CARDIOMETABOLIC RISK PREDICTION IN PATIENTS WITH NON-ALCOHOLIC FATTY LIVER DISEASE COMBINED WITH SUBCLINICAL HYPOTHYROIDISM

Olena Kolesnikova1, Olena Vysotska2, Anna Potapenko1, Anastasia Radchenko1, Anna Trunova2, Nataliya Virstyuk3, Liudmyla Vasylevska-Skupa4, Aliya Kalizhanova5, Nazerka Mukanova6

1Government Institute “L. T. Malaya Therapy Institute of the National Academy of Medical Science of Ukraine”, Kharkiv, Ukraine. 2National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine. 3Ivano-Frankivsk National Medical University, Ivano-Frankivsk, Ukraine. 4Institute of Information and Computational Technologies, Almaty, Kazakhstan. 5Gymnasium No. 159 named after Y. Abylsarin, Almaty, Kazakhstan.

Abstract. One of the most observed diseases of our time is non-alcoholic fatty liver disease (NAFLD). Recently published research results indicate that patients with NAFLD along with traditional risk factors for cardiovascular diseases (CVD) have “new” risk factors such as endothelial dysfunction (ED), carotid intima-media thickness (CIMT), an increase in the CRP level, as well as risk factors combined into the Framingham scale. It is also known that combination of NAFLD with subclinical hypothyroidism (SH) forms an abnormal metabolic phenotype, which is associated with cardiometabolic risk factors. The study of cardiometabolic predictors and vascular markers in patients with NAFLD in combination with SH will provide an opportunity to improve the strategy of cardiovascular events prevention in such comorbid patients.

Keywords: cardiometabolic risk, non-alcoholic fatty liver disease, subclinical hypothyroidism, prediction, binary regression logistic analysis, validation of prognostic models

PRZEWIDYWANIE RYZYKA KARDIOMETABOLICZNEGO U PACJENTÓW Z NIEALKOHOLOWĄ STŁUSZCZENIOWĄ CHORBĄ WĄTROBY W POŁACZENIU Z SUBKLINICZNĄ NIEDOCZYNNOŚCIĄ TARCZYCY

Streszczenie. Niewłaściwe stłuszczenie wątroby (NAFLD) jest jedną z najczęstszych chorób naszych czasów. Ostatnio opublikowane wyniki badań sugerują, że pacjenci z NAFLD, wraz z tradycyjnymi czynnikami ryzyka sercowo-naczyniowych (CVD), mają "nowe" czynniki ryzyka, takie jak dysfunkcja śródbłonka (ED), bruzda błony wewnętrznej i środowej tętnicy szyjnej (CIMT), podwyższony poziom CRP i czynniki ryzyka połączone w skali Framingham. Wiadomo również, że połączenie NAFLD z subkliniczną niedoczynnością tarczy (SH) tworzy nieprawidłowy fenotyp metaboliczny związany z czynnikami ryzyka kardiometabolicznego. Badanie predyktorów kardiometabolicznych i markerów naczyniowych u pacjentów z NAFLD w połączeniu z SH pozwoli na lepsze zapoznanie zdarzeniom sercowo-naczyniowym w takich współistniejących pacjentów.

Słowa kluczowe: ryzyko kardiometaboliczne, niewłaściwe stłuszczenie wątroby, subkliniczna niedoczynność tarczy, rokowanie, binarna analiza regresji logicznej, waliadacja modeli prognozystycznych

Introduction

In recent years, an increasing number of cardiovascular pathologies worldwide makes it necessary to search for new risk factors that could serve as an indicator and determine the likely prognosis of patients with somatic diseases. Currently, the concept of “cardiometabolic risk” is increasingly used in clinical practice. It involves the risk of developing cardiovascular diseases and/or type 2 diabetes, combining both classical (“traditional”) risk factors such as smoking, high levels of low-density lipoprotein cholesterol (LDL-C), arterial hypertension, an increase in glucose level, and factors directly associated with abdomen (especially visceral) obesity, insulin resistance (IR), low high-density lipoprotein cholesterol (HDL-C), hypertension, and an increase in proinflammatory markers [8]. The attention of researchers is attracted by the search for early predictors of cardiometabolic changes and vascular markers – markers of inflammation (C-reactive protein, CRP), dyslipidemia, IR, tumor necrosis factor – α (TNF-α), endothelial dysfunction (ED) indicators such as circulating desquamated endothelial cells (CDECs), vascular endothelial growth factor (VEGF), that could act as a "predictive" [1, 7].

One of the most common diseases of our time is non-alcoholic fatty liver disease (NAFLD) – a chronic liver disease that is considered as a component of some diseases associated with IR, such as metabolic syndrome (MS), obesity, and diabetes. The development of NAFLD is very closely associated not only with abdominal obesity, but also has a significant effect on the formation of cardiometabolic risk factors such as hypertriglyceridemia, a decrease in HDL-C level, hypertension, hyperglycemia, thereby increasing the degree of cardiometabolic risk itself and affecting the prognosis and life expectancy of patients [8, 18].

Recently published research results indicate that patients with NAFLD along with traditional risk factors for cardiovascular diseases (CVD) (obesity, diabetes, MS and others) have "new" risk factors such as endothelial dysfunction (ED), carotid intima-media thickness (CIMT), an increase in the CRP level, as well as risk factors combined into the Framingham scale. It is also known that combination of NAFLD with subclinical hypothyroidism (SH) forms an abnormal metabolic phenotype, which is associated with cardiometabolic risk factors. The study of cardiometabolic predictors and vascular markers in patients with NAFLD in combination with SH will provide an opportunity to improve the strategy of cardiovascular events prevention in such comorbid patients.
popular in Europe, is limited to the age of 45–64 years, while the age range in the PCE scale was much wider – 20–79 years old, but it is recommended to use it in the range of 40–79 years [3, 19].

Solving the problems of medical forecasting often involves the use of regression models. For example, Weibull regression was used in the SCORE scale, Cox regression was used in the second version of the GRACE scale, and the logistic regression was used in the CHAD2VS2–VASc scale [2, 20]. One of the principal limitations of the scales is the impossibility of an individual assessment, since scales can only give a probabilistic risk assessment for a group of patients with given levels of risk factors.

Preventive treatment of CVD is largely based on model creation methods for assessing the "absolute risk" of future cardiovascular events in order to substantiate decisions about therapeutic approaches [5].


Jasnickij L. and Cherepanov F. [17] showed in their article the possibility of creating neuroexpert medical systems, allowing to perform long-term prognosis of the diseases development to predict the occurrence of new diseases in future periods of the patient's life.

The introduction of machine learning allows increasing the accuracy of prediction of cardiovascular events [15]. However, the use of a large array of heterogeneous data instead of selected patient cohort can reduce the accuracy of risk classification and the proximity of the forecast to the frequency of real events [4, 9, 13, 16].

The subject of the present work is creating a model for predicting cardiometabolic risk in patients with NAFLD in combination with SH. The paper presents the results of the GI "L.T. Malaya Therapy Institute of the National Academy of Medical Science of Ukraine" scientific research on the topic: "Influence of factors of cardiovascular development on premature aging" (security number 0117U003031).

1. Materials and methods

71 patients were selected to create a prediction model for cardiometabolic risk in patients with NAFLD in combination with SH. All patients were divided into three groups as follows:

- Group 1 – low risk patients (6 patients);
- Group 2 – patients with moderate risk (43 patients);
- Group 3 – high-risk patients (22 patients).

The following factors were used as potential predictors: age, body mass, waist circumference, thigh circumference, CRP, CDECs, CIMT, VEGF level; telomeres in the blood; telomeres in the buccal epithelium, total cholesterol (TC), triglycerides (TG), VLDL-C, HDL-C, atherogenic coefficient, ALT, aspartate aminotransferase (AST), GGTP, alkaline phosphatase (ALP), HbA1C, TSH, free tyroxine (free T4), free triiodothyronine (free T3), anti-thyroid peroxidase antibodies (TPOAb). All indicators were encoded and placed according to the 27-dimensional vector, which takes into account the absence, which displays the value of each metric.

Mathematical processing of the results was carried out using the SPSS19 application package for Windows.

2. Results of binary regression logistics analysis

49 patients of the first and second groups were analyzed to construct an equation of logistic regression, which determines the probability of having a moderate cardiovascular risk (CVR), taking into account the considered indicators.

Comparison of groups 1 and 2 showed that among all the evaluated factors, the length of the telomeres in the buccal epithelium, the levels of CRP, CDECs had the statistically significant effect on the probability of development of moderate cardiometabolic risk in patients with NAFLD in combination with SH. As a result, the regression function was constructed. It included 3 indicators (table 1).

The binary logistic function selected from the training sample looks like:

$$\hat{P} = \left[1 + e^{-(11.024X_1 + 0.533X_2 + 1.258X_3 - 25.352)}\right]^{-1}$$

where \(\hat{P}\) is the probability of developing moderate CVR in patients with NAFLD in combination with SH.

Table 1. Factors of moderate cardiometabolic risk

<table>
<thead>
<tr>
<th>Code</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_1</td>
<td>Telomere length in the buccal epithelium</td>
</tr>
<tr>
<td>X_2</td>
<td>CRP</td>
</tr>
<tr>
<td>X_3</td>
<td>CDECs</td>
</tr>
</tbody>
</table>

The calculated coefficients of the regression function and the results of checking their significance are presented in table 2. All variables are significant (p < 0.05) and selected correctly according to Wald's statistics.

Table 2. Regression function coefficients

<table>
<thead>
<tr>
<th>Indicators (X_i)</th>
<th>Coefficients (b_i)</th>
<th>Standard errors (S_e)</th>
<th>Wald's criterion (W_0)</th>
<th>Significance (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_1</td>
<td>11.024</td>
<td>5.487</td>
<td>4.037</td>
<td>0.045</td>
</tr>
<tr>
<td>X_2</td>
<td>0.533</td>
<td>0.290</td>
<td>3.384</td>
<td>0.046</td>
</tr>
<tr>
<td>X_3</td>
<td>1.256</td>
<td>0.662</td>
<td>3.599</td>
<td>0.048</td>
</tr>
<tr>
<td>Constant</td>
<td>-25.352</td>
<td>11.348</td>
<td>4.991</td>
<td>0.025</td>
</tr>
</tbody>
</table>

The quality of the regression model approximation was estimated using the similarity function. In our study, \(G = 11.612\) at \(p = 0.001\), that is as a whole the informative indications isolated as independent variables have a significant effect on the development of cardiometabolic risk in patients with NAFLD in combination with SH. The Cox and Schell, and Nagelkerke indices indicate that the dispersion part explained by the developed logistic model stands at 75.8% (table 3).

Table 3. Characteristics of the model of binary logistic regression, created to determine the probability of developing moderate cardiometabolic risk in patients with NAFLD in combination with SH

<table>
<thead>
<tr>
<th>2Log Credibility (G)</th>
<th>R² of Cox and Shell</th>
<th>R² of Nagelkerke</th>
<th>(\chi^2)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.612</td>
<td>0.397</td>
<td>0.758</td>
<td>24.822</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The figure 1 shows the classification diagram. The distribution point was \(P = 0.5\). The closer the value of the predicted probability is to unit, the higher the probability of developing moderate cardiometabolic risk in patients with NAFLD in combination with SH is.
erroneously assigned to a group of low-risk patients. In general 47 cases out of 49 were correctly identified, representing 95.9%.

A general assessment of the agreement between the influence of risk factors found in the model and the actual occurrence of an adverse outcome was carried out using the Hosmer-Lemeshow (HL) test.

The resulting low value of HL = 4.391 at a significance level of p > 0.05 (p = 0.820), indicates the minimum differences between observed and predicted frequencies and the high quality of the selected regression model.

Table 4. Classification results of the model of binary logistic regression, created to determine the probability of moderate CVR in patients with NAFLD in combination with SH

<table>
<thead>
<tr>
<th>Observed groups</th>
<th>Presumed groups</th>
<th>% correctly predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Risk</td>
<td>3</td>
<td>42</td>
</tr>
<tr>
<td>Total percentage</td>
<td>95.9±2.83</td>
<td></td>
</tr>
</tbody>
</table>

The ROC analysis of the obtained model (figure 2) confirmed its effectiveness and excellent predictive quality. The area under curve (AUC) was 0.977 at p = 0.001.

Fig. 2. ROC Curve

Thus, a mathematical model has been developed that includes metabolic risk factors such as the length of telomeres in the buccal epithelium, CRP and CDECs, to determine effectively and qualitatively the probability of forming moderate cardiometabolic risk in patients with NAFLD in combination with SH.

Data from 65 patients in the second and third groups were analyzed for the purpose of constructing a prediction model of cardiometabolic risk of higher gradations in patients with NAFLD in combination with SH. As a result of the binary logistic regression analysis, an equation has been obtained, which determines the likelihood of developing high CVR in patients with NAFLD in combination with SH:

\[
\hat{p} = \left[1 + e^{-\left(4.366X_1 + 0.060X_2 + 0.0009X_3 + 0.871X_4 - 19.391\right)}\right]^{-1}
\]  

(2)

where \(\hat{p}\) is the probability of developing high CVR in patients with NAFLD in combination with SH;  
X₁ – telomere length in the blood;  
X₂ – LDL-C;  
X₃ – VEGF;  
X₄ – TSH.

Comparison of 2nd and 3rd groups therefore showed that of all the factors evaluated, telomere length in blood, LDL-C, VEGF and TSH had a statistically significant effect on the likelihood of high cardiometabolic risk developing in patients with NAFLD in combination with SH.

All variables are significant (p < 0.05) and selected correctly according to Wald’s statistics (table 5).

Table 5. Regression function coefficients

<table>
<thead>
<tr>
<th>Indicators (Xᵢ)</th>
<th>Coefficients (bᵢ)</th>
<th>Standard errors (Sᵢ)</th>
<th>Wald’s criterion (Wᵢ)</th>
<th>Significance (pᵢ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>4.366</td>
<td>1.738</td>
<td>6.311</td>
<td>0.012</td>
</tr>
<tr>
<td>X₂</td>
<td>0.860</td>
<td>0.437</td>
<td>3.872</td>
<td>0.049</td>
</tr>
<tr>
<td>X₃</td>
<td>0.009</td>
<td>0.004</td>
<td>6.661</td>
<td>0.010</td>
</tr>
<tr>
<td>X₄</td>
<td>0.871</td>
<td>0.361</td>
<td>5.820</td>
<td>0.016</td>
</tr>
<tr>
<td>Constant</td>
<td>-19.391</td>
<td>4.787</td>
<td>16.407</td>
<td>0.001</td>
</tr>
</tbody>
</table>

In the study, G = 39.297 at p = 0.001 shows that independent variables have a significant contribution to predict the dependent variable in general. The part of the dispersion, explained by logistic regression according to R² of Cox and Schell stands at 68% (table 6).

Table 6. Characteristics of the model of binary logistic regression, created to determine the probability of developing high cardiometabolic risk in patients with NAFLD in combination with SH

<table>
<thead>
<tr>
<th>Indicator</th>
<th>R² of Cox and Schell</th>
<th>R² of Nagelkerke</th>
<th>X²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAFLD</td>
<td>0.680</td>
<td>43.905</td>
<td>0.001</td>
<td>39.297</td>
</tr>
<tr>
<td>SH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A chart of classification of patients in the second and third study groups is presented in fig. 3 with the help of the developed mathematical model for predicting high cardiometabolic risk in patients with NAFLD in combination with SH.

Fig. 3. Chart of classification. Symbols: 2 – moderate risk patients; 3 – high-risk patients

The value of P = 0.5 served as the limiting point of distribution. The closer the value of the predicted probability is to unity, the higher the degree of cardiometabolic risk in patients with NAFLD in combination with SH is. The classification results presented in table 7 indicate that in general 59 cases were correctly identified out of 65 that is 90.8%.

Table 7. Classification results of the model of binary logistic regression, created for the determination of the probability of development of high CVR in patients with NAFLD in combination with SH

<table>
<thead>
<tr>
<th>Observed groups</th>
<th>Presumed groups</th>
<th>% correctly predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Risk</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td>Total percentage</td>
<td>95.9±2.83</td>
<td></td>
</tr>
</tbody>
</table>

The resulting low value of HL = 10.905 at a level of significance p = 0.143 indicates the high quality of the selected model. The ROC analysis of the resulting model (figure 4) confirmed the excellent quality of the developed model: the AUC value was 0.919 (p = 0.001).
The outlook was unmistakably accurate for 24 patients surveyed according to the results. Thus, the forecast was confirmed in 92.0% of cases.

That means, that the application of synthesized logistic models will allow predicting the development of cardiometabolic risk, which will ensure early diagnosis and appointment of treatment and prevention measures in order to avoid the development of cardiometabolic complications of high gradations.

4. Conclusions

The developed mathematical models allow predicting the degree of cardiometabolic risk in patients with NAFLD in combination with SH based on the of cumulative changes in lipid, carbohydrate states, thyroid function compensation, vascular factors and telomere lengths as a marker for vascular aging. The application of the proposed prognostic models in clinical practice allows to achieve an improvement in the quality of the cardiometabolic risk determination in patients with NAFLD in combination with SH due to the identification of factors that influence the progression of cardiometabolic changes in patients with NAFLD precisely in combination with SH, and the improvement of the applied mathematical apparatus, which in turn will help the physician to prescribe adequate treatment and preventive measures and improve the quality of medical services provision for patients with NAFLD in combination with SH.

References

[12] Ludwig U., Holzner D., Denzer C. et al.: Subclinical and clinical hypothyroidism and non-alcoholic fatty liver disease: a cross-sectional study of a random population sample aged 18 to 65 years. BMC Endocr Disord. 15, 2015, 41.
Prof. Olena Kolesnikova
e-mail: kolesnikova1973@gmail.com
Deputy director for Scientific work, head of the Department for the study of aging processes and prevention of metabolic-associated diseases.  
Government Institution "L.T. Malaya Therapy Institute of the National Academy of Medical Science of Ukraine".  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0001-5606-6621

Prof. Olena Vysotska
e-mail: o.vysotska@khai.edu
Head of the Department of Radio-Electronic and Biomedical Computerized Means and Technologies. National Aerospace University "Kharkiv Aviation Institute".  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0003-3723-9771

Ph.D. Anna Potapenko
e-mail: annav1611sh@gmail.com
Researcher at the Department for the study of aging processes and prevention of metabolic-associated diseases, Endocrinologist of the second qualification category.  
Government Institution "L.T. Malaya Therapy Institute of the National Academy of Medical Science of Ukraine".  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0002-1658-0165

Anastasia Radchenko
e-mail: anastasha.radchenko@gmail.com
Postgraduate student of the department of studying aging processes and prevention of metabolic-associated diseases.  
Government Institution "L.T. Malaya Therapy Institute of the National Academy of Medical Science of Ukraine".

http://orcid.org/0000-0002-9687-8218

Ph.D. Anna I. Trunova
e-mail: a.pcherska@khai.edu
Associate professor of the Department of Radio-Electronic and Biomedical Computerized Means and Technologies National Aerospace University "Kharkiv Aviation Institute".  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0001-7069-0674

Ph.D. Nataliya Virystyuk
e-mail: ndermven@ukr.net
Head of the Department of Pharmacology and Internal Medicine No 3 named after M.M. Berechnytskii Ivano-Frankivsk National Medical University.  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0002-5794-8754

Ph.D. Aliya Kalizhanova
e-mail: kalizhanova_aliya@mail.ru
Candidate of physical and mathematical sciences, professor, University of Power Engineering and Telecommunications, the chief researcher of the Institute of Information and Computational Technologies of the Ministry of Education and Science CS of the Republic of Kazakhstan.  
Scientific interests: mathematical modeling of systems, models of transport systems network analysis, optimization methods, technologies for developing sensor systems for signals receive – transmit, mathematical modeling of Bragg fiber gratings.

http://orcid.org/0000-0002-5979-9756

M.Sc. Nazerka Mukanova
e-mail: mukanova0124@gmail.com
Master of Pedagogical Sciences, teacher of the highest category of computer science of gymnasium No. 159 named after Y. Alynysarin, a prize-winner of several national and international competitions, her scientific articles have been published in several domestic and international publications, and she is a participant in various international teaching-methodical pedagogical conferences.  
Scientific interest: information technologies in medicine.

http://orcid.org/0000-0002-7945-7187