \tag{3}$$

Typical rates of biopotential's changing for samples S_1 and S_2 is shown in Fig. 4 and Fig. 5.

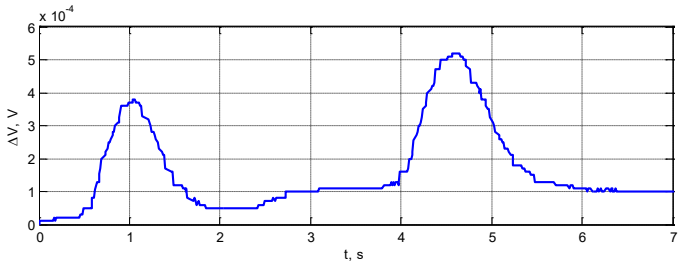

Fig. 3. Example of bioelectric potential response for two variable factor load $P(t)$ changing.

Preliminary analysis of measured and processed data $\Delta V(t)$ for three states S_0, S_1 and S_2 of used liquid biological materials shows the following:

- Positive amplitude $+\Delta V_{max}$ is always more than negative $-\Delta V_{max}$ independently from number of repeatable cycles and the state of the samples.
- The second derivative of the $\tilde{V}(t)$ process $\Delta^2 V(t) = d^2\tilde{V}(t)/dt^2$ has more intersections with the t axis n_- for the negative $V(t)$ compared to the number of intersections n_+ for the positive values $V(t)$ for the states S_1 and S_2.
- State S_0 has $n_+ = n_-$.
- $n_-^{S1} > n_-^{S2}$, where n_-^{S1} is the number of intersections for the second derivative of the $\tilde{V}(t)$ process with the t axis for the negative values $V(t)$ for the states S_1.
- All bioelectric potential responses $V(t)$ independent from the number of cycle and the state are asymmetric, and the asymmetry rates is increasing for changing states from S_0 to S_1 and from S_1 to S_2. It means that the settle time for the process from $-\Delta V_{max}$ to the steady-state value is increasing

All results mentioned above allow us to consider three main information parameters (described below as Y_1, Y_2, Y_3), that meet the next logical requirement: $\Delta V(t) \neq 0$.

The function (4) that depends on rate of change $+\Delta V(t)$, $-\Delta V(t)$ and global/local extremes $+\Delta V_{max}$ and $-\Delta V_{max}$, was chosen as the first information parameter Y_1 describes the biochemical state of the sample.

That parameter is indirectly related to quantum effects of electro-potential transformation in the liquid biological material.

$$Y_1 = F\left(+\Delta V_{MAX}, -\Delta V_{MAX}, +\Delta V(t), -\Delta V(t) \mid \Delta V(t) \neq 0\right) \quad (4)$$

Parameters Y_2 and Y_3 are used as parameters responsible for biochemical states directly connected with quantum effects of electropotential transformation. Parameter Y_2 is a function that depends on number of intersections of $\Delta^2 V(t)$ with t axis:

$$Y_2 = F\left(n_1, n_2, \mid \Delta V(t) <> 0\right) \quad (5)$$

Parameter Y_3 is a function that depends on root-mean square values of phase shift τ_1 and τ_2 of intersection points n_1 and n_2:

$$Y_3 = F\left(\tau_1, \tau_2, \mid \Delta V(t) <> 0\right) \quad (6)$$

IV. Dispersion Analysis of Dynamic Experiment

For planning and analysis of this experimental data we should take into account the following:

- all parameters Y_m (includes Y_1, Y_2, Y_3) are random values due to the internal and external random disturbances;
- every state S_0, S_1, S_2 corresponds to their own set of numerical values (dispersion, mean value, asymmetry and excess coefficients, etc.) of random parameters Y_m.
- A limited amount of experimental data (amount of samples with verified states) allows using only averaged values of analysed parameters Y_1, Y_2, Y_3. Dispersion and high order moments requires additional data in the set.

In such cases, single-factor parametric models for the dispersion analysis can be used, that allows [9]:

- Estimate statistical significance of mean values changing of parameters Y_1, Y_2, Y_3 in the case of transition from state S_k to state S_j $(j \neq k)$, such estimation counts dispersion of the analysed parameters.

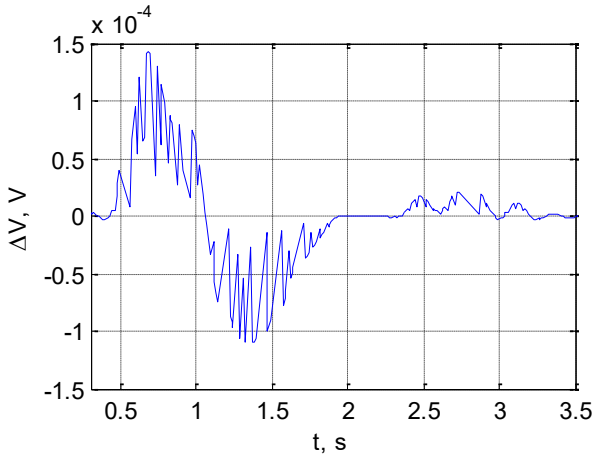

Fig. 4. Rate of biopotential change for the sample S1 (non-pathology whole blood).

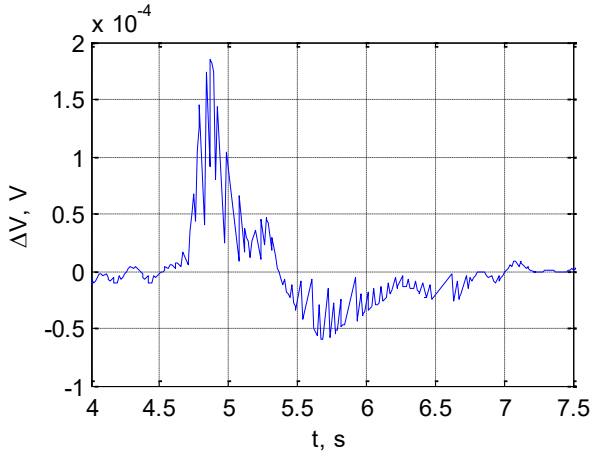

Fig. 7. Rate of biopotential change for the sample S1 (whole blood with pathology – colorectal cancer).

- Estimate information content for any parameter Y_1, Y_2, Y_3 by using test F-statistic. [3] [5]

$$I = 0.5\ln(1+F) \text{ nit} \quad (7)$$

States S_0, S_1, or S_2 are the influence factors for the dispersion analysis. Quality of the analysis is provided by a limited type I errors (significance level $\alpha = 0.05$).

Experiments with several groups of samples have been carried out (number of groups $K = 3$): state S_0 – 8 samples, state S_1 – 5 samples, S_2 – 5 samples, so the total amount of samples are $N = 18$, and $N_1 = 8$, $N_2 = N_3 = 5$.

A. Information Parameters Estimation

Results of estimation of the information parameters will be presented as y_{ji}, where $j = 1…K$; $i = 1…N_j$. A statistical model of such type will be expressed as:

$$y_{ji} = \bar{Y} + \rho_j + z_{ji} \quad (8)$$

where $\bar{Y}$ is a mean value for all N results, ρ_j – deviation, caused by the influence of main factor (pressure), z_{ji} – random deviation of model, caused by uncontrollable and controllable disturbances.

Let's $\bar{Y}_j$ be sat as a mean value for the results in group j, in this case, the dispersion decomposition can be expressed as follows:

$$\sum_{j=1}^{k}\sum_{i=1}^{N_j}\left(y_{ji}-\bar{Y}\right)^2 = \sum_{j=1}^{k}N_j\left(\bar{Y}_j-\bar{Y}\right)^2 + \sum_{j=1}^{k}\sum_{i=1}^{N_j}\left(y_{ji}-\bar{Y}_j\right)^2 \quad (9)$$

Criterial F-statistic as a random value with Fisher-Snedecor's F-distribution with two degrees of freedom $V_1 = K-1 = 2$ and $V_2 = N-K = 15$ will be expressed as the ratio:

$$F_{2;15} = \left(\frac{\sum_{j=1}^{K}N_j\left(\bar{Y}_j-\bar{Y}\right)^2}{\sum_{j=1}^{K}\sum_{i=1}^{N_j}\left(y_{ji}-\bar{Y}_j\right)}\right)\cdot\left(\frac{V_2}{V_1}\right) \quad (10)$$

Table I represents calculation results of information parameters Y_1, Y_2, Y_3 (4)-(6) collected for different states S_0, S_1, S_2 and calculations of criterial F-statistics (10). Critical level of F-statistics for significance level $\alpha = 0.05$ is $F_{CR} = F_{2;15;\alpha} = 3.68$.

As it seen from Table I, the biggest information parameters is Y_3 having F-statistic value $F_{2;15} = 32.98$, the second is Y_2 and finally Y_3. All F-statistics for information parameters exceed critical level, that means Y_1, Y_2, Y_3 can react on changing states of object under test (changing dispersion due to another chemical state of sample $\bar{Y}_j$ 2-9 times more than residual dispersion of random deviations z_{ji}).

Information content of all parameters Y_1, Y_2, Y_3 was calculated by using (7), results are represented in Table II.

B. Estimation of Information Parameters Correlation

Elements of mutual correlation matrices have been calculated to estimate information independence of parameters Y_1, Y_2 and Y_3. The five first columns of Sample information parameter values from Table I ($N_1 = N_2 = N_3 = n = 5$) and two states S_1 and S_2. Obtained matrices for state sets $\{S_1, S_1\}$ and $\{S_2, S_2\}$ are shown below:

$$R_{11} = \begin{bmatrix} 1 & 0.529 & 0.709 \\ 0.529 & 1 & 0.511 \\ 0.709 & 0.511 & 1 \end{bmatrix}$$

$$R_{22} = \begin{bmatrix} 1 & 0.187 & -0.023 \\ 0.187 & 1 & 0.124 \\ -0.023 & 0.124 & 1 \end{bmatrix}$$

If it is assumed the main hypotheses H0 of independence Y_1, Y_2 and Y_3 as $R_{11} = 0$, $R_{22} = 0$, then the criteria F-statistic for any R_{11} and R_{22} elements with correlation R will be represented as follows:

$$F_{1;n-2} = \frac{R^2}{\left(1-R^2\right)}\cdot(n-2) \quad (11)$$

A critical statistic for proving hypothesis H0 will be $F_{CR} = F_{1;3} = 10.13$ for a significance level $\alpha = 0.05$, so the condition of H0 confirmation is:

$$F_{1;n-2} < F_{CF} \quad (12)$$

TABLE I. INFORMATION PARAMETER AND F-STATISTIC VALUES FOR DIFFERENT SAMPLES

Information parameter	State of sample	Sample's information parameter values								F-stat, $F_{2;15}$
Y_1	S_0	1.095	1.364	1.333	1.652	1.227	1.111	1.318	1.292	6.32
	S_1	1.0	1.0	1.34	1.41	1.0	–	–	–	
	S_2	2.433	2.673	2.659	1.238	1.152	–	–	–	
Y_2	S_0	1.667	2.0	4.25	2.0	1.667	1.1	1.444	1.444	12.31
	S_1	2.182	2.0	4.333	3.19	2.125	–	–	–	
	S_2	4.01	9.8	4.75	8.8	4.72	–	–	–	
Y_3	S_0	1	1	1	1	1	1	1	1	32.98
	S_1	3	2	3	3	2	–	–	–	
	S_2	6	7	5	4	3	–	–	–	

TABLE II. INFORMATION CONTENT FOR INFORMATION PARAMETERS

Information parameter	*I, nit*
Y_1	0.995
Y_2	1.294
Y_3	1.763

Finally, by using (11) and (12) calculations confirmation the condition as $R < 0.878$. It is seen from matrices R_{11} and R_{22} that all elements (excluding the main diagonal) satisfied that condition. It means that information parameters Y_1, Y_2 and Y_3 are mutually-independent.

This allows a summary of information obtained from different information parameters, joining them to the integral parameter, as an example, by using linear-additive model of functional transformation [9].

V. CONCLUSION

Three groups of samples having verified chemical and biological states (saline, whole blood and whole blood with colorectal cancer) were chosen for the investigation of electro-physical parameters of liquid biological materials. Three information parameters were selected to describe the rate of biopotential changing that was measured during the dynamic experiments with pressure for different samples. Dispersion analysis of the selected information parameters allows the calculation of F-statistics. Calculation shows that all F-statistics for information parameters exceed critical levels. This means that the selected parameters can react on changing states (normal or pathology) of object under test. So, the results of dynamical experiments and the obtained data analysis show the possibility of abnormal states detection of whole blood. In further research the proposed method of dynamic external pressure and piezo-synthetic effects in whole blood will be extended for bigger sets of samples allowing more accurate statistical analysis and better detection of abnormal states such as cancer cells in whole blood.

ACKNOWLEDGMENT

Authors thanks V. Boyko and P. Zamyatin from State Institution "V.T. Zaytsev Institute of General and Urgent Surgery of NAMS of Ukraine", Kharkiv, Ukraine for their significant help with experiments and biological samples.

This research work was supported by Ukrainian Ministry of Education and Science (Grant № 0115U000542).

REFERENCES

[1] J. Bernal, N. Tajbakhsh at al., "Comparative validation of polyp detection methods in video colonoscopy: Results from the MICCAI 2015 Endoscopic Vision Challenge," *IEEE Trans. on Medical Imaging*, vol. 36, no. 6, pp. 1231–1249, June 2017.

[2] C.D. Johnson, M.-H. Chen et al., "Accuracy of ct colonography for detection of large adenomas and cancers," *New England Journal of Medicine*, vol. 359, no. 12, pp. 1207–1217, Sept. 2008.

[3] T.B. Bevers, B.O. Anderson et al., "Clinical practice guidelines in oncology: Breast cancer screening and diagnosis," *National Comprehensive Cancer Network*, vol. 7, no. 10, pp. 1060–1096, Nov. 2009.

[4] K.M. Kelly, J. Dean, W.S. Comulada, and S.J. Lee, "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts," *European Radiology*, vol. 20, no. 3, pp. 734–742, Mar. 2010.

[5] P. Filipczuk, T. Fevens, A. Krzyżak, and R. Monczak, "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," *IEEE Trans. on Medical Imaging*, vol. 32, no. 12, pp. 2169–2178, July 2013.

[6] K. Machhale, H.B. Nandpuru, V. Kapur, and L. Kosta, "MRI brain cancer classification using hybrid classifier (SVM-KNN)," in *Proc. Int. Conf. on Industrial Instrumentation and Control*, 28-30 May 2015, Pune, India, pp. 60–65.

[7] P.R.C. Gascoyne, X.-B. Wang, Y. Huang, and F.F. Becker, "Dielectrophoretic separation of cancer cells from blood," *IEEE Trans. on Industry Applications*, vol. 33, no. 3, pp. 670–678, May/Jun 1997.

[8] E. Ostrovsky, U. Zelig, I. Gusakova, S. Ariad, S. Mordechai, I. Nisky, and J. Kapilushnik, "Detection of cancer using advanced computerized analysis of infrared spectra of peripheral blood," IEEE *Trans. on Biomedical Engineering*, vol. 60, is. 3, pp. 343–353, Feb. 2013.

[9] Z. Chang, G. Pop, and G. Meijer "A comparison of two- and four-electrode techniques to characterize blood impedance for the frequency range of 100 Hz to 100 MHz," *IEEE Trans. on Biomedical Engineering*, vol. 55, no. 3, pp. 1247–1249, Mar. 2008.

[10] G.A. Pop, Z. Chang, C.J. Slager, B. Kooij, E.D. van Deel, L. Moraru, J. Quak, G.C. Meijer, and D.J. Duncker, "Catheter-based impedance measurements in the right atrium for continuously monitoring hematocrit and estimating blood viscosity changes; An in vivo feasibility study in swine," *Biosensors and Bioelectronics*, vol. 19, no. 12, pp. 1685–1693, July 2004.

[11] P.F. Shchapov, and O.G. Avrunin, *Increasing the Reliability of the Control and Diagnostic Facilities in the Face of Uncertainty*. Kharkiv: HNADU, 2011.

[12] V.V. Boyko, and P.N. Zamyatin, "Changes in the electrical parameters of the cell membranes of biological tissues under influences of mechanical factor," *Kharkiv Surgical School*, vol. 56, no. 5, pp. 9–12, 2012.