

Geltser B. I.<sup>1</sup>, Shahgeldyan K. J.<sup>2</sup>, Rublev V. Yu.<sup>3</sup>,  
Kotelnikov V. N.<sup>1</sup>, Kriger A. B.<sup>2</sup>, Shirobokov V. G.<sup>4</sup>

<sup>1</sup> Far Eastern federal university. School of biomedicine. Vladivostok, Russia

<sup>2</sup> Institute of Information Technologies, Vladivostok State University of Economics and Service, Vladivostok, Russia

<sup>3</sup> Regional Clinical Hospital #1, Vladivostok, Russia

<sup>4</sup> MISIS national research technological University, Moscow, Russia

## MACHINE LEARNING METHODS FOR PREDICTION OF HOSPITAL MORTALITY IN PATIENTS WITH CORONARY HEART DISEASE AFTER CORONARY ARTERY BYPASS GRAFTING

<i>Aim</i>	To compare the accuracy of predicting an in-hospital fatal outcome for models based on current machine-learning technologies in patients with ischemic heart disease (IHD) after coronary bypass (CB) surgery.
<i>Material and methods</i>	A retrospective analysis of 866 electronic medical records was performed for patients (685 men and 181 women) who have had a CB surgery for IHD in 2008–2018. Results of clinical, laboratory, and instrumental evaluations obtained prior to the CB surgery were analyzed. Patients were divided into two groups: group 1 included 35 (4%) patients who died within the first 20 days of CB, and group 2 consisted of 831 (96%) patients with a beneficial outcome of the surgery. Predictors of the in-hospital fatal outcome were identified by a multistep selection procedure with analysis of statistical hypotheses and calculation of weight coefficients. For construction of models and verification of predictors, machine-learning methods were used, including the multifactorial logistic regression (LR), random forest (RF), and artificial neural networks (ANN). Model accuracy was evaluated by three metrics: area under the ROC curve (AUC), sensitivity, and specificity. Cross validation of the models was performed on test samples, and the control validation was performed on a cohort of patients with IHD after CB, whose data were not used in development of the models.
<i>Results</i>	The following 7 risk factors for in-hospital fatal outcome with the greatest predictive potential were isolated from the EuroSCORE II scale: ejection fraction (EF) <30%, EF 30–50%, age of patients with recent MI, damage of peripheral arterial circulation, urgency of CB, functional class III–IV chronic heart failure, and 5 additional predictors, including heart rate, systolic blood pressure, presence of aortic stenosis, posterior left ventricular (LV) wall relative thickness index (RTI), and LV relative mass index (LVRMI). The models developed by the authors using LR, RF and ANN methods had higher AUC values and sensitivity compared to the classical EuroSCORE II scale. The ANN models including the RTI and LVRMI predictors demonstrated a maximum level of prognostic accuracy, which was illustrated by values of the quality metrics, AUC 93%, sensitivity 90%, and specificity 96%. The predictive robustness of the models was confirmed by results of the control validation.
<i>Conclusion</i>	The use of current machine-learning technologies allowed developing a novel algorithm for selection of predictors and highly accurate models for predicting an in-hospital fatal outcome after CB.
<i>Keywords</i>	Machine-learning methods; prognostic models; coronary bypass; in-hospital fatal outcome
<i>For citation</i>	Geltser B. I., Shahgeldyan K. J., Rublev V. Yu., Kotelnikov V. N., Kriger A. B., Shirobokov V. G. Machine Learning Methods for Prediction of Hospital Mortality in Patients with Coronary Heart Disease after Coronary Artery Bypass Grafting. <i>Kardiologiya</i> . 2020;60(10):38–46. [Russian: Гельцер Б.И., Шахгельдян К.И., Рублев В.Ю., Котельников В.Н., Кригер А.Б., Ширококов В.Г. Методы машинного обучения в прогнозировании летальных исходов в стационаре у больных ишемической болезнью сердца после коронарного шунтирования. <i>Кардиология</i> . 2020;60(10):38–46]
<i>Corresponding author</i>	Shahgeldyan K. J. E-mail: carinash@vvsu.ru

Coronary artery disease (CAD) is a leading cause of disability and mortality in most countries worldwide. Coronary artery bypass grafting (CABG) is one of the most common techniques of restoring coronary blood flow. Thus, there is a growing interest in exploring

the role of factors that influence the risk of adverse outcomes of surgical interventions. According to the American Heart Association, the in-hospital mortality rate after CABG surgery is 1–3% in patients under 70 yrs and 5–6% in those who are older than 70 yrs [1].

In recent years, universal adverse event prediction tools based on large prospective trials have been increasingly used in clinical cardiology.

The EuroSCORE II and Society of Thoracic Surgeons (STS) Score are the classic models of in-hospital death prediction, i.e., within 30 days after CABG. The EuroSCORE II includes 18 predictors characterizing the patient's clinical and functional status before the intervention, and the types and urgency of cardiac surgery [2]. The level of evidence of this model is rated IIa, and the level of confidence is rated B. The 2014 STS risk stratification model uses 40 clinical and 2 angiographic predictors, with level of evidence IB. The European Society of Cardiology (ESC) and the European Association of Cardiothoracic Surgeons (EACTS) working group on myocardial revascularization confirmed in their 2018 guidelines that the EuroSCORE II and STS Score have comparable predictive value in assessing the risk of in-hospital death after CABG. Both prediction systems demonstrated an acceptable validity when used in different cohorts of patients who underwent CABG [3]. Several publications recognized a higher prediction capacity of EuroSCORE II and overestimated risks as calculated using the STS Score [4].

The lack of ideal prediction technologies has stimulated an increasing number of studies that use modern machine learning techniques to develop approaches to improve existing models and to develop new models that stratify the risk of in-hospital death after CABG. Machine learning methods such as decision tree, random forest (RF), support vector machine, naive Bayes classifier, Cox regression, artificial neural network (ANN), and others are used to build new models in addition to the logit regression (LR) procedure used in the classical scores [5, 6]. Improved prediction and risk management should be achieved through interdisciplinary collaboration between physicians and data scientists, possibly resulting in information systems that support making more informed medical decisions.

## Objective

The study aimed to evaluate the accuracy of in-hospital mortality prediction models developed using modern machine learning technologies in patients with CAD after CABG.

## Material and Methods

The study involved a retrospective analysis of the database of 866 electronic medical records (685 male and 181 female patients) aged 35 to 81 yrs (median

(Me) 63 yrs), who underwent CABG for CAD during 2008–2018 in the Cardiovascular Surgery Department of Primorsk Regional Clinical Hospital No. 1, Vladivostok, Russia. All CABG surgeries were performed during cardiopulmonary bypass.

Two groups of patients were identified in the study cohort. Group 1 included 35 (4%) patients who died within the first 30 days after CABG, and Group 2 consisted of 831 (96%) patients with a favorable outcome of the surgery. The cause of death was intra- and post-operative myocardial infarction (MI, 26 patients) and post-operative acute renal failure (6 patients). In other cases, deaths were associated with pancreonecrosis (1), subarachnoid hemorrhage (1), and mediastinitis (1).

At the first stage of the study, a comparative analysis of 99 factors including clinical and laboratory findings in patients of Groups 1 and 2 before CABG, was performed to identify potential predictors that are correlated linearly with in-hospital death using the Student's t-test, Mann-Whitney test, chi-square analysis, and Fisher's exact test. Parameters with statistically significant intergroup differences were selected for further analysis. At the second stage, one-factor, in-hospital mortality models were used to estimate the predictive potential of the identified factors, based on the calculation of weighting coefficients that characterize the impact of individual measures on the risk of death. These models were developed based on normalized data. At the third stage, the study cohort was analyzed for the EuroSCORE II's prediction accuracy using three metrics: sensitivity and specificity, which were calculated with respect to the risk of death >5%, and the area under the ROC-curve (AUC). As there were no patients who were subjected to hemodialysis, who were critical, low-mobility, had active endocarditis, or had severe neurological dysfunction before CABG, the corresponding EuroSCORE II predictors were excluded from the study. At this third stage, original models of in-hospital mortality prediction were developed based on multifactor LR using the EuroSCORE II predictors and additional factors identified in the first and second stages of the study. Cross-validation of the models was performed by multiple randomization of the study cohort into training and test samples at the ratio of 75% and 25%, respectively. The same number of diseased patients were included in the test samples: 8 randomly selected patients. The control validation was also made in a cohort of 45 patients with CAD (20 died within 30 days after CABG and 25 with a favorable outcome of the surgery) who underwent CABG in 2015–2019 in the Cardiovascular Surgery Department of the

Medical Center of the Far Eastern Federal University. The sample of patients with a favorable outcome of the surgery was randomly selected.

At the fourth stage of the study, models of in-hospital lethal outcomes were developed based on the RF and ANN methods. The original prediction accuracy of the models was increased by expanding the range of predictors beyond those included in the EuroSCORE II and by selecting different model parameters. The main parameters of the RF model were 1,000 trees for voting and 6 to 8 signs for splitting. ANN networks had a multi-layered architecture of 3 hidden layers, each containing 7–10 neurons. The models were trained on normalized data. The improvement of the models' quality was the main prerequisite for the inclusion of new predictors. Given the imbalanced samples, the model's improved prediction accuracy was determined under two conditions: increased value of at least one of the three quality metrics or with a stable or increased level of the general metric obtained as a mean between sensitivity and specificity. In our study, the generalized metric corresponded to the model selection result obtained using the Matthews correlation coefficient, which is often used in imbalanced sample tasks [7]. Sensitivity and specificity were tested with respect to the risk of death >5%, which corresponded to high-risk level, according to EuroSCORE II.

Body mass index (BMI), Charlson comorbidity index, and echocardiographic indicators of left ventricular (LV) hypertrophy, i.e., relative wall thickness (RWT) of the LV posterior wall and LV mass index (LVMI), were calculated for all patients. LVMI was normalized to the upper limit of its sex-associated reference value to exclude the influence of sex: 115 g/m<sup>2</sup> for male patients, 95 g/m<sup>2</sup> for female patients.

Data analysis and model development were performed in the R language in R-studio v.1.0.153 and in Python v.3.7.4. The Keras and Tensorflow libraries were used to build the ANN network [13, 14].

## Results

Statistical analyses of the factors characterizing the clinical and functional status of patients with CAD of Groups 1 and 2 before CABG showed that most factors did not differ significantly (Table 1). For example, only 7 of the 18 continuous and categorical variables of EuroSCORE II had significant intergroup differences, including age, left ventricular ejection fraction (LVEF) <30%, LVEF 30–50%, history of MI, peripheral atherosclerosis, emergency CABG, and chronic heart failure (CHF) of functional class (FC) III–IV. To build prediction models, the in-hospital risk factors previously

included in EuroSCORE II were supplemented by new indicators with the highest predictive potential according to the data pre-processing results. These factors included systolic blood pressure (SBP), pulse pressure (PP), Charlson comorbidity index, and the presence of aortic stenosis. The most significant odds ratios (OR), which characterizes levels of influence of certain factors on the probability of in-hospital death, were found in LVEF <30% (OR 11.8, 95% CI 1.4–66.5), LVEF 30–50% (OR 3.0, 95% CI 1.5–6.9), recent history of MI (OR 5.6, 95% CI 2.8–11.4), emergency CABG (OR 5.4, 95% CI 1.7–14.3), FC III–IV CHF (OR 5.1, 95% CI 2.5–10.3), and the presence of aortic stenosis (OR 4.5, 95% CI 1.44–11.73). The influence of other categorical indicators was lower and statistically insignificant. For example, the combination of CAD with chronic obstructive pulmonary disease (COPD) or type 2 diabetes mellitus did not increase the risk of in-hospital death.

The combination of CAD with aortic stenosis increased the likelihood of an adverse outcome, but the combination of CAD with mitral, aortic, or tricuspid valve failure did not affect the immediate results of CABG. Simultaneously, the Charlson index reflecting the severity of concomitant pathology that is associated with older age was significantly higher in patients of Group 1, which is taken into account in the calculation of this indicator. The statistical analyses also showed no inter-group differences in BMI, smoking, HR, diastolic blood pressure (DBP), LV hypertrophy, serum creatinine levels, and creatinine clearance. At the same time, the level of SBP and PP in patients of Group 1 was significantly lower than in Group 2, which can be explained by a greater prevalence of aortic stenosis (14.3% versus 3.6%, respectively) and a lower percentage of hypertension (80% versus 92.4%, respectively).

At the second stage of the study, we constructed single-factor models of in-hospital death and calculated weighting coefficients, which characterize the predictive value of the analyzed indicators, in order to verify the effect of individual risk factors on the risk of in-hospital death (Table 2). This approach significantly enhances the possibilities for processing and analyzing information by a more detailed estimation of potential predictors' influence on the resulting variable. According to this analysis, the maximum weighting coefficient (2.44) is associated with LVEF <30%. Such variables as the recent history of MI, emergency surgery, FC III–IV CHF, LVEF 30–50%, the presence of aortic stenosis, were comparable to it in magnitude and level of confidence. Peripheral artery disease, age, HR, SBP, PP,

**Table 1.** Clinical and functional characteristics of patients with CAD before CABG with reference to the EuroSCORE II and additional predictors

Parameters	Sample size	Group 1, n = 35	Group 2, n = 831	OR (95% CI)	p
Total patients	866	35 (4)	831 (96)	–	–
Female	866	10 (28.6)	171 (20.6)	1.56 (0.7–3.2)	0.35
Male	866	25 (71.4)	660 (79.4)	0.64 (0.31–1.44)	0.36
Age, yrs	866	66.9±5.5	62.9±7.5	–	0.001
Creatinine clearance ≤50 ml/min	687	7 (20)	65 (10%)	2.3 (0.88–5.2)	0.109
Creatinine clearance 50–85 ml/min	687	21 (60)	373 (57.2%)	1.11 (0.53–2.4)	0.97
Peripheral artery disease	840	18 (51.4)	238 (29.6)	2.6 (1.3–5.4)	0.0066
History of heart surgery	840	5 (14.3)	75 (9.3)	1.66 (0.54–4.1)	0.49
Recent history of MI	839	19 (55.9)	140 (17.4)	5.6 (2.8–11.4)	< 0.0001
LVEF 30–50%	797	16 (45.7)	168 (22)	3 (1.5–6.9)	0.0023
LVEF <30%	797	2 (5.7)	4 (0.52)	11.8 (1.4–66.5)	0.013
CHF FC III-IV	854	17 (48.6)	130 (15.6)	5.1 (2.5–10.3)	< 0.0001
Unstable angina	862	16 (45.7)	460 (55.6)	0.67 (0.33–1.33)	0.33
Stable angina FC IV	862	3 (8.6)	30 (3.6)	2.6 (0.56–7.9)	0.3
mPAP 31–55 mmHg	778	7 (20)	111 (14.9)	1.4 (0.56–3.2)	0.56
mPAP ≥55 mmHg	778	1 (2.8)	13 (1.7)	1.8 (0.07–9.8)	1
Emergency CABG	861	5 (14.3)	25 (3)	5.4 (1.7–14.3)	0.002
DM	866	7 (20)	28 (23.6)	0.82 (0.32–1.8)	0.77
COPD	866	6 (17.1)	90 (10.8)	1.7 (0.63–4)	0.37
Serum creatinine*, μmol/l	768	111.13±30	103.5±21	–	0.15
LVEF*, %	797	50.2±11.8	57.7±9.6	–	0.0007
Comorbidity index*	864	5.5±1.61	4.7±1.66	–	0.0049
BMI*, kg/m <sup>2</sup>	771	27.6±6.2	28.8±5.2	–	0.28
Smoking status*	859	12 (34.3)	262 (31.5)	1.1 (0.54–2.3)	0.87
HR*	866	74.5±16.6	70±10	–	0.12
SBP*, mm Hg	866	125.7±17.6	133.2±18.2	–	0.018
DBP*, mm Hg	866	76.7±9.7	79.2±8.1	–	0.15
PP*, mm Hg	866	48.9±10.3	54±13.9	–	0.0078
Hypertension*	861	28 (80)	768 (92.4)	0.3 (0.13–0.78)	0.011
LV posterior wall RWT*	786	0.39±0.1	0.42±0.09	–	0.123
Relative LVMI*	694	1.12±0.346	1.1±0.9	–	0.8
Aortic stenosis*	866	5 (14.3)	30 (3.6)	4.5 (1.44–11.73)	0.0069

Data are M±SD or n (%). The table includes the EuroSCORE II predictors and additional factors with the most significant predictive potential. OR was calculated only for categorical variables; \*, additional factors. CAD, coronary artery disease; CABG, coronary artery bypass grafting; OR, odds ratio; CI, confidence interval; LVEF, left ventricular ejection fraction; CHF, chronic heart failure; FC, functional class; mPAP, mean pulmonary artery pressure; COPD, chronic obstructive pulmonary disease; BMI, body mass index; HR, heart rate; SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure; LV, left ventricle; IRWT, relative wall thickness; LVMI, left ventricular mass index.

and comorbidity index had less evident, yet significant, effects on the immediate outcome of CABG.

At the third stage, the authors developed two optional, multifactor models of in-hospital death (Table 3). In the first model, seven EuroSCORE II predictors were used, and two new factors (HR and aortic stenosis) were added in the second model. The high predictive potential of these two models was confirmed during the model validation. When the

second model was constructed, its predictors, which were continuous variables (age, serum creatinine level, and EF) had better prediction properties than their categorical analogs used in EuroSCORE II. This was confirmed by the quality metrics. At the same time, such EuroSCORE II predictors as sex, CHF FC II, stable angina FC IV, COPD, type 2 DM, mPAP were statistically insignificant in the original multifactor models. The predictive value of a recent history of

MI was insufficient in EuroSCORE II ( $p>0.05$ ), by contrast acquired a significant predictive value in the original model ( $p<0.0001$ ). This can be explained by a larger percentage of patients with a history of MI in Group 1 than in Group 2 (55.9% vs 17.4%). It should also be noted that EuroSCORE II used a combination of linearly dependent indicators of age and creatinine clearance that caused challenging multicollinearity and limited its predictive efficacy. To meet this challenge, linearly independent predictors (age and serum creatinine levels) were used in the original model of in-hospital mortality, which improved the accuracy of prediction. To exclude two correlative signs in a single in-hospital mortality model, similar measures were taken for SBP and PP, which were directly correlated with HR, and with the linearly dependent factors of age and comorbidity index. In the original models, all weighting coefficients of the predictors, except for EF, were positive. This showed an increased risk of in-hospital death if these signs were present or increased (see Table 3). For example, age, serum creatinine levels, or HR are associated with an increased risk of in-hospital death. In contrast, when a negative value of the EF weighting coefficient is increased, it indicates a lower probability of an unfavorable outcome of CABG.

Comparative analysis of indicators of prediction accuracy showed some differences between EuroSCORE II models and the original multifactor LR models (Table 4). It was found that the quality of prognosis by AUC, sensitivity, and specificity for the classic EuroSCORE II with 18 predictors was 0.73, 0.25, and 0.92, respectively. These findings indicate that it lacks accuracy when tested in the analyzed cohort and that it needs to be improved. The original model LR-I, which was designed using seven statistically significant EuroSCORE II predictors, provided some increase in sensitivity (0.74) and AUC (0.83) with lower specificity (0.78) in the test sample. Model LR-II, with additional parameters of HR and aortic stenosis, increased the specificity and AUC to 0.8 and 0.85, respectively. Based on other machine learning methods, the model RF and ANN were developed to obtain a more accurate prognosis. These models were constructed using the EuroSCORE II predictors and their combinations and with additional factors. The RF-I model had worse prediction effectiveness with respect to LR-I and LR-II, but better than the classic EuroSCORE II model, due to its higher sensitivity (0.69 versus 0.25). After the inclusion of additional predictors (HR, SBP, and aortic stenosis), the RF-II model's sensitivity increased to 0.82. The ANN models had significantly higher accuracy as illustrated by best quality metrics. The

**Table 2.** Weighting coefficients of single factor logistic regression models for the assessment of in-hospital mortality risk

Parameter	Coef- ficient	Standard error	P
Age	0.0788	0.0257	0.0021
Female	0.4343	0.3839	0.25793
Creatinine clearance ≤50 ml/min, n (%)	0.8143	0.4423	0.0656
Creatinine clearance 50-85 mL/min, n (%)	0.1151	0.5398	0.7451
Peripheral artery disease	0.9253	0.3469	0.0077
History of heart surgery	0.4838	0.4980	0.3313
History of MI	1.7285	0.3518	0.0000009
LVEF 30-50%	1.0851	0.3503	0.0020
LVEF <30%	2.4423	0.8841	0.0057
FC III-IV CHF	1.6105	0.3515	0.000005
Unstable angina	-0.3977	0.3465	0.2510
Stable angina FC IV	0.9125	0.318	0.1486
mPAP 31–55 mmHg	0.3546	0.4349	0.4148
mPAP ≥55 mmHg	0.5031	1.0525	0.6326
Emergency CABG	1.6752	0.5240	0.0014
Other surgeries excluding heart surgeries	-12.4024	840.2741	0.9882
Type 2 diabetes mellitus	-0.2108	0.4304	0.6243
COPD	0.5327	0.4622	0.2491
Creatinine*, μmol/l	0.01331	0.0065	0.0413
Comorbidity index*	0.2722	0.0948	0.0041
BMI*	-0.0456	0.0355	0.1993
Smoking*	0.0043	0.0097	0.6603
HR*	0.0276	0.0116	0.0172
SBP*	-0.0294	0.0122	0.0160
DBP*	-0.0386	0.0220	0.0796
PP*	-0.0348	0.0163	0.0331
RWT*	-2.2966	2.2428	0.3058
Relative LVMI*	0.0194	0.1734	0.9109
Aortic stenosis*	1.4929	0.5176	0.0039

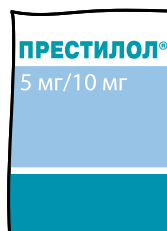
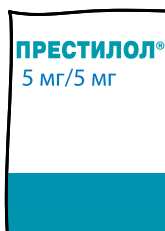
CAD, coronary artery disease; CABG, coronary artery bypass grafting; OR, odds ratio; CI, confidence interval; LVEF, left ventricular ejection fraction; CHF, chronic heart failure; FC, functional class; mPAP, mean pulmonary artery pressure; COPD, chronic obstructive pulmonary disease; BMI, body mass index; HR, heart rate; SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure; LV, left ventricle; RWT, relative wall thickness; LVMI, left ventricular mass index.

maximum level of the prognosis accuracy was reached in the ANN-III model with RWT and relative LVMI included in its structure. In these cases, the AUC was 0.93, sensitivity was 0.90, and specificity was 0.96. The results of the control validation of the original models

Бисопролол / Периндоприл

1 ТАБЛЕТКА\*  
ДЛЯ КОНТРОЛЯ  
АД И ЧСС!<sup>1,2</sup>

БИСОПРОЛОЛ  
+  
ПЕРИНДОПРИЛ



## Артериальная гипертензия

## Стабильная ишемическая болезнь сердца

## Стабильная хроническая сердечная недостаточность\*\*

**1** таблетка  
**В ДЕНЬ**

\* Для компонентов препарата доказано влияние на АД (периндоприл и бисопролол) и ЧСС (бисопролол)

**\*\*Со сниженной систолической функцией левого желудочка у взрослых пациентов, которым показана терапия бисопрололом и периндоприлом в соответствующих дозах**

1. Инструкция по медицинскому применению лекарственного препарата Престилол. Регистрационное удостоверение ЛП-004521 от 31.10.2017.

2. Madej A et al. *Int J Clin Pharmacol Ther*. 2009;47(11): 686-694.

**КРАТКАЯ ИНСТРУКЦИЯ ПО МЕДИЦИНСКОМУ ПРИМЕНЕНИЮ ПРЕПАРАТА ПРЕСТИЛОЛ®**[illegible]

\*Для получения полной информации, пожалуйста, обратитесь к инструкции по медицинскому применению лекарственного препарата.

АО «Сервье». 125196, г. Москва, ул. Лесная, д.7, этаж 7/8/9. Тел.: (495) 937-0700, факс: (495) 937-0701, [www.servier.ru](http://www.servier.ru)  
Материал предназначен для специалистов здравоохранения



**Table 3. Weighting coefficients of the predictors of multifactor LR models for prediction of in-hospital mortality after CABG**

Predictor	Original LR model with the EuroSCORE II predictors		Original LR model with the EuroSCORE II and additional predictors	
	Coefficient (95% CI)	p	Coefficient (95% CI)	p
Age	0.056 (0.0024–0.11)	0.04	0.059 (0.0054–0.12)	0.04
Serum creatinine	0.015 (0–0.03)	0.05	0.015 (0–0.03)	0.07
Peripheral artery disease	0.89 (0.12–1.66)	0.023	1 (0.22–1.81)	0.014
Recent history of MI	1.76 (0.98–2.53)	< 0.0001	1.93 (1.13–2.77)	< 0.0001
EF	-0.042 (-0.078– -0.0065)	0.02	-0.044 (-0.08– -0.008)	0.026
Emergency CABG	1.71 (0.38–3.04)	0.011	1.59 (0.18–2.87)	0.011
FC III-IV CHF	1.69 (0.89–2.53)	< 0.0001	1.73 (0.83–2.53)	< 0.0001
HR*	–	–	0.032 (0.004–0.06)	0.021
Aortic stenosis*	–	–	1.61 (0.25–2.83)	0.0074
Constant	-7.5 (-11.6– -3.31)	0.000045	-10.13 (-15.4– -5.3)	0.00007

Creatinine, EF, and HR are presented as continuous factors; \*, additional factors; EF, ejection fraction; CHF, chronic heart failure; FC, functional class; HR, heart rate. LR, logistic regression; CABG, coronary artery bypass grafting; CI, confidence interval; MI, myocardial infarction.

**Table 4. Estimation of accuracy of models for the prediction of in-hospital mortality after CABG**

Method	Predictor	Test samples			Validation sample		
		AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
LR EuroSCORE II	EuroSCORE II (18)	0.75	0.25	0.92	0.62	0.24	1
LR-I	EuroSCORE II (7)	0.83	0.74	0.78	0.62	0.85	0.53
LR-II	EuroSCORE II (7) + HR + aortic stenosis	0.85	0.7	0.8	0.71	0.67	0.69
RF-I	EuroSCORE II (7)	0.71	0.69	0.71	0.62	0.71	0.53
RF-II	EuroSCORE II (7) + HR + aortic stenosis + SBP	0.78	0.82	0.74	0.67	0.71	0.69
ANN-I	EuroSCORE II (7)	0.85	0.798	0.895	0.71	0.73	0.72
ANN-II	EuroSCORE II (7) + HR + aortic stenosis + SBP	0.89	0.83	0.92	0.74	0.76	0.72
ANN-III	EuroSCORE II (7) + HR + aortic stenosis + SBP + RWT + relative LVMI	0.93	0.90	0.96	0.803	0.857	0.75

CABG, coronary artery bypass grafting; HR, heart rate; SBP, systolic blood pressure; RWT, relative wall thickness; LVMI, left ventricular mass index. The number of the EuroSCORE II predictors are given in parentheses.

in a cohort of patients from another hospital showed their prediction stability, which was confirmed by the quality metrics (see Table 4). At the same time, the ANN-III model had the maximum prediction accuracy, which confirmed the high predictive potential of the indicators of LV hypertrophy.

## Discussion

Machine learning methods are among the main artificial intelligence tools that are increasingly used in various fields of clinical physiology and medicine, including prognostic studies [8]. One of the main advantages of modern machine learning technologies with respect to the classic statistical methods is their ability to process and analyze many variables, identify hidden or unobvious patterns, and extract new knowledge, i.e., data mining. The interest in this area has increased significantly in recent years due to the general trend of digitalizing personalized medicine [9].

In this study, the probability of in-hospital death after CABG was predicted, based on the analysis of patients' pre-operative status, which corresponds to the concept of the EuroSCORE II screening system. The high quality of in-hospital death predictors was provided in our study by a multi-step selection procedure. This procedure included an estimation of their informative value by using the weighting coefficients produced when constructing single-factor of LR models, as well as by traditional statistical methods. This algorithm allowed rating the individual predictors according to their level of influence on the probability of in-hospital death. Correlated variables were excluded from the structure of multifactor LR models due to their challenging multicollinearity, which impaired their predictive potential [10].

These conditions allowed us to increase the predictive value of the models, which was evidenced by the positive changes in quality metrics. However,

it is known that multicollinearity is not a limitation for the RF- and ANN-based models [11]. When constructing the ANN-based models, this allowed use a wider range of additional predictors and not to separate them. For example, HR, SBP, aortic stenosis, indicators of LV hypertrophy, RWT, and relative LVMI, were included in the ANN-III model, as well as the seven factors of EuroSCORE II. Their inclusion in this model allowed us to maximize the model's AUC, specificity, and sensitivity. The non-accidental nature of the combination of these predictors in a single model can be explained by the natural pathophysiological correlation between aortic stenosis and indicators of LV hypertrophy and lower SBP, which was characteristic of patients with an unfavorable outcome of CABG. The clinical value of LV hypertrophy in predicting in-hospital death is associated with a limitation of coronary blood flow reserve, as the number of microvessels per myocardial tissue unit decreases, along with an increase in intracardiac vascular resistance. Moreover, RWT and relative LVMI characterize not only the severity of LV hypertrophy but also types of remodeling that can be used in the future to specify the in-hospital death prognosis. An increase in relative LVMI has been shown previously in patients with CAD to have a higher predictive value for mortality than increased BP, with the worst prognosis associated with concentric LV remodeling [12].

In this study, the high accuracy of the novel LR, RF, and ANN models was confirmed by the multi-step validation process on the test samples, which was shown by increasing values of quality metrics, and by the control validation in the cohort of patients from another hospital. The EuroSCORE II risk assessment system demonstrated a less accurate prognosis than the new models, which was mainly due to its low sensitivity. According to several authors, models have the best prediction capacity in the populations in which

the initial data were obtained [13]. The less effective prognosis in other populations may be due to regional features of the healthcare resource allocation and to other factors. Thus, the availability and regular updating of regional CAD registers and the use of modern machine learning techniques for big data processing and analysis will contribute to more accurate predictions and improvement of special medical care.

## Conclusion

The algorithm developed for processing and analyzing large-scale data characterizing the clinical and functional status of patients with CAD before CABG provided high-quality selection of predictors of in-hospital death. Prediction models based on logistic regression methods, random forest, and artificial neural networks had advantages over the classic EuroSCORE II in accurately calculating the risk of adverse outcomes of CABG. Heart rate, systolic blood pressure, and the presence of aortic stenosis were the comprehensive predictors that determined high prediction accuracy, as well as the seven factors of EuroSCORE II. Artificial neural network model III demonstrated the most significant predictive value due to the predictors «relative wall thickness of the left ventricular posterior wall» and «relative left ventricular mass index.» Relevant research in the future should benefit from the expansion of training samples by collecting data from other treatment facilities. This will increase the models' accuracy.

## Funding

The study was supported by the RFBR as a part of research projects 182903131 and 192901077.

*No conflict of interest is reported.*

**The article was received on 12/05/2020**

## REFERENCES

1. Bokeria L.A., Sokolskaya N.O., Kopylova N.S., Alshibaya M.M. Echocardiographic predictors of the severity of the early postoperative period in patients after surgical myocardial revascularization. Russian Journal of Anaesthesiology and Reanimatology. 2015;60(5):8–11. [Russian: Бокерия Л.А., Сокольская Н.О., Копылова Н.С., Алшибая М.М. Эхокардиографические предикторы тяжести течения раннего послеоперационного периода у больных после хирургической реваскуляризации миокарда. Анестезиология и реаниматология. 2015;60(5):8–11]
2. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR et al. EuroSCORE II. European Journal of Cardio-Thoracic Surgery. 2012;41(4):734–45. DOI: 10.1093/ejcts/ezs043
3. Ad N, Holmes SD, Patel J, Pritchard G, Shuman DJ, Halpin L. Comparison of EuroSCORE II, Original EuroSCORE, and The Society of Thoracic Surgeons Risk Score in Cardiac Surgery Patients. The Annals of Thoracic Surgery. 2016;102(2):573–9. DOI: 10.1016/j.athoracsur.2016.01.105
4. Yamaoka H, Kuwaki K, Inaba H, Yamamoto T, Kato TS, Dohi S et al. Comparison of modern risk scores in predicting operative mortality for patients undergoing aortic valve replacement for aortic stenosis. Journal of Cardiology. 2016;68(2):135–40. DOI: 10.1016/j.jcc.2015.08.017
5. Allyn J, Allou N, Augustin P, Philip I, Martinet O, Belghiti M et al. A Comparison of a Machine Learning Model with EuroSCORE II in Predicting Mortality after Elective Cardiac Surgery: A Decision Curve Analysis. PLOS ONE. 2017;12(1):e0169772. DOI: 10.1371/journal.pone.0169772
6. Kilic A, Goyal A, Miller JK, Gjekmarkaj E, Tam WL, Gleason TG et al. Predictive Utility of a Machine Learning Algorithm in Estimating Mortality Risk in Cardiac Surgery. The Annals of Thoracic Surgery. 2020;109(6):1811–9. DOI: 10.1016/j.athoracsur.2019.09.049

7. Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PLOS ONE*. 2017;12(6):e0177678. DOI: 10.1371/journal.pone.0177678
8. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. *PLOS ONE*. 2018;13(8):e0202344. DOI: 10.1371/journal.pone.0202344
9. Cuocolo R, Perillo T, De Rosa E, Ugga L, Petretta M. Current applications of big data and machine learning in cardiology. *Journal of Geriatric Cardiology*. 2019;16(8):601–7. DOI: 10.11909/j.issn.1671-5411.2019.08.002
10. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Logistic regression. *Perspectives in Clinical Research*. 2017;8(3):148–51. DOI: 10.4103/picr.PICR\_87\_17
11. Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9(1):307. DOI: 10.1186/1471-2105-9-307
12. Luan L, Hu H, Li S-C. A Review of Studies of Quality of Life for Chinese-Speaking Patients with Ischemic Heart Disease. *Value in Health Regional Issues*. 2018;15:82–90. DOI: 10.1016/j.vhri.2017.08.013
13. Alizadehsani R, Abdar M, Roshanzamir M, Khosravi A, Kebria PM, Khozimeh F et al. Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Computers in Biology and Medicine*. 2019;111:103346. DOI: 10.1016/j.compbiomed.2019.103346