



## Возможности методов машинного обучения в стратификации операционного риска у больных ишемической болезнью сердца, направляемых на коронарное шунтирование

Голухова Е. З., Керен М. А., Завалихина Т. В., Булаева Н. И., Акатов Д. С., Сигаев И. Ю., Яхьяева К. Б., Колесников Д. А.

**Цель.** Разработка и оценка эффективности моделей прогнозирования летального исхода после операции коронарного шунтирования, полученных с помощью методов машинного обучения на основании преоперационных данных.

**Материал и методы.** В рамках когортного исследования проводилось ретроспективное прогнозирование вероятности госпитальной летальности после коронарного шунтирования (КШ) у 2182 больных со стабильной ишемической болезнью сердца. Пациенты были разделены на 2 выборки: обучающую (80%, n=1745) и тренировочную (20%, n=437). Исходное соотношение выживших (n=2153) и умерших (n=29) пациентов в общей выборке свидетельствовало о выраженном дисбалансе классов, для преодоления которого в тренировочной выборке был использован метод передискретизации. Для построения прогностических моделей риска использовали пять алгоритмов машинного обучения (МО): логистическая регрессия (Logistic regression), случайный лес (Random Forest), CatBoost, LightGBM, XGBoost. Для каждого из данных алгоритмов на тренировочной выборке проводили кросс-валидацию и поиск гиперпараметров. В результате получили пять прогностических моделей с наилучшими параметрами. Полученные прогностические модели были применены к обучающей выборке, после чего сравнивались их производительность с целью определения наиболее эффективной модели.

**Результаты.** Прогностические модели, реализованные на ансамблевых классификаторах (CatBoost, LightGBM, XGBoost), демонстрировали лучшие результаты, в сравнении с моделями на основе логистической регрессии и случайного леса. Наилучшие метрики качества были получены для моделей на основе CatBoost и LightGBM (Precision — 0,667, Recall — 0,333, F1-мера — 0,444, ROC AUC — 0,666 для обеих моделей). Общими высокоранговыми параметрами для принятия решения об исходе для обеих моделей являлись: уровни креатинина и глюкозы крови, фракция выброса левого желудочка, возраст, критическое поражение (>70%) каротидных артерий и магистральных артерий нижних конечностей.

**Заключение.** Ансамблевые методы МО демонстрируют более высокие возможности прогнозирования исходов в сравнении с традиционными методами МО, например, логистической регрессией. Полученные в исследовании прогностические модели для дооперационного прогнозирования госпитальной летальности у больных, направляемых на КШ, могут служить основой для разработки систем поддержки принятия врачебных решений у больных ишемической болезнью сердца.

**Ключевые слова:** методы машинного обучения, коронарное шунтирование, госпитальная летальность после коронарного шунтирования.

**Отношения и деятельность:** нет.

ФГБУ Научный медицинский исследовательский центр сердечно-сосудистой хирургии им. А. Н. Бакулева Минздрава России, Москва, Россия.

Голухова Е. З. — академик РАН, д.м.н., профессор, директор, ORCID: 0000-0002-6252-0322, Керен М. А.\* — д.м.н., с.н.с. отделения хирургии сочетанных заболеваний коронарных и магистральных артерий, ORCID: 0000-0003-2428-1559, Завалихина Т. В. — к.м.н., зам. главного врача по амбулаторно-клинической работе ИКХ им. В. И. Бураковского, ORCID: 0000-0002-3754-3469, Булаева Н. И. — к.б.н., зав. отделом координации и сопровождения научно-исследовательской деятельности и проведения тематических мероприятий, ORCID: 0000-0002-5091-0518, Акатов Д. С. — к.м.н., врач сердечно-сосудистый хирург отделения реконструктивной хирургии новорожденных и детей первого года жизни с ВПС, ORCID: 0000-0002-8401-2556, Сигаев И. Ю. — д.м.н., профессор, зав. отделением хирургии сочетанных заболеваний коронарных и магистральных артерий, ORCID: 0000-0002-1323-8072, Яхьяева К. Б. — ординатор 2-го года по специальности "кардиология" отделения хирургии сочетанных заболеваний коронарных и магистральных артерий, ORCID: 0000-0001-7928-2247, Колесников Д. А. — ведущий специалист автоматизированной истории болезни, ORCID: 0000-0002-1550-2041.

\*Автор, ответственный за переписку (Corresponding author):

milenamailru@mail.ru

ИБС — ишемическая болезнь сердца, КШ — коронарное шунтирование, МО — машинное обучение, ПМ — прогностическая модель, LR — логистическая регрессия.

Рукопись получена 07.09.2022

Рецензия получена 13.10.2022

Принята к публикации 19.11.2022



**Для цитирования:** Голухова Е. З., Керен М. А., Завалихина Т. В., Булаева Н. И., Акатов Д. С., Сигаев И. Ю., Яхьяева К. Б., Колесников Д. А. Возможности методов машинного обучения в стратификации операционного риска у больных ишемической болезнью сердца, направляемых на коронарное шунтирование. *Российский кардиологический журнал*. 2023;28(2):5211. doi:10.15829/1560-4071-2023-5211. EDN EPCOIT

## Potential of machine learning methods in operational risk stratification in patients with coronary artery disease scheduled for coronary bypass surgery

Golukhova E. Z., Keren M. A., Zavalikhina T. V., Bulaeva N. I., Akatov D. S., Sigaev I. Yu., Yakhyaeva K. B., Kolesnikov D. A.

**Aim.** To develop and evaluate the effectiveness of models for predicting mortality after coronary bypass surgery, obtained using machine learning analysis of preoperative data.

**Material and methods.** As part of a cohort study, a retrospective prediction of in-hospital mortality after coronary artery bypass grafting (CABG) was performed in 2182 patients with stable coronary artery disease. Patients were divided into 2 following samples: learning (80%, n=1745) and training (20%, n=437). The initial ratio of surviving (n=2153) and deceased (n=29) patients in the total sample indicated a pronounced class imbalance, and therefore the resampling method was used in the training sample. Five machine learning (ML) algorithms

were used to build predictive risk models: Logistic regression, Random Forest, CatBoost, LightGBM, XGBoost. For each of these algorithms, cross-validation and hyperparameter search were performed on the training sample. As a result, five predictive models with the best parameters were obtained. The resulting predictive models were applied to the learning sample, after which their performance was compared in order to determine the most effective model.

**Results.** Predictive models implemented on ensemble classifiers (CatBoost, LightGBM, XGBoost) showed better results compared to models based on logistic regression and random forest. The best quality metrics were obtained for CatBoost and LightGBM based models (Precision — 0,667, Recall — 0,333, F1-score —

0,444, ROC AUC — 0,666 for both models). There were following common high-ranking parameters for deciding on the outcome for both models: creatinine and blood glucose levels, left ventricular ejection fraction, age, critical stenosis (>70%) of carotid arteries and main lower limb arteries.

**Conclusion.** Ensemble machine learning methods demonstrate higher predictive power compared to traditional methods such as logistic regression. The prognostic models obtained in the study for preoperative prediction of in-hospital mortality in patients referred for CABG can serve as a basis for developing systems to support medical decision-making in patients with coronary artery disease.

**Keywords:** machine learning methods, coronary artery bypass grafting, inhospital mortality after coronary artery bypass grafting.

**Relationships and Activities:** none.

A. N. Bakulev National Medical Research Center for Cardiovascular Surgery, Moscow, Russia.

Golukhova E. Z. ORCID: 0000-0002-6252-0322, Keren M. A.\* ORCID: 0000-0003-2428-1559, Zavalikhina T. V. ORCID: 0000-0002-3754-3469, Bulaeva N. I. ORCID: 0000-0002-5091-0518, Akatov D. S. ORCID: 0000-0002-8401-2556, Sigaev I. Yu. ORCID: 0000-0002-1323-8072, Yakhyeva K. B. ORCID: 0000-0001-7928-2247, Kolesnikov D. A. ORCID: 0000-0002-1550-2041.

\*Corresponding author:  
milenamailru@mail.ru

**Received:** 07.09.2022 **Revision Received:** 13.10.2022 **Accepted:** 19.11.2022

**For citation:** Golukhova E. Z., Keren M. A., Zavalikhina T. V., Bulaeva N. I., Akatov D. S., Sigaev I. Yu., Yakhyeva K. B., Kolesnikov D. A. Potential of machine learning methods in operational risk stratification in patients with coronary artery disease scheduled for coronary bypass surgery. *Russian Journal of Cardiology*. 2023;28(2):5211. doi:10.15829/1560-4071-2023-5211. EDN EPCOIT

### Ключевые моменты

- Возможность персонального прогнозирования риска смерти при выполнении коронарного шунтирования (КШ) способна значительно облегчить принятие оптимального решения как специалисту, так и пациенту.
- Методы машинного обучения позволяют эффективно прогнозировать возможные сценарии развития событий, в т.ч. оценивать риск развития негативных госпитальных исходов после операции КШ.
- Разработанные прогностические модели на основе ансамблевых классификаторов позволяют обеспечить более эффективное дооперационное прогнозирование госпитальной летальности у больных, направляемых на КШ. Полученные результаты могут служить основой для разработки систем поддержки принятия врачебных решений у больных, направляемых на реваскуляризацию миокарда.

На сегодняшний день в Российской Федерации выполняется ~210 тыс. операций реваскуляризации миокарда, из которых на операции коронарного шунтирования (КШ) приходится ~15% (~30 тыс. в год), что значительно меньше необходимого с учетом текущего роста больных ишемической болезнью сердца (ИБС) [1]. Одна из причин — сложность в принятии решения о необходимости выполнения вмешательства, в т.ч. из-за сниженной доступности медицинской помощи, квалифицированного медицинского персонала и недостаточной осведомленности врачей о потенциальном риске послеоперационных осложнений.

Средние показатели госпитальной смертности после КШ составляют 1,5-3%, но варьируют в зависимости

### Key messages

- The possibility of individual death prediction in coronary artery bypass grafting (CABG) can greatly facilitate the optimal decision-making for both the specialist and the patient.
- Machine learning methods make it possible to effectively predict possible scenarios, including unfavorable in-hospital outcomes after CABG.
- The developed prognostic models based on ensemble classifiers allow for more effective preoperative prediction of in-hospital mortality in patients referred for CABG. The results obtained can serve as a basis for developing systems to support medical decision-making in patients referred for myocardial revascularization.

от срочности выполнения операции, тяжести поражения коронарного русла, возраста, общего состояния, коморбидности пациента и пр. Потенциальные риски, сопровождающие кардиохирургические вмешательства, не всегда точно прогнозируемы, что в сочетании с индивидуальными обстоятельствами негативно сказывается на принятии оптимального решения о дальнейшей тактике лечения. Возможность прогнозирования послеоперационных исходов значительно облегчает задачу специалиста в оценке потенциальных пользы/риска вмешательства у пациента, рассматриваемого в качестве кандидата на реваскуляризацию миокарда.

В настоящее время для практического использования доступно множество шкал стратификации риска смерти после кардиохирургических вмешательств. Среди них шкалы EuroScore 2 и STS являются наиболее изученными и рекомендованы к рутинному использованию в практике [2]. Однако, несмотря на широкое применение, в ряде исследова-

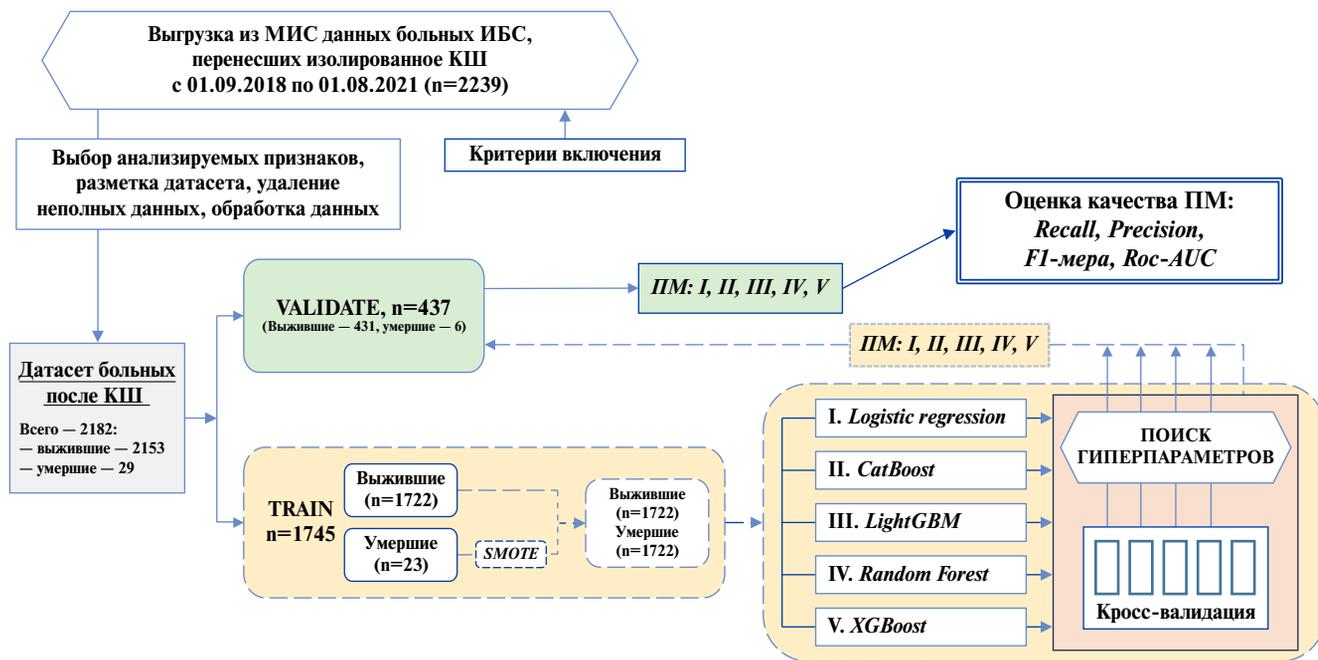


Рис. 1. Концепция разработки ПМ.

**Сокращения:** ИБС — ишемическая болезнь сердца, КШ — коронарное шунтирование, МИС — медицинские информационные системы, ПМ — прогностическая модель.

ний были показаны существенные ограничения этих шкал, заключающиеся в переоценке риска смерти у пациентов низкого риска и недооценке риска у пациентов с высоким риском [3-5].

К основным причинам, отрицательно влияющим на прогностическую точность, можно отнести: 1) низкую локальную адаптацию шкал, т.е. шкалы могут не отражать региональную реальность; 2) "универсальность" шкал, что не позволяет учитывать специфику оперативного вмешательства; 3) "устаревание" клинических баз данных, использованных для создания прежних шкал; 4) неполный учет прогностически значимых признаков; 5) погрешности и ограничения стандартных математических методов расчета риска.

В свою очередь, набирающие популярность методы машинного обучения (МО) имеют значительные преимущества над традиционными методами математического анализа, поскольку позволяют искать как линейные, так и нелинейные зависимости между исходными медицинскими данными и целевым признаком. Методы МО относятся к разновидности искусственного интеллекта и способны прогнозировать возможные сценарии развития событий, понимать их значение, формировать шаблоны распознавания и рассчитывать модели индивидуального риска, что является значительным шагом в сторону развития персонализированной медицины и позволит минимизировать негативные исходы, сопряженные с индивидуальными особенностями пациента. В ряде зарубежных исследований использование методов МО показало более высокую точность

прогнозирования краткосрочной смертности после кардиохирургических операций в сравнении с традиционными шкалами [6-8]. Среди отечественных публикаций также имеются единичные исследования, демонстрирующие результаты применения методов МО для оценки риска при реваскуляризации миокарда [9]. Однако крупных исследований, адаптированных к отечественным условиям оказания кардиохирургической помощи, не проводилось. Целью нашего исследования явилась разработка и оценка эффективности моделей прогнозирования вероятности летального исхода после операции КШ, полученных с помощью методов МО на основании анализа предоперационных данных.

### Материал и методы

Концепция разработки и валидирования прогностических моделей (ПМ) представлена на рисунке 1. В рамках одноцентрового когортного исследования нами был сформирован датасет, содержащий ретроспективный пул медицинских данных 2239 больных ИБС, перенесших КШ с сентября 2018г по август 2021г. Критерии включения больных: выполненное изолированное КШ, как результат госпитализации у больных со стабильным течением ИБС. В исследование не включали больных, перенесших КШ одновременно с протезированием клапанов, резекцией аневризмы левого желудочка, реваскуляризацией других сосудистых бассейнов и пр. Данные пациентов были выгружены из медицинских информационных систем и обезличены. После формирования

Таблица 1

## Исходная характеристика исследуемых групп

	Все больные после КШ, n=2182 (%)	Выжившие после КШ, n=2153 (%)	Умершие после КШ, n=29 (%)	p-value
Возраст, годы	64 (58-69)	64 (58-69)	65 (61-69)	0,318
Женский пол	442 (20,3)	437 (20,3)	5 (17,2)	0,684
Постинфарктный кардиосклероз	1268 (58,1)	1247 (57,9)	21 (72,4)	0,116
Стенокардия ФК:				
1 ФК	8 (0,04)	8 (0,4)	0 (0,0)	0,711
2 ФК	422 (19,3)	418 (19,4)	4 (13,8)	
3 ФК	1518 (69,6)	1495 (69,4)	23 (79,3)	
4 ФК	234 (10,7)	232 (10,8)	2 (6,9)	
Артериальная гипертензия	2061 (94,5)	2033 (94,4)	28 (96,6)	0,619
ОХС, ммоль/л	4,16 (3,50-5,09)	4,17 (3,50-5,10)	3,86 (3,59-4,29)	0,259
ЛНП, ммоль/л	2,33 (1,80-3,10)	2,34 (1,80-3,11)	2,29 (1,81-2,7)	0,778
СД 2 типа, всего	534 (24,5)	527 (24,5)	7 (24,1)	0,966
— инсулинозависимый СД 2 типа	57 (2,6)	55 (2,6)	2 (6,9)	0,308
Средний ИМТ, кг/м <sup>2</sup>	28,8 (26,07-32,05)	28,9 (26,12-32,1)	27,31 (24,5-29,5)	0,037
ИМТ ≥30, кг/м <sup>2</sup>	888 (40,7)	882 (41,0)	6 (20,7)	0,027
Курильщики	751 (34,4)	740 (34,4)	11 (37,9)	0,689
Фракция выброса левого желудочка, %	56 (52-59)	56 (52-59)	48 (40-50)	0,001
Критический стеноз (≥70%) внутренних сонных артерий	242 (11,1)	236 (11,0)	6 (20,7)	0,097
Критический стеноз (≥70%) магистральных артерий нижних конечностей	286 (13,1)	275 (12,8)	11 (37,9)	<0,001
Мультифокальный атеросклероз	767 (35,2)	751 (34,9)	16 (55,2)	0,023
Креатинин, мкмоль/л	79,10 (70-93,7)	79,0 (70-93,3)	101,2 (70,6-157)	0,012
Глюкоза, ммоль/л	6,0 (5,0-7,0)	6,0 (5,0-7,0)	6,0 (5,0-7,0)	0,499
СКФ, мл/мин/1,75 м <sup>2</sup>	85,7 (70,25-100,41)	85,9 (70,4-100,4)	66,9 (30,95-97,7)	0,006
КШ в анамнезе	50 (2,0)	46 (2,1)	4 (13,8)	<0,001
ЧКВ в анамнезе	649 (29,7)	638 (29,6)	11 (37,9)	0,332
Однососудистое поражение КА	195 (8,9)	194 (9,0)	1 (3,4)	0,297
Двухсосудистое поражение КА	706 (32,4)	701 (32,6)	5 (17,2)	0,080
Трёхсосудистое поражение КА и/или поражение ствола ЛКА	1278 (58,6)	1256 (58,3)	22 (75,9)	0,057
EuroSCORE I (n=779)	3,0 (2,0-5,0)	3,0 (2,0-5,0)	3,0 (2,75-3,25)	0,452
EuroSCORE II (n=1403)	1,44 (0,89-2,48)	1,44 (0,89-2,43)	2,96 (1,55-6,22)	<0,001

**Сокращения:** ИМТ — индекс массы тела, КА — коронарная артерия, КШ — коронарное шунтирование, ЛКА — левая коронарная артерия, ЛНП — липопротеиды низкой плотности, ОХС — общий холестерин, СД — сахарный диабет, СКФ — скорость клубочковой фильтрации, ФК — функциональный класс, ЧКВ — чрескожное коронарное вмешательство.

и разметки датасета проводили удаление всех случаев, содержащих неполные данные.

Итоговый вариант включил 2182 больных, перенесших КШ. В нашем исследовании проблема выраженного дисбаланса классов (число выживших (n=2153) значительно превосходило число умерших (n=29)) являлась серьезным барьером для получения эффективных ПМ, поскольку приводила к их переобучению. В таких ситуациях качество моделей может быть неудовлетворительным, т.к. граница решения, вероятно, будет резко смещаться из-за несбалансированных данных [10]. Для решения проблем несбалансированной выборки в первую очередь прибегают к методам передискретизации [11], среди которых наиболее эффективным методом является SMOTE (Synthetic Minority Over-sampling Technique) [12]. SMOTE способен устранять дисбаланс классов методом увеличения числа наблюдений в минори-

тарном классе за счет создания новых синтетических выборок, образованных путем искусственного формирования случаев с похожим набором наблюдаемых признаков, значения которых лежат в границах таковых для представителей миноритарного класса (интерполяция между второстепенными выборками). По мнению авторов, проведение передискретизации снижает вероятность переобучения и повышает эффективность обобщения классификатора на тестовой выборке [7].

Исходно общая группа больных (n=2182) была случайно разделена на 2 выборки в соотношении 80%:20% — для обучения (TRAIN) (n=1745, из них 1722 — выжившие и 23 — умершие) и для тестирования моделей (VALID) (n=437, из них 431 — выжившие, 6 — умершие) (рис. 1).

На этапе обучения в выборке TRAIN была проведена передискретизация с помощью SMOTE для

Таблица 2

Метрики качества ПМ, полученных после применения к обучающей выборке (VALID)

ПМ	Методы МО	Precision	Recall	F1-мера	ROC-AUC
ПМ I	Logistic Regression	0,143	0,333	0,200	0,653
ПМ II	Catboost	0,667	0,333	0,444	0,666
ПМ III	LightGBM	0,667	0,333	0,444	0,666
ПМ IV	Random forest	0,500	0,167	0,250	0,582
ПМ V	XGBoost	0,400	0,333	0,364	0,663

Сокращения: МО — машинное обучение, ПМ — прогностическая модель.

миноритарного класса (умершие), в результате чего класс "умерших" количественно сравнялся с классом "выжившие". После передискретизации датасет содержал данные 3444 больных (1722 — выжившие, 1722 — умершие (синтетически сформированная выборка)).

Для разработки моделей прогнозирования летального исхода после КШ мы использовали пять алгоритмов МО: логистическую регрессию (LR), Random Forrest, CatBoost, LightGBM, XGBoost.

В выборке TRAIN для каждого из алгоритмов МО проводили 5-фолдовую кросс-валидацию (k-Fold=5) и поиск гиперпараметров (параметров, рекомендуемых программой на основании математических расчетов, для оптимального управления процессом обучения модели). В качестве потенциальных гиперпараметров рассматривались исходные клинические признаки (представлены в таблице 1).

Для проведения кросс-валидации обучение моделей проводили на 80% (n=2755) и тестировали на 20% (n=689) выборки TRAIN. Результаты, полученные для каждого из алгоритмов, усредняли. В итоге на выборке TRAIN для каждого из пяти алгоритмов получали ПМ с усредненными гиперпараметрами.

Таким образом, для прогнозирования исхода после КШ нами было построено 5 ПМ: ПМ I — реализована с помощью LR, ПМ II — с помощью CatBoost, ПМ III — с помощью LightGBM, ПМ IV — с помощью Random forest, ПМ V — с помощью XGBoost.

Далее проводили тестирование всех моделей в группе VALID (n=437, из которых 431 — выжившие и 6 — умершие) (истинная группа, без применения передискретизации).

Результаты эффективности полученных моделей сравнивались между собой путем анализа метрик качества, рекомендованных к применению при несбалансированных данных: 1. Recall (полнота) — отражает долю предсказанных алгоритмом объектов положительного класса из всех объектов класса; 2. Precision (точность) — показывает долю объектов, названных алгоритмом положительными и при этом действительно являющимися положительными; 3. F1-мера — отображает гармоническое среднее на

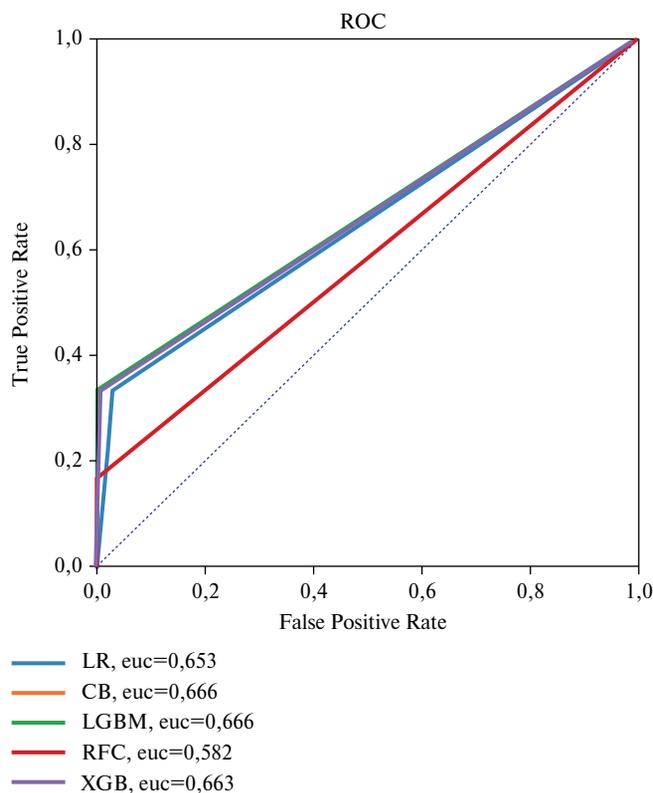


Рис. 2. Roc-кривые полученных моделей на обучающей выборке (VALID).

Примечание: цветное изображение доступно в электронной версии журнала.

основе Precision/Recall. Также для оценки зависимости верно классифицируемых объектов положительного класса от ложноположительных классифицируемых объектов негативного класса рассчитывалось значение ROC-AUC с построением Roc-кривых. Для объективной визуализации результатов работы каждой из моделей МО проводили построение матриц несоответствия (матрицы ошибок).

Построение ПМ, оценка их производительности и визуализация данных проводились с использованием пакета программ JUPYTER Notebook с помощью интерпретатора Python 3.9.12 на операционной системе Windows 10 Pro.

**Статистика.** При анализе исходных данных количественные показатели описывались с помощью ме-

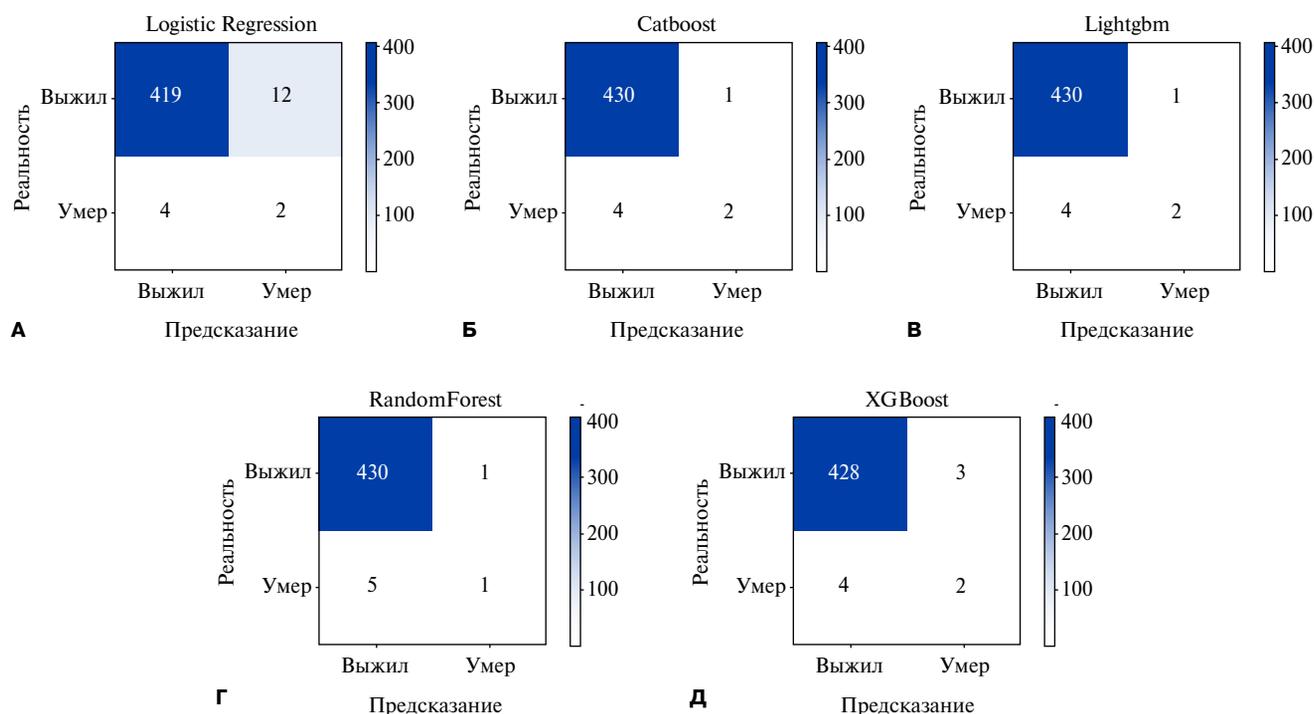


Рис. 3. Матрица ошибок при прогнозировании риска смерти после КШ с помощью различных алгоритмов МО (А, Б, В, Г, Д), полученных на обучающей выборке (VALID).

дианы (Me) и нижнего и верхнего квартилей (Q1-Q3), а категориальные показатели — с указанием абсолютных значений и процентных долей. Сравнение двух групп по количественному показателю, распределение которого отличалось от нормального, выполнялось с помощью U-критерия Манна-Уитни. Сравнение процентных долей при анализе четырехпольных таблиц сопряженности выполнялось с помощью критерия хи-квадрат Пирсона. Статистический анализ проводился с использованием программы StatTech v.2.8.8 (Статтех, Россия).

Исследование было выполнено в соответствии со стандартами надлежащей клинической практики (Good Clinical Practice) и принципами Хельсинкской декларации. Протокол исследования был одобрен Этическим комитетом НМИЦ ССХ им. А. Н. Бакулева. До включения в исследование у всех участников было получено письменное информированное согласие.

### Результаты

Госпитальная смертность после КШ составила 1,32% (n=29). Пациенты, умершие после КШ, реже страдали ожирением (p=0,027), имели более низкую общую фракцию выброса левого желудочка (p<0,001), чаще имели критическое поражение артерий нижних конечностей (p<0,001), мультифокальный атеросклероз (p=0,023), более высокий уровень креатинина (p=0,012) и более низкие значения скорости клубочковой фильтрации (p=0,006), чаще подвергались повторному КШ (p<0,001) в сравне-

нии с выжившими. Средний риск смерти после КШ в группе умерших больных был статистически значимо выше по шкале EuroSCORE II (p<0,001) (табл. 1).

Эффективность ПМ, полученных после применения алгоритмов к обучающей выборке, представлена в таблице 2. Большинство моделей демонстрировали сходную способность к прогнозированию исходов положительного класса среди общего класса (Recall — 0,333), за исключением ПМ IV, имевшей более низкий Recall — 0,167 и как следствие более низкое значение F1-меры (F1-мера — 0,25). Наилучшие метрики качества были получены для ПМ II и ПМ III. При этом обращало на себя внимание, что они продемонстрировали сходную производительность (табл. 2). Метрики качества ПМ V оказались несколько хуже, чем у ПМ II и ПМ III. В свою очередь, ПМ I значительно уступала остальным моделям в прогнозировании исходов положительного класса среди истинно положительного класса (Precision — 0,143), что привело к наиболее низкому показателю F1-меры (F1-мера — 0,2) среди всех моделей (табл. 2).

На рисунке 2 показаны ROC-кривые полученных моделей. Наиболее высокие значения ROC-AUC были получены для ПМ II и ПМ III (AUC — 0,666 для обеих моделей). Самые низкие значения ROC-AUC наблюдались у ПМ IV (AUC — 0,582).

На рисунке 3 показаны возможности ПМ в прогнозировании исходов после КШ на основании матрицы несоответствия алгоритмов МО.

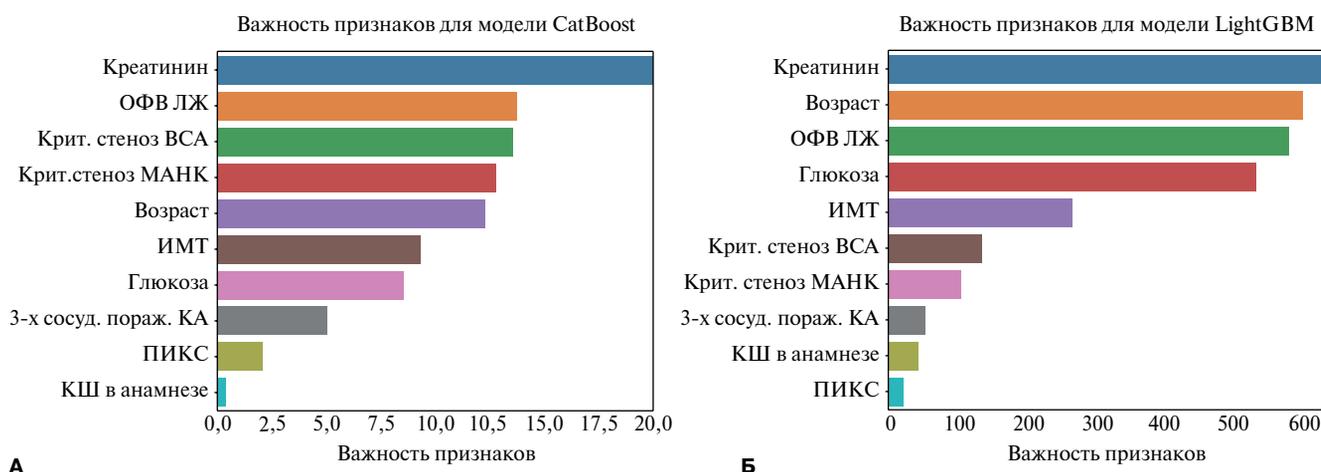


Рис. 4. Прогностическая важность признаков, обеспечивающих принятие решения для ПМ II (А) и ПМ III (Б).

**Сокращения:** ВСА — внутренняя сонная артерия, ИМТ — индекс массы тела, КШ — коронарное шунтирование, МАНК — магистральные артерии нижних конечностей, ОФВ ЛЖ — общая фракция выброса левого желудочка, 3-х сосуд. поражение КА — трехсосудистое поражение магистральных коронарных артерии и/или стеноз ствола левой коронарной артерии, ПИКС — постинфарктный кардиосклероз.

В целом все модели, кроме ПМ IV (рис. 3 Г), показали схожую эффективность к прогнозированию смертельных исходов, однако различались в прогнозировании исходов у выживших пациентов. Однако у ПМ I (рис. 3 А) наблюдалась наиболее выраженная тенденция к переоценке риска смерти, в сравнении с остальными моделями. Наиболее сбалансированный результат по совокупному числу ошибок 1-го и 2-го рода в прогнозировании послеоперационных исходов показали ПМ II (рис. 3 Б) и ПМ III (рис. 3 В).

Таким образом, данные таблицы 2 и рисунок 3 демонстрируют, что ансамблевые алгоритмы CatBoost, LightGBM и XGBoost имели более высокую прогностическую эффективность в сравнении с более простыми методами (LR, RandomForest).

На рисунке 4 показана значимость клинических признаков, определяющих точность прогнозирования в матрице ошибок для моделей с наилучшей производительностью ПМ II и ПМ III. При этом обе модели отнесли такие показатели, как уровень креатинина крови, фракцию выброса левого желудочка, возраст, критическое поражение (>70%) каротидных артерий и магистральных артерий нижних конечностей, трехсосудистое поражение коронарных артерий, перенесенное КШ в анамнезе и постинфарктный кардиосклероз, к наиболее значимым параметрам для принятия решения об исходе.

### Обсуждение

Активное внедрение методов МО для реализации поддержки принятия клинических решений набирает обороты в связи с формирующимся запросом на развитие цифровых технологий и курсом на персонализированную медицину. В современной кардио-

логии и кардиохирургии стратификация предоперационного риска, при наличии показаний к операции реваскуляризации миокарда, играет важную роль для определения оптимальной для пациента тактики решения, а именно, принятия решения в пользу хирургического или эндоваскулярного метода или полного отказа от хирургического лечения.

Целью данного исследования стало изучение возможностей методов МО в прогнозировании риска госпитальной смерти после КШ на основании дооперационных данных. Полученные результаты свидетельствуют об удовлетворительной способности моделей МО прогнозирования исхода после КШ. В нашем исследовании наилучшие метрики качества наблюдались при применении методов МО, реализованных при помощи ансамблевых классификаторов (Catboost и LightGBM). Для данных моделей были выявлены признаки, определяющие неблагоприятный исход после КШ, что может иметь большое значение для концептуального перехода от традиционной модели к персонализированной оценке риска.

Наиболее часто используемыми шкалами стратификации в кардиохирургии являются EuroScore II и STS [2], реализация которых основана на применении LR. Однако, по мнению ряда авторов, данные шкалы обладают низкой дискриминационной способностью, что снижает эффективность их применения [13]. Также, по мнению исследователей, ПМ, локально реализованные с помощью методов МО, имеют большую эффективность в сравнении с EuroScore II и STS, что делает их более предпочтительными в использовании [6, 7, 14].

В нашем исследовании более сложные методы МО превосходили возможности традиционной статистической модели риска на основе LR. Важно под-

черкнуть, что модель, реализованная на основе LR (ПМ I), в сравнении с остальными методами МО, обладала сопоставимой способностью предсказывать исход в группе "умерших", однако значительно завышала риски в группе "выживших", что подтверждает мнение о более высокой частоте ложноотрицательных ошибок у моделей на основе LR [3-5], и в итоге проявляется в переоценке риска смерти в общей выборке. Важно учитывать, что LR наиболее подходит для решения задач, где вероятность развития события линейно связана с независимыми факторами риска. Однако LR неспособна учитывать нелинейные связи между признаками, что делает более перспективными другие методы МО, выявляющие подобные закономерности.

**Ограничения исследования.** К ограничениям данного исследования следует отнести: 1) лимитированное число наблюдений, что являлось критичным в первую очередь для миноритарной группы умерших после КШ и привело нас к необходимости применения передискредитации для борьбы с дисбалансом данных; 2) лимитированное число доступных для анализа медицинских данных, что было обусловлено сложностью интерпретации и автоматического перевода текстовых данных из электронных медицинских карт в цифровой формат, необходимый для анализа; 3) проведение одноцентрового исследования — полученные данные могут отражать предоперационные риски КШ в условиях центра, реализовавшего модель, однако при внешней валидации

моделей (в условиях других кардиохирургических клиник) эффективность модели, вероятно, будет снижена.

### Заключение

Для прогнозирования вероятности смерти после КШ с помощью методов МО нами было разработано пять ПМ на основе LR, случайного леса, CatBoost, LightGBM и XGBoost. Наилучшую производительность демонстрировали ПМ, построенные на основе ансамблевых классификаторов (CatBoost и LightGBM), что в дальнейшем может быть использовано для реализации нового подхода к персонализированной оценке операционного риска для больных, имеющих показания в реваскуляризации миокарда.

Для внедрения этих результатов в практику кардиохирургических стационаров необходимо проведение более масштабных исследований со значительным увеличением числа больных и расширением доступных для анализа данных, а также последующей валидацией полученных моделей на внешней выборке пациентов. Полученные ПМ могут служить основой для реализации систем поддержки принятия врачебных решений у больных, нуждающихся в кардиохирургической помощи по поводу ИБС.

**Отношения и деятельность:** все авторы заявляют об отсутствии потенциального конфликта интересов, требующего раскрытия в данной статье.

### Литература/References

1. Bokeria LA, Milievskaia EB, Pryanishnikov VV, et al. Cardiovascular surgery — 2020. Diseases and congenital anomalies of the circulatory system. Moscow: Bakulev Scientific Center for Cardiovascular Surgery, 2021. p.294. (In Russ.) Бокерия Л.А., Милевская Е.Б., Прянисников В.В. и др. Сердечно-сосудистая хирургия — 2020. Болезни и врожденные anomalies системы кровообращения. М.: НМИЦ ССХ им. А.Н. Бакулева Минздрава России; 2021. p.294. ISBN: 978-5-7982-0425-0.
2. ESC/EACTS recommendations on myocardial revascularization 2018. Russian Journal of Cardiology. 2019;(8):151-226. (In Russ.) Рекомендации ESC/EACTS по реваскуляризации миокарда 2018. Российский кардиологический журнал. 2019;(8):151-226. doi:10.15829/1560-4071-2019-8-151-226.
3. Kunt AG, Kurtcephe M, Hidiroglu M, et al. Comparison of original EuroSCORE, EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort. Interact Cardiovasc Thorac Surg. 2013;16:625-9. doi:10.1093/icvts/ivt022.
4. Mejia OAV, Borgomoni GB, Zubelli JP, et al. REPLICCAR Study Group. Validation and quality measurements for STS, EuroSCORE II and a regional risk model in Brazilian patients. PloS One. 2020;10;15(9):e0238737. doi:10.1371/journal.pone.0238737.
5. Provenchère S, Chevalier A, Ghodbane W, et al. Is the EuroSCORE II reliable to estimate operative mortality among octogenarians? PloS One. 2017;12:e0187056. doi:10.1371/journal.pone.0187056.
6. Nistal-Nuño B. Machine learning applied to a Cardiac Surgery Recovery Unit and to a Coronary Care Unit for mortality prediction. J Clin Monit Comput. 2021;36(3):751-63. doi:10.1007/s10877-021-00703-2.
7. Fernandes MPB, Armengol de la Hoz M, Rangasamy V, et al. Machine learning models with preoperative risk factors and intraoperative hypotension parameters predict mortality after cardiac surgery. J Cardiothorac Vasc Anesth. 2021;35:857-65. doi:10.1053/j.jvca.2020.07.029.
8. Allyn J, Allou N, Augustin P, 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:e0169772. doi:10.1371/journal.pone.0169772.
9. Geltser BI, Shahgeldyan KJ, Rublev VY, et al. Machine Learning Methods for Prediction of Hospital Mortality in Patients with Coronary Heart Disease after Coronary Artery Bypass Grafting. Kardiologia. 2020;60(10):38-46. (In Russ.) Гельцер Б.И., Шахгельдян К.И., Рублев В.Ю. и др. Методы машинного обучения в прогнозировании летальных исходов в стационаре у больных ишемической болезнью сердца после коронарного шунтирования. Кардиология. 2020;60(10):38-46. doi:10.18087/cardio.2020.10.n1170.
10. Douzas G, Vacao F, Last F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. Inf Sci. 2018;465:1-20. doi:10.1016/j.ins.2018.06.056.
11. Branco P, Torgo L, Ribeiro RP. A survey of predictive modeling on imbalanced domains. ACM Comput Surv. 2016;11:49(2):1-50. doi:10.1145/2907070.
12. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res. 2002;01;16:321-57. doi:10.1613/jair.953.
13. Borde D, Gandhe U, Hargave N, et al. The application of European system for cardiac operative risk evaluation II (EuroSCORE II) and Society of Thoracic Surgeons (STS) risk-score for risk stratification in Indian patients undergoing cardiac surgery. Ann Card Anaesth. 2013;16(3):163-6. doi:10.4103/0971-9784.114234.
14. Singh N, Gimpel D, Manikavasagar V, et al. Performance of the AusSCORE II and STS Score for Coronary Artery Bypass Grafting in a New Zealand Population. Heart Lung Circ. 2021;30(4):600-4. doi:10.1016/j.hlc.2020.08.021.