PRIA JOURNAL SPECIAL ISSUE XXI NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE WITH INTERNATIONAL PARTICIPATION (CAI-2023)/SECTION 7 "APPLIED BIOMEDICAL SYSTEMS"

Methods of Prognostic Analysis for the Prediction of In-Hospital Mortality in Patients with Acute ST-Elevation Myocardial Infarction after Percutaneous Coronary Interventions

K. J. Shakhgeldyan[a],*, N. S. Kuksin[a],, I. G. Domzhalov[b],***, and B. I. Geltser[a],****

[a] *Vladivostok State University, Vladivostok, 690014 Russian Federation*

[b] *Primorsky Krai Clinical Hospital No. 1, Vladivostok, 690091 Russian Federation*

** e-mail: carinash@vvsu.ru*

*** e-mail: kuksin.ns@dvfu.ru*

**** e-mail: igor@domzhalov.ru*

***** e-mail: boris.geltser@vvsu.ru*

Abstract—The aim of this study was to develop an explainable machine learning model for predicting in-hospital mortality (IHF) in patients with ST-elevation myocardial infarction (STEMI) after percutaneous coronary intervention (PCI). We analyzed data from 4681 electronic medical records of patients with STEMI and identified 12 risk factors for IHF. The predictive models were developed based on multivariate logistic regression, random forest, and stochastic gradient boosting methods. The search for threshold values on the grid while maximizing the area under the ROC-curve and their validation by Shapley's additive explanation method made it possible to verify the risk factors for IHF. The model, whose parameters were risk factors, was superior in accuracy to the best model with continuous predictors based on stochastic gradient boosting. The use of IHF risk factors as predictors makes it possible to explain the obtained prognosis and reduce the risk of adverse events after PCI.

Keywords: machine learning predictive models, risk factors, Shapley's additive explanation, stochastic gradient boosting, predictor thresholding

DOI: 10.1134/S1054661824700676

1. INTRODUCTION

Cardiovascular diseases persist as a paramount public health issue across the globe [17]. Among these, ischemic heart disease (IHD) commands a significant portion of cardiovascular mortality, being responsible for 20% of all such deaths [9]. ST-elevation myocardial infarction (STEMI) is one of the most critical clinical manifestations of IHD. One of the most effective therapeutic interventions for STEMI is myocardial revascularization via percutaneous coronary intervention (PCI), which necessitates timely administration following the onset of symptoms. Despite enhancements in PCI technologies, the in-hospital mortality (IHM) rates post-procedure continue to be substantial, fluctuating between 4 and 7%.

In the realm of clinical practice, a variety of scoring systems are utilized to stratify the risk of IHM among patients based on clinical presentations, diagnostic findings, and laboratory results. Prominent among these risk assessment tools are the TIMI, GRACE, PAMI, and CADILLAC scales [2, 7, 10, 21, 22]. The application of these metrics facilitates the determination of the most efficacious therapeutic strategies, thereby optimizing clinical outcomes. The GRACE scale (Global Registry of Acute Coronary Events), developed through Cox regression analysis, is predominantly employed and has been the subject of extensive research aimed at its enhancement in recent years [3, 4, 14, 19, 24]. To refine the predictive accuracy of models based on this scale, their framework is often expanded with new variables. Concurrently, the employment of more sophisticated machine learning algorithms, such as random forests (RF) and stochastic gradient boosting (SGB), has become more prevalent, offering improved performance metrics for these models [13]. Nevertheless, the integration of novel predictors frequently lacks comprehensive explanations and evaluations of their influence on the clinical endpoints, thus constraining their practical utility.

The goal of this study was to perform a meticulous analysis of the dataset describing the conditions of patients with STEMI treated with PCI, to identify novel predictors of adverse events, assess their predictive capacity, and develop multifactorial models for

Received January 16, 2023; revised February 12, 2024; accepted February 27, 2024

forecasting IHM. This includes a detailed evaluation of the influence exerted by individual predictors on the outcome.

2. MATERIALS AND METHODS

2.1. Data

This study entails a retrospective analysis of data extracted from 4681 patient medical records. These patients underwent treatment at the Regional Vascular Center of the Primorsky Regional Clinical Hospital No. 1, Vladivostok, from 2015 to 2021. On the first day of hospital admission, all patients received invasive coronary angiography followed by transluminal balloon angioplasty and stenting of the arteries implicated in the myocardial infarction. Data extraction was performed from the DOCA+ medical information system by parsing html-files, culminating in a comprehensive dataset. The cohort consisted of 3207 males and 1474 females, with a median age of 63 years. Patients were classified into two groups: 318 individuals (6.8%) who succumbed to in-hospital mortality and 4363 individuals (93.2%) who survived post percutaneous coronary intervention (PCI).

Clinical and functional status were assessed on the first day of hospitalization, utilizing 136 variables. The analyzed data encompassed demographic and historical information, clinical blood analysis results, serum creatinine concentration (Cr), glomerular filtration rate (GFR), neutrophil and eosinophil counts, the international normalized ratio (INR), thrombin time (TT), prothrombin index (PTI), activated partial thromboplastin time (aPTT), and fibrinogen level (Fg). Echocardiographic assessments were conducted to determine the transverse (LA1) and longitudinal (LA2) dimensions of both the left and right atria (RA1 and RA2), the end-systolic (ESD) and end-diastolic (EDD) dimensions of the left ventricle (LV), and the left ventricular ejection fraction (EF) using the Teicholz method. The primary endpoint was defined as the in-hospital mortality (IHM) rate post-PCI, categorized as a binary variable indicating either the "absence" or "development" of mortality.

Due to the urgent nature of hospital admissions for patients with ST-elevation myocardial infarction (STEMI), the majority of laboratory tests were conducted immediately post-PCI. To substantiate the hypothesis that these data could serve as valid preoperative indicators, we conducted statistical tests comparing dependent samples across the dataset and revealed no significant differences in samples with both pre- and post-PCI data. Missing values within the dataset were not replaced with synthetic data to preserve the integrity of the original clinical data.

2.2. Data Selection and Risk Factor Determination

The process of selecting predictors and determining risk factors unfolded across five methodical stages. Initially, intergroup comparison tests between two cohorts, that is, deceased and surviving patients following percutaneous coronary intervention (PCI) with ST-elevation myocardial infarction (STEMI), facilitated the compilation of a list of potential predictors. The inclusion criterion was established at a p-value threshold of <0.05. Subsequently, the second stage involved quantifying the impact of these isolated potential predictors on the clinical endpoint (refer to Table 2). This was achieved by assessing the weight coefficients derived from univariate logistic regression models applied to normalized datasets. The definitive ensemble of predictors underwent validation within multivariate models. These models were iteratively refined by sequentially incorporating factors with the highest weight coefficients from the univariate logistic regression analysis, followed by a meticulous assessment of the integrity of the resulting model. The prognostic validity of each predictor was corroborated if its inclusion resulted in an enhancement of the area under the receiver operating characteristic curve (AUC).

In the penultimate stage, variables that comprised the optimal prognostic model were converted into risk factors for IHM. These risk factors were identified by isolating unique values of the variables and were determined using four distinct methods: the maximum odds ratio, the minimum p-value ascertained through the $\chi 2$ test, the greatest value of AUC in the univariate logistic regression model, and centroid determination [18]. Comparative analyses were conducted against diagrams of SHAP values to corroborate these findings.

2.3. Development of Predictive Models

The foundational model employed was the GRACE scale, which incorporates five principal predictors: patient age, acute heart failure classification according to Killip, heart rate, serum creatinine levels, and systolic blood pressure [7]. Two other predictors within the GRACE framework, that is, cardiac arrest upon patient admission and significant elevations in cardio-specific enzymes, were excluded due to their lack of variability in the analyzed sample. The predictor of "ST-segment deviation" on the electrocardiogram was considered for all patients with STEMI, attributing 28 points to this parameter. Each predictor within the GRACE scale was segmented into ranges linked to defined risk levels and corresponding scores. The aggregate score was interpreted as denoting low (less than 126 points), moderate (126 to 154 points), and high (more than 154 points) IHM risk [5]. The primary analytical model was a univariate logistic

regression, where the sole predictor was the cumulative score of the GRACE scale.

In this investigation, models were constructed using three advanced algorithms: multivariate logistic regression (MLR), random forest (RF), and stochastic gradient boosting (SGB). The dataset was bifurcated into a training and cross-validation subset (80%) and a final testing subset (20%). The training and cross-validation process was implemented through the stratified k-fold technique across 10 folds. Aggregated metrics of AUC, sensitivity (Sen), and specificity (Sp) were employed to identify the superior model, facilitating the adjustment of predictors and hyperparameters. The threshold for determining sensitivity and specificity was identified by seeking an equilibrium between these metrics. For RF and SGB models, hyperparameters were optimized through an exhaustive grid search methodology. The concluding testing phase involved educating the models with optimal configurations on the primary subset (80%) and evaluating them on the secondary subset (20%).

The final phase of the research culminated in the establishment of an MLR predictive model utilizing the corroborated risk factors.

2.4. Statistical Analysis Methods

In this study, conformity to a normal distribution was assessed utilizing the Shapiro–Wilk test. Due to the non-normal distribution of the data, nonparametric statistical methods were employed, including intergroup comparison tests. Metrics were presented as median values (Me) and their 95% confidence intervals (95% CI) for continuous variables, and as frequencies for categorical variables. The Mann–Whitney test was utilized for continuous variables, while the chi-squared ($\chi 2$) test was used for categorical variables. Odds ratios (OR) and their 95% CI were computed using Fisher's exact test. Models were developed using multiple logistic regression (MLR), random forest (RF), and stochastic gradient boosting (SGB) techniques. Model quality was evaluated based on three metrics: area under the curve (AUC), sensitivity (Sen), and specificity (Sp). To assess the impact of predictors on the clinical endpoint, Shapley Additive Explanations (SHAP) were utilized. Analyses were conducted in Python, open-source version 3.9.16.

3. RESULTS

3.1. Patient Cohort Characteristics and Predictor Selection

The initial stages of analysis involved median and frequency estimations, as well as intergroup comparison tests (Table 1). Table 1 presents the critical metrics.

3.2. Model Training and Validation

For indicators exhibiting statistically significant intergroup differences, we determined the weight coefficients using univariate logistic regression (LR) (Table 2). The GRACE score total serves as an indicator of IHM, with increases associated with a higher probability of adverse outcomes postpercutaneous coronary intervention (PCI). Similar relationships with IHM were noted for the levels of neutrophils, creatinine, heart rate, white blood cells, pulmonary artery systolic pressure, patient age, glucose, and class of acute heart failure, among others. For predictors such as systolic blood pressure (SBP), diastolic blood pressure (DBP), left ventricular ejection fraction (LVEF), prothrombin index, eosinophils, and hemoglobin, an inverse relationship was observed: higher values of these variables were associated with a higher probability of favorable PCI outcomes.

In addition to the baseline univariate LR model based on the total GRACE score, we developed models using multivariate logistic regression (MLR), random forest (RF), and stochastic gradient boosting (SGB) with five factors from the GRACE scale (Table 3). The best results were achieved by expanding the spectrum of predictors to include LVEF, PASP, hemoglobin, hematocrit, relative levels of neutrophils and eosinophils, and patient height. Data analysis demonstrated a consistent improvement in the quality of prognosis for the three types of models (MLR, RF, and SGB) both during cross-validation and final testing. The RF and SGB models exhibited superior prognostic properties compared to MLR, with AUCs of 0.901 and 0.903 versus 0.899, respectively.

3.3. Risk Factors

Identification of risk factors for IHM was performed using a grid search method to find the optimal solution. The optimal solution was determined using several objective functions: the maximum odds ratio (OR) or area under the curve (AUC), the minimum p-value, and centroid calculation [18]. Threshold values with the highest predictive potential were determined for the following parameters: age, Killip class, heart rate, creatinine, systolic arterial pressure, left ventricular ejection fraction, lactate dehydrogenase, hemoglobin, plateletcrit, relative neutrophil count, and relative eosinophil count. These threshold values were then classified as risk factors (Table 4). A risk factor is assigned a value of "1" if the predictor value exceeds the threshold with a "+" postfix, or falls below it with a "–" postfix. In cases where the indicator has a value of "0," it is not considered a risk factor (Table 4).

The risk factors for IHM derived through objective function optimization were ascertained by evaluating the isolated impact of each predictor on the primary endpoint. However, the threshold values for these factors may shift when considering their combined influ-

Table 1. The clinical and functional characteristics of patients

Predictor	Group 1 (n = 318)	Group 2 (n = 4363)	OR (95%) CI	p-value
Gender: F, abs. (%)	142 (44.65%)	1332 (30.5%)	1.8 [1.5; 2.3]	<0.000001
Age, years	71 [70; 72]	62 [62; 62]	–	<0.000001
Height, cm	168 [167; 169]	170 [169.8; 170.3]	–	0.000001
Weight, kg	78 [76; 80]	80 [79.5; 80.5]	–	0.000005
BMI, kg/m^2	26.8 [26.2; 27.4]	27.7 [27.5; 27.8]	–	0.039489
SBP, mm Hg	110 [106; 114]	130 [129; 131]	–	<0.000001
DBP, mm Hg	72 [69.9; 74.1]	80 [79.6; 80.4]	–	<0.000001
Heart rate, bpm	86 [83.3; 88.7]	72 [71.6; 72.4]	–	<0.000001
Creatinine, μmol/L	130 [115; 144]	97 [95.7; 98.2]	–	<0.000001
Killip class				
I	71 (22.33)	2745 (62.5%)	0.2 [0.13; 0.22]	<0.000001
II	58 (18.2%)	870 (20%)	0.9 [0.67; 1.20]	0.508052
III	66 (20.75%)	479 (11%)	2.1 [1.59; 2.8]	<0.000001
IV	123 (38.7%)	269 (6.17%)	9.6 [7.4; 12.4]	<0.000001
GRACE score	203 [197; 211]	147 [147; 149]	–	<0.000001
LVEF, %	46.5 [44.7; 48.3]	56 [55.7; 56.3]	–	<0.000001
LVEDD, cm	5 [4.9; 5.1]	5 [4.98; 5.02]	–	0.356921
LVESD, cm	3.7 [3.6;3.8]	3.4 [3.38;3.42]	–	<0.000001
PASP, mm Hg	35 [33.1;36.9]	28 [27.8; 28.2]	–	<0.000001
LA1, cm	4.10 [4.02; 4.18]	3.90 [3.89; 3.91]	–	<0.000001
LA2, cm	5.20 [5.09; 5.31]	4.90 [4.88; 4.92]	–	<0.000001
RA1, cm	3.80 [3.71; 3.89]	3.60 [3.59; 3.61]	–	<0.000001
RA2, cm	4.80 [4.70; 4.90]	4.70 [4.68; 4.72]	–	0.00004
RBC, 1012/L	4.25 [3.96; 4.53]	4.48 [4.46; 4.50]	–	<0.000001
Hb, g/L	132 [129; 135]	141 [140; 142]	–	<0.000001
PLT, 109/L	228 [215; 241]	221 [219; 223]	–	0.020104
Neutrophils, %	81.3 [79.9; 82.7]	66.7 [66.4; 67.1]	–	<0.000001
Eosinophils, %	0.1 [–0.02; 0.22]	0.90 [0.85; 0.95]	–	<0.000001
Glu, mmol/L	7.95 [6.89; 9.01]	5.79 [5.71; 5.87]	–	<0.000001
Urea, μmol/L	12.12 [7.8; 16.5]	6.7 [6.38; 7.02]	–	<0.000001
Plateletcrit, %	0.22 [0.21;0.23]	0.20 [0.20; 0.20]	–	0.0012
WBC, 109/L	14 [13.2; 14.8]	10.5 [10.3; 10.6]	–	<0.000001
PT, %	75.5 [72.2; 78.8]	89.3 [88.8; 89.9]	–	<0.000001
INR, units	1.26 [1.10; 1.42]	1.06 [1.05; 1.07]	–	<0.000001
TT, s	21.9 [21.3; 23.3]	21.4 [21.3; 21.7]	–	0.012
aPTT, s	39.7 [32.9; 46.5]	36.5 [35.1; 37.9]	–	0.000026
Anterior MI, abs. (%)	177 (55.66%)	2023 (46.37%)	1.5 [1.15; 1.83]	0.001647
Atrial fibr., abs. (%)	129 (40.57%)	772 (17.69%)	3.2 [2.51; 4.02]	<0.000001
Type 2 diabetes, abs. (%)	99 (31.13%)	831 (19.05%)	1.9 [1.50; 2.46]	<0.000001
CKD, abs. (%)	83 (26.1%)	677 (15.5%)	1.97 [1.5; 2.6]	<0.000001

LVEF, left ventricular ejection fraction; LVEDD, left ventricular end-diastolic dimension; LVESD, left ventricular end-systolic dimension; PASP, pulmonary artery systolic pressure; LA2, longitudinal dimension of the left atrium; RA1, transverse dimension of the right atrium; RA2, longitudinal dimension of the right atrium; RBC, red blood cells; Hb, hemoglobin; PLT, thrombocytes; Glu, glucose; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; WBC, white blood cells; PT, prothrombin time; INR, international normalized ratio; TT, thrombin time; aPTT, activated partial thromboplastin time; Cr, creatinine; AF, atrial fibrillation; T2D, type 2 diabetes; CKD, chronic kidney disease.

Table 2. Weight coefficients in univariate logistic regression models

Predictor	Coefficient	p-value
GRACE score total	7.507 [7.503; 7.511]	<0.000001
Neutrophils (relative value)	6.485 [6.481; 6.489]	<0.000001
SBP	–6.280 [–6.283; –6.277]	<0.000001
DBP	–6.226 [–6.228; –6.224]	<0.000001
Cr	6.102 [6.100; 6.103]	<0.000001
Heart Rate	5.573 [5.570; 5.575]	<0.000001
Lymphocytes (relative value)	–5.349 [–5.353; –5.345]	<0.000001
LVEF	–5.225 [–5.229; –5.221]	<0.000001
WBC	5.067 [5.064; 5.070]	<0.000001
SII	4.932 [4.930; 4.934]	<0.000001
PASP	4.639 [4.637; 4.642]	<0.000001
Glu	4.385 [4.383; 4.386]	<0.000001
Prothrombin Index	–4.349 [–4.352; –4.345]	<0.000001
Eosinophils (relative value)	–4.206 [–4.208; –4.204]	<0.000001
Age	4.171 [4.167; 4.176]	<0.000001
INR	3.788 [3.786; 3.789]	<0.000001
Killip class	3.583 [3.575; 3.590]	<0.000001
PLR	3.133 [3.131; 3.136]	<0.000001
Hb	–3.088 [–3.091; –3.086]	<0.000001
La1	3.028 [3.026; 3.031]	<0.000001
RDW-CV	2.990 [2.988; 2.992]	<0.000001
Hematocrit	–2.818 [–2.819; –2.816]	<0.000001
Ra1	2.705 [2.702; 2.708]	<0.000001
La2	2.583 [2.580; 2.586]	<0.000001
Red blood cells	–2.417 [–2.420; –2.414]	<0.000001
Monocytes (absolute value)	–2.396 [–2.400; –2.392]	<0.000001
LVESD	2.268 [2.264; 2.271]	<0.000001
Plateletcrit	2.063 [2.060; 2.065]	<0.000001

ence on the resultant variable. To delineate threshold values within a multifactorial model, the Shapley Additive Explanations (SHAP) method was employed to assess predictor importance. SHAP value plots were generated for each predictor within the optimal XGBoost multifactorial model (Fig. 1).

It is noteworthy that the risk factor thresholds obtained through maximizing the AUC of univariate logistic regression corresponded with the results of the SHAP analysis. Specifically, SHAP values surpassing 0 indicated an elevated probability of IHM. As an example, AUC maximization yielded a threshold value of 75% for neutrophils as a risk factor, which was associated with an 11-fold increase in IHM risk (Table 4). This threshold is further illustrated in Fig. 1, where a sharp escalation in SHAP values exceeding 0 is observed when the blood neutrophil content surpasses 75%. Analogous correlations were observed for other predictors.

Utilizing the risk factors derived from AUC maximization, a multifactorial IHM prediction model was developed, which has cross-validated performance metrics of AUC = 0.917, sensitivity = 0.821, and specificity = 0.854. In the final testing phase, the model achieved AUC = 0.903, sensitivity = 0.775, and specificity = 0.851, demonstrating parity with the best XGBoost model. The salient advantage of the risk fac-

Table 3. The performance metrics of predictive models

	Age, Killip Class, Cr, SAD, HR			Age, Height, Killip Class, HR, Cr, SAD, LVEF, SDLA, Hemoglobin (Hb), Hematocrit, Neutrophils, and Eosinophils		
	MLR	RF	SGB	MLR	RF	SGB
Cross-validation metrics on validation samples						
AUC	0.855	0.853	0.859	0.915	0.919	0.915
Se	0.774	0.710	0.742	0.840	0.720	0.760
Sp	0.778	0.844	0.819	0.828	0.910	0.895
Performance metrics on final test sample						
AUC	0.823	0.836	0.839	0.899	0.901	0.903
Se	0.711	0.711	0.689	0.771	0.686	0.714
Sp	0.772	0.836	0.831	0.846	0.899	0.896

tor-based model lies in its interpretability for clinicians.

The identified risk factors for IHM in post-PCI STEMI patients include: blood neutrophil levels exceeding 75.4% and eosinophil levels surpassing 0.3%; LVEF below 51%, blood hemoglobin levels below 144 g/L; blood plateletcrit levels exceeding 0.22%, LDH exceeding 34 mm Hg; height below 173 cm, systolic arterial pressure below 112 mm Hg; blood creatinine exceeding 123 μmol/L, heart rate exceeding 79 bpm; age surpassing 65 years; and Killip class III–IV.

4. DISCUSSION

This investigation used data mining techniques on a patient dataset with ST-segment elevation myocardial infarction (STEMI) postcardiac surgery. Our objective was to uncover and authenticate new predictors of IHM, to confirm fatal outcome risk factors, and to engineer a prognostic model that not only exceeds the precision of the well-established GRACE model but also incorporates explanatory components. The array of potential predictors encompassed five metrics from the GRACE scale in addition to a spectrum of novel indicators not previously considered in this predictive capacity. Utilizing three sophisticated machine learning algorithms, alongside rigorous cross-validation and testing protocols, we substantiated their prognostic relevance for IHM. The apex of prediction accuracy was realized through the application of the GBM model. Notably, all three methodologies, that is, multiple logistic regression, support vector machines, and GBM, demonstrated a progressive enhancement in predictive performance during both cross-validation and subsequent final evaluations.

The application of the Shapley additive explanations method found a significant impact exerted by the newly discerned predictors on IHM, specifically highlighting LVEF, SPAP, neutrophil levels, eosinophils, and thrombocrit (see Fig. 2). The impact magnitude of hemoglobin on the endpoint paralleled that of patient age, while the influence attributable to patient height was minimal.

By employing the maximization of the AUC of a unifactorial logistic regression, corroborated by results from Shapley additive explanations, we delineated predictor threshold values. Deviations from these thresholds were designated as IHM risk factors. This methodological approach facilitates real-time risk assessment by clinicians, enabling prompt intervention and adjustment of risk factors.

LVEF emerges as the predominant predictor augmenting the GRACE scale in influencing the endpoint, a finding corroborated by several other studies [1, 6, 11, 20, 23]. Although most literature cites a critical threshold of LVEF < 40%, our findings indicate a heightened IHM risk commencing at LVEF < 50%. The role of neutrophil levels exceeding 75%, although not identified by other researchers, has been implicated as a predictive factor of IHM in conjunction with lymphocyte levels in numerous studies [16, 20]. An elevated SPAP above 35 mm Hg is similarly recognized as a risk factor [8]. The interrelation between platelet volume, hemoglobin levels, and IHM has been a subject of discussion in the literature [15, 20, 23]. The factor with the least contribution to IHM prediction identified in this study is patient height; our analysis confirms that a stature less than 173 cm is a risk factor for IHM among patients with STEMI following cardiac surgery. The prognostic potential of this parameter in relation to cardiovascular disease progression is currently under vigorous investigation [12].

Table 4. Identification of IHM risk factors using different methods

Predictor	Method	Threshold pen_spark	*p*-value	OR (95% CI)	AUC
Neutrophils	max(OR)	94.2+	<0.00001	23.01 [4.6; 114.8]	0.513
	min(*p*-value)	78.8+	<0.00001	9.12 [6.63; 12.53]	0.730
	max(AUC)	75.4+	<0.00001	11.28 [7.83; 16.2]	0.774
	Centroid	74.0+	<0.00001	9.92 [6.86; 14.36]	0.751
Eosinophils	max(OR)	1.3–	<0.00001	9.79 [5.3; 18.08]	0.666
	min(*p*-value)	0.3–	<0.00001	7.87 [5.62; 11.03]	0.741
	max(AUC)	0.3–	<0.00001	7.87 [5.62; 11.1]	0.741
	Centroid	0.5–	<0.00001	7.630 [5.239; 11.112]	0.722
LVEF	max(OR)	31.0–	<0.00001	19.693 [12.08; 32.11]	0.514
	min(*p*-value)	31.0–	<0.00001	19.693 [12.08; 32.11]	0.594
	max(AUC)	51.0–	<0.00001	4.850 [3.504; 6.712]	0.690
	Centroid	51.25–	<0.00001	4.885 [3.514; 6.792]	0.683
Hemoglobin	max(OR)	94.0–	<0.00001	5.304 [3.278; 8.582]	0.540
	min(*p*-value)	94.0–	<0.00001	5.304 [3.278; 8.582]	0.541
	max(AUC)	144.0–	<0.00001	2.285 [1.703; 3.068]	0.614
	Centroid	136.5–	<0.00001	2.250 [1.728; 2.928]	0.585
Plateletcrit	max(OR)	0.36+	<0.00001	4.490 [2.743; 7.350]	0.500
	min(*p*-value)	0.36+	<0.00001	4.490 [2.743; 7.350]	0.524
	max(AUC)	0.22+	0.00002	1.816 [1.379; 2.391]	0.598
	Centroid	0.21+	0.0009	1.629 [1.238; 2.144]	0.576
LDH	max(OR)	34.0+	<0.00001	6.401 [4.680; 8.753]	0.500
	min(*p*-value)	34.0+	<0.00001	6.401 [4.680; 8.753]	0.728
	max(AUC)	34.0+	<0.00001	6.401 [4.680; 8.753]	0.728
	Centroid	31.5+	<0.00001	5.902 [4.298; 8.103]	0.701
Height	max(OR)	152.0–	0.00003	4.174 [2.118; 8.228]	0.513
	min(*p*-value)	173.0–	<0.00001	1.890 [1.451; 2.464]	0.594
	max(AUC)	173.0–	<0.00001	1.890 [1.451; 2.464]	0.594
	Centroid	169.0–	0.00004	1.667 [1.307; 2.126]	0.559
SAP	max(OR)	60.0–	<0.00001	31.538 [10.89; 91.35]	0.523
	min(*p*-value)	92.0–	<0.00001	11.014 [8.28; 14.65]	0.613
	max(AUC)	112.0–	<0.00001	5.361 [4.237; 6.783]	0.685
	Centroid	120.0–	<0.00001	4.777 [3.781; 6.035]	0.676
Creatinine	max(OR)	427.0+	<0.00001	30.523 [10.16; 91.74]	0.518
	min(*p*-value)	188.6+	<0.00001	13.054 [9.44; 18.04]	0.625
	max(AUC)	122.9+	<0.00001	5.774 [4.477; 7.448]	0.701
	Centroid	113.3+	<0.00001	4.605 [3.559; 5.960]	0.693
HR	max(OR)	150.0+	<0.00001	41.635 [4.32; 401.42]	0.500
	min(*p*-value)	94.0+	<0.00001	6.225 [4.860; 7.972]	0.663
	max(AUC)	79.0+	<0.00001	4.019 [3.170; 5.096]	0.672
	Centroid	79.0+	<0.00001	4.019 [3.170; 5.096]	0.672

Table 4. (Contd.)

Predictor	Method	Threshold pen_spark	*p*-value	OR (95% CI)	AUC
Age	max(OR)	85.0+	<0.00001	6.194 [3.407; 11.264]	0,513
	min(*p*-value)	70.0+	<0.00001	4.109 [3.259; 5.181]	0.644
	max(AUC)	65.0+	<0.00001	3.735 [2.923; 4.772]	0.667
	Centroid	66.5+	<0.00001	3.646 [2.872; 4.628]	0.660
Killip class	max(OR)	3.0+	<0.00001	9.600 [7.423; 12.416]	0.665
	min(*p*-value)	3.0+	<0.00001	9.600 [7.423; 12.416]	0.665
	max(AUC)	2.0+	<0.00001	7.081 [5.585; 8.977]	0.714
	Centroid	2.0+	<0.00001	7.081 [5.585; 8.977]	0.714

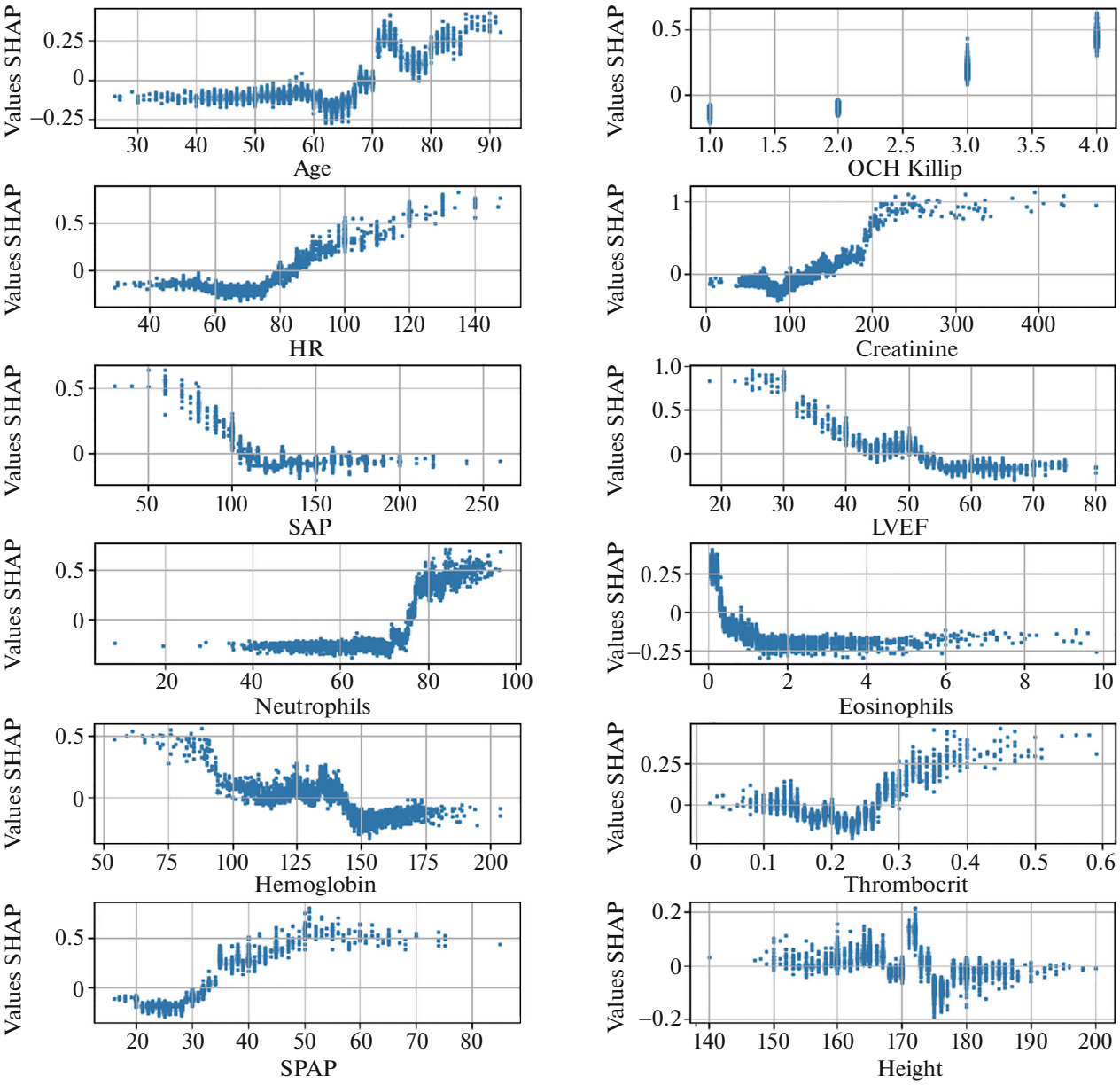

Fig. 1. The influence of predictors in the multifactorial model on IHM.

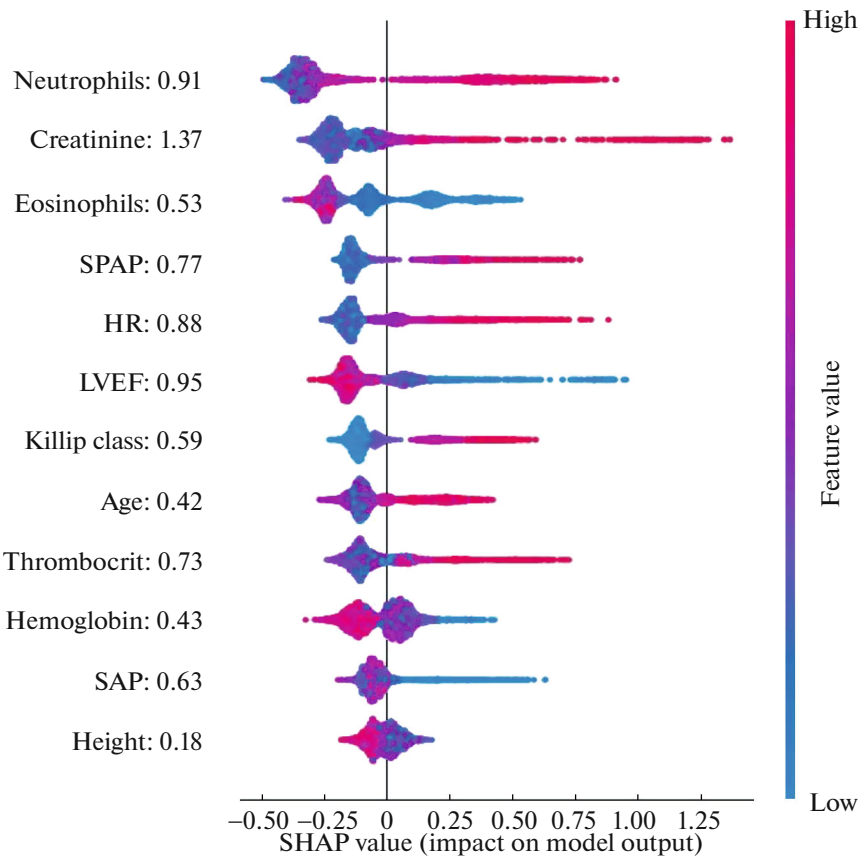

Fig. 2. Chart of SHAP values by GBM model.

Figure 2 displays a graphical representation of SHAP values, illustrating the significance of predictors that constitute the optimal GBM model. As indicated, elevated parameters such as creatinine, heart rate, neutrophils, SPAP, thrombocrit, and age are associated with an increased probability of mortality postsurgery in patients with STEMI. Additionally, it is pertinent to highlight that an increase in the severity of heart failure, as classified by T.Killip, precipitates a substantial increase in the probability of IHM. Conversely, increases in LVEF, eosinophil counts, hemoglobin levels, as well as patient height and systolic arterial pressure, inversely affect the endpoint: with an increase of these variables, the likelihood of mortality diminishes.

The precision of the prognostic model developed herein, which incorporates these IHM risk factors, notably surpasses outcomes documented by previous researchers [13, 20].

CONCLUSIONS

In the present study we delineated and validated risk factors for IHM among patients experiencing STEMI who underwent cardiac surgery. The prognostic model formulated on the basis of these identified factors exhibited enhanced predictive capabilities when contrasted with the traditional GRACE model and other models advanced by fellow researchers. Employing these IHM risk factors as predictors in the model not only promotes computational transparency but also empowers clinicians to promptly enact interventions aimed at mitigating the risk of fatal outcomes.

The advancement of this research necessitates an expansion of the existing dataset and the exploration of novel predictors for in-hospital mortality, thereby broadening the scope and applicability of our findings in clinical settings.

FUNDING

This research was supported by the Russian Science Foundation (project no 23-21-00250, https://rscf.ru/project/23-21-00250/).

CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. M. Alnaser, I. V. Sychev, Ya. A. Pushkina, and L. N. Goncharova, "Comparative assessment of short-term prognosis in patients with ST-segment elevation acute coronary syndrome without percutaneous coronary intervention based on the GRACE, TIMI, RECORD, PREDICT scales," Kardiovaskulyarnaya Terapiya i Profilaktika **21** (1), 2850 (2022). https://doi.org/10.15829/1728-8800-2022-2850
2. X.-F. Bai, Y.-P. Zhang, J. Zhou, Yu. Wu, R.-F. Li, L.-Z. Sun, Q.-Q. Ma, B.-W. Lou, B.-W. Zhai, M.-P. Liu, L.-L. Cheng, X.-N. Tong, and Z.-Y. Yuan, "Combination of the CYP2C19 metabolizer and the GRACE risk score better predicts the long-term major adverse cardiac events in acute coronary syndrome undergoing percutaneous coronary intervention," Thromb. Res. **170**, 142–147 (2018). https://doi.org/10.1016/j.thromres.2018.08.016
3. I. S. Bessonov, V. A. Kuznetsov, S. S. Sapozhnikov, E. A. Gorbatenko, and A. A. Shadrin, "The risk score for in-hospital mortality in patients with ST-segment elevation myocardial infarction," Kardiologiia **61** (9), 11–19 (2021). https://doi.org/10.18087/cardio.2021.9.n1720
4. X. Chen, M. Shao, T. Zhang, W. Zhang, Yo. Meng, H. Zhang, H. Hai, and G. Li, "Prognostic value of the combination of GRACE risk score and mean platelet volume to lymphocyte count ratio in patients with ST-segment elevation myocardial infarction after percutaneous coronary intervention," Exp. Ther. Med. **19**, 3664–3674 (2020). https://doi.org/10.3892/etm.2020.8626
5. K. A. A. Fox, K. A. Eagle, J. M. Gore, P. G. Steg, and F. A. Anderson, "The Global Registry of Acute Coronary Events, 1999 to 2009-GRACE," Heart **96**, 1095–1101 (2010). https://doi.org/10.1136/hrt.2009.190827
6. S. M. Ghorashi, M. Salarifar, H. Poorhosseini, S. Sadeghian, A. Jalali, H. Aghajani, A.-M. Haji-Zeinali, and N. Omidi, "Predictors of in-hospital mortality in diabetic patients with non-ST-elevation myocardial infarction," Egypt. Heart J. **74**, 20 (2022). https://doi.org/10.1186/s43044-022-00256-y
7. Ch. B. Granger, R. J. Goldberg, O. Dabbous, K. S. Pieper, K. A. Eagle, C. P. Cannon, et al., "Predictors of hospital mortality in the global registry of acute coronary events," Arch. Int. Med. **163**, 2345 (2003). https://doi.org/10.1001/archinte.163.19.2345
8. M. L. A. Haeck, G. E. Hoogslag, H. Boden, M. A. Velders, S. Katsanos, I. Al Amri, P. Debonnaire, M. J. Schalij, H. W. Vliegen, J. J. Bax, N. A. Marsan, and V. Delgado, "Prognostic implications of elevated pulmonary artery pressure after ST-segment elevation myocardial infarction," Am. J. Cardiol. **118**, 326–331 (2016). https://doi.org/10.1016/j.amjcard.2016.05.008
9. B. Ibanez, S. James, et al., "The task force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the european society of cardiology (ESC) 2017 esc guidelines for themanagement of acutemyocardial infarction in patients presenting with ST-segment elevation," Rossiiskii Kardiologicheskii Zh. **23** (5), 103–158 (2018). https://doi.org/10.15829/1560-4071-2018-5-103-158
10. Y.-T. Kao, Y.-C. Hsieh, C.-Y. Hsu, C.-Y. Huang, M.-H. Hsieh, Y.-K. Lin, and J.-S. Yeh, "Comparison of the TIMI, GRACE, PAMI and CADILLAC risk scores for prediction of long-term cardiovascular outcomes in Taiwanese diabetic patients with ST-segment elevation myocardial infarction: From the registry of the Taiwan Society of Cardiology," PLoS One **15**, e0229186 (2020). https://doi.org/10.1371/journal.pone.0229186
11. T. Kırıs, E. Avcı, and A. Çelik, "Combined value of left ventricular ejection fraction and the model for end-stage liver disease (MELD) score for predicting mortality in patients with acute coronary syndrome who were undergoing percutaneous coronary intervention," BMC Cardiovasc. Disord. **18**, 44 (2018). https://doi.org/10.1186/s12872-018-0782-8
12. S. Krieg, K. Kostev, M. Luedde, A. Krieg, T. Luedde, Ch. Roderburg, and S. H. Loosen, "The association between the body height and cardiovascular diseases: a retrospective analysis of 657,310 outpatients in Germany," Eur. J. Med. Res. **27**, 240 (2022). https://doi.org/10.1186/s40001-022-00881-y
13. R. Li, L. Shen, W. Ma, B. Yan, W. Chen, J. Zhu, L. Li, J. Yuan, and Ch. Pan, "Use of machine learning models to predict in-hospital mortality in patients with acute coronary syndrome," Clin. Cardiol. **46**, 184–194 (2023). https://doi.org/10.1002/clc.23957
14. X.-J. Liu, Z.-F. Wan, N. Zhao, Y.-P. Zhang, L. Mi, X.-H. Wang, D. Zhou, Ya. Wu, and Z.-Y. Yuan, "Adjustment of the GRACE score by HemoglobinA1c enables a more accurate prediction of long-term major adverse cardiac events in acute coronary syndrome without diabetes undergoing percutaneous coronary intervention," Cardiovasc. Diabetology **14**, 110 (2015). https://doi.org/10.1186/s12933-015-0274-4
15. Ö. Şatıroğlu, M. E. Durakoğlugil, H. A. Uydu, H. Duman, M. Çetin, Yü. Çiçek, and T. Erdoğan, "The relationship between mean platelet volume and in-hospital mortality in geriatric patients with ST segment elevation myocardial infarction who underwent primary percutaneous coronary intervention," Cardiovasc. Innovations Appl. **4**, 135–141 (2019). https://doi.org/10.15212/cvia.2019.0015
16. F. Tavares, P. I. M. Moraes, J. M. Souza, A. H. Barbosa, E. M. Santos, J. A. Marcondes, M. D. Oliveira, I. Gonçalves, C. V. A. Espírito Santo, H. T. Bianco, F. A. H. Fonseca, V. Moisés, and A. Caixeta, "Prognostic role of neutrophil-to-lymphocyte ratio in patients with ST-elevation myocardial infarction undergoing to pharmaco-invasive strategy," Cardiovasc. Revasc. Med. **34**, 99–103 (2022). https://doi.org/10.1016/j.carrev.2021.01.027
17. The World Health Organization The top 10 causes of death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. Cited May 2, 2023.
18. F. Valente, J. Henriques, S. Paredes, T. Rocha, P. De Carvalho, and J. Morais, "A new approach for interpretability and reliability in clinical risk prediction: Acute coronary syndrome scenario," Artif. Intell. Med. **117**, 102113 (2021). https://doi.org/10.1016/j.artmed.2021.102113

19. J. J. Wang, Y. Fan, Y. Zhu, J. D. Zhang, S. M. Zhang, Z. F. Wan, H. L. Su, and N. Jiang, "Biomarkers enhance the long-term predictive ability of the KAMIR risk score in Chinese patients with ST-elevation myocardial infarction," Chin. Med. J. **132** (1), 30–41 (2019). https://doi.org/10.1097/CM9.0000000000000015

20. Yu. Wang, W. Wang, Sh. Jia, M. Gao, Sh. Zheng, J. Wang, Yi. Dang, Yi. Li, and X. Qi, "Development of a nomogram for the prediction of in-hospital mortality in patients with acute ST-elevation myocardial infarction after primary percutaneous coronary intervention: A multicentre, retrospective, observational study in Hebei province, China," BMJ Open **12**, e056101 (2022). https://doi.org/10.1136/bmjopen-2021-056101

21. L. Wang, Zh. Zhang, and T. Hu, "Effectiveness of LODS, OASIS, and SAPS II to predict in-hospital mortality for intensive care patients with ST elevation myocardial infarction," Sci. Rep. **11**, 23887 (2021). https://doi.org/10.1038/s41598-021-03397-3

22. R. S. Wilson, P. Malamas, B. Dembo, S. K. Lall, N. Zaman, and B. R. Peterson, "The CADILLAC risk score accurately identifies patients at low risk for in-hospital mortality and adverse cardiovascular events following ST elevation myocardial infarction," BMC Cardiovasc. Disord. **21**, 533 (2021). https://doi.org/10.1186/s12872-021-02348-0

23. M. F. Yılmaz, A. Karagöz, G. Zeren, İ. İ. Avcı, B. Timur, M. A. Sungur, S. Kalkan, B. Şimşek, F. Can, E. Acar, İ. H. Tanboğa, and C. Y. Karabay, "Relationship between in-hospital mortality and creatinine/albumin in patients with ST-elevation myocardial infarction without standard modifiable risk factors," Biomarkers Med. **16**, 1043–1053 (2022). https://doi.org/10.2217/bmm-2022-0241

24. M. V. Zykov, N. V. Dyachenko, R. M. Velieva, V. V. Kashtalap, and O. L. Barbarash, "Combined use of the GRACE ACS risk score and comorbidity indices to increase the effectiveness of hospital mortality risk assessment in patients with acute coronary syndrome," Ter. Arkhiv **94**, 816–821 (2022). https://doi.org/10.26442/00403660.2022.07.201742

Translated by V. Tereshchenko

Publisher's Note. Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Karina J. Shakhgeldyan. Doctor of Technical Sciences, Associate Professor, Director of the Institute of Information Technologies, Vladivostok State University (VSU). Research interests include medical data mining, explainable artificial intelligence, and interpretable machine learning models. Author of 250 scientific publications.

Nikita S. Kuksin. Master's student at Vladivostok State University (VSU). Research interests focus on the development of prognostic models, categorization of continuous variables, and the creation of explainable artificial intelligence models. Author of six scientific publications.

Igor G. Domzhalov. Physician in the Department of Resuscitation and Intensive Care at the Regional Vascular Center, Primorsky Krai Clinical Hospital No. 1. Research interests: identifying predictors of mortality in cardiology, clinical interpretation of machine learning models, studying the factors of cardiogenic shock development, developing prognostic models, categorizing continuous variables, and creating explainable artificial intelligence models. Author of five scientific publications.

Boris I. Geltser. Doctor of Medical Sciences, Professor, Corresponding Member of the Russian Academy of Sciences, Senior Research Scientist at the Laboratory of Data Mining, Vladivostok State University (VSU). Research interests include artificial intelligence in medicine, clinical and pathological pathophysiology of circulation and respiration, internal diseases. Author of 653 scientific publications.