

# The Portraits Creating Method by Correlation Analysis of Hormone-Producing Cells Data

Olha Ryabukha<sup>1</sup>[0000-0001-6220-4381] and Ivanna Dronyuk<sup>2</sup>[0000-0003-1667-2584]

<sup>1</sup> Lviv Medical Institute, Polishchuka str. 76, 79018 Lviv, Ukraine  
oriabuha@ukr.net

<sup>2</sup> Lviv Polytechnic National University, S. Bandera str. 12, 79013 Lviv, Ukraine  
ivanna.droniuk@gmail.com

**Abstract.** Method of using the correlation analysis for studies of the hormone-producing cells activity is presented as exemplified by the follicular thyrocyte. Electronograms of the thyroid ultrathin sections are studied, prepared using routine technologies. Studying any type of follicular thyrocyte's activities (synthesis, secretion, transport of hormones, energy supply of these processes) involves the successive application of such stages: research objective formulation and isolation from the entire set of organelles' cells ultrastructures – implementers of the respective activity (the studied cluster formation); transformation of linguistic (qualitative) information on the cluster elements into quantitative (numeric) indices; determining of intra-cluster (intra-system) interconnections and interdependencies; graphic representation of correlation links in correlation portraits and their analysis. In need to deepen the study, a module can be formed of several clusters in which the elements most relevant for solving the task are determined. Subsequently, the correlation analysis establishes the connections between the elements of the studied clusters, which are reflected in the correlation portraits, analyze the information obtained and develop an integrative model of the module's activities. The suggested method is not connected with the strict determinism of the input information, but it permits to correctly transform qualitative/binary indices into quantitative ones and to generalize the results obtained. Establishment of interdependencies between elements of investigated clusters is an integral research approach that enables the use of correlation portraits as the basis for the further development of an expert system for solving problems in cytomorphology and cytophysiology.

**Keywords:** correlation analysis, correlation portrait, follicular thyrocyte.

## 1 Introduction

The human body is the object of multi-directional influences that are acting simultaneously. Thus, in particular, the state of a sick person's organism is caused by a number of internal and external factors: the features of premorbid development and life, working conditions, living conditions, the state of all body systems during a disease, etc. By virtue of this, the application of mathematical methods in the study of the

organism/system/organ/cell activity is a very complicated process, because it involves the need to consider a large number of various factors whose significance is unequal [1].

The task of medical diagnostics is the search for the expression

$$X^* = (x_1^*, x_2^*, \dots, x_n^*) \rightarrow d_j \in D = (d_1, d_2, \dots, d_m),$$

where  $X^*$  is a set of a particular patient's state parameters, and  $D$  is a set of diagnoses inherent in the given field of medicine. Development of mathematical approaches to the study of biological systems requires the use of a mathematical apparatus, which is an integral part of formalized cognition. The present stage of the medicine development as a science is characterized by quite frequent implementation of mathematical methods into the practice of medical diagnostics, being intended to objectivize the results obtained and to serve as the basis for making decisions on the belonging of pathological manifestations observed in the patient to this or that nosological group [2, 3]. Difficulties associated with solving diagnostic tasks are due to a number of reasons, the most important of which is the necessity to know a large number of the patient's condition parameters, which is steadily increasing with the development of medical science. This makes it quite problematic to remember them and reduces the opportunity of quickly taking them into account when diagnosing, even the by the most skilled expert. At the same time, the analytical dependencies between the parameters of the patient's condition and the diagnosis in their classical sense are usually absent, and these parameters in particular may be of a various nature: quantitative (age, height, body weight, content of certain substances, etc.), qualitative (the nature of pain sensations, mood sensations, etc.) or binary (presence or absence of a state or a process – yes/no).

The method of correlation analysis has a long history of application. Since the task of the method is to establish connections between the parameters under study, it is successfully used in various branches of biology and medicine. The peculiarity of the correlation analysis lies in the search of interdependence between two or more indices, which nature and pronouncement are established by the pair correlation coefficient, calculated by the formula. Durable application of the method contributed to accumulating the information on biomedical processes at their multiple levels (genetic, phenotypic, physiological, etc.), which predicated the development of tools to facilitate its visualization, analysis and interpretation [4]. An example of such scientific tools is the correlation grids, permitting to trace the interdependencies in large amounts of quantitative data. Despite its widespread use, the method is not designed for work with qualitative or binary data on the studied biosystem status.

The concept of fuzzy sets arose in response to the requirements of the classical theory of systems to provide artificial precision, which can not be achieved in the biological objects life process. The theory of fuzzy sets is a means of formalizing uncertainties [5, 6], arising in the biological system, and methods for solving problems inherent in a living organism.

It applies qualitative data, relatively simple mathematical methods, namely the notions of membership function and the highest and the least expressiveness of any sign (max-y and min-y). For example, in the study of the thyroid gland pathology, the signs may have the following form: hypothyroidism may be non threatening, moder-

ate, severe, etc.; cytoplasmic reticulum may be narrowed, moderately expressed, expanded, etc.

The principle of linguistic diagnostic data implies that the causal relationships between the biological system status parameters, which can be the organism/organ/cell (cause), and the diagnosis (consequence) are initially described by words of the language used, thereafter they are formalized as a collection of fuzzy logical statements in the “if... – then” sort. This principle may be exemplified by the following description of the thyrocyte morpho-functional status: “If the cell’s shape is cubic, the electron density of the cytoplasm is moderate, the electron density of the colloid is moderate, apical microvilli are thin, short, and their number is moderate; mitochondria in sufficient quantities, the elements of the granular cytoplasmic reticulum are moderately expressed, the Golgi complex elements are expressed moderately, the number of free and bound ribosomes and the polysomes is moderate, the number of lysosomal bodies is moderate, the number of apical secretory granules is moderate, then the thyrocyte’s functional status is balanced“. This permits to take into account the status quality, which considerably broadens the the researcher’s opportunities. Diagnosing on the principles of Fuzzy Logic requires special training of the researcher for carrying out mathematical transformations, which significantly impedes the final result obtaining. In addition, in order to avoid excessive complications in the biological processes simulation, one has to neglect a sufficiently large number of actual facts. However, in biological systems, there are no minor or unnecessary processes, therefore their neglecting leads to reduction in the reliability of conclusions based on such incomplete data.

Therefore, both of the described methods do not permit studying peculiarities of changes in cell structures involved in the hormonopoiesis to the full extent. The necessity in formalizing qualitative and binary information on the studied biological object/system status is acute in cytology traditionally applying such a heuristic method as a linguistic description. In this case, the completeness of the information obtained and its interpretation depend on the researcher’s qualification, thus subjectivizing the final conclusion. The use of mathematical statistics for the quantitative parameters processing, despite the possibility of comparison, does not allow to carry out an expanded analysis of the data obtained to the full extent and to generalize them, which is very important for establishing the regularities of the cells activity and highlighting peculiarities of their changes in response to the various factors’ effects.

## **2 Purpose of the study**

The purpose of the study was to develop, based on follicular thyrocyte, a method for analyzing the morpho-functional status of hormone-producing cells, permitting to transform qualitative (linguistic)/binary information into objective (quantitative) indices, which can be subjected to further mathematical transformations, analysis and generalizing.

### **3 Materials and methods**

The object of the study were electron diffraction patterns of the thyroid glands tissue ultrathin (4-6  $\mu\text{m}$ ) sections of white outbred male rats, prepared for electron microscopic studies using routine technologies, which were studied using methods of electron diffraction patterns semi-quantitative analysis and determination of hormone-producing cells special possibilities, by means of separate elements of such components as mathematical statistics, phase interval, Fuzzy logic, correlation analysis [7]; forming separate clusters of ultrastructural elements is carried out based on the cytophysiology data on the functional role of each cell organelle [8].

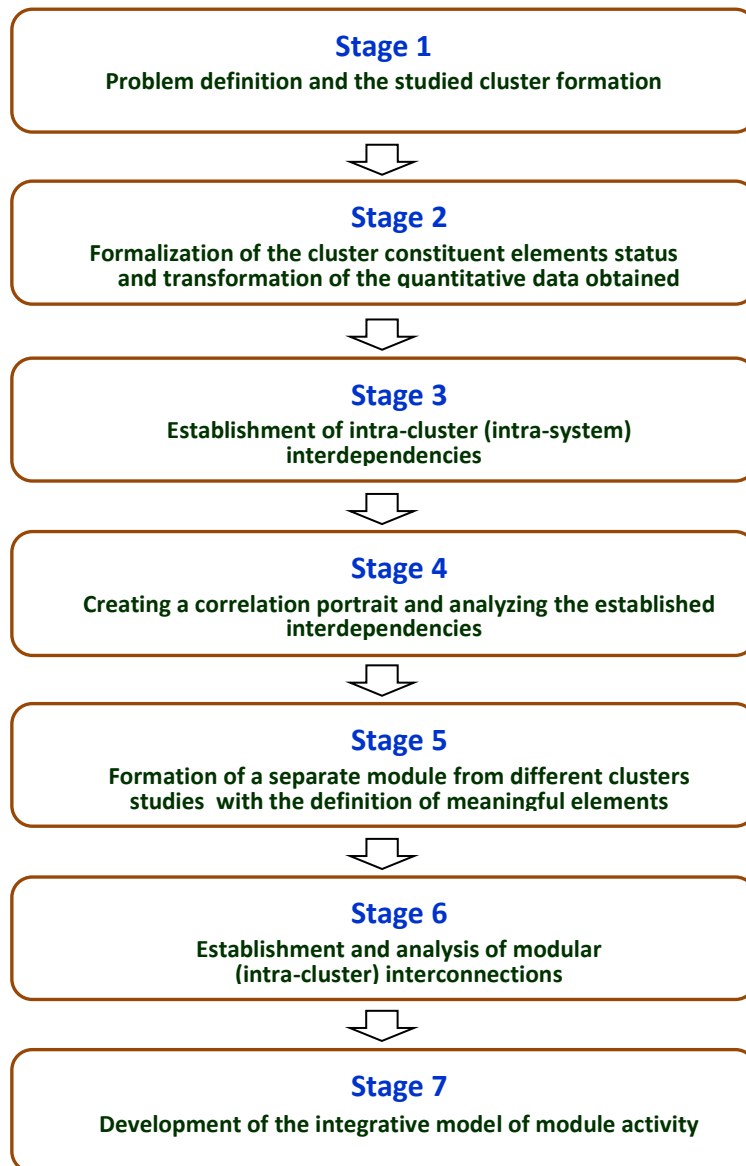
At all stages of the study, international requirements for the humane treatment of vertebrate animals were observed in accordance with the "Guidelines for Accommodation and Care of Animals" (Strasbourg, 2006, Annex 4) and Helsinki Declaration on humane endpoints to experiment animals.

### **4 The portraits creating method by correlation analysis of hormone-producing cells data: description**

In our opinion, the promising trend of studies in cytology, should be based on a combination of adequate methods, which are elements of mathematical statistics, cluster and correlation analysis, the phase interval method, the concept of fuzzy sets, all available quantitative, qualitative and binary data being used for the research needs. In this case, the research process will be comprehensive and objective. We have developed a method for studying morphological and functional characteristics of hormone-producing cells, including seven stages (see Fig. 1).

At the first stage, according to the task of studying a certain field of the cell activity (synthetic, secretory, transport, energy) create separate clusters of the studied structural elements/signs/manifestations that are implementers of this field. For example, the status of the intrafollicular colloid electronic density, apical microvilli, lysosomal bodies, secretory granule indicates the secretory cluster of the thyrocyte activity (see Table 1).

At the second stage, each ultrastructural/substructural element (form, number, sub-cellular location, etc.) of the cluster and its state (reduced, moderate, increased) is assigned an alphanumeric index. Using the principle of the phase interval method for comparing the studied system status with two diametrically opposite reference standards, which can be conventionally defined as "health" and "studied pathological process/disease", a digital assessment of each cluster element is carried out, i.e., qualitative or binary features are transformed into quantitative ones. Special tables being applied, the results obtained are expressed in percentage or points (see Table 2). The cluster components are studied and valued in several fields of sight (see Table 3, Table 4).



**Fig. 2.** Structure of the correlation portraits creating method for study the hormone-producing cells activity.

Averaged digital data is used for mathematical transformations at the third stage of the study, when, using a correlation analysis, the existence of links is established between the studied cluster constituent elements. For this purpose, the correlation coefficients, being calculated by the Pearson formula, are used.

**Table 5.** The cluster of ultrastructures-realizators of follicular thyrocytes secretory potential.

Ultrastructural element	The researched feature of the ultrastructural element	The quality of the researched ultrastructural element feature	Quality designation of the researched ultrastructural element feature
Intrafollicular colloid	electron density	insignificant	E <sub>1</sub>
		moderate	E <sub>2</sub>
		significant	E <sub>3</sub>
Microvilli of the apical cytosolic membrane	quantity	insignificant	H <sub>1</sub>
		moderate	H <sub>2</sub>
		significant	H <sub>3</sub>
	density, length	insignificant	H <sub>4</sub>
		moderate	H <sub>5</sub>
		significant	H <sub>6</sub>
Lysosome bodies	quantity	insignificant	G <sub>1</sub>
		moderate	G <sub>2</sub>
		significant	G <sub>3</sub>
	size	small	G <sub>4</sub>
		medium	G <sub>5</sub>
		big	G <sub>6</sub>
	electron density	insignificant	G <sub>7</sub>
		moderate	G <sub>8</sub>
		significant	G <sub>9</sub>
Secretory granules	quantity	insignificant	M <sub>1</sub>
		moderate	M <sub>2</sub>
		significant	M <sub>3</sub>
	electron density	insignificant	M <sub>4</sub>
		moderate	M <sub>5</sub>
		significant	M <sub>6</sub>
	allocation	apical cellular pole along the whole cytosolic membrane	M <sub>7</sub>
			M <sub>8</sub>
	topographic connection with lysosome bodies	present	M <sub>9</sub>
		absent	M <sub>10</sub>

**Table 6.** Scale of the evaluation of the features severity in the semi-quantitative analysis of electronograms.

Feature severity degree	Graphic symbol	Numerical assessment	
		(points)	(percentage)
Feature absent	-	0	0
weak	+	1	25
moderate	++	2	50
significant	+++	3	75
maximal	++++	4	100

Note. 0 points - state of unattended pathology under study ("disease"); 4 points - state of the studied pathology complete absence ("health").

**Table 7.** Results of transforming qualitative data of follicular thyrocyte secretory activity cluster into quantitative indices.

		Designation / numerical rating (points)												
		E1	E2	E3	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
<b>A</b>	#1	0	1	3	3	1	0	0	3	1	2	2	1	3
	#2	0	1	2	3	1	0	0	3	1	3	3	2	2
	#3	0	0	3	3	2	0	0	4	2	3	3	2	3
	#4	0	1	2	3	1	0	0	3	1	2	2	1	3
	#5	0	1	3	2	1	0	0	3	1	2	2	1	3
	<b>mean values</b>	<b>0</b>	<b>0,8</b>	<b>2,6</b>	<b>2,8</b>	<b>1,2</b>	<b>0</b>	<b>0</b>	<b>3,2</b>	<b>1,2</b>	<b>2,4</b>	<b>2,4</b>	<b>1,4</b>	<b>2,8</b>
<b>B</b>	#6	0	1	3	3	1	0	0	0	4	1	3	4	0
	#7	0	1	2	3	1	0	0	0	3	1	4	3	0
	#8	0	0	3	3	0	0	0	0	4	1	4	4	0
	#9	0	1	2	4	1	0	0	0	4	0	3	4	0
	#10	0	1	3	4	1	0	0	0	4	1	3	4	0
	<b>mean values</b>	<b>0</b>	<b>0,8</b>	<b>2,6</b>	<b>3,4</b>	<b>0,8</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3,8</b>	<b>0,8</b>	<b>3,4</b>	<b>3,8</b>	<b>0</b>

Note. A - series with adding 100 µg of organic iodine into the white male rats ratio under the conditions of thyroid hyperthyroidism; B - series with adding 100 µg of inorganic iodine into the white male rats ratio under the conditions of thyroid hyperthyroidism; # - studied electronograms.

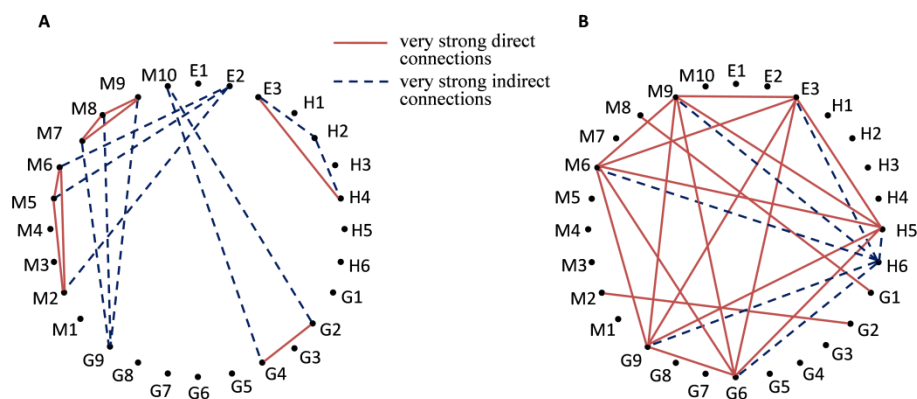
**Table 8.** Results of transforming qualitative data of follicular thyrocyte secretory activity cluster into quantitative indices.

		Designation / numerical rating (points)														
		H1	H2	H3	H4	H5	H6	G1	G2	G3	G4	G5	G6	G7	G8	G9
<b>A</b>	#1	3	1	0	3	1	0	3	1	0	1	0	3	0	0	4
	#2	2	2	0	2	1	0	3	2	0	2	0	4	0	0	3
	#3	2	1	0	3	1	0	3	1	0	1	0	2	0	0	3
	#4	2	2	0	2	2	0	2	1	0	1	0	3	0	0	4
	#5	3	1	0	3	1	0	2	1	0	1	0	3	0	0	4
	<b>mean values</b>	<b>2,4</b>	<b>1,4</b>	<b>0</b>	<b>2,6</b>	<b>1,2</b>	<b>0</b>	<b>2,6</b>	<b>1,2</b>	<b>0</b>	<b>1,2</b>	<b>0</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>3,6</b>
<b>B</b>	#6	0	1	3	0	1	3	3	1	0	0	0	4	0	0	4
	#7	0	1	4	0	0	4	4	1	0	0	0	3	0	0	3
	#8	0	1	3	0	1	3	4	0	0	0	0	4	0	0	4
	#9	0	1	4	0	1	3	3	1	0	0	0	4	0	0	4
	#10	0	2	3	0	1	3	3	1	0	0	0	4	0	0	4
	<b>mean values</b>	<b>0</b>	<b>1,2</b>	<b>3,4</b>	<b>0</b>	<b>0,8</b>	<b>3,2</b>	<b>3,4</b>	<b>0,8</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3,8</b>	<b>0</b>	<b>0</b>	<b>3,8</b>

Note. A - series with adding 100 µg of organic iodine into the white male rats ratio under the conditions of thyroid hyperthyroidism; B - series with adding 100 µg of inorganic iodine into the white male rats ratio under the conditions of thyroid hyperthyroidism; # - studied electronograms.

The fourth stage is devoted to designing the intra-cluster (intra-system) correlation portraits, which permits to visualize the traced correlations. Analysis of the established connections is performed taking into account their strength, quantity and direction (sign). The positive value of the pair correlation coefficient indicates the same direction of change in the studied indices, the negative means that with an increase in one of the indices another indice associated with it reduces; the value  $r_{xy} = 1.0$  indicates the existence of a direct proportional feedback between the x and y,  $r_{xy} = -1.0$  means inversely proportional feedback. In the structural organization of the interrelations between the indices, the most significant are considered very strong and strong connections, which on the Chaddok correlation scale are respectively within  $1.0 \geq |r| > 0.9$  and  $0.9 \geq |r| \geq 0.7$ ; in the absence of such connections, the noticeable ( $0.7 \geq |r| > 0.5$ ) and moderate ( $0.5 \geq |r| > 0.3$ ) connections are studied. We find it irrelevant to analyze the weak ( $0.3 \geq |r| > 0.1$ ) connections, since it can distort results because of the fact that the biological system's functioning as a whole is due to the presence of various connections between its constituent parts. Functional analysis of the traced correlations is carried out based on cytophysiology [8], taking into account the role of each cluster ultrastructure and its status significance for the organelle activity.

As an example of the cytophysiological information visualizing possibilities, we report the study on the follicular thyrocytes secretory capacity in white male rats under the conditions of administering iodine of different chemical nature against the background of thyreoidin hyperthyroidism (see Fig. 3).



**Fig. 4.** Graphic representation of the correlation portraits structure for the secretory features of the thyroid glands follicular thyrocytes in white male rats receiving 100 µg of organic (A) and inorganic (B) iodine in the model conditions of thyreoidin hyperthyroidism.

The results processing was performed using the software: for the digital parameters - StatSoft Statistica v6.0 package, for correlation tables and portraits – Microsoft Office 2010 package (Microsoft Excel spreadsheet and Microsoft Word editor, respectively).

If it is necessary to carry out an in-depth intra-cluster/intra-system study, several more steps are to be added. For this purpose the module of these clusters have been formed. At the fifth stage, the elements significant for implementing the research in



the studied area of the cell activity are determined, in particular, those elements of the studied object, which in the studied conditions, are the most sensitive to the studied influence ("point of reference") and elements of the surrounding systems, changes in which under studied conditions are the most pronounced ("point of attraction"). For example, in the case of taking iodine in some pathological conditions (hypothyroidism, etc.), the dose of iodine consumed ("the point of action"), affecting the state of ultrastructure of a certain protein-synthesizing organelle ("the point of response"), causes changes in the structure of the "attraction points", which are other cells' organelles of the thyroid gland, functionally associated with follicular thyrocytes.

At the sixth stage, the presence/absence of correlation dependencies between "points of response" and "points of attraction" is determined and analyzed, the results obtained are interpreted.

The seventh, final, stage is devoted to the development of integrative mathematical complexes/models as a result of the interaction between "points of response" and "points of attraction" that are present in the studied biological object [9]. The peculiarity of intra-thyroid control of the thyroid gland activity as an organ is the simultaneous presence of cells in its parenchyma that produce specific thyroid hormones, and the cells producing the hormones-antagonists. This can cause some difficulties in the development of integrative mathematical research models. At the same time, application of the suggested approach permits to simultaneously take into account the features of cell activity in different areas. In this case, the "response system" are cells of the follicular epithelium, and the "system of action" and the "target of action" are certain substances and their doses affecting the follicular thyrocyte, the "object of response" is the thyrocyte's structures producing a specific hormone, the "points of response" are the form of thyrocyte and the form and degree of its nucleus electron density, the "attraction system" is the C-cell as a functional antagonist of the thyrocyte, "attraction points" in the C-cell are its cytoplasm and nucleus (the number of organelles in the cytoplasm, the nomenclature of prevailing organelles and their status, the electron density of nuclear chromatin, etc.) [9]. Subsequent analysis, processing and integration of the data obtained can give reason to suppose that an increase in the number of C-cells and their ultrastructural elements, such as granular cytoplasmic reticulum and the Golgi complex, under the condition of a certain extension of the said organelles' substructures at flattening of thyrocytes and their nuclei, will indicate a reduced background of the thyrocyte functional activity. The experimental proof of these theoretical assumptions creates the prospect of adjusting the activity of the thyroid gland by affecting the C-cells activity.

The presented method of studying hormone-producing cells was implemented by us into studying the features of follicular thyrocytes synthetic capacity, in particular the search for markers of changes in their morpho-functional status under the influence of iodine-containing compounds [10]. In general, the use of mathematical technologies as a tool for studying the biological system permits to take into account a significant number of its qualitative characteristics, which significantly expands the capabilities of the researcher-theorist to determine its status and possible changes, and the researcher-clinician – to establish the correct diagnosis and choose a strategy of treatment.

Thus, the methods of analysis, on which the mathematical support of biomedical diagnostics is based, particularly in cytology, are not sufficiently adapted to work with high-quality information. At the same time, it is linguistic (non-numeric) information that is a prerequisite for increasing the amount of information obtained about the studied object, since it contains the greatest information about all the subtle aspects of its condition. Our approach, suggested for work with linguistic information while performing cell status analysis, is in line with the idea of expediency of using step-by-step algorithms for the medical data analysis [11].

The similar idea is suggested by [12] who believe that the process of medical diagnosis should take place stagewise, with the first stage being processing of linguistic information obtained from the patient. To some extent, the approach presented is consistent with the views of [13] concerning the necessity of all the morpho-functional module components joint activities, aimed at achieving a general beneficial result.

At testing the informativity of the studies results, obtained with their help, and clarifying the availability of the results that was carried out in our work [7], carried out from the standpoint of cybernetic insight into the cell as a complex self-regulating system, it was established that the presented method of studying hormone-producing cells is not associated with the strict determinism of the input information. Instead, it permits the correct transformation of qualitative/binary indices into quantitative ones and integration of the results obtained.

## 5 Conclusion

The suggested objectivizing of the database with a collection of morpho-functional information on the thyrocyte as a hormone producing cell is carried out using StatSoft Statistica and Microsoft Office (Word and Excel) software packages in accordance with our method's stages developed with a detailed description of each constituent element in the studied clusters. Establishment of interdependencies between them is an integral research approach that enables the creation of correlation portraits as the basis for the further development of an expert system for solving problems in cytomorphology and cytophysiology.

## References

1. Avtandilov, G.G.: Medical morphometry: a guide. Meditsina, Moscow (1990). (in Russian)
2. Murray, J.D.: Mathematical biology II: spatial models and biomedical applications (interdisciplinary applied mathematics). 3rd edn., 18. Springer, New York (2003).
3. Amigó, J.M., Small, M.: Mathematical methods in medicine: neuroscience, cardiology and pathology. Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences 375(2096), (2017).

4. Batushansky, A., Toubiana, D., Fait, A.: Correlation-based network generation, visualization, and analysis as a powerful tool in biological studies: a case study in cancer cell metabolism. *BioMed Research International* 2016(8313272), 9 (2016).
5. Syavavko, M., Rybytska, O.: *Mathematical modeling in conditions of uncertainty*. Lviv, Ukrainski Tekhnolohii (2000). (in Ukrainian)
6. Zadeh, L.A.: Can mathematics deal with computational problems which are stated in a natural language? *Logic Colloquium, UC Berkeley*, <http://www.springer.com/lncs>, last accessed 2011/09/11.
7. Ryabukha, O.I.: Perspectives of applying new approaches to the implementation of mathematical technologies in the study of cell activity. *Medical Informatics and Engineering* 1, 67–75 (2018). doi: 10.11603/mie.1996-1960.2018.1.8894 (in Ukrainian)
8. Lutsenko, M.T.: *Cytophysiology: a guide*. SB RAMS, Novosibirsk – Blagoveshchensk (2011). (in Russian)
9. Ryabukha, O.I.: Substantiation of conceptual apparatus for mathematical studies on the hormone-producing cell activity. *Bulletin of problems in biology and medicine* 3, 1(145), 234–237 (2018). doi 10.29254/2077-4214-2018-3-145-234-237 (in Ukrainian)
10. Ryabukha, O.I.: Search for markers of changes of the synthetic activity of thyrocyte under the influence of iodine reception in iodine deficiency conditions. *World of Medicine and Biology* 3(65), 179–185 (2018). doi 10.26724/2079-8334-2018-3-65-179-185
11. Goiko, O.V., Mokhnachov, S.I.: Technology of processing and analysis of medical data in view of their features. *Zaporozhye Medical Journal* 1(76), 82–83 (2013). (in Ukrainian)
12. Das, S., Guha, D., Dutta, B.: Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic. *Applied Intelligence* 45(3), 850–867 (2016).
13. Mintser, O.P., Voronenko, Yu.V., Vlasov, V.V.: *Processing of clinical and experimental data in medicine: textbook for students of higher education institutions*. Vyshcha shkola, Kyiv (2003). (in Ukrainian)