

Predictors of genetic polymorphisms in patients with arterial hypertension and ischemic heart disease

Svitlana PIDRUCHNA¹, Nadiya YAREMA², Iryna KUZMAK¹, Olena PROKOPOVYCH³, Oksana KOTSIUBA⁴, Andrii SVERSTUK⁵, Oksana BAHRII-ZAIATS⁶, Petro LYKHATSKYI¹, Alla MUDRA¹, Lylya PALYTSYA¹, Nataliya LETNIAK¹, Oksana OSTRIVKA¹, Tetyana YAROSHENKO¹, Nadija VASYLYSHYN¹

¹Department of Medical Biochemistry, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine; ²Department of Internal Medicine №1, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine; ³Emergency Medical Care, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine; ⁴Therapeutics and Family Medicine, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine; ⁵Department of Medical Informatics, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine; ⁶Department of Medical Physics of Diagnostic and Therapeutic Equipment, I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine
E-mail: pidruchna@tdmu.edu.ua

Objective. This study was aimed to develop a mathematical model for predicting the presence of pathological alleles of the angiotensin II receptor type 1 (AGTR1) (C1166C) and endothelial nitric oxide synthase (eNOS) (T786C) genes in patients with arterial hypertension using readily available clinical and demographic parameters without the need for laboratory genetic analysis.

Methods. The study included 86 patients with arterial hypertension, 45 of whom had also clinical signs of ischemic heart disease and 30 healthy controls. Clinical and demographic data, including age, sex, body mass index (BMI), smoking and alcohol habits, and systolic and diastolic blood pressure, were collected. A predictive model of genetic polymorphism was constructed using a multivariate regression analysis in Statistica 10.0. Model quality was assessed via residual analysis (histogram, normal probability plot, scatter plot), ANOVA, and the Nagelkerke coefficient of determination (R^2).

Results. Eight significant predictors of pathological gene variants were identified: age, sex, disease duration, BMI, smoking, alcohol consumption, systolic blood pressure, and diastolic blood pressure. The regression model explained 98.42% of the variance in gene polymorphism ($R^2=0.9842$). Residuals were normally distributed and randomly scattered confirming the model adequacy. ANOVA analysis demonstrated high statistical significance ($p<0.001$) indicating that the model performed significantly better than predictions based on the mean values. The model allows inference of the probable presence of C alleles in AGTR1 and eNOS genes using simple clinical information.

Conclusions. The proposed mathematical model demonstrates high predictive capability and statistical stability providing a cost-effective tool for preliminary genetic screening in patients with hypertension. It enables identification of individuals at risk for carrying pathological gene variants and supports the implementation of a personalized approach in clinical practice. Further validation on independent datasets is recommended.

Keywords: arterial hypertension, AGTR1 (C1166C), eNOS (T786C), gene polymorphism, predictive model, multivariate regression

Arterial hypertension (AH) is one of the leading yet modifiable risk factors for the development of cardiovascular diseases including ischemic heart disease, myocardial infarction, stroke, heart failure, peripheral artery disease, and atrial fibrillation (Williams et al. 2018). Arterial hypertension has a complex etiology that includes both genetic predisposition and environmental factors (Choi et al. 2020; Vaura et al. 2021; Abu-El-Haija et al. 2023; Pidruchna et al. 2024a).

In recent years, increasing attention has been paid to personalized medicine, which involves considering a patient's genetic characteristics when selecting diagnostic and therapeutic strategies. A promising direction in this context is the study of gene polymorphisms that may influence the development of cardiovascular, metabolic, and other chronic diseases. In particular, structural variants of genetic polymorphism in the angiotensin II receptor type 1 (AGTR1) and endothelial nitric oxide synthase (eNOS) genes have been associated with an increased risk of hypertension, atherosclerosis, and endothelial dysfunction (Wang and Wang 2019).

The identified genetic variants of eNOS (T786C) are critical markers not only for hypertension, but also for coronary artery disease, as they impair nitric oxide (NO) bioavailability leading to accelerated atherosclerosis and myocardial ischemia.

However, genetic testing remains costly and is not widely accessible in routine clinical practice. Therefore, the search for alternative methods of predicting the presence of pathological alleles is particularly relevant, specifically through the mathematical modeling based on readily available clinical

and demographic indicators (Yarema et al. 2020; Pidruchna et al. 2024b; Shmanko et al. 2024).

The aim of this study was to develop a mathematical model that allows the prediction of the probability of developing arterial hypertension based on clinical and demographic factors [age, sex, body mass index (BMI), habits, blood pressure] in the presence of a specific allele of the AGTR1 gene (e.g., C1166C) or eNOS gene (e.g., T786C) without the need for laboratory genetic analysis.

This approach is particularly valuable for patients with ischemic heart disease, as predicting the AGTR1 (C1166C) polymorphism is essential due to its association with increased sensitivity to angiotensin II, potentially worsening ventricular remodeling and cardiac workload.

Material and Methods

The predictive model of gene polymorphism was developed using multiple regression analysis. Statistical processing of the obtained results was performed using the Statistica 10.0 software package and Microsoft Excel 2007. Normality of distribution of variables was assessed using skewness and kurtosis coefficients, as well as the Kolmogorov–Smirnov and Shapiro–Wilk tests. Differences between the compared groups were considered significant at $p < 0.05$.

Results

The multifactorial regression analysis was implemented by Statistica 10.0 software to identify

Table 1
Results of identifying statistically significant factors for predicting genetic polymorphism obtained via multivariate regression analysis in Statistica 10.0

Regression Summary for Dependent Variable: Pred_B_ (1 in Model) R=0.99207405; R ² =0.98421092; Adjusted R ² =0.98303043; F(8, 107)=833.73; p<0.0000; SE of estimate: 0.83164						
N=116	b*	S.E. of b*	b	S.E. of b	t (107)	p-value
Intercept			-22.0367	0.822671	-26.7867	0.000000
Age	0.169319	0.019368	0.0624	0.007142	8.7424	0.000000
Gender	0.086722	0.017173	1.1160	0.221004	5.0498	0.000002
Duration of disease	0.116916	0.022301	0.0779	0.014865	5.2427	0.000001
BMI	0.118152	0.015267	0.1298	0.016775	7.7392	0.000000
Smoking	0.303219	0.024727	4.4018	0.358962	12.2625	0.000000
Alcohol abuse	0.250325	0.017042	4.8845	0.332532	14.6888	0.000000
SBP	0.206426	0.021705	0.0591	0.006214	9.5106	0.000000
DBP	0.203590	0.018654	0.0837	0.007672	10.9141	0.000000

Abbreviations: BMI – body mass index; DBP – diastolic blood pressure; SBP – systolic blood pressure; SE – Standard Error.

significant clinical and demographic predictors of the presence of a pathological gene variant, specifically the C allele in the AGTR1 and eNOS genes. The regression analysis method was chosen as the mathematical model allowing prediction of the probability of gene polymorphism by analyzing regression coefficients and the values of risk factors with significant influence. This approach enables the identification of relationships between risk factors and gene polymorphism.

The model lies at the intersection of genetics, clinical medicine, and machine learning. It is designed to predict the polymorphism of the AGTR1 and eNOS genes or at least the presence of specific alleles using minimal patient information, such as age, sex, BMI, presence of harmful habits, and systolic and diastolic blood pressure (parameters that can be obtained directly during a clinical examination). This allows the clinician to infer the potential presence of a pathological “C” allele of each gene without costly laboratory testing.

For the construction of the predictive model, probable factors influencing gene polymorphism were selected. Linear regression analysis identified eight of the most significant risk factors for predicting genetic polymorphism: X1 – age, X2 – sex, X3 – duration of disease, X4 – BMI, X5 – smoking, X6 – alcohol consumption, X7 – systolic blood pressure (SBP, mmHg), and X8 – diastolic blood pressure (DBP, mmHg).

Based on the obtained results presented in Table 1, a mathematical model was developed to predict genetic polymorphism. The formula for determining the gene polymorphism prediction coefficient (GPPC) is:

$$\text{GPPC} = 0.0624 \cdot X_1 + 1.1160 \cdot X_2 + 0.0779 \cdot X_3 + 0.1298 \cdot X_4 + 4.4018 \cdot X_5 + 4.8845 \cdot X_6 + 0.0591 \cdot X_7 + 0.0837 \cdot X_8 - 22.0367$$

To assess the quality of the regression model, residual deviations were analyzed. As shown in the histogram (Figure 1), the residuals are symmetrically distributed and closely approximate a normal distribution curve confirming the model's adequacy for predicting gene polymorphism.

To further confirm the conformity of the residuals to the normal distribution, a normal probability plot was constructed (Figure 2). Analysis of this plot reveals the absence of systematic deviations from the normal probability line. This allows to conclude that the residuals are normally distributed.

To examine the dependence of residuals on the predicted values, a scatter plot was constructed (Figure 3). Based on the obtained results, the residuals

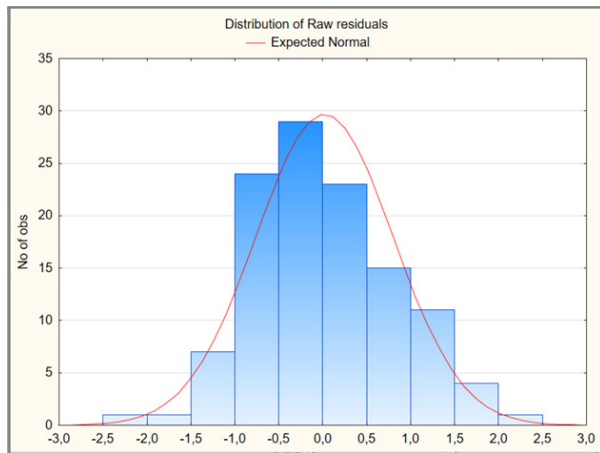


Figure 1. Histogram of residuals of the multivariate regression model for predicting genetic polymorphism.

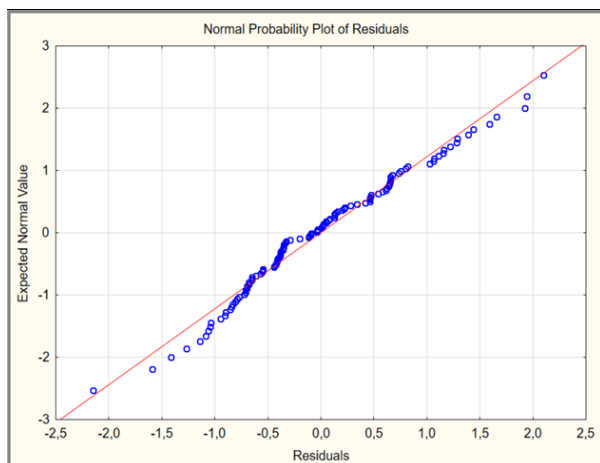


Figure 2. Normal probability plot of residuals of the multivariate regression model for predicting genetic polymorphism.

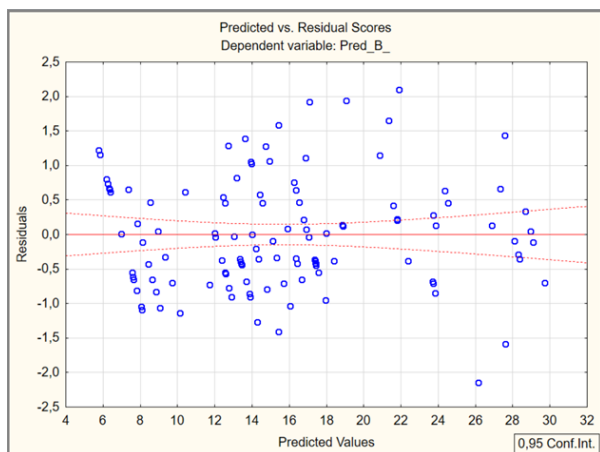


Figure 3. Scatter plot of residuals of the multivariate regression model for predicting genetic polymorphism.

relative to the predicted values are randomly scattered indicating the absence of dependence on the predicted genetic polymorphism values. The histogram and the normal probability plot confirm the conformity of the residuals to a normal distribution. Therefore, the developed genetic polymorphism prediction model is of high quality and is considered adequate.

The next step involved assessing the overall adequacy of the model, for which an ANOVA analysis was performed (Table 2). Analysis of the obtained data indicates a high level of overall adequacy of the gene structure prediction model, as confirmed by the ANOVA results, since the significance level is $p < 0.001$. This demonstrates that the proposed model performs significantly better than a simple forecast based on mean values.

To further assess the quality of the mathematical model for gene structure prediction, the Nagelkerke coefficient of determination (R^2) was analyzed, which indicates the proportion of factors accounted for in the prediction. It is considered a universal measure of the relationship between one random variable and others. The coefficient of determination ranges from 0 to 1 with values approaching 1 indicating a higher quality of the multivariate regression model. In the proposed mathematical model for gene structure prediction, the coefficient of determination is $R^2 = 0.9842$ (in Statistica 10.0, $R^2 = 0.9842$; Table 1). Thus, in our case, 98.42% of the factors are accounted for in the gene structure prediction model. The coefficient of determination indicates the extent, to which the observed data support the mathematical model.

Discussion

In this study, a multivariate regression analysis method was implemented to identify clinical and demographic predictors influencing the presence of pathological variants of genetic polymorphism, particularly the C alleles in the AGTR1 and eNOS genes. The developed model allows prediction of genetic polymorphisms using data that can be easily

collected during a standard medical examination.

Significant predictors included age, sex, disease duration, BMI, the presence of harmful habits (smoking, alcohol consumption) as well as systolic and diastolic blood pressure levels. This is consistent with the literature, which highlights the relationship between these factors and genetic predisposition to cardiovascular diseases (Warren et al. 2017; Padmanabhan et al. 2024).

Polymorphism of the NOS3 (eNOS) and AGTR1 genes is associated with early development and complicated course of cardiovascular pathology. The combination of these polymorphisms in patients with high BMI increases the risk of complications in hypertension (Potaskalova et al. 2022).

The high predictive value of our model is particularly relevant for understanding the endocrine mechanisms underlying hypertension. The AGTR1 (C1166C) polymorphism is intrinsically linked to the hyperactivation of the renin-angiotensin-aldosterone system (RAAS). This genetic variant leads to an increased sensitivity of type 1 angiotensin II receptors promoting chronic vasoconstriction and adverse cardiovascular remodeling. Furthermore, our findings emphasize the synergistic impact of metabolic factors specifically BMI on endothelial function. In patients carrying the eNOS (T786C) variant, reduced NO bioavailability is exacerbated by the pro-inflammatory state associated with high BMI. This metabolic-genetic interaction significantly accelerates endothelial dysfunction and increases the risk of complications in hypertensive patients.

The constructed model demonstrated high predictive capability, as evidenced by the coefficient of determination $R^2 = 0.9842$ indicating that nearly 98.5% of the variation in gene polymorphism can be explained by the included factors. This is an exceptionally high level of model quality for biomedical research, where the complexity of interactions among variables typically makes achieving such accuracy challenging.

Analysis of residuals (histogram, normal probability plot, scatter plot) indicates a normal

Table 2
ANOVA results for assessing the statistical significance and overall adequacy of the multivariate regression model

Analysis of Variance; DV: Pred_B_ (1 inModel)					
Effect	Sums of Squares	df	Mean Squares	F	p-value
Regression	4613.056	8	576.6320	833.7297	0.00
Residual	74.004	107	0.6916		
Total	4687.060				

distribution and the absence of systematic errors confirming that the developed gene polymorphism prediction model is of high quality and statistically stable.

The clinical importance of identifying the AGTR1 (C1166C) variant is further supported by recent findings showing that this polymorphism is independently associated with severe ($\geq 90\%$) diameter stenosis of the coronary arteries in patients with acute myocardial infarction (Tran et al. 2024). This underscores the potential of our predictive model to identify individuals at high risk for critical ischemic complications.

The significance level of $p < 0.001$ obtained from ANOVA analysis further confirms the high statistical significance of the model. Importantly, the model performs significantly better than a simple prediction strategy based on mean values, further emphasizing its practical utility.

The results have important clinical implications: the proposed model allows clinicians to perform preliminary genetic screening using simple clinical parameters, potentially reducing the need for costly molecular-genetic testing. This is particularly relevant in clinical settings with limited resources.

It should be noted, however, that the model is based on retrospective data and requires further validation on independent datasets. Additionally, the model could be expanded by including additional biomarkers or employing more advanced machine learning techniques to enhance predictive accuracy.

By using this model, clinicians can identify high-risk individuals for major adverse cardiovascular events associated with silent genetic predispositions, providing a window for early intervention in patients with both arterial hypertension and ischemic heart disease.

Conclusions

Using multivariate regression analysis, a mathematical model was developed to predict the presence of pathological genetic polymorphism, specifically the C alleles in the AGTR1 and eNOS genes based on the clinical and demographic characteristics.

Eight most significant factors were identified that have a substantial impact on gene polymorphism including age, sex, disease duration, BMI, smoking, alcohol consumption, and blood pressure levels.

The developed model allows prediction of the probability of the presence of a pathological gene variant, which may significantly reduce the need for costly laboratory testing.

Statistical analysis confirms the adequacy and quality of the model with residuals following a normal distribution.

The coefficient of determination ($R^2 = 0.9842$) reflects high explanatory power meaning that 98.42% of the variation in genetic polymorphism is accounted for by the included factors.

The model is a valuable tool for early risk stratification of patients with comorbid arterial hypertension and ischemic heart disease.

References

- Abu-El-Haija A, Reddi HV, Wand H, Rose NC, Mori M, Qian E, Murray MF; ACMG Professional Practice and Guidelines Committee. Electronic address: documents@acmg.net. The clinical application of polygenic risk scores: A points to consider statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 25, 100803, 2023.
- Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 15, 2759–2772, 2020.
- Padmanabhan S, du Toit C, Dominiczak AF. Genetics of hypertension and heart failure. In: *Hypertension and Heart Failure: Epidemiology, Mechanisms and Treatment* (Eds. Dorobantu M, Voicu V, Grassi G, Agabiti-Rosei E, Mancia G), pp. 37–51, Springer, Cham, 2024.
- Pidruchna S, Shmanko V, Hnizdyukh R, Sverstiuk A, Lykhatskyy P, Kuzmak I, Yaroshenko T, Bandas I, Vasylyshyn N, Ostrivka O, Mudra A, Palytsia L, Letnyak N, Tokarskyy O. Statistically verified methods for determining predictors of development of arterial hypertension depending on endothelial nitric oxide synthase T786C gene promoter polymorphism using lipid profile indicators. *Endocr Regul* 58, 138–143, 2024a.
- Pidruchna S, Shmanko V, Zakharchuk U, Tokarskyy O, Hnizdyukh R, Lynkhatskyi P, Kuzmak I, Yaroshenko T, Bandas I, Vasylyshyn N, Ostrivka O, Mudra A, Palytsya L, Letniak N, Pohorielova O. Changes in adipokine indicators depending on A1166C polymorphism of the angiotensin II type 1 receptor gene as a predictor of the arterial hypertension. *Endocr Regul* 58, 153–157, 2024b.

- Potaskalova VS, Khaitovych NV, Natrus LV, Abdriakhimova TB, Lunov VY, Kleban KI, Korop OA. The effect of NOS3 and AGTR1 genotypes on the course of the arterial hypertension for overweight or obese patients. *Wiad Lek* 75, 533–540, 2022.
- Shmanko VV, Pidruchna SR, Hnizdyukh RV, Mudra AE, Zakharchuk UM, Kuzmak IP, Yaroshenko TY, Palytsya LM, Bandas IA, Dzhyvak VH, Nikitina IM, Zavarynska IF. Changes in lipid profile parameters depending on the a1166c polymorphism of the angiotensin II type I receptor gene as a predictor of arterial hypertension. *Wiad Lek* 77, 1554–1561, 2024.
- Tran DC, Le LHG, Thai TT, Van Hoang S, Do MD, Truong BQ. Effect of AGTR1 A1166C genetic polymorphism on coronary artery lesions and mortality in patients with acute myocardial infarction. *PloS One* 19, e0300273, 2024.
- Vaura F, Kauko A, Suvila K, Havulinna AS, Mars N, Salomaa V, Cheng FGS, Niiranen T. Polygenic risk scores predict hypertension onset and cardiovascular risk. *Hypertension* 77, 1119–1127, 2021.
- Wang Y, Wang JG. Genome-wide association studies of hypertension and several other cardiovascular diseases. *Pulse* 6, 169–186, 2019.
- Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet* 49, 403–415, 2017.
- Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, Clement D, Coca A, De Simone G, Dominiczak A, Kahan T, Mahfoud F, Redon J, Ruilope L, Zanchetti A, Kerins M, Kjeldsen S, Kreutz R, Laurent S, Lip GYH, McManus R, Narkiewicz K, Ruschitzka F, Schmieder R, Shlyakhto E, Tsioufis K, Aboyans V, Desormais I; List of authors/Task Force members:. 2018 Practice Guidelines for the Management of Arterial Hypertension of the European Society of Hypertension and the European Society of Cardiology: ESH/ESC Task Force for the Management of Arterial Hypertension. *J Hypertens* 36, 2284–2309, 2018.
- Yarema N, Kotsiuba O, Krytskyy T, Marushchak M, Krynytska I. Peculiarities of arterial hypertension in postmenopausal women with bone mineral density disorders and dyslipidemia. *Pol Merkur Lekarski* 48, 5–9, 2020.