

Prediction of the progression of endometrial hyperplasia in women of premenopausal and menopausal age based on an analysis of clinical and anamnestic indicators using multiparametric neural network clustering

PETRO SELSKYY^{1, A, C-F}, ANDRII SVERSTIUK^{2, D-F}, ANDRII SLYVA^{1, B, E, F}, BORYSLAV SELSKYI^{3, D-F}

ORCID ID: 0000-0001-9778-2499

ORCID ID: 0000-0001-8644-0776

ORCID ID: 0000-0003-2385-4001

ORCID ID: 0000-0001-6787-4843

¹ Department of Pathological Anatomy with Autopsy Course and Forensic Medicine, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine

² Department of Medical Informatics, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine

³ Department of Surgery № 2, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine

A – Study Design, **B** – Data Collection, **C** – Statistical Analysis, **D** – Data Interpretation, **E** – Manuscript Preparation, **F** – Literature Search, **G** – Funds Collection

Summary Background. A number of studies are aimed at solving the problems of implementing innovative medical information technologies, but the issues of family medicine informatisation have not been fully resolved. This is especially important to optimise the diagnosis of the most common diseases.

Objectives. The aim of our study was to develop a methodology for predicting the progression of endometrial hyperplasia at the primary healthcare level based on an analysis of clinical and anamnestic indicators using multiparameter neural network clustering.

Material and methods. Clinical and anamnestic data was obtained based on the results of a retrospective analysis of the inpatient charts of 52 patients with non-atypical endometrial hyperplasia. For a deeper analysis and clustering, the neural network approach was used with the NeuroXL Classifier add-in application for Microsoft Excel.

Results. The results of the cluster analysis showed that when predicting the progression of endometrial hyperplasia based on an analysis of clinical and anamnestic indicators, it is important to take into account the combination of the use of intrauterine contraception, as well as infertility, obesity and diseases of the gastrointestinal tract in patients. At the same time, the probability of progression of endometrial hyperplasia also increases with an increase in the number of operative obstetric and gynaecological interventions.

Conclusions. In order to effectively and objectively assign patients to the risk group for the progression of endometrial hyperplasia according to the indicators obtained during observation, neural network clustering was used, which allows one to determine the value of combined changes of certain parameters for the prognosis of the progression of the disease.

Key words: primary health care, endometrial hyperplasia, women.

Selskyy P, Sverstiuk A, Slyva A, Selskyi B. Prediction of the progression of endometrial hyperplasia in women of premenopausal and menopausal age based on an analysis of clinical and anamnestic indicators using multiparametric neural network clustering. *Fam Med Prim Care Rev* 2023; 25(2): 184–189, doi: <https://doi.org/10.5114/fmpcr.2023.127679>.

Background

The volume of medical data is constantly increasing due to the description of new pathological conditions, syndromes, nosologies, diagnostic methods and medicines [1–3]. To ensure proper quality of medical care, it is necessary to accumulate and analyse data over a long period of time [3]. All things considered, information technologies and methods in medicine are gaining in importance.

For family medicine, when organising the provision of primary medical care, it is especially important to provide information on early diagnosis and disease prevention, timely provision of medical care to patients of various risk groups and persons with socially significant diseases, decision-making regarding conducting medical examinations, as well as quality personnel training [3].

Scientific studies on the economic efficiency of the implementation of medical information systems shows a significant saving of medical workers' time for record keeping, which takes about 60% of their time [1, 4]. National European healthcare

informatisation programmes are aimed at uniting medical facilities into one network and ensuring electronic storage of patients' medical data, followed by their effective use.

A number of studies are aimed at solving the problems of implementing innovative medical information technologies [1, 5–7], but the issues of family medicine informatisation have not been fully resolved. This is especially important to optimise the diagnosis of the most common diseases using inexpensive and convenient methodologies for common use.

One important problems is timely and effective diagnosis of endometrial hyperplasia (EH), the relevance of which is due to a number of aspects. First is the wide prevalence of this disease in the population, which makes up 15–50% in the incidence of gynaecological pathology. Secondly, EH is the background for the development of malignant transformation of the uterine mucosa. Thirdly, the morphological diagnosis of histological variants of hyperplasia creates certain difficulties [8–10].

Timely diagnosis and treatment of EH is an important medical and social problem caused not only by the high frequency of metrorrhagia and endometritis but also by the risk of malignancy, since in 30–70% of cases, EH plays a key role in the de-



velopment of endometrial cancer in perimenopause and postmenopause [10–12]. Considering this, it is especially necessary to optimise the prediction of the progression of this disease to ensure timely diagnostic and therapeutic measures.

Objectives

The aim of our study was to develop a methodology for predicting the progression of EH at the primary healthcare level based on an analysis of clinical and anamnestic indicators using multiparameter neural network clustering.

Material and methods

Study design

Studies of patients with EH were carried out using standard statistical methods, as well as neural network clustering. According to the results of the study, a methodology was developed to optimise the prediction of disease progression at the level of primary medical care.

Setting

Clinical and anamnestic data was obtained based on the results of a retrospective analysis of the inpatient charts of patients with non-atypical endometrial hyperplasia who received inpatient treatment in Ternopil Hospital No. 2 and in district hospitals in the Ternopil Region (Ukraine) in 2014–2017. Histological studies of endometrium biopsies were conducted using archival material of the Ternopil Regional Pathology Bureau over a specified period of time.

Participants

We conducted a study of 52 patients with non-atypical endometrial hyperplasia. The inclusion criteria in the research study: pre-menopausal and menopausal age of patients (> 38 years), presence of histologically confirmed non-atypical endometrial hyperplasia, informed consent to participation in the research study. The control group included 12 healthy women > 38 years of age who did not suffer from EH.

Variables

The analysis included women who suffered from EH and participants of the control group. A separate research study of groups of patients' indicators with and without signs of disease progression was also conducted. In the comparative analysis, indicators by which the study groups differed considerably were determined. Studies of combined changes of indicators during disease progression were also carried out.

Bias

Patients were included in the general research study group by random sampling. Random sampling was also performed in a control group of women who did not suffer from EH. The research used indicators obtained under the same conditions.

Data sources

This research study used anamnesis data regarding socio-economic factors, obstetric and gynaecological anamnesis and diagnostic conclusions regarding comorbid pathology, which were included in the inpatient charts by hospital medical staff during treatment. The specified documents also contained the personal data of the patients, information about the performed diagnostic and treatment procedures and their results.

Study size

The study size was calculated by a representative sampling of the general group. From this group, the groups of research were formed. The next stage of the study was carried out taking into account the results of the previous study. By generalisation of the results of use of statistical methods and neural network clustering, an algorithm of the decision-making was developed.

Quantitative variables

The following data of quantitative variables was processed in the analysis: shares in percentages of study groups, relative values of age, socio-economic factors, indicators of obstetric and gynaecological history, clinical manifestations of climacteric syndrome and existing extragenital pathology in each group. A study of the parameters of the neural clustering algorithm proposed by the NeuroXL Classifier program was also conducted.

Statistical methods

Statistical analysis of the data was conducted using the software package "Microsoft Excel" (Microsoft Office 2013). The statistical significance of differences between relative values was estimated by Student's *t*-Test (*t*).

For a deeper analysis and clustering of patients in a group in order to optimise the prognosis of pathology, the neural network approach was used with the NeuroXL Classifier add-in application for Microsoft Excel. The neural network is a proven and very common technology to solve complex classification problems. They are modelled on the basis of the human brain and composed of interconnected networks of independent processors that arrange by changing the connections (a process known as training) through learning and problem-solving processes. The NeuroXL Classifier program (developed by AnalyzerXL) creates the self-organising neural networks, which perform categorisation by studying trends and correlations within the data. Despite the high performance, neural networks are not often used because of their complexity and extensive need for training, which are necessary for their proper implementation. NeuroXL Classifier eliminates such barriers, using the complexity of methods based on already known neural networks and taking advantage of the Microsoft Excel computational engine. The key advantages of using NeuroXL Classifier are simplicity in operation, optional in-depth knowledge database of the neural networks, integration with Microsoft Excel, provision of reasonable neural network technology for high-precision classification and determination of trends and correlations that cannot be shown by traditional methods [13–15].

Results

Participants

We conducted a retrospective analysis of 52 women's inpatient charts. In this study, we analysed the following factors: age of patients, socio-economic factors, obstetric and gynaecological history, clinical features of the climacteric syndrome and comorbid conditions in women with various types of non-atypical endometrial hyperplasia.

According to the histological examination, the 1st group consisted of 28 women (53.8%) who showed no signs of disease progression; the 2nd group consisted of 24 women (46.2%) with progression of disease. The average age of patients in the 1st group was 47.0 ± 0.7 years of age, in the 2nd group – 56.4 ± 1.2 years of age. The control group consisted of 12 healthy women 48.5 ± 0.8 years of age.

Descriptive data

All patients, including the control group, were rural and urban residents of different occupations who permanently lived

in the respective places of residence. Among the occupational hazards, we identified significant physical load, as well as the influence of substances of chemical origin such as pesticides, herbicides and insecticides.

Outcome data

Histological examination of endometrial biopsies of all patients revealed signs of non-atypical endometrial hyperplasia [16]. The 1st group included patients with the type of EH in which an increased number of glandular and stromal elements without structural rearrangement of the endometrium (simple hyperplasia) was histologically detected.

The 2nd group includes patients with the type of EH with a changed shape of the glands and a decrease in the expression of the stromal component, i.e. with the presence of structural changes, which indicated the progression of the disease (complex hyperplasia) [16–18].

Data on the analysed socio-economic factors, together with the obstetric and gynaecological anamnesis, is summarised in Table 1.

A moderate and severe degree of climacteric syndrome was found in patients of the 1st group (17.9 ± 7.4% and 7.1 ± 5.0%, respectively) and in patients of the 2nd group (25.0 ± 9.0% and 16.7 ± 11.2%, respectively). A mild degree was observed in 75.0 ± 8.3% of women of the 1st group and in 58.3 ± 10.3% of the 2nd study group. Women in the control group only had mild (83.3 ± 11.2%) and medium (16.7 ± 11.2%) degrees of climacteric syndrome.

All patients exhibited dysfunctional uterine bleeding. Percentages of reported clinical signs of climacteric syndrome, such as irritability, fatigue, depression, headache, hot flashes and obesity, are presented in Table 2.

Analysis of extragenital pathology in patients with EH showed that the following diseases were observed in both study groups: dyshormonal pathology of the mammary glands, cardiovascular system diseases, endocrine system diseases, pathology of the gastrointestinal tract, urinary system diseases and hepatobiliary system diseases (Table 3).

Table 1. Socio-economic factors, indicators of obstetric and gynaecological anamnesis in patients with EH

Indicator of anamnesis	1 st group (n = 28)		2 nd group (n = 24)		Control group (n = 12)	
	Number of females	Percentage, % (p ± m)	Number of females	Percentage, % (p ± m)	Number of females	Percentage, % (p ± m)
Rural areas	20	71.4 ± 8.7	17	70.8 ± 9.5	8	66.7 ± 14.2
Harmful working conditions	5	17.9 ± 7.4	6	25.0 ± 9.0	1	8.3 ± 8.3
Smoking	3	10.7 ± 6.0	3	12.5 ± 6.9	1	8.3 ± 8.3
≥ 3 pregnancies	13	46.4 ± 9.6*	11	45.8 ± 10.4	2	16.7 ± 11.2
≥ 4 obstetric and gynaecological interventions	17	60.7 ± 9.4*	16	66.7 ± 9.9*	3	25.0 ± 13.1
Intrauterine contraception (IUD)	6	21.4 ± 7.9	9	37.5 ± 10.1*	1	8.3 ± 8.3
Infertility	2	7.1 ± 5.0	1	4.2 ± 4.17	–	–

* Indicates statistical significance compared to control at p < 0.05.

Table 2. Clinical signs of climacteric syndrome in patients with EH

Indicator of anamnesis	1 st group (n = 28)		2 nd group (n = 24)		Control group (n = 12)	
	Number of females	Percentage, % (p ± m)	Number of females	Percentage, % (p ± m)	Number of females	Percentage, % (p ± m)
Irritability	25	89.3 ± 6.0	22	91.7 ± 5.8*	7	58.3 ± 14.9
Fatigue	21	75.0 ± 8.3	20	83.3 ± 7.8*	6	50.0 ± 15.1
Depression	18	64.3 ± 9.2	18	75.0 ± 9.0	8	66.7 ± 14.2
Headache	17	60.7 ± 9.4	17	70.8 ± 9.5	5	41.7 ± 14.9
Hot flashes	19	67.9 ± 9.0	18	75.0 ± 9.0	5	41.7 ± 14.9
Obesity	8	28.6 ± 8.7	7	29.2 ± 9.5	2	16.7 ± 11.2

* Indicates statistical significance compared to control at p < 0.05.

Table 3. Extragenital pathology in patients with EH

Indicator of anamnesis	1 st group (n = 28)		2 nd group (n = 24)	
	Number of females	Percentage, % (p ± m)	Number of females	Percentage, % (p ± m)
Dyshormonal pathology of the mammary glands	9	32.1 ± 9.0	8	33.3 ± 9.8
Cardiovascular system diseases	6	21.4 ± 7.9	5	20.8 ± 8.5
Endocrine system diseases	5	17.9 ± 7.4	4	16.7 ± 7.8
Pathology of the gastrointestinal tract	3	10.7 ± 6.0	5	20.8 ± 8.5
Urinary system diseases	4	14.3 ± 6.7	6	25.0 ± 9.0
Hepatobiliary system diseases	8	28.6 ± 8.7	7	29.2 ± 9.5

Main results

There was no statistically significant difference ($p > 0.05$) between the socio-economic indicators of both groups, nor was any significant difference found between each of these groups and the control group. The analysis of obstetric and gynaecological anamnesis showed that the percentage of women with three or more pregnancies in the 1st group was higher compared to the control group ($p < 0.05$). The proportion of women with four or more obstetric and gynaecological interventions in both groups was significantly higher compared to the control group ($p < 0.05$). The proportion of women who used intrauterine contraception (IUD) was significantly ($p < 0.05$) higher in women of the 2nd group than in women without hyperplasia. In the patients of the 1st group, there was no significant difference in this indicator in comparison with the patients of the 2nd group and the patients of the control group ($p > 0.05$).

The analysis of the clinical course of the perimenopausal period established that there was no significant difference between the rates of moderate and severe climacteric syndrome in female patients ($p > 0.05$). In women of the control group, there was only a mild and moderate degree of climacteric syndrome with a significant difference ($p < 0.05$) with the corresponding indicators of the 1st and the 2nd study groups. A statistically significant difference ($p > 0.05$) in the clinical signs of climacteric syndrome, depending on the type of hyperplasia, was not found.

The percentage of patients in the 2nd group who complained about bouts of irritability significantly ($p < 0.05$) prevailed corresponding to the control group. The percentage of the women

in the 2nd study group who complained about fatigue was also significantly higher ($p < 0.05$) compared to the women of the control group.

Analysis of the proportions of extragenital pathology did not reveal a significant difference ($p > 0.05$) between the groups of patients with EH.

In order to determine the value of the combined changes of certain parameters for predicting the progression of the disease, we performed neural network clustering. The study was conducted on the basis of the following clinical and anamnestic indicators of the 1st and 2nd groups of patients: 1 – age, 2 – place of residence (city/village), 3 – number of previous pregnancies, 4 – number of obstetric and gynaecological interventions, 5 – working conditions (harmless/harmful), 6 – smoking, 7 – intrauterine contraception, 8 – infertility, 9–12 – severity of menopause syndrome (9 – mild, 10 – medium, 11 – severe), 12 – dysfunctional uterine bleeding, 13 – irritability, 14 – fatigue, 15 – depression, 16 – headache, 17 – hot flashes, 18 – obesity, 19 – dyshormonal pathology of the mammary glands, 20 – diseases of the cardiovascular system, 21 – diseases of the endocrine system, 22 – pathology of the gastrointestinal tract, 23 – diseases of the urinary system, 24 – diseases of the hepatobiliary system. The indicator of disease progression number 25 was given a value for the female patients of the 1st group as “1” and for the 2nd group female patients – as “2”.

For the neural network clustering algorithm, a number of clusters equal to three was chosen. 21.15% of patients belonged to the first cluster, 53.85% to the second, and 25.00% to the third (Figure 1a). The largest relative share of patients with disease progression was found in the 2nd cluster.

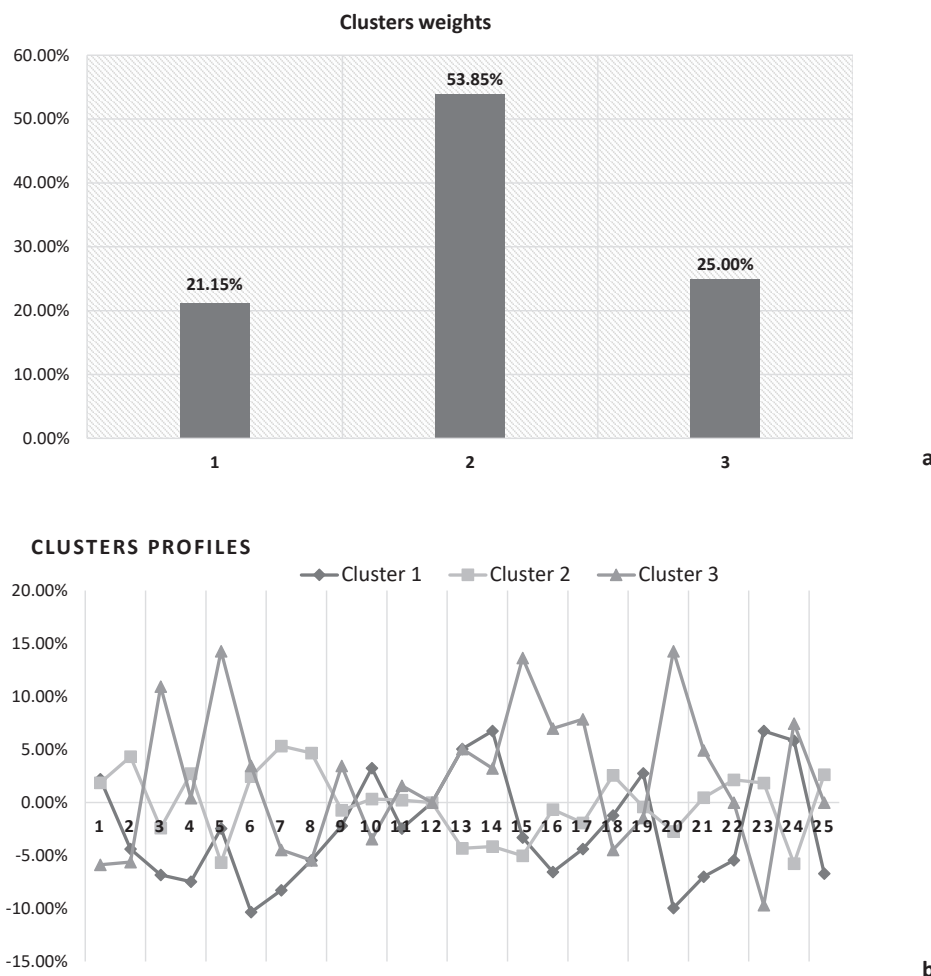


Figure 1. The results of clinical and anamnestic clustering indicators of patients with various types of endometrial hyperplasia: a) cluster portrait – values of parameters, including clinical and anamnestic indicators and an indicator of disease progression; b) shares of clusters – percentages of patients who fell into a certain cluster

With the help of the cluster portrait (Figure 1b), we found out that this cluster also includes most patients who lived in rural areas (2) and patients with the highest number of obstetric and gynaecological surgeries (4). The second cluster includes the largest number of patients who used intrauterine contraception in the anamnesis (7), patients with infertility (8), obesity (18) and diseases of the gastrointestinal tract (22).

Discussion

Key results

A comparative analysis of the relative shares of socio-economic factors, indicators of obstetric and gynaecological anamnesis, manifestations of climacteric syndrome and comorbid pathology in patients revealed certain tendencies but did not show statistically significant differences between the 1st and 2nd groups. Significant differences were established only in relation to individual indicators in comparison with the control group.

As the results of the study show, the analysis of average and relative values makes it impossible to establish the value of combined changes of certain parameters for predicting the course of the disease, in particular, the progression of the disease. At the same time, neural network clustering can effectively and objectively assign patients to the appropriate category – either with progression of the disease or with a stable result.

Thus, the results of the cluster analysis showed that when predicting the progression of EH based on the analysis of clinical and anamnestic indicators, it is important to take into account the combination of the use of intrauterine contraception, as well as infertility, obesity and diseases of the gastrointestinal tract in female patients. At the same time, the probability of EH progression also increases with an increase in the number of operative obstetric and gynaecological interventions and is higher for residents of rural areas.

The following method of step-by-step analysis of examination indicators of patients with EH is proposed in order to optimise the prognosis of disease progression for improvement of examination and treatment. During the examination at the primary level, the necessary indicators (average and relative values) are recorded and analysed. Cluster analysis of obser-

vational data in the region is carried out to determine groups of indicators important for forecasting and for their combined changes, in particular, using the NeuroXL Classifier program.

Limitations of the study

The proposed approach (using examples of patients with EH) can be widely used to optimise the prediction of the course of diseases at the primary level of medical care. At the first stage of getting to know the family doctor, it will be necessary to get advice from an IT specialist. However, after conducting neural network clustering using the original dataset, the results can be applied in the future, according to the proposed methodology, without the need for an expert opinion for each clinical case.

Conclusions

This article describes the methodology of predicting the progression of EH, which includes an analysis of relative age values, socio-economic factors, indicators of obstetric and gynaecological history, clinical manifestations of climacteric syndrome, existing extragenital pathology and algorithms of network clustering.

In order to effectively and objectively assign patients to the risk group for the progression of EH according to the indicators obtained during observation, neural network clustering was used, which allows one to determine the value of combined changes of certain parameters for the prognosis of the progression of the disease.

A technique for optimising the prediction of disease progression at the primary level has been developed in order to make corrections to examination and treatment using neural network clustering.

Generalisability

The proposed technique can be used to determine the risk of progression of various diseases that are most common in one or another region. This technique is easy to use and does not require large financial investments, which is important in organising the work of family-medicine doctors.

Source of funding: This work was funded from the authors' own resources.

Conflicts of interest: The authors declare no conflicts of interest.

References

1. Epstein RM, Franks P, Fiscella K, et al. Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. *Soc Sci Med* 2005; 61: 1516–1528.
2. Batko Y, Batryn N, Melnyk G, et al. Development of algorithms for biomedical image segmentation based on preliminary markup and texture attributes. *Eastern-European Journal of Enterprise Technologies* 2017; 90(6/4): 35–44.
3. Mintzer OP, Woronenko YV, Babintseva LY, et al. Health Informatization Concept in Ukraine. *Medical Informatics and Engineering* 2012; 3: 5–29.
4. Wong D, Galleqos Y, Weinger M, et al. Changes in intensive care unit nurse task activity after installation of a third-generation intensive care unit information system. *Crit Care Med* 2003; 31(10): 2488–2494.
5. Vostrov GN, Mintzer OP, Pavlov OO, et al. Information model of remote medical services. The first message. *Medical Informatics and Engineering* 2010; 3: 3747.
6. Hartzband P, Groopman J. Off the record – avoiding the pitfalls of going electronic. *N Engl J Med* 2008; 358: 1656–1658.
7. Florczak B, Scheurich A, Croghan J, et al. An observational study to assess an electronic point-of-care wound documentation and reporting system regarding user satisfaction and potential for improved care. *Ostomy Wound Manage* 2012; 58(3): 46–51.
8. Kalugina LV, Tatarchuk TF, Zanko OV. Genitourinary menopausal syndrome is a problem for women of elegant age and not only. *Reproductive Endocrinology* 2016; 28: 68–74.
9. Slyva A. Endometrial hyperplastic processes in perimenopausal women. Clinical and anamnestic analysis by using neural network clustering. *J Educ Health Sport* 2019; 9: 1014–1024.
10. Yoshihama S, Roszik J, Downs I, et al. NLRC5/MHC class I transactivator is a target for immune evasion in cancer. *Proc Natl Acad Sci USA* 2016; 113: 5999–6004.
11. Cormio A, Guerra F, Cormio G, et al. Mitochondrial DNA content and mass increase in progression from normal to hyperplastic to cancer endometrium. *BMC Res Notes* 2012; 5: 279.

12. Benyuk VO, Altibaeva DM, Goncharenko VM, et al. The role of immune mechanisms in the development of benign pathology of the uterus. *Women's Health* 2016; 5(111): 30–32.
13. Martsenyuk V. On an indirect method of exponential estimation for a neural network model with discretely distributed delays. *Electronic Journal of Qualitative Theory of Differential Equations* 2017; 23: 1–16.
14. Selskyy PR, Vakulenko DV, Televiak AT, et al. On an algorithm for decision-making for the optimization of disease prediction at the primary health care level using neural network clustering. *Fam Med Prim Care Rev* 2018; 20(2): 171–175.
15. Martsenyuk V. Indirect method of exponential convergence estimation for neural network with discrete and distributed delays. *Electronic Journal of Differential Equations* 2017; 246: 1–12.
16. Sobczuk K, Sobczuk A. New classification system of endometrial hyperplasia WHO 2014 and its clinical implications. *Prz Menopauzalny* 2017; 16 (3): 107–111.
17. Boichuk AV, Shadrina VS, Vereshchahina TV. Hyperplasia of endometrium – a modern systemic-pathogenetic view of the problem (literature review). *Actual Problems of Pediatrics, Obstetrics and Gynecology* 2019; 1: 67–72.
18. Hapangama DK, Kamal AM, Bulmer JN. Estrogen receptor b: the guardian of the endometrium. *Hum Reprod Update* 2015; 21: 174–193.

Tables: 3

Figures: 1

References: 18

Received: 01.12.2022

Reviewed: 02.02.2023

Accepted: 24.03.2023

Address for correspondence:

Prof. Petro Selskyy, MD, PhD, DSc

I. Horbachevsky Ternopil National Medical University

Maydan Voli 1

46000 Ternopil

Ukraine

Tel.: +38 0673985279, +38 0352253998

E-mail: selskyy@tdmu.edu.ua