УДК 57.041

Нейросетевая оценка накопления металлов в организме жителей крупного города в результате полиметаллического загрязнения среды обитания

Тунакова Ю. А.[1], доктор хим. наук, профессор;
Новикова С. В.[1], доктор техн. наук, доцент; Валиев В. С.[2], Новикова К. Н.[2]
[1]Казанский национальный исследовательский технический университет им. А. Н. Туполева – КАИ
420126, Россия, г. Казань, ул. Четаева, 18,
[2]Институт проблем экологии и недропользования Академии наук Республики Татарстан
420087, Россия, г. Казань, ул. Даурская, 28

Аннотация. Прямая оценка ретенции чрезвычайно сложна как из-за множества внутренних механизмов, обеспечивающих динамику содержания металлов, так и из-за огромного разнообразия органов, тканей, процессов, обеспечивающих их перераспределение, транспорт и накопление. Содержание металлов в аккумулирующей среде организма является интегральным показателем, суммирующим мультисредовое воздействие и учитывающим все пути поступления металлов в организм. Разработана сложная нейросетевая модель задержки металлов в организме, включающая ансамбли нейросетевых регрессионных моделей, рассчитывающих уровни задержки металлов в зависимости от места проживания детей-подростков и поступления не только с потребляемой питьевой водой, но и с вдыхаемым воздухом. В результате исследований сформирована база данных концентраций металлов в атмосферном воздухе, потребляемой питьевой воде, крови и моче с учетом физиологических особенностей тестируемой чувствительной группы детей-подростков с целевой территориальной привязкой анализируемых проб. Упрощенная структура модели нейросетевой регрессии (уменьшение количества входов) дает достаточную точность, а сокращение количества нейронных сетей повышает адекватность моделей.

Ключевые слова: металлы, поступление, организм, удержание, кумуляция, биосубстраты, нейронные сети.

Neural network assessment of metal accumulation in the body of inhabitants of a large city as a result of polymetallic pollution of the habitat

Tunakova Y. A.[1], Novikova S. V.[1], Valiev V. S.[2], Novikova K. N.[2]
[1]Kazan National Research Technical University named after A. N. Tupolev
[2]Institute of Problems of Ecology and Subsoil Use Academy of Sciences of the Republic of Tatarstan

Annotation. Direct assessment of retention is extremely complex, both due to the many internal mechanisms that ensure the dynamics of the metal content, and due to the huge variety of organs, tissues, processes that ensure their redistribution, transport and accumulation. The content of metals in the accumulating body environment is an

integral indicator that summarizes the multi-environment impact and takes into account all the ways in which metals enter the body. A complex neural network model of metal retention in the body was developed, including ensembles of neural network regression models that calculate the levels of metal retention, depending on the place of residence of adolescent children and intake not only with consumed drinking water, but also with inhaled air. As a result of the studies, a database of metal concentrations in the atmospheric air, consumed drinking water, blood and urine was formed, taking into account the physiological characteristics of the tested sensitive group of adolescent children, with targeted territorial reference of the analyzed samples. The simplified structure of the neural network regression model (reducing the number of inputs) gives sufficient accuracy, and the reduction of neural networks increases the adequacy of the models.
Keywords: metals, intake, organism, retention, cumulation, biosubstrates, neural networks.

Introduction. The response marker of the human body to environmental objects pollution is accumulation of various pollutants in the blood, urine and tissues of the body. At the same time, both the pollutants themselves and specific compounds, called endogenous metabolites, are accumulated, the variability of which is affected by the assessed pollution. In this case, metals have a special meaning, since they belong to both substances' groups. Metals are of the greatest interest as components of the internal environment of living organisms, being direct participants in their vital activity and metabolism [1–6].

The elemental composition stability for the body is one of the most important and mandatory conditions for its normal functioning. Each metal has its own safe concentrations range in tissues, which makes it possible to adequately maintain the physiological state of the body and its biochemical functions. At the same time, metals can exhibit toxicity if they accumulate excessively in the body [7–9].

Metals that enter the human body orally and with inhaled air are redistributed among organs and tissues, excreted and partially accumulated in its biosubstrates. Therefore, one of the most reliable methods characterizing the polymetallic impact on public health is the assessment of the metals content in biosubstrates. At the same time, those biosubstrates that are involved in the processes of metal accumulation should be considered the most informative or diagnostic [10–15].

The key results feature of such researches should be recognized not as the usual statement about fact of change in the metals concentrations, but the special kinetics of this change, leading to the retention of these substances in the body [16–18].

Most researchers note a direct correlation between the metals content in biosubstrates of the body and environmental objects [15–18].

In average statistical terms, the intake of metals by the water-food route into the body of schoolchildren within a single urban area can be taken as a constant value. However, the quality of atmospheric air can vary significantly due to the high dynamism of the air environment. In this regard, the relationship assessment between the content of atmospheric impurities and metals in hair is of particular interest, realized by our experimental studies.

Main part. At the first stage, we conducted the metals content study in hair of adolescent children living in various districts of the Nizhnekamsk city. The representative group of examined children consisted of 48 people from 8 to 14 years old. They lived in the coverage areas of automatic air pollution control stations (AAPCS) located in different parts of the city. As cumulative biosubstrates, we used hair samples taken as part of a special screening study.

Atomic absorption spectrophotometry (AAS) was used to evaluate the content of 9 metals in children's hair: zinc, copper, manganese, cadmium, nickel, lead, chromium, iron and strontium. Results statistical evaluation was carried out by STATISTICA 12.6 program. The study results are presented in table 1.

Table 1. Metal content (M ±m, μg/g) in children's hair for different study areas

Metal	Zones/number of examined				
	AAPCS 12 $n = 7$	AAPCS 14 $n = 7$	AAPCS 13 $n = 10$	AAPCS 15 $n = 10$	AAPCS 11 $n = 14$
Zn	92.4 ±7.8	110.7 ±7.1	114.5 ±7.1	104.5 ±7.5	123.5 ±12.7
Cd	0.64 ±0.11	0.61 ±0.14	0.72 ±0.16	0.71 ±0.13	0.87 ±0.14
Cu	9.23 ±1.21	10.6 ±0.89	11.4 ±0.92	11.3 ±1.25	11.3 ±0.99
Mn	1.046 ±0.28	1.005 ±0.16	1.087 ±0.06	1.12 ±0.18	0.97 ±0.12
Ni	0.82 ±0.17	1.24 ±0.27	1.42 ±0.26	1.13 ±0.16	1.25 ±0.22
Pb	4.27 ±0.3	10.47 ±1.72	7.69 ±1.07	8.35 ±0.81	6.6 ±1.79
Cr	0.418 ±0.11	1.43 ±0.35	0.87 ±0.09	1.026 ±0.15	1.32 ±0.16
Fe	22.2 ±4.3	22.6 ±1.52	20.6 ±1.79	17.7 ±1.12	28.2 ±3.79
Sr	11.9 ±3.68	6.85 ±1.006	9.65 ±1.34	6.55 ±0.81	8.61 ±1.14

When comparing the average values of various metals content in the study areas, the following areas are statistically significant (significance level $p < 0.05$, according to Student criterion): for cadmium, the AAPCS 11 coverage zones (significantly differ from the AAPCS 12 and AAPCS 14 zones); for lead – AAPCS 14, AAPCS 15 (significantly differ from AAPCS 12); for chromium – AAPCS 14 and AAPCS 11 (significantly differ from AAPCS 12). For other metals, no statistically significant differences were noted, although there is a significant range of average values.

The best way to scale different-dimensional variables is to correlate each negative factor to its threshold value. This is a manifestation of an event whose probability is described by Bayes' theorem, which takes into account both a priori and a posteriori probability [20].

We used the ratio of actual concentrations to background (reference) concentrations: C_{act} / C_{ref}, which we will refer to as the hazard ratio (HR) in what follows, to assess the facts of exceeding the background indicator by specific values. Background metals concentrations (M ±m, μg/g) are given in our previous studies [22–24].

We took into account the ratio of a posteriori and a priori probabilities. This ratio estimates the probability of a risk event in a particular area of space relative to information about how often this event occurred in all areas. In what follows, we will refer to this as Bayesian probability. We used the following algorithm for determining Bayesian probability:

D – event consisting in the fact that we consider the full set of experimentally measured indicators tuples throughout the territory. According to the task $P(D) = 1$.

D_i – an event consisting in the fact that we consider a set of experimentally measured indicators tuple within a given zone i ($i = \overline{1 \ldots n}$).

S – an event consisting in the fact that, during consideration, an excess of the permissible threshold value was found for at least one indicator in at least one data tuple.

The probability of event S can be calculated as the total probability using the formula:

$$P(S) = \sum_{i=1}^{n} p(D_i)\, p(S \mid D_i). \tag{1}$$

We will assume that we are examining the entire set of data: $P(D) = 1$. Then the events D_i form a complete group:

$$\sum_{i=1}^{n} p(D_i) = P(D) = 1. \tag{2}$$

It follows from formula (4) that to calculate the posteriori probability of detecting a threshold exceeding when considering a specific zone *i*, we can apply the Bayes formula:

$$P(D_i|S) = \frac{p(D_i)\, p(S|D_i)}{p(S)}. \tag{3}$$

Note that due to (3) and (4):

$$P(D_i|S) = \frac{p(D_i)\, p(S \mid D_i)}{\sum_{i=1}^{n} p(D_i)\, p(S \mid D_i)} = \frac{p(D_i)\, p(S \mid D_i)}{\sum_{i=1}^{n} p(S \mid D_i)}, \tag{4}$$

here $p(S|D_i)$ is a prior probability that if the excess occurred, then it happened in the zone D_i.

We calculate the probability $p(D_i)$ using the classical formula (7) for determining the probability as the ratio of the conditional powers of the sets D_i and the entire set *D* under consideration:

$$P(D_i) = \frac{n_i}{N}. \tag{5}$$

If n_i is the area of the territory of the region D_i, then *N* is the entire area of the region under study.

In this way, we can calculate the probabilities of impurity concentrations exceeding their threshold values in different zones, and as a generalizing indicator, we propose to use the probability of exceeding the threshold for at least one of the estimated parameters.

As calculations result, we obtained the following probabilistic characteristics for integral HR distribution according to metals content in the hair of children living in different coverage areas of AAPCS on the Nizhnekamsk sity territory (table 2).

Thus, the highest probability of simultaneous excess of concentrations in hair over the background for at least half of the analyzed metals is highest in the AAPCS 11 coverage area (21 %), and the lowest in the AAPCS 12 coverage area (4.8 %).

If we consider the probability of exceeding the conventional metal content in hair by 1.5 times, then the descending series of values is built as follows: AAPCS 11, (0.267) > AAPCS 13 (0.2) > AAPCS 14 (0.133) > AAPCS 15 (0.067) > AAPCS 12 (conditionally 0).

Table 2. Values of integral Hazard Ratio and their probabilities in different study areas*

Areas	HR values and their probabilities (P_{Re})				
	0.5	1	1.5	2	2.5
AAPCS 11	0.200	0.210	0.267	0.700	1.00
AAPCS 13	0.143	0.161	0.200	0.000	0.00
AAPCS 14	0.157	0.145	0.133	0.000	0.00
AAPCS 15	0.143	0.129	0.067	0.000	0.00
AAPCS 12	0.071	0.048	0.000	0.000	0.00

* Note: probabilities ≤0.0001 were evaluated as zero

Due to the fact that the content of impurities of pollutants in the environment, in particular, in the atmospheric air, directly affects the failure of the organism's adaptation, it is possible to group data series covering the probability of metals accumulation and impurities concentrations, with the corresponding metals concentrations in biosubstrates. Thus, during building our models we took into account the metals accumulation in the hair as a way of expressing the retention of metals in the body, generalized as the probability that the accumulation of at least one metal in the hair will go beyond the background value and indicated in our model as the retention probability (P_{Re}).

The main indicator widely used in assessing the risk to public health as a result of environmental pollution is the hazard coefficient (HQ), expressed as the ratio of the actual exposure concentration (or dose) of a chemical to its safe (reference) value [25]:

$$HQ = \frac{AC}{RfC}, \tag{6}$$

here AC – average concentration, mg/m^3; RfC – reference (safe) concentration, mg/m^3.

At the same time, we assume that if the calculated hazard coefficient (HQ) of a particular substance does not exceed one, then the probability of developing harmful effects in a person with daily intake of the substance at a given concentration (dose) is insignificant and such an effect is characterized as acceptable. In this study, the list of substances for priority evaluation was selected from the general list of impurities controlled at AAPCS in Nizhnekamsk, the concentrations of which exceed the reference values.

For each listed substance, hazard quotients (HQ) were calculated, which are the ratios of actual and reference concentrations for these substances. By adding the hazard coefficients of all impurities, for each calculated point, hazard indices (IHQ) were obtained.

The data were grouped by study areas (areas linked to AAPCS posts). The values of the obtained IHQ indices are presented in table 3.

Table 3. IHQ values for the totality of substances

Areas	IHQ
AAPCS 13	3.39
AAPCS 12	0.64
AAPCS 11	5.45
AAPCS 14	0.88
AAPCS 15	0.64

The analysis of received IHQ allows to exclude sites in the coverage area of AAPCS 12, AAPCS 14, AAPCS 15, focusing on the coverage areas of AAPCS 13 and AAPCS 11.

When we compared the values of the adverse event probability obtained by assessing the contamination of children's hair samples with metals and assessment of air pollution calculated for different parts of the city, we saw a general trend in their variability. This allows us to consider these indicators as part of the integral assessment of the aerogenic impact.

Based on the foregoing, we designed a hybrid intelligent model consisting of two cascaded neural networks. The first level neural network model calculates the probability P_{Re} depending on the content of pollutant impurities set in the air. The second-level neural network predicts the toxic metals content (strontium, lead and chromium) in the hair of children living in different territory parts.

Comparing the values of priority impurities concentrations with the metal's accumulation indicators in the hair, it is possible to form predictive values P_{Re}. Since the probability P_{Re} is directly related to disturbances in the content of metals in the body when air with pollutants is inhaled, it indirectly reflects the intake of metals with inhaled air.

The pollutants concentration values in the air are fed to the input of the first level neural network model. For second level neural network models, the probability values P_{Re} calculated by the first level neural network of the cascade and the average value of HQ over all impurities are used as inputs. The HQ was calculated for impurities that were fed to the input of the first level model. The output of the model is the content of metals in the hair of children.

The general operation scheme for cascade model is as follows [26–30]:

Structure of the first level neural network model for calculating the retention probability P_{Re} value depending on the content of priority pollutants in the atmospheric air:

– number of input layer neurons – corresponds to the number of priority substances (up to 10);

– number of output layer neurons – 1;

– number of hidden layers – 1;

– number of neurons in the hidden layer – 6;
– activation function of hidden layer neurons – hyperbolic tangent with slope 3;
– activation function of output layer neurons – linear with saturation with slope 3;
– Method of normalization – reduction to the interval [–1; +1].

At the second level, there are three independent neural network models for calculating the content of chromium, strontium and lead in the hair of children, respectively. The inputs for all three models are the P_{Re} values obtained at the previous level, as well as the HQ values calculated relative to the concentrations of pollutants in the atmosphere. Outputs – metals concentrations in children's hair.

Graphically, the structure of developed two-level cascade model is shown in fig. 1.

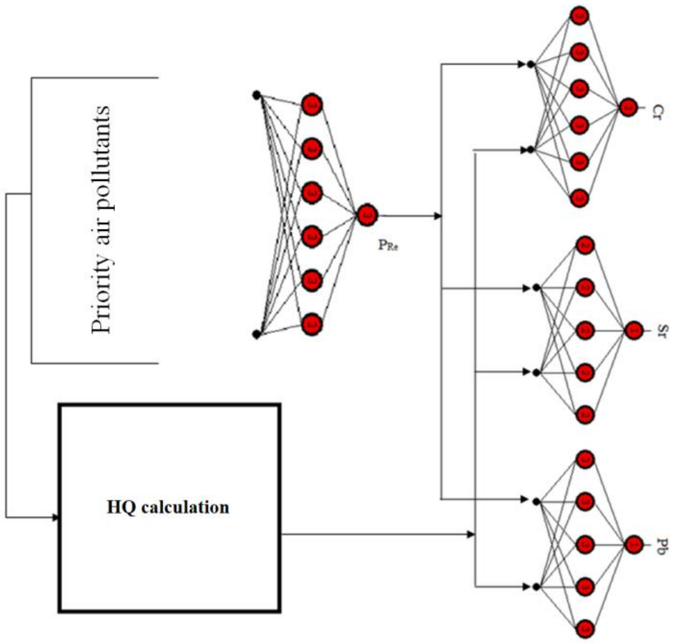

Fig. 1. Scheme of sequential neural network cascade for calculating the metals content in the hair of adolescent children

The individual neural network model’s errors for assessing the metals content in the hair of adolescent children seemed to be high only for the content of chromium (an error of more than 30 %), with respect to other metals, the error was 5–10 %.

Conclusions. As a result of this research, we have proposed a method for assessing the intake of metals into the body based on their content in bio substrates and taking into account the retention of metals in the body using a cascade hybrid intelligent system. This system does not require expensive laboratory tests and allows us to quantify the metals balance in the body. We have built a neural networks cascade that forms a closed circuit for metals content homeostatic assessment in the biological substrates of adolescent children, depending on the intake of priority pollutants with atmospheric air. The second level of the sequential neural network cascade for calculating the metals content in the hair of children and adolescents should be organized as separate predictive neural networks. The results of this study can be used, among other things, for patient-oriented diagnosis of environmentally conditioned microelementoses in individual locations.

Acknowledgments. Scientific research was carried out with the financial support of the Ministry of Education and Science of the Russian Federation as part of the fulfillment of obligations under the Agreement No. 075-03-2023-032 dated 16.01.2023 (topic number FZSU-2023-0005).

References

1. Qu, C. S. Human Exposure Pathways of Heavy Metals in a Lead-Zinc Mining Area / C. S. Qu, Z. W. Ma, J. Yang // PLoS One. – 2012. – No 7 (11).

2. The levels of calcium and magnesium, and of selected trace elements, in whole blood and scalp hair of children with growth retardation / H. Ozmen [et al] // Iran J Pediatr. – 2013. – Apr. 23 (2). – P. 125–130.

3. Li, J. Identification of human age using hair trace element concentration and support vector machine method / J. Li, W. Wang, F. Zhang // Biol. Trace Elem. Res. – 2011. – Dec. 143 (3). – P. 1441–1450.

4. Levels of Cd, Pb, As, Hg, and Se in hair / Y. Li [et al] // Bull. Environ. Contam. Toxicol. – 2012. – Jul 89 (1). – P. 125–128.

5. Lonnerdal, B. Trace element absorption in infants as a foundation to setting upper limits for trace elements in infant formulas / B. Lonnerdal // J. Nutr. – 1989. – Vol. 119, No 123. – P. 1839–1844.

6. Metal(loid) levels in biological matrices from human populations exposed to mining contamination-Panasqueira Mine (Portugal) / P. Coelho [et al] // J. Toxicol. Environ. Health. A. – 2012. – No 75 (13–15). – P. 893–908.

7. In search of decoding the syntax of the bioelements in human hair – A critical overview / B. Momčilović // J Trace Elem Med Biol. – 2018. – Dec. 50. – P. 543–553.

8. Reference Values for Chemical Elements Concentration in Hair of Adults in the Republic of Tatarstan / N. A. Agadzhanyan [et al] // Human Ecology. – 2016. – No 4. – P. 38–44.

9. Indicator ability of biosubstrances in moderate occupational exposure to toxic metals / A. R. Grabeklis [et al] // Journal of Trace Elements in Medicine and Biology. – 2011. – Vol. 25. – P. 41–44.

10. Bridges, C. C. Molecular and ionic mimicry and the transport of toxic metals / C. C. Bridges, R. K. Zalups // Toxicol Appl Pharmacol. – 2005. – May, 204 (3). – P. 274–308.

11. Reference Values for Chemical Elements Concentration in Hair of Adults in the Republic of Tatarstan / N. A. Agadzhanyan [et al] // Ekologiya cheloveka [Human Ecology]. – 2016. – № 4. – P. 38–44.

12. Concentration and reference interval of trace elements in human hair from students living in Palermo, Sicily (Italy) / G. Dongarra [et al] // Environ. Toxicol. – 2011. – No 32. – P. 27–34.

13. Clinical Toxicology / M. D. Ford [et al]. – W. B. Saunders Company, 2001.

14. Hanif, S. Statistical Evaluation of Trace Metals, TSH and T4 in Blood Serum of Thyroid Disease Patients in Comparison with Controls / S. Hanif, A. Ilyas, M. Shah // Biological Trace Element Research. – 2018. – May, 183 (1). – P. 58–70.

15. Levels of heavy metals in urine samples of school children from selected industrial and non-industrial areas in Dar es Salaam, Tanzania / A. M. John [et al] // African Health Sciences. – 2018. – Dec, 18 (4). – P. 1226–1235.

16. Klaassen, C. D. Casarett and Doull's Essentials of Toxicology / C. D. Klaassen, J. B. Watkins. – MC Graw-Hill, 2004.
17. The Relationship between Dietary, Serum and Hair Levels of Minerals (Fe, Zn, Cu) and Glucose Metabolism Indices in Obese Type 2 Diabetic Patients / E. Kryl [et al] // Biological Trace Element Research. – 2019. – May, 89 (1). – P. 34–44.
18. Assessment of the trace element blood condition in children with congenital deformities of the thoracic and lumbar spine (Preliminary report) / T. V. Lobachevskaya [et al] // Pediatric Traumatology, Orthopaedics and Reconstructive Surgery. – 2019. – Vol. 7. – No. 1. – P. 51–56.
19. Blaurock-Busch, E. Heavy metals and trace elements in hair and urine of a sample of arab children with autistic spectrum disorder / E. Blaurock-Busch, O. Amin, T. Rabah // Maedica (Buchar). – 2011. – Oct, 6 (4). – P. 47–57.
20. Michalak, I. Exposure to nickel by hair mineral analysis / I. Michalak, M. Mikulewicz, K. Chojnacka // Environ Toxicol Pharmacol. – 2012. – Nov. 34 (3). – P. 27–34.
21. Vanaelst, B. Mineral concentrations in hair of Belgian elementary school girls: reference values and relationship with food consumption frequencies / B. Vanaelst, I. Huybrechts, N. Michels // Biol Trace Elem Res. – 2012. – Dec. 150 (1–3). – P. 56–67.
22. Judgment under Uncertainty: Heuristics and Biases / D. Kahneman [et al]. – Cambridge University Press, 2005. – 555 p.
23. Buhrmester, V. Analysis of explainers of black box deep neural networks for computer vision: a survey / V. Buhrmester, D. Münch, M. Arens // Machine Learning and Knowledge Extraction. – 2021. – Vol. 3, No. 4. – P. 966–989.
24. Finding the optimal topology of an approximating neural network / K. Yotov [et al] // Mathematics. – 2023. – Vol. 11, No. 217. – P. 1–18.
25. Nonlinear Poisson regression using neural networks: a simulation study / N. Fallah [et al] // Neural Computing and Applications. – 2009. – Vol. 18, No. 8. – P. 939–943.
26. Method for determining regional reference values of metal content in biological substrates and their intake into the body via drinking water / Y. Tunakova [et al] // International Journal of Environmental Research and Public Health. – 2021. – No 18. – P. 1–9.
27. Modeling of processes of formation of the microelement status of an organism of children taking in to account the waterway of receipt / Yu. A. Tunakova [et al] // Messenger Kazan. Tekhnolog. un. – 2014. – Vol. 17, No. 16. – P. 145–147.
28. P 2.1.10.1920-04 Human Health Risk Assessment from Environmental Chemicals. – Moscow : Federal Center for State Sanitary and Epidemiological Surveillance of the Russian Ministry of Health, 2004. – 143 p.
29. Novikova, S. V. Use of neural network technologies for the forecast of high level of pollution of atmospheric air in the cities / S. V. Novikova, Yu. A. Tunakov // Health and safety. – 2011. – No. 1. – P. 21–28.