

## ORIGINAL ARTICLE

# APPLICATION OF NEURAL NETWORK TECHNOLOGIES IN THE DENTAL CARIES FORECAST

DOI: 10.36740/WLek202007135

**Oleksandr A. Udod<sup>1</sup>, Hanna S. Voronina<sup>1</sup>, Olena Yu. Ivchenkova<sup>2</sup>**<sup>1</sup>DONETSK NATIONAL MEDICAL UNIVERSITY, LYMAN, UKRAINE<sup>2</sup>DONBASS STATE ENGINEERING ACADEMY, KRAMATORSK, UKRAINE

## ABSTRACT

**The aim:** of the work was to develop and apply in the clinical trial a software product for the dental caries prediction based on neural network programming.

**Materials and methods:** Dental examination of 73 persons aged 6-7, 12-15 and 35-44 years was carried out. The data obtained during the survey were used as input for the construction and training of the neural network. The output index was determined by the increase in the intensity of caries, taking into account the number of cavities. To build a neural network, a high-level Python programming language with the NumPy extension was used.

**Results:** The intensity of carious dental lesions was the highest in 35-44 years old patients –  $6.69 \pm 0.38$ , in 6-7 years old children and 12-15 years old children it was  $3.85 \pm 0.27$  and  $2.15 \pm 0.24$ , respectively ( $p < 0.05$ ). After constructing and training the neural network, 61 true and 12 false predictions were obtained based on these indices, the accuracy of predicting the occurrence of caries was 83.56%. Based on these results, a graphical user interface for the “CariesPro” software application was created.

**Conclusions:** The resulting neural network and the software product based on it permit to predict the development of dental caries in persons of all ages with a probability of 83.56%.

**KEY WORDS:** caries, prognosis, neural network, software

Wiad Lek. 2020;73(7):1499-1504

## INTRODUCTION

Innovative and promising research areas include the use of intelligent systems, in particular, those based on neural networks [1,2,3]. In health care, such systems are used in human genome research, drug development, diagnosis and prognosis of certain diseases, such as cancer, cardiovascular diseases, etc. [2,3,4,5]. In the processing of medical data, information systems based on neural networks have several advantages over traditional methods of mathematical statistics. Neural networks are capable of simultaneous processing a large array of different parameters, which is extremely important in the diagnosis and prognosis of diseases, particularly if the patient has several diseases. In such systems, after their creation, there is a possibility for “extra training” of the initial model, for example, with the emergence of new or more accurate data of additional studies or “extra training” of the model on the initial data in changing external conditions, for example, in the case of new drugs prescribing during treatment [1,4,5].

The work of neural networks is based on algorithms that are, in their turn, based on the principles and patterns of signal transmission through neurons and synapses of the human nervous system. Neural networks are capable of taking decisions and returning results based on the hidden patterns found in the large data array. The peculiarity of neural networks is that they are not programmed and do not use any rules to draw a conclusion, but they are trained to do it by examples [1,6]. Thus, neural network

technologies can be applied to solve difficult and monumental challenging tasks with numerous parameters and characteristics in various fields of healthcare.

State of the present-day health care is drawing attention to preventive care. Against this background, it is quite promising and, certainly, important to predict the development and course of diseases, in particular dental ones, among which, undoubtedly, the leading place is occupied by dental caries and its complications [7,8]. To date, the prevalence and intensity of this disease in pediatric and adult populations in different regions of Ukraine remain at a fairly significant level [8,9,10]. Integrated caries prevention measures applied to different segments of the population, including organized contingents, taking into account regional peculiarities, do not always show high efficiency due to, first of all, the lack of an individualized approach. Caries prevention, based on the principles of selectivity, i.e. sampling, and timeliness, based on determining the individual disposition toward caries, creates the preconditions for improving its efficacy, for more rational use of the dentist's working time as well as therapeutic and preventative means. In this sense, the predictive field of cariesology, i.e. prediction of this disease course and the possible results of preventive measures for a particular patient, seems quite appropriate and necessary.

There are various approaches to the prediction of dental caries in children and adults, well-known are, for example, mathematical models, as well as numerous predictive

methods based on the determining oral fluid properties, the state of the dental hard tissues, hygienic parameters of the oral cavity, etc. [11]. Clinical and laboratory studies of the tooth enamel acid resistance and its restitution after acid exposure are considered to be the most informative and prognostically valuable. These include determination of structural and functional acid resistance of tooth enamel by the enamel resistance test, clinical assessment of enamel remineralization rate, CRT-test (color, reaction, time), but they have recently been used infrequently enough [11,12].

Unfortunately, most of the proposed prediction methods are technically sophisticated, sometimes multi-step and time-consuming, and quite frequently their implementation is not possible in daily clinical dental practice. Therefore, it is important to develop fundamentally new approaches to the prediction of dental caries based on software products, especially since computer technologies are being actively implemented into the medical-diagnostic process, particularly in private-owned dental institutions.

## THE AIM

The aim of the work was to develop and apply a software product to predict dental caries on the basis of neural network programming.

## MATERIALS AND METHODS

Dental examination of 73 persons was carried out, which was divided into three age groups. The first group included 26 children aged 6-7 years, the second group included 34 12-15 years old children, the third group comprised 13 persons aged 35-44 years. During the examination, the caries intensity in temporary (if present) and permanent teeth was determined according to the index of cfc (caries, filling, cavity) + cfcex (extracted teeth) or CFcEX, taking into account the number of tooth decay cavities, oral hygiene condition according to the simplified hygienic Green-Vermillion index (OHI-S), structurally-functional acid resistance of tooth enamel (SFARE) according to the enamel resistance test (TER) and the functional component of the structural and functional acid resistance of enamel ( $\Delta$ TER) [13,14]. After 1 year, a repeated clinical examination was performed, each patient was re-assessed by cfc + CFcEX or CFcEX index and the increase in the intensity of carious lesions was calculated according to the  $\Delta$ cfc + CFcEX or  $\Delta$  CFcEX index. The results of the study were recorded in the medical card of the dental patient (form No. 043 / o). We used a standard package of statistical software for Microsoft Excel to prepare the patient source data tables for and to carry out calculations.

The following indices were used as input (X) data, on the basis of which the caries probability was predicted: age of the patient (X1), dental caries intensity index taking into account the number of cavities (X2), hygienic index OHI-S (X3), structural- functional acid resistance of tooth enamel (X4) and functional component of structural and functional acid resistance of enamel (X5), i.e., five indices

in total. The output, i.e. final, index (Y) was determined by the increase in the intensity of dental caries, taking into account the number of cavities. The obtained database was divided into two samples – the training and the control ones, which included indices of 65 and 8 patients, respectively.

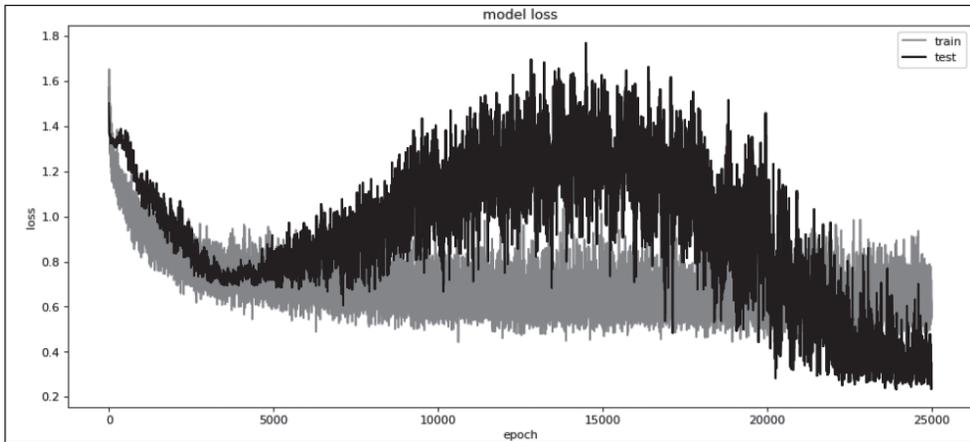
To build a neural network, we used a high-level Python programming language with the NumPy extension. The following software products were used during the work: Pandas, TensorFlow, Keras, Scikit-learn, Matplotlib, Jupyter Notebook [15].

At the first of the successive stages series, the data were processed and prepared. Initially, using the Pandas electronic library for data analysis and processing, the generated database was read from Excel files, after which the downloaded data was divided into input (X) and output (Y). Further, in order to further normalize the data for each of the factors used, maximum values were determined using the standard NumPy function, according to which the normalization was performed. This was necessary to improve the next neural network training process and to reduce the input to a single computing system.

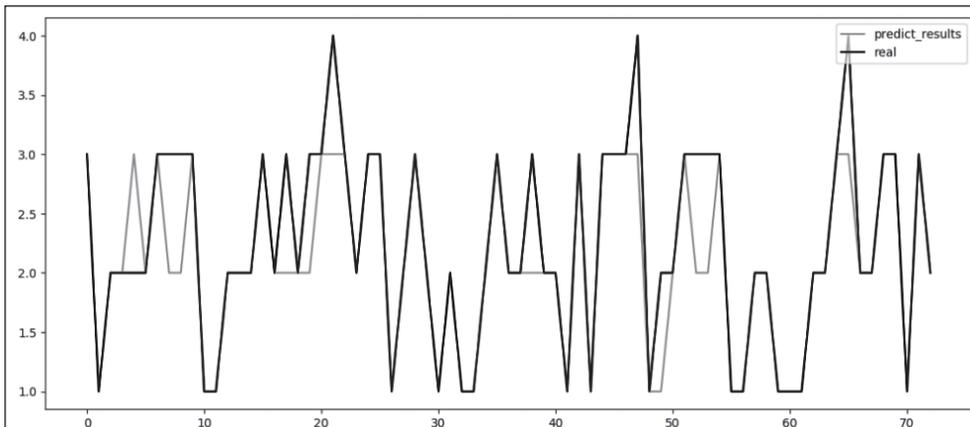
The final step of this stage was to convert the output coordinate to the unitary code (one-hot encoding) using the standard Scikit-learn function. The use of unitary code in the construction of neural networks is required in cases where the number of outputs corresponds to the number of possible output coordinate classes. When a neural network receives a test sample of factors, the maximum value will be only set to one of the outputs.

At the second stage of the neural network construction its future structure and training was drafted. This was done using the sequential method of describing the neural network using Sequential software. During the work, we created a neural network containing five neurons in the input layer (based on the number of available dental indices), one hidden fully connected layer and an output layer with four neurons. The number of synaptic weights and neurons in the hidden layer was determined according to the Kolmogorov – Arnold – Hecht – Nielsen theorem [15]. A ReLU was selected as the activation function for the input layer, and the softmax logistic function – for activation of the output layer. In order to avoid the problem of the neural network “retraining”, which may cause the dependence of the output index on the input data to occur sharply nonlinear, i.e. the result will be significantly and unpredictably changed with slight fluctuations of the input data, the method of “thinning” was applied, better known as Dropout. Its essence is that instead of training a single neural network, there is a simultaneous training of several ones, with further generalization of the results obtained. The RMSprop algorithm was used as the optimization method, for the quality metric of the constructed neural network – Accuracy.

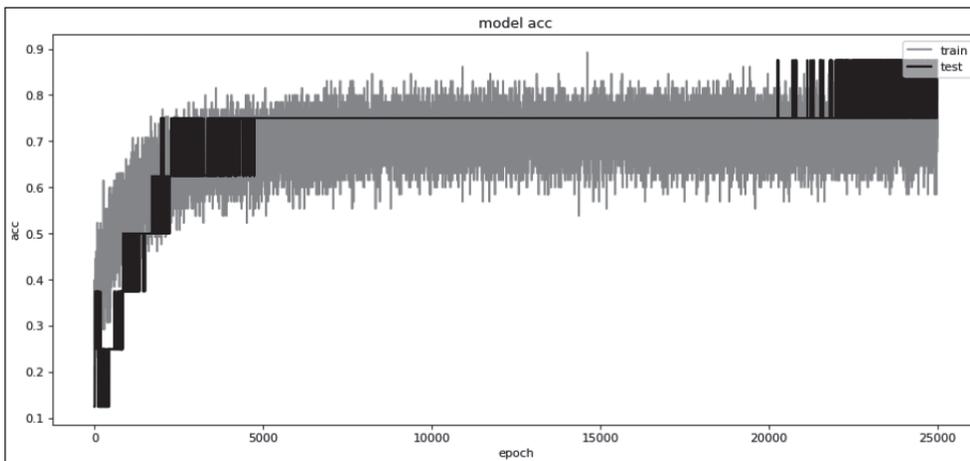
The duration of the neural network training process was 25,000 epochs, with 10% of the available sample data not participating in the process, but used for test purposes only. After training, the neural network and the software



**Fig. 1.** Diagram of loss-function based on training data and test data.



**Fig. 2.** Diagrams for visualization of the dental caries intensity growth rate prediction based on the developed neural network.



**Fig. 3.** Visualization of changes in accuracy (acc metrics) in the training process of the resulting neural network.

**Fig. 4.** Graphical interface to the main user form of the “CariesPro” application.

product based on it, called CariesPro, was used to achieve the set purpose, i.e., predicting dental caries. Matplotlib e-library was used to visualize the training outcomes and the results obtained.

All the studies were carried out in accordance with the principles of the Declaration of Helsinki, adopted by the General Assembly of the World Medical Association (1997-2000) and approved by the University Bioethics Commission.

## RESULTS

In order to create an input database, a primary examination of persons was carried out, during which it was found that the intensity of tooth decay by  $cfc + CFcEX$  index, taking into account the number of cavities in children aged 6-7 years, was  $3.85 \pm 0.27$ , which, according to the WHO standard criteria is considered medium. Statistically significant ( $p < 0.05$ ), this index was lower in 12-15 years old children, in whom it was equal to  $2.15 \pm 0.24$ , which corresponds to the low level according to the WHO recommendations. The intensity of carious lesions in teeth according to the  $CFcEX$  index was very high by the WHO criteria in patients aged 35-44 years, it was  $6.69 \pm 0.38$  and statistically significantly ( $p < 0.05$ ) higher than all the others.

Assessment of the oral cavity hygienic condition by the simplified hygienic index OHI-S revealed its unsatisfactory level in individuals of all three age groups, and their mean values were not statistically significantly different ( $p > 0.05$ ), being  $1.98 \pm 0.05$  points,  $1.87 \pm 0.03$  points and  $1.86 \pm 0.05$  points, respectively.

The structural and functional acid resistance of the teeth enamel according to the enamel resistance test was the worst in children aged 6-7 and 12-15 years, in particular, their mean values were  $4.96 \pm 0.33$  points and  $5.00 \pm 0.34$  points ( $p > 0.05$ ), respectively. Somewhat better, but with no statistically significant difference ( $p > 0.05$ ), structural and functional acid resistance of enamel (SFARE) was in patients of the older age group, making  $4.15 \pm 0.66$  points.

After carrying out hemostimulation of the dental pulp, which was performed by rubbing hypertonic sodium chloride solution into the palatal surface of the tested tooth for 10 minutes, the most significant dynamics of structural and functional acid resistance was recorded in 6-7 and 12-15 years old children, who had a functional SFARE component,  $1.73 \pm 0.15$  points and  $1.71 \pm 0.12$  points ( $p < 0.05$ ), respectively [13].

Statistically significantly ( $p < 0.05$ ) less, in particular, almost by 11.5 times, this index was found in persons of 35-44 years –  $0.15 \pm 0.10$  points. Such differences are undoubtedly related to the age patterns regarding the dynamics of the enamel acid resistance and the dominance of its structural component in the elderly patients [14].

After 1 year, the intensity of dental caries in children aged 6-7 years increased by 1.6 times to  $6.27 \pm 0.35$ , in 12-15 years old children this index increased almost by 2 times to  $4.23 \pm 0.24$ , in persons aged 35-44 years – up to  $8.92 \pm 0.52$ , i.e. only by 1.3 times. In all the age groups surveyed,

the indices in the re-survey were statistically significantly ( $p < 0.05$ ) different from the primary ones. The increase in the intensity of dental caries in these individuals by the indices  $\Delta cfc + CFcEX$  or  $\Delta CFcEX$  was  $2.42 \pm 0.15$ ;  $2.09 \pm 0.15$  and  $2.23 \pm 0.25$ , respectively, that is, indices did not differ statistically ( $p > 0.05$ ).

The results obtained during the two surveys were used to train the neural network. The training outcomes were visualized by constructing a loss function in Matplotlib. This function was used to calculate the error between the existing and the obtained result, i.e. its ultimate purpose is to reduce this error. Thus, the loss function of both in the training data that participated in the training process and in the test data that did not participate in the training process was minimized (fig. 1). This testified to the ability of the neural network to learn. Significant fluctuations in loss functions during the transition from epoch to epoch occurred due to the use of Dropout.

After the process of training the neural network, a whole selection of the reported factors was provided. In response, it has generated a category of numbers for the individual increase in caries intensity (Y), among which index 4 was never met. This is due to the imbalance of the training sample and to the very poor representation of this result in it.

The graphical overlay of the predicted results and the real results diagrams, which were constructed in the same coordinate system, indicated a coincidence of the results (fig. 2).

Visualization of the accuracy of the constructed Accuracy (acc) neural network, which was calculated using Keras, has confirmed the correct course of the neural network training process, and oscillations during the transition from epoch to epoch were due to the use of Dropout (fig. 3).

As a result of calculating work accuracy index of the Accuracy (acc) neural network, 61 true and 12 false predictions were obtained from the entire sample, i.e. the prediction accuracy of the constructed and trained neural network was 83.56%.

Later, using the high-level Delphi programming language, a graphical user interface was created (fig. 4).

In order to work with the presented software, it is necessary to fill in the patient's passport data and enter the five dental indices obtained during the clinical dental examination. The following buttons are present in this window:

“Network calculation” – when this button is pressed, the value of the probable number of carious cavities index of the patient is calculated according to the set values in the analysis data;

“Save” – pressing this button saves patient data and estimated values;

“Monthly visit report” – pressing this button generates a report on the status of teeth and oral hygiene in patients who visited a dentist in the last month;

“Conclusion” – pressing this button generates a conclusion that contains data on the condition of the teeth and oral hygiene during the current and past examinations of the patient, as well as the results of predicting the probable number of caries cavities, based on the obtained dental indices.

## DISCUSSION

As of today, there is a large number of computer software to predict dental caries, most of which are based on standardized multifactorial questionnaires [16]. The most famous of them are NusCra (National University of Singapore Caries Risk Assessment), CARIOGRAM, PreViser, ADA (caries risk assessment by the American Dental Association), CAMBRA (Caries Management By Risk Assessment), CAT (America Academy of Pediatric Dentistry's Caries Assessment Tool). Each software takes into account nine to twelve factors, among which, besides clinical, are socio-demographic (age, ethnicity), behavioral factors (diet) and others.

One of the most widely used software in the world, called "CARIOGRAM", is based on the multifactorial analysis of factors that, from the developers' point of view, are leading in the caries development [17]. The software takes into account ten factors, including the salivation rate, the buffer capacity of the oral fluid, the content of lactobacilli and streptococci in it, etc. Identifying some of them requires laboratory testing, additional equipment, and therefore unnecessary financial and time costs. The significance of each factor is conditionally estimated from 0 to 3 points, without relying on any quantitative criteria, which should certainly be considered subjective enough.

The result is a graphical representation of a sectoral diagram including five parts of different colors and only demonstrating the influence of etiological factors. The risk of developing caries, according to the image, is rated as high, medium or low, but the percentages given do not answer the question of the number of possible dental caries lesions in a particular patient over time.

The developed "CariesPro" computer software does not have the above disadvantages. In order to perform full-time work with this software, one needs to identify only five, which is important, exclusively clinical indices that can be established during a single dental examination, so that additional examinations and visits are not required. In addition, the software takes into account the exact figures regarding the condition of the teeth and oral hygiene, which are specific to the particular individual.

The final index in working with the "CariesPro" software is the prediction of a certain number of carious cavities probability after a certain time, which permits to plan and, if necessary, timely carry out the most individualized measures of the selective caries prevention. "CariesPro's" high level of forecasting performance is in line with existing software products, and the said software has the ability to constantly learn and improve.

## CONCLUSION

The resulting neural network obtained and the software product based on it permit to predict the development of dental caries in persons of all ages with a probability of 83.56%. Thanks to the developed form, the presented software can be installed in a personal computer, which, in turn, permits its efficient use for predicting dental caries in daily clinical practice.

## REFERENCES

1. Rethinavel Subramanian A. Neural Philosophy in Medical Applications. *Int. Journal of Engineering Research and Applications*. 2014;4(1):237-241.
2. Park W.J., Park J.B. History and application of artificial neural networks in dentistry. *Eur J Dent*. 2018;12(4): 594–601. doi:10.4103/ejd.ejd\_325\_18.
3. Hwang J.J., Jung Y.H., Cho B.H. et al. An overview of deep learning in the field of dentistry. *Imaging Sci Dent*. 2019;49(1):1-7. doi: 10.5624/isd.2019.49.1.1.
4. Nindrea R.D., Aryandono T., Lazuardi L. et al. Diagnostic Accuracy of Different Machine Learning Algorithms for Breast Cancer Risk Calculation: a Meta-Analysis. *Asian Pac J Cancer Prev*. 2018;19(7):1747-1752.
5. Itchhaporia D., Snow P.B., Almassy R.J. et al. Artificial neural networks: current status in cardiovascular medicine. *J Am Coll Cardiol*. 1996;28(2):515-21.
6. Han S.H., Kim K.W., Kim S. et al. Artificial Neural Network: Understanding the Basic Concepts without Mathematics. *Dement Neurocogn Disord*. 2018;17(3):83-89. doi: 10.12779/dnd.2018.17.3.83.
7. Frencken J.E., Sharma P., Stenhouse L. et al. Global epidemiology of dental caries and severe periodontitis – a comprehensive review. *J Clin Periodontol*. 2017;44(18):94-105. doi: 10.1111/jcpe.12677.
8. Kassebaum N.J., Smith A.G.C., Bernabé E. et al. Global, Regional, and National Prevalence, Incidence, and Disability-Adjusted Life Years for Oral Conditions for 195 Countries, 1990–2015: A Systematic Analysis for the Global Burden of Diseases, Injuries, and Risk Factors. *J Dent Res*. 2017;96(4):380-387. doi: 10.1177/0022034517693566.
9. Bezvushko E.V., Lahoda L.S. Stan tverdykh tkanyn zubiv u ditei mista Lutska [Condition of hard teeth tissues in children of Lutska]. *Aktualni problemy suchasnoi medytsyny: Visnyk ukraïnskoi medychnoi stomatolohichnoi akademii*. 2017;17(58):232-235. (in Ukrainian).
10. Zadorozhna I.V., Povorozniuk V.V. Poshyrenist ta intensyvni kariiisu zubiv u ditei Ukrainy: rezultaty kliniko-epidemiolohichnoho obstezhennia [Prevalence and intensity of dental caries in children of Ukraine: results of clinical and epidemiological examination]. *Bil. Suhloby. Khrebet*. 2014;1-2(13-14):26-29. (in Ukrainian).
11. Udod O.A., Voronina H.S. Suchasni tendentsii ta pidkhody do prohnuzovannia kariiisu zubiv [Modern trends and approaches to predicting dental caries]. *Novyny stomatolohii*. 2019;1(98):S. 75-79. (in Ukrainian).
12. Brudevold F., McCann H.G. Enamel solubility tests and their significance in regard to dental caries. *Ann N Y Acad Sci*. 2006;153(1):20-51. doi: 10.1111/j.1749-6632.1968.tb11725.x.
13. Pat. 118523 Ukrainy na korysnu model. Sposib vyznachennia funktsionalnoi komponenty strukturno-funktsionalnoi kyslotostiikosti emali zubiv [Method of determining functional component of structural and functional acid resistance of tooth enamel]. Udod OA, Voronina HS-Nº u201702237; zaiavl. 10.03.2017; opubl. 10.08.2017, Biul. Nº15. (in Ukrainian).
14. Okushko V.R. Osnovy fiziologii zuba [Fundamentals of tooth physiology]. Tiraspol': Izd-vo Pridnestr. un-ta. 2005. 240 s. (in Russian).
15. Rudenko O., Bezsonov O., Romanyk O. Neural network time series prediction based on multilayer perceptron. *Development Management*. 2019;5(1):23-34. doi:10.21511/dm.5(1).2019.03.
16. Senneby A., Mejare I., Sahlin N.E. et al. Diagnostic accuracy of different caries risk assessment methods. A systematic review. *J Dent*. 2015;43:1385–1393. doi: 10.1016/j.jdent.2015.10.011.

17. Bratthall D., Hansel-Petersson G. Cariogram – multifactorial risk assessment model for multifactorial disease. *Community Dent. Oral Epidemiol.* 2005;33:256-264.

*The present work is a fragment of the Stomatology Department No. 1 of Donetsk national medical university “Clinical and laboratory substantiation of improving technologies for diagnostics, treatment, prognosis and prevention of dental diseases” (state registration No. 0119 U 001447).*

**ORCID and contributionship:**

Oleksandr A. Udod: 0000-0001-6790-1936 <sup>A,E,F</sup>

Hanna S. Voronina: 0000-0002-3301-7808 <sup>B,D,E</sup>

Olena Yu. Ivchenkova: 0000-0003-4739-0192 <sup>C</sup>

**Conflict of interest:**

*Authors declare no conflict of interest.*

---

**CORRESPONDING AUTHOR**

**Hanna S. Voronina**

Donetsk National Medical University  
39 Mashinobudivnikiv blvd, 84313 Lyman, Ukraine

tel: 0509438625

e-mail: annavoronina2812@gmail.com

**Received:** 27.08.2019

**Accepted:** 16.04.2020

---

**A** – Work concept and design, **B** – Data collection and analysis, **C** – Responsibility for statistical analysis, **D** – Writing the article, **E** – Critical review, **F** – Final approval of the article