Materials Physics and Mechanics 9 (2010) 228-235 Received: March 17, 2010

WAVELETS OF THE IMMUNOFLUORESCENCE DISTRIBUTIONS; SHANNON ENTROPIES, THEIR CENTRAL MOMENTS AND FRACTAL DIMENSIONS FOR MEDICAL DIAGNOSTICS

Nikolay E. Galich

Department of Experimental Physics, St. Petersburg State Polytechnical University

Polytekhnicheskaya 29, St. Petersburg, 195251, Russia

E-mail: n.galich@mail.ru

Abstract. Communication contains the immunology data treatment. New nonlinear methods of immunofluorescence statistical analysis of peripheral blood neutrophils have been developed. We used technology of respiratory burst reaction of DNA fluorescence in the neutrophils cells nuclei due to oxidative activity. The histograms of photon count statistics for the radiant neutrophils populations in flow cytometry experiments are considered. Distributions of the fluorescence flashes frequency as functions of the fluorescence intensity are analyzed. Square of deviations on the mean level of fluorescence flashes number considered as the probability measure of immunofluorescence fluctuations. Shannon information entropy, the square of fluctuations, their wavelet and wavelet spectra are investigated for medical diagnostics. Immunofluorescence fractal structure is analyzed for Hurst exponent and Shannon Weaver indices of biodiversity in the spaces of fluctuations, their wavelet and wavelet spectra. Biodiversity of fluorescence neutrophil populations has increasing tendency for oncology diseases, with anomalous increase of correspondence fractal dimension. We observe the universal exponential distribution of the central moments for information entropies. Parameters of exponential decrease are various for different health statuses. Therefore we have new tools to conduct diagnostics and monitoring of health statuses with the exponentially high sensitivity, precision and accuracy. Health or illness criteria are connected with statistics features of immunofluorescence histograms. Neutrophils populations fluorescence presents the sensitive clear indicator of health status.

1. Introduction

Flow cytometry is one of the main methods of modern analytical biology [1]. The method is based on registrations of rather big collections of photon count statistics for fluorescence cells and DNA. This technological process is performed at rates of thousands of cells per second. There are many applications of flow cytometry to immunology, cell cycle kinetics, cell kinetics, genetics, molecular biology, microbiology, parasitology, bioterrorism, biological oceanology, husbandry and others.

We used flow cytometry for immunology. The statistical instabilities of local immunofluorescence intensity distributions provide the main reason of difficulties for cytometrical histogram interpretations. Average value of fluorescence intensity is smallest than dispersion, asymmetry and others higher statistical moments of intensity fluctuations. We observe the universality of strong exponential growth for central moments of immunofluorescence intensity fluctuations in any fluorescence conditions and various cells [2]. Rapid exponential increase of the statistical moments forms the basic question about the methods of the correct

description of very informative neutrophils fluorescence distributions. Now we present examples for one of the opportunity approaches, which interconnected with simplest information entropy of wavelet transform description for immunofluorescence distributions. Other nonlinear approaches to classification of immunofluorescence for health statuses diagnostics were discussed in [3, 4].

We present the experimental data treatment of highly sensitive quantitative method for registration the inflammatory reactions of organism founded on collective features of cytofluorescence. The method based on the flow cytometrical measurement of the capability of the peripheral blood neutrophils for the reaction of respiratory burst or oxidizing explosion [3, 5]. We used hydroethidine addition with concentration 150 μg/ml for fluorescence initiation. Details of experimental procedures are described in [3]. The fragments of nuclear and mitochondrial DNA with oxidizing metabolism activity and oxidants [2, 3, 5] determine the basic place for localization of the fluorescent dye [1-3, 5], its distribution and concentration, and consequently intensity and statistics of fluorescence. The heterogeneous fluorescence of chromosomes reflects simultaneously the genetic special, individual features and immune response to the pathogenic actions.

It is difficult to distinguish the closely-related signs of different young and old fluorescence neutrophils for the large populations $\sim 10^5$ of cells in the patient blood. Immunofluorescence histograms represent many statistically closely allied and very similar features of our objects i.e. the fluorescence of nuclear DNA for blood neutrophils of various patients. Therefore it is necessary to investigate the distributions of high order statistical moments. High statistical moments allow distinguish various statistical peculiarities of immunofluorescence fluctuations [2]. This important feature is interconnected with a fatal, unavoidable and inevitable noise of physical, biological and technical nature [2, 4]. Some of universal distributions of various correlations and moments for immunofluorescence were noted in [2, 4]. A difference between parameters and special features of these universal distributions gives clear division of all immunofluorescence histograms into three large groups. First histograms group belongs to healthy donors. Two other groups belong to donors with autoimmune and inflammatory diseases. Therefore we have method for diagnostics and monitoring of disease with the exponentially high sensitivity [2, 4]. Some of the illnesses are not diagnosed by standards biochemical methods. Medical standards and statistical data of the immunofluorescence histograms for identifications of health and illnesses are interconnected.

We hope to extract new additional information from immunofluorescence histograms. We will intend to obtain the new kit of means and tools of analysis. Wavelet transformation technique, which accomplishes filtration of high-frequency components of immunofluorescence noise, gives additional information and new tools. We're expanding on a study of immunofluorescence to the wavelet space and to the space of their wavelet spectra, just as when using the traditional Fourier analysis.

We analyze the immunofluorescence wavelets and their wavelet spectra. Nonlinear statistical descriptions of information entropies for central moments of the immunofluorescence flashes number distributions for wavelets are represented in this article. We observe new features of immunofluorescence such as different Shannon entropy density distributions and their various fractal dimensions for medical diagnostics. This complements the original classification of diagnostics [3]. The purpose and direction of these approaches associated with the analysis of real genes expression in real living cells, for a full set of chromosomes and mitochondria. This is actual ‘terra incognita’. It is very important to note that we analyze only long-range correlations of oxidative activity of DNA. Marks of these regions are designated by fluorescence dye.

2. Wavelet transform of fluorescence flashes number

Distributions for deviations from their mean level. Three typical examples of the initial histograms are shown in Fig.1. The area under the final histograms of P(I) normalized to unit; rhomb points correspond to bronchial asthma. Total number of flashes is $N_0 = 76623$; quadrate points correspond to a healthy donor. Common number of flashes is $N_0 = 40109$; triangle points correspond to the oncology disease. Common number of flashes is $N_0 = 40752$.

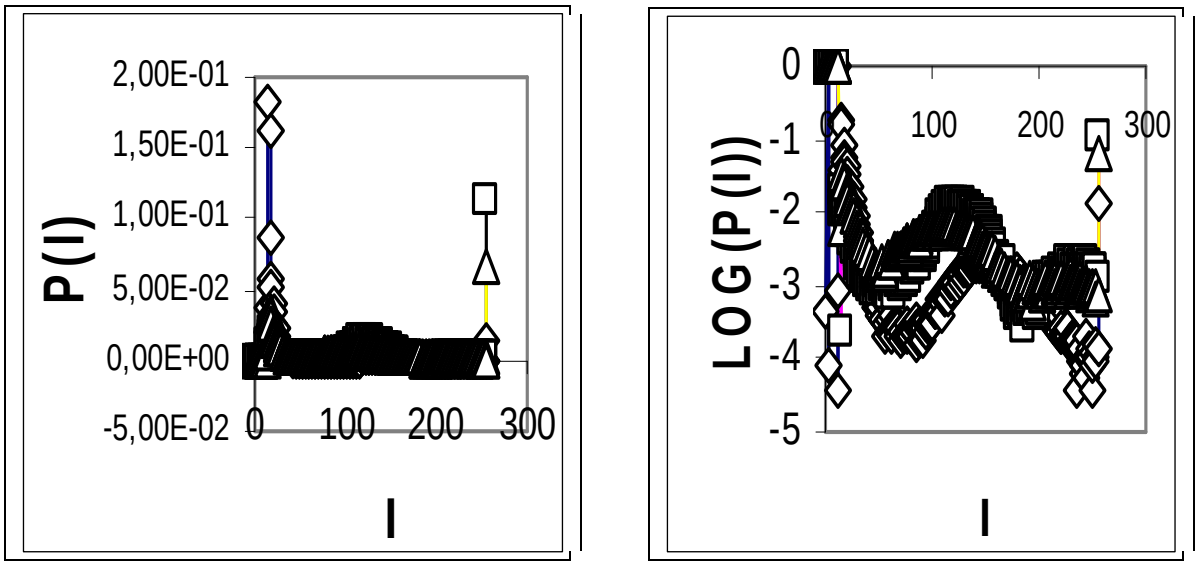

Fig. 1. Dependence of normalized spontaneous fluorescence flashes number P (I) on their intensity I (left), functions of LOG(P(I)) (right).

DNA of neutrophils nuclei absorbs dye. The heterogeneous fluorescence of all chromosomes in the cells reflects simultaneously the genetic special, individual features and immune response to the pathogenic actions due to oxidative activity of DNA. These features are illustrated by histograms individualities in Fig.1. We observe the increase of common number of flashes N_0 for transition from a healthy donor to a donor with autoimmune and to donors with inflammation diseases. Total number of flashes N_0 has shown only one of essential signs of diseases.

Let us consider examples of the central moment distributions for immunofluorescence intensity fluctuations. We use the centered random variables and their centered moments for all procedures on any step. This approach is necessarily for the exception of uncontrollable and systematic errors, instabilities of algorithmic procedures and corresponding drift of averages. Let us consider the relative deflections of fluorescence flashes number from their average level

$$n = (P_l - < P >) = < P > \left(\frac{N}{< N >} - 1 \right), \qquad (1)$$

where $P_l = N(l)/N_0$ is the probability distribution density of the flashes number, l is the number of channels, $l = 1,2,\ldots 256$; N_0 is the total number of flashes; $< P > = (I_{max} - I_{min})^{-1}$ is the mean probability value; $N = N(l)$ is the number of flashes with the assigned intensity $I = l$, for the dimensionless intensity I coincides with the number of channels l; $< N > = N_0 \, x < P >$ is the average value of flashes number; symbol <…> denotes statistical average of the fluorescence fluctuations for all 256 channels of intensity measurement. The mean probability value $< P > = (I_{max} - I_{min})^{-1}$ is equal to 0.03906 for 256 channels of intensity measurement.

Some of the special features of immunofluorescence noise can be separated by means of wavelets technique. More suitable are Doubetchies wavelets. Doubetchies wavelet $iwave(n_l)$ and wavelet spectra $wave(n_l)$ are shown in Fig. 2. Calculations of wavelet conversions are based on MATHCAD technique.

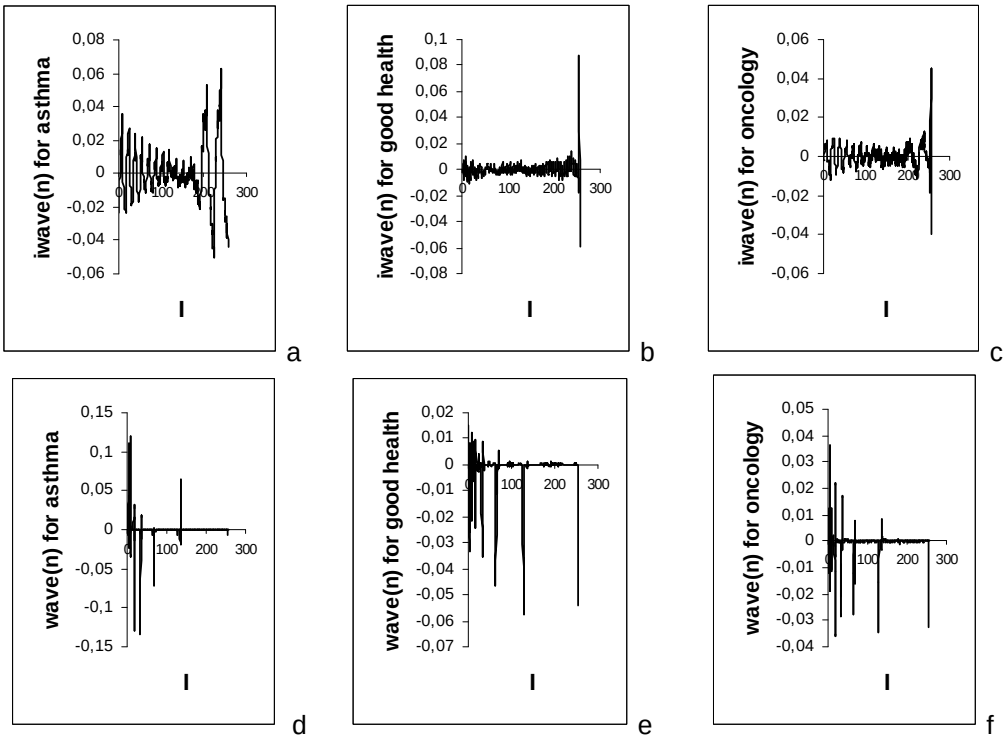

Fig. 2. Wavelets $iwave(n_l)$ and wavelet spectra $wave(n_l)$ of the immunofluorescence distributions in Fig. 1; Here $iwave(n_l)$ for asthma (a), healthy human (b), and oncology (c); $wave(n_l)$ for asthma (d), healthy human (e), and oncology (f). Let us note strong ragged wavelet in a, d for strong inflammation (asthma). Ragged is gentler for oncology in c, f. Histogram for healthy man has dominance of negative wavelet spectra in e.

3. Information entropies, their central moments and fractal

Dimentions of fluorescence flashes distributions and wavelets for deviations from their mean level. Let us consider the square of the relative deviations of relative flashes numbers from the average level as the probability measure of the immunofluorescence fluctuations. Then the probability of the immunofluorescence fluctuations for square of deviations is

$$p_l(n) = (n_l)^2 / \nu_n \quad , \quad \nu_n = \sum_{l=1}^{l=256} (n_l)^2 \qquad (2)$$

Let us introduce Shannon information entropy

$$H_S = -\sum_{l=1}^{l=256} p_l \times Ln(p_l) \qquad (3)$$

Probability density distribution of Shannon information entropy is

$$h_l = -p_l \times Ln(p_l) / H_S \qquad (4)$$

Deviation from the average level of Shannon entropy probability density is

$$r_l = (p_l \times Ln(p_l) - < p_l \times Ln(p_l) >) / H_S \quad , \quad < p_l \times Ln(p_l) > = < P > \times H_S \qquad (5)$$

The central moments of Shannon entropy distributions of the order k is

$$M(r,k) = < r^k > = < P > \times \sum_{l=1}^{l=256} r_l{}^k \qquad (6)$$

Let us denote $wave(n_l) = w_l$ and $iwave(n_l) = W_l$. Let us consider the square of the relative deviations $(wave(n_l))^2 = w_l{}^2$ and the square of relative flashes numbers $(iwave(n_l))^2 = W_l{}^2$ from the average level as the probability measure of wavelet ($W_l{}^2$) and wavelet spectra ($w_l{}^2$) fluctuations. Then correspondence probabilities of wavelet fluctuations for square of deviations are

$$p_l(w) = (w_l)^2 / \nu_w \;,\; \nu_w = \sum_{l=1}^{l=256} (w_l)^2 \text{ and } p_l(W) = (W_l)^2 / \nu_W \;,\; \nu_W = \sum_{l=1}^{l=256} (W_l)^2 \qquad (7)$$

Definitions of p_l, H_S, h_l, r_l, r, $M(r,k)$ in formulas (2)-(6) are dependent on probability measure or given variables n_l, w_l, W_l as $H_S(n_l)$, $H_S(w_l)$, $H_S(W_l)$, $M(r(n),k)$, $M(r(w),k)$, $M(r(W),k)$ and so on and so others. Therefore we have three sets of information distributions for the variable spaces of n, w, W.

Let us consider, briefly, the basic statistical special features of the task analysis, comparing the results of investigating three initial immunofluorescence histograms for different donors, shown in Fig. 1. Results are shown in Figs. 3-6.

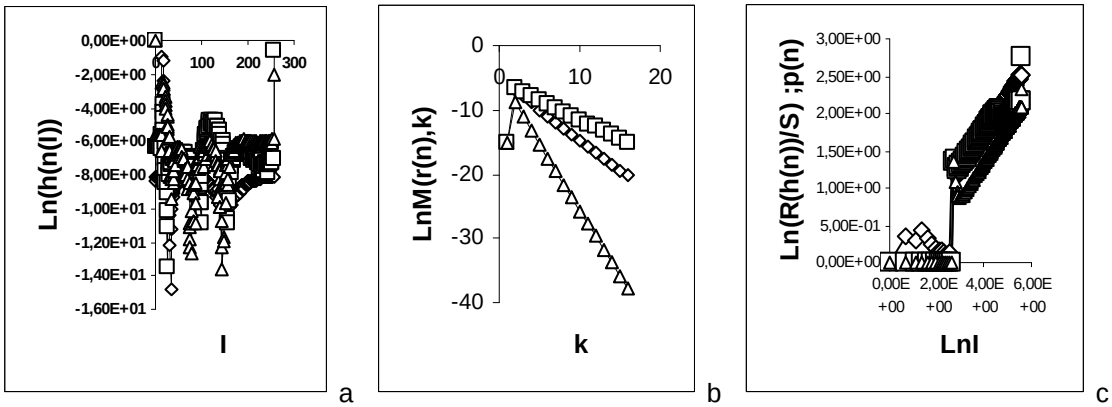

Fig. 3. Logarithmic distributions of Shannon entropies (a); their central moments M(r(n),k) (b); and Hurst exponent $H_H(R(r(n))/S)$, $H_H = \partial Ln(R/S)/\partial LnI$ (c); rhomb points correspond to bronchial asthma, quadrate points correspond to the healthy donor, triangle points correspond to the oncology disease. Initial histograms are shown in Fig. 1.

In our case Hurst's index or Hurst exponent H_H [6] is determined by means of regression equation

$$Ln(R/S) = H_H \times LnI + const \qquad (8)$$

where R/S is the rescaled range $(R=S)$, R is the range or maximal deviation of $h(I)$ from a local mean level, S is the standard deviation of $h(I)$. Hurst index H_H of Shannon entropy distribution $h(I)$ corresponds to its fractal dimension D_H [6] if

$$D_H = 2 - H_H \qquad (9)$$

In particular, the exponent H_H indicates persistent ($H_H > \frac{1}{2}$) or anti-persistent ($H_H < \frac{1}{2}$) behavior of the trend. In our case we have persistent behaviors of $H_H(R(h(n))/S = 0.596$ for asthma, $H_H(R(h(n))/S = 0.618$ for healthy human and $H_H(R(h(n))/S = 0.537$ for oncology. These fractal dimensions are $1 < D_H < 1.5$. Shannon–Weaver index [7] for biodiversity of

cells in our case is $D_S = H_S$ and coincides with the determination of Shannon entropy (3).Three examples of fractal dimensions $D_S(p(n))$ are $D_S(asthma) = 1.74$, $D_S(health) = 1.82$, and anomalous $D_S(oncology) = 2.65$; $D_S > 1.5$.

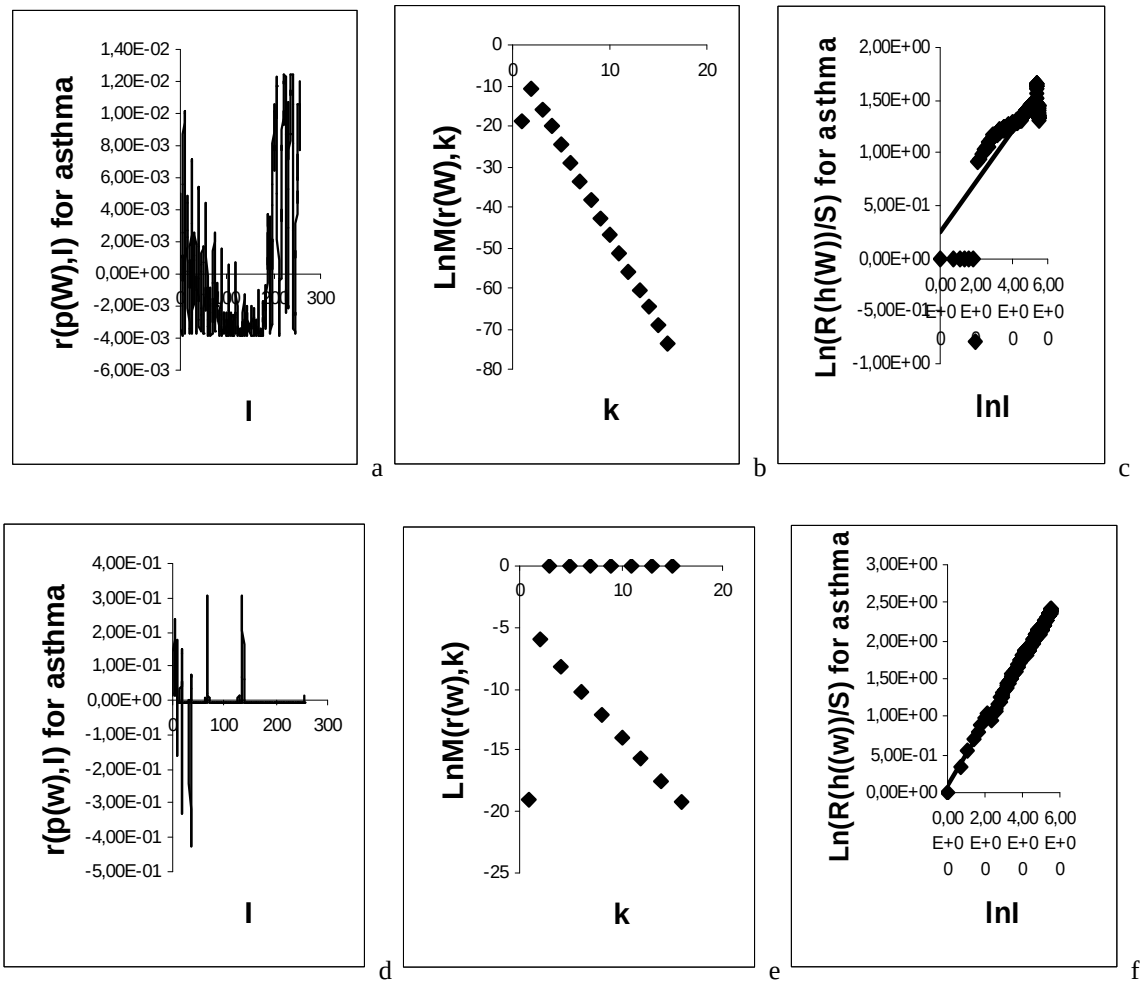

Fig. 4. Distributions of Shannon entropies of wavelet and wavelet spectra distributions for asthma in Figs. 2.

Here density r(W) (a); logarithm of central moments M(r(W),k) (b); Hurst index $H_H(R(h(n))/S = 0.234$ (c); density r(w) (d); logarithm of central moments M(r(w),k) (e); Hurst index $H_H(R(h(n))/S = 0.417$ (f). Initial histograms are shown in Fig. 2a, 2d; $H_H(R(h(n))/S = 0.424$ (c); density r(w) (d); logarithm of central moments M(r(w),k) (e); Hurst index $H_H(R(h(n))/S = 0.076$ (f); Initial histograms are shown in Fig. 2b, 2e.

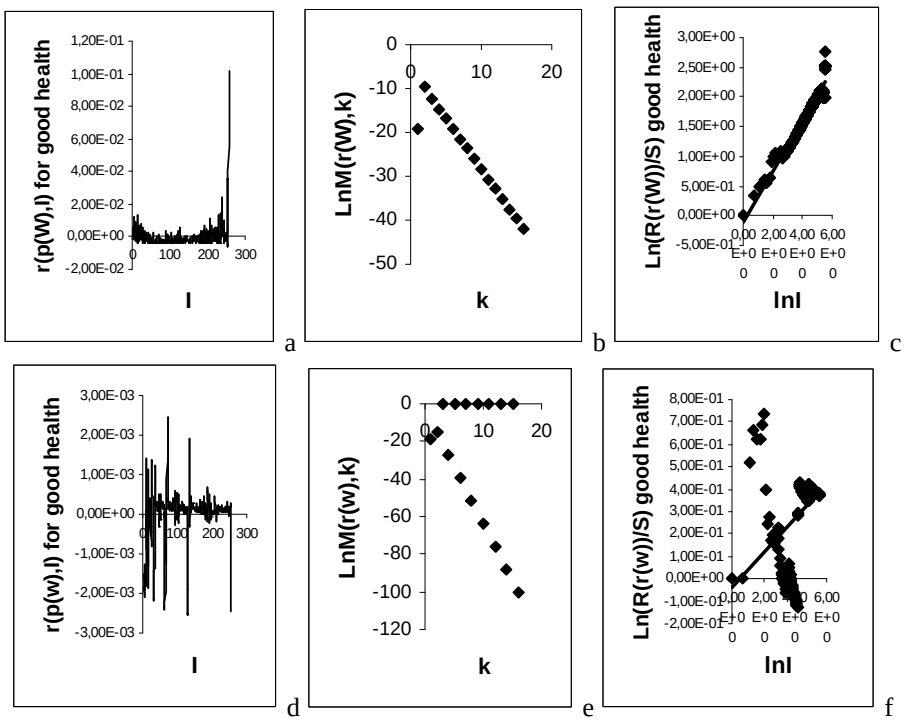

Fig. 5. Distributions of Shannon entropies of wavelet and wavelet spectra distributions for good health in Fig. 2; density r(W) (a), logarithm of central moments M(r(W),k) (b), Hurst index (c), density r(w) (d); logarithm of central moments M(r(w),k) (e); Hurst index (f).

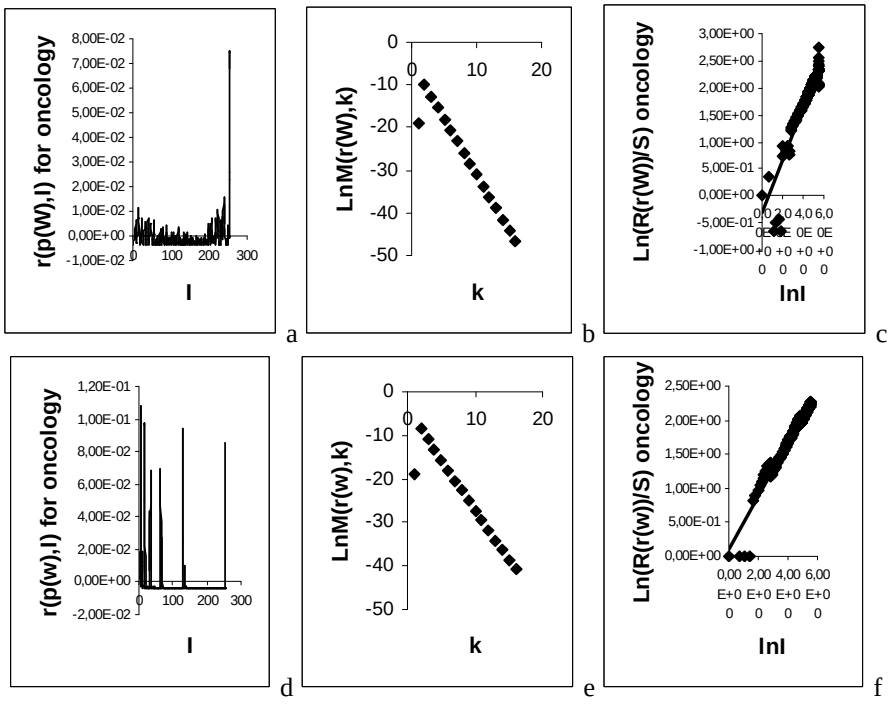

Fig. 6. Distributions of Shannon entropies of wavelet and wavelet spectra distributions for asthma in Fig. 2; density r(W) (a), logarithm of central moments M(r(W),k) (b), Hurst index $H_H\,(R(h(n))/S = 0.479$ (c); density r(w) (d); logarithm of central moments M(r(w),k) (e); Hurst index $H_H\,(R(h(n))/S = 0.392$ (f). Initial histograms are shown in Fig. 2c, 2f.

4. Brief discussion and conclusions

We strengthen of steady differences for various diseases during transitions of Shannon entropies from n to w and W variables in Figs. 3-6 (a, d). We observe the universal exponential distribution of the central moments for information entropies in Figs. 3-6 (b, e). Parameters of exponential decreases are various for the different health statuses. Therefore we have new tools to conduct diagnostics and monitoring of health statuses with the exponentially high sensitivity, precision and accuracy. We observe anti-persistent, $H_H < 0.5$ distributions of Hurst's indices for information entropy of wavelet and wavelet spectra in Figs. 3-6 (c, f). In the latter case wavelet analysis of information entropies in Figs. 4-6 does not have advantages in comparison with the analysis of the persistent Hurst indices for the initial information entropies in Fig. 3c. Most unexpected is a noticeable, sometimes two-fold increase, Shannon–Weaver index $D_S > 2$ upon transfer to the oncology diseases. Large values of $D_S > 2$ reflect the noticeably more complex networks of correlations in the populations of neutrophils and more complex genes regulation, than for the healthy people.

Acknowledgement

Specially thanks to M. Filatov for the kindly furnished experimental data.

References

[1] H.M. Shapiro, *Practical Flow Cytometry* (John Wiley & Sons, New York, 2003).
[2] N.E. Galich, M.V. Filatov // *Proc. SPAS* **11** (2007) 1123-1.
[3] C.F. Bassoe, Li Nianyu, K. Ragheb, G. Lawler, J. Sturgis, J.P. Robinson *// Cytometry Part B (Clinical Cytometr*y) **51B** (2003) 21.
[4] N.E. Galich, M.V. Filatov // *Proc. SPIE* **6597** (2007) 6597OL.
[5] M.V. Filatov, E.Y. Varfolomeeva, E.A. Ivanov *// Biochemical and Molecular Medicine* **55** (1995) 116.
[6] J. Feder, *Fractals* (Plenum Press, NewYork, 1988).
[7] W. Weaver, C.E. Shannon, *The Mathematical Theory of Communication* (University of Illinois, Urbana, 1949).