

## Фенотипирование факторов риска и прогнозирование внутригоспитальной летальности у больных ишемической болезнью сердца после коронарного шунтирования на основе методов объяснимого искусственного интеллекта

Гельцер Б. И.<sup>1</sup>, Шахгельдян К. И.<sup>1,2</sup>, Рублев В. Ю.<sup>1</sup>, Домжалов И. Г.<sup>1</sup>, Циванюк М. М.<sup>1</sup>, Шекунова О. И.<sup>1</sup>

**Цель.** Разработать прогностические модели внутригоспитальной летальности (ВГЛ) у больных ишемической болезнью сердца после коронарного шунтирования (КШ) с учетом результатов фенотипирования дооперационных факторов риска.

**Материал и методы.** Проведено ретроспективное исследование по данным 999 электронных историй болезни пациентов (805 мужчин, 194 женщины) в возрасте от 35 до 81 года с медианой (Me) 63 года, которым выполнялось плановое изолированное КШ в условиях искусственного кровообращения. Выделено две группы пациентов, первая из которых была представлена 63 (6,3%) больными, умершими в стационаре в течение первых 30 сут. после КШ, вторая — 936 (93,7%) с благоприятным исходом операции. Дооперационный клинико-функциональный статус оценивали с помощью 102 факторов. Для обработки и анализа данных использовали методы хи-квадрат, Фишера, Манна-Уитни. Пороговые значения предикторов определялись комплексом методов, включающих максимизацию отношений истинно положительных случаев ВГЛ к ложноположительным. Для разработки прогностических моделей применяли многофакторную логистическую регрессию (МЛР). Точность моделей оценивали по 3 метрикам: площадь под ROC-кривой (AUC), чувствительность (Sen) и специфичность (Spec).

**Результаты.** Анализ показателей дооперационного статуса пациентов позволил выделить 28 факторов риска ВГЛ, объединённых в 7 фенотипов. Последние формировали признаковое пространство прогностической модели ВГЛ, в котором каждый признак демонстрирует соответствие пациента определённому фенотипу факторов риска. Авторская модель МЛР отличалась высокими метриками качества (AUC — 0,91; Sen — 0,9 и Spec — 0,85).

**Заключение.** Разработанный алгоритм обработки и анализа данных обеспечил высокое качество выделения дооперационных факторов риска и прогнозирования ВГЛ после КШ. Перспективы дальнейших исследований по данной проблеме связаны с совершенствованием технологий объяснимого искусственного интеллекта, позволяющих разрабатывать информационные системы по управлению рисками, востребованные в повседневной клинической практике.

**Ключевые слова:** коронарное шунтирование, внутригоспитальная летальность, фенотипирование факторов риска, прогностические модели, объяснимый искусственный интеллект.

**Отношения и деятельность.** Работа выполнена при финансовой поддержке проекта FZNS-2023-0010 Госзадания ДВФУ.

<sup>1</sup>ФГАОУ ВО Дальневосточный федеральный университет, Школа медицины, Владивосток; <sup>2</sup>ФГБОУ ВО Владивостокский государственный университет, Институт информационных технологий, Владивосток, Россия.

Гельцер Б. И. — д.м.н., профессор, член-корр. РАН, зам. директора, ORCID: 0000-0002-9250-557X, Шахгельдян К. И. — д.т.н., доцент, директор, зав. лабораторией анализа больших данных в биомедицине и здравоохранении, ORCID: 0000-0002-4539-685X, Рублев В. Ю. — ассистент департамента клинической медицины, врач сердечно-сосудистый хирург, ORCID: 0000-0001-7620-4454, Домжалов И. Г. — аспирант, врач анестезиолог-реаниматолог, ORCID: 0000-0002-6722-2535, Циванюк М. М.\* — м.н.с. департамента клинической медицины, врач по рентгенохирургическим методам диагностики и лечения, кардиолог, ORCID: 0000-0003-3545-3862, Шекунова О. И. — к.м.н., доцент, ORCID: 0000-0003-3054-3797.

\*Автор, ответственный за переписку (Corresponding author): m\_tsivanyuk@list.ru

АЧТВ — активированное частичное тромбопластиновое время, ВГЛ — внутригоспитальная летальность, ГСД — градиент систолического давления, ИБС — ишемическая болезнь сердца, ИИ — ишемический инсульт, ИМ — инфаркт миокарда, ИММ — индекс массы миокарда, ИОТ — индекс относительной толщины, КДО — конечный диастолический объем, КК — клиренс креатинина, КШ — коронарное шунтирование, ЛЖ — левый желудочек, МЛР — многофакторная логистическая регрессия, МО — машинное обучение, НИМ — недавний инфаркт миокарда, ОИИ — объяснимый искусственный интеллект, ОИММ — относительный индекс массы миокарда, ПТИ — протромбиновый индекс, РЛП1 — продольный размер левого предсердия, РЛП2 — поперечный размер левого предсердия, РПП1 — продольный размер правого предсердия, РПП2 — поперечный размер правого предсердия, СрДЛА — среднее давление легочной артерии, ТВ — тромбиновое время, ФВ — фракция выброса, ФК — функциональный класс, ХСН — хроническая сердечная недостаточность, AUC — площадь под ROC-кривой, FP — ложноположительный случай (False Positive), Sen — чувствительность, Spec — специфичность, TP — положительный случай (True Positive).

Рукопись получена 30.11.2022

Рецензия получена 08.01.2023

Принята к публикации 20.02.2023



**Для цитирования:** Гельцер Б. И., Шахгельдян К. И., Рублев В. Ю., Домжалов И. Г., Циванюк М. М., Шекунова О. И. Фенотипирование факторов риска и прогнозирование внутригоспитальной летальности у больных ишемической болезнью сердца после коронарного шунтирования на основе методов объяснимого искусственного интеллекта. *Российский кардиологический журнал*. 2023;28(4):5302. doi:10.15829/1560-4071-2023-5302. EDN UGDPSE

## Phenotyping of risk factors and prediction of inhospital mortality in patients with coronary artery disease after coronary artery bypass grafting based on explainable artificial intelligence methods

Geltser B. I.<sup>1</sup>, Shakhgelyan K. I.<sup>1,2</sup>, Rublev V. Yu.<sup>1</sup>, Domzhalov I. G.<sup>1</sup>, Tsivanyuk M. M.<sup>1</sup>, Shekunova O. I.<sup>1</sup>

**Aim.** To develop predictive models of inhospital mortality (IHM) in patients with coronary artery disease after coronary artery bypass grafting (CABG), taking into account the results of phenotyping of preoperative risk factors.

**Material and methods.** This retrospective study was conducted based on the data of 999 electronic health records of patients (805 men, 194 women) aged 35 to 81

years with a median (Me) of 63 years who underwent on-pump elective isolated CABG. Two groups of patients were distinguished, the first of which was represented by 63 (6,3%) patients who died in the hospital during the first 30 days after CABG, the second — 936 (93,7%) with a favorable outcome. Preoperative clinical and functional status was assessed using 102 factors. Chi-squares, Fisher, Mann-Whitney methods

were used for data processing and analysis. Threshold values of predictors were determined by methods, including maximizing the ratio of true positive IHM cases to false positive ones. Multivariate logistic regression (MLR) was used to develop predictive models. Model accuracy was assessed using 3 following metrics: area under the ROC curve (AUC), sensitivity (Sens), and specificity (Spec).

**Results.** An analysis of preoperative status of patients made it possible to identify 28 risk factors for IHM, combined into 7 phenotypes. The latter formed the feature space of IHM prognostic model, in which each feature demonstrates the patient's compliance with a certain risk factor phenotype. The author's MLR model had high quality metrics (AUC-0,91; Sen-0,9 and Spec-0,85).

**Conclusion.** The developed data processing and analysis algorithm ensured high quality of preoperative risk factors identification and IHM prediction after CABG. Prospects for further research on this issue are related to the improvement of explainable artificial intelligence technologies, which allow developing information systems for managing clinical practice risks.

**Keywords:** coronary artery bypass grafting, in-hospital mortality, risk factor phenotyping, predictive models, explainable artificial intelligence.

**Relationships and Activities.** The work was supported by the FZNS-2023-0010 project of the Far Eastern Federal University State Assignment.

<sup>1</sup>Far Eastern Federal University, School of Medicine, Vladivostok; <sup>2</sup>Vladivostok State University, Institute of Information Technologies, Vladivostok, Russia.

Geltser B. I. ORCID: 0000-0002-9250-557X, Shakhgeldyan K. I. ORCID: 0000-0002-4539-685X, Rublev V. Yu. ORCID: 0000-0001-7620-4454, Domzhalov I. G. ORCID: 0000-0002-6722-2535, Tsivanyuk M. M.\* ORCID: 0000-0003-3545-3862, Shekunova O. I. ORCID: 0000-0003-3054-3797.

\*Corresponding author:  
m\_tsivanyuk@list.ru

**Received:** 30.11.2022 **Revision Received:** 08.01.2023 **Accepted:** 20.02.2023

**For citation:** Geltser B. I., Shakhgeldyan K. I., Rublev V. Yu., Domzhalov I. G., Tsivanyuk M. M., Shekunova O. I. Phenotyping of risk factors and prediction of in-hospital mortality in patients with coronary artery disease after coronary artery bypass grafting based on explainable artificial intelligence methods. *Russian Journal of Cardiology*. 2023;28(4):5302. doi:10.15829/1560-4071-2023-5302. EDN UGDPS

### Ключевые моменты

- Фенотипирование факторов риска является новым подходом к разработке прогностических моделей на основе методов машинного обучения.
- Использование фенотипов в качестве предикторов прогностических моделей позволяет более точно оценить вероятность неблагоприятных исходов коронарного шунтирования.
- Применение технологий объяснимого искусственного интеллекта создает условия для разработки информационных систем по управлению рисками, востребованными в повседневной клинической практике.

Ишемическая болезнь сердца (ИБС) доминирует среди причин инвалидизации и смертности населения в большинстве стран мира. Операции коронарного шунтирования (КШ) относятся к одним из наиболее распространенных видов реваскуляризации миокарда, в связи с чем неуклонно возрастает интерес к изучению роли факторов, влияющих на риск развития неблагоприятных исходов оперативного лечения. По данным ряда исследований, частота внутригоспитальной летальности (ВГЛ) после КШ составляет 1-3% среди пациентов моложе 70 лет и 5-6% среди больных старше 70 лет [1]. В последние годы в клинической кардиологии все чаще используются прогностические инструменты, позволяющие оценить риск развития неблагоприятных событий после КШ. Наиболее известными из них являются шкалы EuroSCORE II и STS Score (The Society of Thoracic Surgery), которые демонстри-

### Key messages

- Phenotyping of risk factors is a new approach to the development of predictive models based on machine learning methods.
- The use of phenotypes as predictors of prognostic models allows a more accurate prediction of adverse outcomes of coronary bypass surgery.
- The use of explainable artificial intelligence technologies creates conditions for the development of risk management information systems that are in demand in everyday clinical practice.

ровали приемлемую и сопоставимую точность прогноза во многих исследованиях [2, 3]. Согласно рекомендациям Европейского общества кардиологов по реваскуляризации миокарда данные шкалы могут использоваться только в качестве дополнительной информации для обсуждения тактики ведения пациентов командой кардиологов и кардиохирургов [4]. Отсутствие "идеальных" прогностических технологий инициирует все большее количество исследований, в которых на основе методов машинного обучения (МО) реализуются подходы, направленные на разработку новых моделей, структура которых представлена факторами с более высоким предиктивным потенциалом, чем в "классических" шкалах, что обеспечивает повышение качества прогнозирования [5, 6]. К перспективным инструментам для отбора таких предикторов относят алгоритмы на основе объяснимого искусственного интеллекта (ОИИ). Последние включают определение диапазонов значений или критериальных границ анализируемых

факторов, обладающих наибольшей предсказательной ценностью, а также расчет весовых коэффициентов, характеризующих степень влияния отдельных показателей на конечную точку исследования [7]. Концепция ОИИ базируется на возможности клинического обоснования и интерпретации тех или иных решений, полученных в результате моделирования, оценке их значимости и точности генерируемых заключений. Одним из барьеров для реализации указанных принципов является многофакторность прогностических моделей, когда влияние на конечную точку оказывает совокупность разнородных данных, ассоциированных с различными причинами фатальных событий. Непрозрачность причинно-следственных связей отдельных факторов или их комбинаций с ВГЛ может быть частично преодолена за счет фенотипирования предикторов фатального исхода, позволяющего детализировать соотношения показателей дооперационного клинико-функционального статуса больных ИБС с исходом КШ, повысить объяснимость и надежность прогностических систем.

Цель исследования состояла в разработке прогностических моделей ВГЛ у больных ИБС после КШ с учетом результатов фенотипирования дооперационных факторов риска.

### Материал и методы

Исследование выполнено по результатам ретроспективного анализа показателей базы данных "Прогностическая оценка клинико-функционального статуса пациентов с ИБС после КШ"<sup>1</sup>, включающей сведения о 999 больных (805 мужчин, 194 женщины) в возрасте от 35 до 81 года с медианой (Me) 63 года, которым в период с 2008 по 2021гг в ГБУЗ "Приморская краевая клиническая больница № 1" г. Владивостока выполнялось плановое изолированное КШ в условиях искусственного кровообращения. Среди обследованной когорты было выделено 2 группы пациентов. В 1 из них вошли 63 (6,3%) больных, умерших в стационаре в течение первых 30 сут. после КШ, во 2 — 936 (93,7%) выживших пациентов.

Причиной смерти у 46 больных первой группы был интра- и послеоперационный инфаркт миокарда (ИМ), у 8 больных — послеоперационная острая почечная недостаточность. В остальных случаях летальные исходы были связаны с развитием ишемического инсульта (ИИ) (2), субарахноидального кровоизлияния (3), панкреонекроза (2) и медиастинита (2). Конечная точка исследования была представлена бинарным признаком: фатальным событием или его отсутствием.

По результатам электрокардиографии, выполненной перед КШ, оценивали продолжительность ин-

тервалов PQ, QT, QRS, RR в стандартных отведениях. Эхокардиографические измерения проводились на аппарате GE "Vivid 7" согласно стандартному протоколу [8]. Определяли толщину межжелудочковой перегородки, индекс относительной толщины (ИОТ) задней стенки левого желудочка (ЛЖ), индекс массы миокарда (ИММ) ЛЖ. Для исключения влияния гендерного фактора ИММ ЛЖ нормировали на верхнюю границу его референсных значений, ассоциированных с полом: 115 гр/м<sup>2</sup> — для мужчин и 95 гр/м<sup>2</sup> — для женщин с вычислением относительного ИММ (ОИММ) ЛЖ. Регистрировали линейные размеры полостей сердца, в т.ч.: продольный и поперечный размеры правого (РПП1 и РПП2) и левого (РЛП1 и РЛП2) предсердий, их объемы, фракцию выброса (ФВ) ЛЖ, градиент систолического давления (ГСД) между ЛЖ и аортой, среднее давление в легочной артерии (СрДЛА). Рассчитывали индекс массы тела и индекс коморбидности Чарлсона. Определяли тромбиновое время (ТВ), протромбиновый индекс (ПТИ), активированное частичное тромбопластическое время (АЧТВ), клиренс креатинина (КК), концентрацию глюкозы в крови.

Дизайн исследования включал несколько этапов. На первом из них с помощью методов математической статистики (тесты Манна-Уитни, хи-квадрат, Фишера) анализировали 102 показателя, характеризующих дооперационный клинико-функциональный статус больных. По результатам межгруппового сравнения были выделены потенциальные предикторы, которые в изолированном виде ассоциировались с ВГЛ. Второй шаг включал поиск пороговых значений этих показателей, позволяющих повысить их предсказательную ценность и отнести к факторам риска фатальных событий. Эта задача решалась с помощью оптимизации 4 целевых функций, 3 из которых (минимизация величины p-value в соответствующих тестах на межгрупповые сравнения, максимизация отношения шансов и площади под ROC-кривой (AUC)) ранее применялись для бинаризации данных [9]. В качестве 4-ой целевой функции был апробирован новый метод поиска пороговых значений предикторов за счет максимизации отношений истинно положительных случаев ВГЛ (True Positive — TP) к ложноположительным (False Positive — FP). Оптимизация целевых функций осуществлялась методом перебора значений показателей с шагом  $\Delta = (\max(x) - \min(x)) / 1000$ . Выделенные таким образом дихотомические показатели относили к факторам риска ВГЛ, комбинации которых фенотипировали на 3 этапе исследования с помощью процедур перебора. При достижении близких к нулю значений FP, полученные сочетания факторов риска рассматривали как отдельные фенотипы. Фенотипирование признаков и выделение пороговых значений предикторов относятся к методам извлечения правил ОИИ [9].

<sup>1</sup> Рублев В. Ю., Гельцер Б. И., Шахгельдян К. И. ДВФУ. Свидетельство о государственной регистрации № 2022621907, опублик. 02.08.2022, бюл. № 8.

Таблица 1

Показатели клинично-функционального статуса больных ИБС до операции КШ, Ме (Q1; Q3)

Показатели	Умершие, n=63	Выжившие, n=936	p-value	ОШ
Возраст, лет	68 [63; 72]	63 [58; 69]	0,00003	–
Женщины, n (%)	18 (28,6)	204 (21,8)	0,21	1,44 [0,8; 2,5]
Мужчины, n (%)	45 (71,4)	732 (78,2)		
ИМТ, кг/м <sup>2</sup>	26,4 [23,7; 30,7]	28,4 [25,2; 31,6]	0,003	–
ХСН III-IV ФК, n (%)	21 (33,3)	140 (15,1)	0,00056	2,8 [1,6; 4,9]
Стенокардия III-IV ФК, n (%)	32 (50,8)	321 (34,4)	0,0137	1,97 [1,2; 3,3]
ЭКА, n (%)	31 (49,2)	286 (31,4)	0,0051	2,1 [1,27; 3,54]
НИМ, n (%)	35 (55,6)	148 (16,2)	<0,00001	6,5 [3,8; 10,9]
ИИ, n (%)	7 (11,1)	59 (6,3)	0,18	1,9 [0,8; 4,3]
ФП, n (%)	13 (20,6)	97 (10,4)	0,02	2,2 [1,2; 4,3]
НТК, n (%)	20 (31,7)	141 (15,1)	0,0012	2,6 [1,5; 4,6]
ИЧ, усл. ед.	5 [4; 6]	4 [4; 6]	0,001	–
ФВ ЛЖ, %	52 [43; 60]	60 [52; 64]	<0,0001	–
СрдЛА, мм рт.ст.	27 [22; 38,7]	25 [22; 28]	0,129	–
РПП1, мм	42 [40; 45,5]	37 [34; 41]	<0,0001	–
РПП2, мм	46,1 [42; 53]	41 [37; 46]	<0,0001	–
РЛП1, мм	40 [37; 43]	40 [36; 45]	<0,0001	–
РЛП2, мм	46 [42; 52,8]	41 [37; 47]	<0,0001	–
КДО ЛЖ, мл	126,5 [116,3; 168,5]	123,8 [107,5; 147,4]	0,04	–
ОИММ ЛЖ, усл. ед.	1 [0,9; 1,2]	1 [0,9; 1,19]	0,09	–
ИОТ ЛЖ, усл. ед.	0,42 [0,37; 0,49]	0,42 [0,37; 0,46]	0,26	–
ГСД, мм рт.ст.	8 [5,2; 10]	6 [5; 8]	0,372	–
QRS, мс	80 [80; 100]	95 [80; 100]	0,0045	–
QT, мс	38 [36; 42]	40 [36; 42]	0,178	–
PQ, мс	16 [14; 18]	15 [14; 18]	0,16	–
RR, мс	1000 [880; 1020]	920 [800; 1080]	0,45	–
КК, мл/мин	60,7 [50,8; 78,4]	73,9 [60,7; 89,7]	<0,0001	–
ТВ, с	17,55 [14,5; 20,1]	19,6 [16,3; 21,2]	<0,0001	–
ПТИ, %	89,7 [81,8; 96]	93,6 [86,9; 99,4]	<0,0001	–
АЧТВ, с	35,4 [31,2; 39,1]	35 [31,8; 38,4]	0,29	–
СОЭ, мм/ч	15 [9,5; 28,5]	13 [8; 22]	0,411	–
Глюкоза, ммоль/л	5,73 [4,95; 7,27]	6,64 [5,12; 6,38]	0,2	–

**Сокращения:** АЧТВ — активированное частичное тромбопластиновое время, ГСД — градиент систолического давления, ИИ — ишемический инсульт, ИМТ — индекс массы тела, ИОТ — индекс относительной толщины, ИЧ — индекс коморбидности Чарлсона, КДО — конечный диастолический объем, КК — клиренс креатинина, ЛЖ — левый желудочек, НИМ — недавний инфаркт миокарда, НТК — недостаточность трикуспидального клапана, ОИММ — относительный индекс массы миокарда, ОШ — отношение шансов, ПТИ — протромбиновый индекс, РЛП1 — продольный размер левого предсердия, РЛП2 — поперечный размер левого предсердия, РПП1 — продольный размер правого предсердия, РПП2 — поперечный размер правого предсердия, СОЭ — скорость оседания эритроцитов, СрдЛА — среднее давление легочной артерии, ТВ — тромбиновое время, ФВ — фракция выброса, ФК — функциональный класс, ФП — фибрилляция предсердий, ХСН — хроническая сердечная недостаточность, ЭКА — экстракардиальная артериопатия.

На четвертом этапе исследования с помощью относящейся к базовым методам МО многофакторной логистической регрессии (МЛР) была разработана прогностическая модель ВГЛ, где в качестве предикторов использовались фенотипы факторов риска ВГЛ, кодированные бинарными переменными. Для верификации модели использовали кросс-валидацию методом K-Fold на 10 выборках. Качество моделей оценивали по трем метрикам: AUC, чувствительность (Sen) и специфичность (Spec). Полученные результаты сопоставляли с прогнозом ВГЛ по шкале EuroSCORE II. Обработка и анализ данных

выполнялись на языке Python. Работа выполнена при финансовой поддержке проекта FZNS-2023-0010 Госзадания ДВФУ.

**Результаты**

Межгрупповой анализ 102 факторов, характеризующих дооперационный клинично-функциональный статус умерших и выживших больных ИБС, демонстрировал достоверность различий по 17 из них (табл. 1). При этом максимальный уровень статистической значимости (p<0,0001) фиксировался у 10 показателей: возраст больных, недавний ИМ (НИМ),

ФВ ЛЖ, РПП1, РПП2, РЛП1, РЛП2, КК, ТВ, ПТИ. Высокой степенью достоверности отличались показатели хронической сердечной недостаточности (ХСН) III-IV функционального класса (ФК) ( $p=0,00056$ ), индекс массы тела ( $p=0,003$ ), экстракардиальная артериопатия ( $p=0,0051$ ), индекс коморбидности Чарлсона ( $p=0,001$ ) и QRS ( $p=0,0045$ ). Менее заметные, но статистически значимые межгрупповые различия ассоциировались с наличием стенокардии III-IV ФК ( $p=0,0137$ ), фибрилляции предсердий ( $p=0,02$ ), конечным диастолическим объемом (КДО) ЛЖ ( $p=0,04$ ). По данным предварительного анализа гендерная принадлежность, ИИ в анамнезе,

уровень СрДЛА, ОИММ ЛЖ, ИОТ, ГСД, АЧТВ, скорость оседания эритроцитов, продолжительность интервалов QT, PQ, RR, концентрация глюкозы в крови не имели линейных взаимосвязей с ВГЛ.

На втором этапе исследования с помощью комплекса методов для всех показателей определяли пороговые значения с наибольшим предиктивным потенциалом. В таблице 2 представлены только те из них, которые имели по крайней мере для одного порогового значения статистическую значимость на уровне  $p\text{-value} < 0,05$ . Для категориальных факторов, которые имели более двух значений, была выполнена бинаризация переменных: например, ХСН III-

Таблица 2

## Определение пороговых значений потенциальных предикторов ВГЛ с помощью различных методов

Показатели	Порог	p-value	AUC	ОШ	ТР, абс.	FP, абс.	ТР/FP, %
Возраст, лет	>67	0,00013	0,621	2,74 [1,6; 4,6]	34	280	12,1
	>79	0,0022	0,524	23,4 [3,8; 142]	3	2	150
ИМТ, кг/м <sup>2</sup>	>25,4	0,011	0,576	0,5 [0,29; 0,84]	36	620	5,8
	>37,5	0,14	0,523	1,99 [0,81; 4,9]	5	43	11,6
СрДЛА, мм рт.ст.	>34	<0,00001	0,617	4,4 [2,5; 7,9]	20	83	24,1
	>38	<0,00001	0,602	5,3 [2,8; 9,9]	16	53	24,4
КДО ЛЖ, мл	>89	0,053	0,537	5,9 [0,8; 43]	61	788	7,7
	>158	0,00059	0,6	2,7 [1,6; 4,7]	23	153	15
	>163	0,00033	0,593	3 [1,7; 5,3]	20	119	16,8
	>209	0,016	0,53	3,5 [1,4; 8,7]	6	26	23,1
ОИММ ЛЖ, усл. ед.	>1,79	0,0078	0,532	5,1 [1,76; 14,9]	4	13	30,8
	>1,81	0,0158	0,526	5,2 [1,6; 17,2]	1	2	50
ИОТ ЛЖ, усл. ед.	<0,21	0,0011	0,523	43,5 [4,5; 500]	3	1	300
	>0,49	0,09	0,54	1,7 [0,94; 3]	17	158	11
	>0,67	0,34	0,49	2,8 [0,32; 24,4]	1	5	20
КК, мл/мин	<30,7	0,0072	0,52	12,6 [2,5; 64]	3	3	100
	<57,7	0,00002	0,63	3,3 [1,9; 5,6]	28	150	18,7
	<63,7	0,00005	0,63	3 [1,8; 5,1]	35	224	15,8
ТВ, с	<11,8	0,00013	0,53	53 [5,9; 484,8]	4	1	400
	<13,4	<0,00001	0,57	18,7 [6,7; 52,1]	8	6	133
	<18,3	0,0037	0,6	2,2 [1,3; 3,8]	34	281	12,1
	>27	0,042	0,52	4,9 [1,3; 18,8]	3	8	37,5
ПТИ, %	<64,4	0,0258	0,536	2,8 [1,2; 6,6]	7	34	20,6
	<93	0,00029	0,622	2,8 [1,6; 4,9]	43	354	12,1
	<106	0,0227	0,543	6,9 [0,96; 50,6]	60	685	8,8
АЧТВ, с	<20,5	0,0054	0,5167	–	2	0	–
	<24	0,089	0,513	5,1 [0,98; 27,1]	2	5	40
СОЭ, мм/ч	>25	0,0042	0,575	2,3 [1,3; 4]	21	148	14,2
	>39	0,027	0,536	2,5 [1,14; 5,67]	8	45	17,8
	>70	0,31	0,489	3,3 [0,37; 30,4]	1	4	25
РПП2, мм	>39	<0,00001	0,653	5,77 [2,6; 12,8]	55	499	11
	>43	0,000001	0,66	37,6 [21,7; 65]	42	310	13,5
	>55	0,00052	0,562	44 [21; 94]	10	36	28
РЛП2, мм	>40	<0,00001	0,676	59 [29; 121]	53	433	12,2
	>55	0,00052	0,558	49 [22; 109]	9	29	31
ФВ ЛЖ, %	<29	0,0044	0,523	15,2 [3; 76,8]	3	3	100
	<50	0,000001	0,639	3,9 [2,3; 6,6]	28	155	18,1

Таблица 2. Продолжение

Показатели	Порог	p-value	AUC	ОШ	ТР, абс.	FP, абс.	ТР/FP, %
ГСД, мм рт.ст.	>6,8	0,0001	0,63	3 [1,7; 5,3]	45	358	12,6
	>12,9	0,02	0,547	2,4 [1,17; 4,75]	11	64	17,2
	>20,2	0,17	0,52	2,2 [0,83; 6]	5	29	17,2
QT, мс	<32	0,0125	0,5307	4,3 [1,55; 12]	5	18	27,8
QRS, мс	<61	0,061	0,515	30,2 [2,7; 338]	2	1	200
	<91	0,00103	0,605	2,5 [1,42; 4,4]	45	461	9,7
PQ, мс	>22	0,054	0,521	3,2 [1,05; 9,75]	4	19	21,05
RR, мс	>870	0,021	0,573	2,1 [1,13; 3,8]	49	577	8,5
	>1400	0,0043	0,521	15,3 [3; 77,3]	3	3	100
	>1500	0,00025	0,523	–	3	0	–
НИМ	–	<0,00001	0,702	6,9 [4,1; 11,7]	36	148	24,3
ИИ	–	0,18	0,524	1,9 [0,81; 4,3]	7	59	11,9
ЭКА	–	0,0051	0,589	2,1 [1,3; 3,5]	31	286	10,8
ХСН III-IV ФК	–	0,00056	0,585	2,8 [1,6; 4,9]	21	140	15
Стенокардия III-IV ФК	–	0,0137	0,578	1,96 [1,18; 3,28]	32	321	10
ФП до операции	–	0,0199	0,552	2,2 [1,18; 4,29]	13	97	13,4
НТК	–	0,00124	0,588	2,6 [1,5; 4,6]	20	141	14,2

**Сокращения:** АЧТВ — активированное частичное тромбопластиновое время, ГСД — градиент систолического давления, ИИ — ишемический инсульт, ИМТ — индекс массы тела, ИОТ — индекс относительной толщины, КДО — конечный диастолический объем, КК — клиренс креатинина, ЛЖ — левый желудочек, НИМ — недавний инфаркт миокарда, НТК — недостаточность трикуспидального клапана, ОИММ — относительный индекс массы миокарда, ОШ — отношение шансов, ПТИ — протромбиновый индекс, РПП2 — поперечный размер левого предсердия, РПП2 — поперечный размер правого предсердия, СОЭ — скорость оседания эритроцитов, СрДЛА — среднее давление легочной артерии, ТВ — тромбиновое время, ФВ — фракция выброса, ФК — функциональный класс, ФП — фибрилляция предсердий, ХСН — хроническая сердечная недостаточность, ЭКА — экстракардиальная артериопатия, AUC — площадь под ROC-кривой, FP — ложноположительный случай (False Positive), TP — положительный случай (True Positive).

IV ФК и стабильная стенокардия III-IV ФК сопоставлялись с ФК I-II. Для части показателей, которые не имели статистически значимых различий в непрерывной форме (табл. 1), были выделены пороговые значения, которые обеспечивали достоверное влияние соответствующего фактора на конечную точку. Так, например, показатель СрДЛА в непрерывном виде был статистически не значим ( $p=0,129$ ), но приобретал высокую значимость ( $p<0,0001$ ) при пороговом значении  $>34$  мм рт.ст. Другими примерами прогностической трансформации факторов при переходе от непрерывной к категориальной форме являются показатели КДО ЛЖ, ОИММ ЛЖ, ИОТ ЛЖ, QT, PQ, RR, АЧТВ. Кроме того, при поиске пороговых значений использовалась двух-фазная кодировка, когда на первой фазе единица ассоциировалась со значениями выше порога, а на второй — ниже. Были выделены также показатели, которые имели статистически значимые пороговые значения для обеих фаз. Так, показатель ТВ являлся фактором риска ВГЛ при значениях  $<13,4$  с и  $>27$  с. Для дальнейшего анализа выбирали показатели, имеющие по крайней мере одно статистически значимое пороговое значение. Исключение составил категориальный признак ИИ в анамнезе, который при дальнейшем исследовании демонстрировал прогностическую значимость при сочетании с предиктором РПП2  $>55$  мм (табл. 3).

Результаты оценки прогностического потенциала отдельных индикаторов с использованием метода минимизации FP и максимизации TP/FP показали, что за исключением показателей АЧТВ  $<20,5$  с и RR  $>1500$  мс ни один из анализируемых факторов в изолированной форме не обладал способностью к выделению пациентов исключительно из группы ВГЛ (табл. 2). Так, например, фактор КК  $<30,7$  мл/мин фиксировался у 3 пациентов из группы ВГЛ и у 3 выживших больных (FP=3), а РПП2  $>55$  мм имел место у 46 больных, в т.ч. только у 10 — из группы ВГЛ (FP=36). Эти данные указывают на то, что в изолированной форме анализируемые факторы не обеспечивали корректную оценку прогноза ВГЛ.

Задачей третьего этапа исследования являлось фенотипирование факторов риска ВГЛ путем формирования их комбинаций, позволяющих максимально ограничить их ассоциацию с больными 2 группы (табл. 3). Реализация данной задачи осуществлялась методом перебора ранее выделенных признаков со значениями FP, близкими к нулю. Использование данного подхода позволило сформировать 7 фенотипов факторов риска ВГЛ после КШ. Ключевыми индикаторами первого фенотипа были показатели ТВ  $<13,4$  с и АЧТВ  $<20,5$  с, указывающие на возрастающий риск тромбообразования. Сочетание ТВ  $<13,4$  с с признаками дилатации ле-

Таблица 3

## Фенотипы факторов риска ВГЛ после КШ

№	Фенотипы	TP   FP	TP	FP
1	ТВ <13,4 с + РПП2 >55 мм	10   0	3	0
	ТВ <13,4 с + РЛП2 >55 мм		3	0
	ТВ <13,4 с + КК <57,7 мл/мин		6	0
	АЧТВ <20,5 с		2	0
2	ТВ >27 с + СОЭ >39 мм/ч	10   3	2	0
	(ТВ >27 с или ПТИ <64,4%) + (РЛП2 >55 мм или РПП2 >55 мм)		6	0
	(ТВ >27 с или ПТИ <64,4%) + СрДЛА >38 мм рт.ст. + НТК		6	3
3	ИИ + РПП2 >55 мм	4   0	4	0
4	ФП + Возраст >79 лет	2   0	2	0
5	КДО ЛЖ >209 мл + НИМ	8   0	5	0
	КДО ЛЖ >209 мл + ИОТ >0,49 усл. ед.		2	0
	КДО >163 мл + ИОТ >0,49 усл. ед. + ОИММ ЛЖ >1,78 усл. ед.		2	0
6	КК <31 мл/мин + НИМ	14   4	2	1
	КК <57 мл/мин + ТВ <13,4 с		6	0
	КК <40 мл/мин + ОИММ ЛЖ >1,5 усл. ед. + НИМ		1	1
	КК <57,7 мл/мин + RR >1500 мс		3	0
	КК <57,7 мл/мин + РПП2 >55 мм + Билирубин >21 ммоль/л		4	2
7	НИМ + ХСН III-IV + ЭКА + Стенокардия III-IV ФК + (ИЧ >9 усл. ед. или ФВ ЛЖ <50% или ГСД >20,5 мм рт.ст.)	8   0	5	0
	ФВ ЛЖ <30% + ЭКА + НИМ		3	0

**Сокращения:** АЧТВ — активированное частичное тромбопластиновое время, ГСД — градиент систолического давления, ИИ — ишемический инсульт, ИОТ — индекс относительной толщины, ИЧ — индекс коморбидности Чарлсона, КДО — конечный диастолический объем, КК — клиренс креатинина, ЛЖ — левый желудочек, НИМ — недавний инфаркт миокарда, НТК — недостаточность трикуспидального клапана, ОИММ — относительный индекс массы миокарда, ОШ — отношение шансов, ПТИ — протромбиновый индекс, РЛП2 — поперечный размер левого предсердия, РПП2 — поперечный размер правого предсердия, СОЭ — скорость оседания эритроцитов, СрДЛА — среднее давление легочной артерии, ТВ — тромбиновое время, ФВ — фракция выброса, ФК — функциональный класс, ФП — фибрилляция предсердий, ХСН — хроническая сердечная недостаточность, ЭКА — экстракардиальная артериопатия, FP — ложноположительный случай (False Positive), TP — положительный случай (True Positive).

вого и правого предсердий, снижения клубочковой фильтрации обеспечивало надежный прогноз ВГЛ (FP=0). Второй фенотип характеризовался повышением риска геморрагических осложнений, что иллюстрировалось уровнем ТВ >27 с или ПТИ <64,4%. К субфенотипам этой комбинации относились сочетания индикаторов гипокоагуляции с дилатацией предсердий, наличием недостаточности трикуспидального клапана, легочной гипертензии, увеличением скорости оседания эритроцитов >39 мм/ч. Третий фенотип факторов риска ВГЛ ассоциировался с ИИ в анамнезе и увеличением размеров правого предсердия, а четвертый — с наличием фибрилляции предсердий до КШ у больных старше 79 лет. Пятый фенотип был представлен признаком значительного увеличения КДО ЛЖ у больных с НИМ или его сочетанием с индикаторами гипертрофии ЛЖ (ИОТ >0,49 усл. ед. и ОИММ ЛЖ >1,78 усл. ед.). Шестой фенотип ассоциировался с маркерами снижения клубочковой фильтрации и сердечной недостаточности (ФВ ЛЖ <30%, РПП2 >55 мм, RR >1500 мс, НИМ, ОИММ ЛЖ >1,5 усл. ед.). Седьмой фенотип характеризовался признаками тяжелой сердечно-сосудистой коморбидности с поражением периферических артериальных бассейнов.

На четвертом этапе исследования нами была разработана прогностическая модель ВГЛ после КШ на основе МЛР, где в качестве предикторов использовали 7 признаков, фиксирующих отношение пациентов к одному из фенотипов факторов риска. Предиктор принимал значение 1, если пациент относился к соответствующему фенотипу, и 0 — если не относился. В тех случаях, когда пациент ассоциировался с несколькими фенотипами, значение 1 принимали предикторы, относящиеся к данным фенотипам. Таким образом, главной особенностью этой модели являлась её структура, которая была представлена признаковым пространством, где каждый признак демонстрирует соответствие пациента определенному фенотипу факторов риска. Сравнительный анализ метрик качества модели EuroSCORE II и авторской модели МЛР при пороге отсечения 0,05 демонстрировал различия в их прогностической точности. Было установлено, что качество прогноза по индикаторам AUC, Sen и Spec для классической шкалы EuroSCORE II составляло 0,62, 0,85 и 0,53, соответственно, что свидетельствует о её недостаточной точности при апробации на анализируемой когорте. Авторская модель на основе МЛР демонстрировала значительно более высо-

кую точность прогноза ВГЛ после КШ: AUC — 0,91, Sen — 0,9 и Spec — 0,85.

### Обсуждение

Одним из наиболее востребованных и перспективных направлений Национального проекта "Здравоохранение" является создание информационных систем оценки рисков сердечно-сосудистых заболеваний и их осложнений, в т.ч. ассоциированных с реваскуляризацией миокарда [10-13]. Реализация этой задачи базируется на применении современных методов МО, с помощью которых осуществляется интеллектуальный анализ больших данных и разрабатываются модели для предсказания неблагоприятных событий в послеоперационном периоде и смертности на различных горизонтах наблюдения. В последние годы совершенствование прогностических технологий в клинической медицине, в т.ч. в кардиологии, связано с использованием методов ОИИ, позволяющих создавать интерпретируемые модели МО, повышать прозрачность и объяснимость генерируемых решений [5, 14, 15]. В настоящем исследовании для оценки рисков ВГЛ после КШ были использованы технологии ОИИ, которые включали определение пороговых значений дооперационных клинико-функциональных показателей, усиливающих их предиктивный потенциал, и выделение фенотипов факторов риска ВГЛ.

Результаты анализа показали, что у 56 (89%) больных с фатальным исходом КШ имели место дооперационные факторы риска ВГЛ, включенные в сформированные фенотипы. При этом у 13 (23,2%) пациентов фиксировались комбинации признаков из нескольких фенотипов, что может свидетельствовать о более тяжелой коморбидности этих больных и возрастающей вероятности неблагоприятного исхода КШ. Вместе с тем 7 (11%) больных из группы ВГЛ не ассоциировались ни с одним из выделенных фенотипов, что может быть обусловлено наличием более широкого спектра факторов риска ВГЛ, не относящихся к тем, которые были представлены в настоящем исследовании. Важно отметить, что часть выделенных предикторов (РПП2, РЛП2, КДО ЛЖ, ИОТ, ОИММ ЛЖ, RR, ТВ, ПТИ, АЧТВ) не входили в структуру модели EuroSCORE II. Кроме того, ряд предикторов, присутствующих в обеих моделях, имели различные

пороговые значения. Например, КК в авторской модели использовался с пороговыми значениями <40 и <57 мл/мин, в то время как в EuroSCORE II его уровень составлял <50 и 50-85 мл/мин.

Необходимо также отметить, что при фенотипировании признаков в отличие от их кластеризации, при которой используется единый набор атрибутов, в настоящей работе применялись различные наборы показателей, что позволяло формировать пересекающиеся группы и верифицировать пациентов, относящихся к одному или нескольким фенотипам факторов риска. Это, в свою очередь, создает условия для более точной оценки прогноза ВГЛ, учитывающей фенотипические особенности факторов риска и возможности управления ими. Использование фенотипов в качестве предикторов является новым подходом к разработке интерпретируемых прогностических моделей. В нашем исследовании целесообразность его применения подтверждалась высокой точностью прогнозирования ВГЛ после КШ (AUC — 0,91, Sen — 0,9, Spec — 0,85).

### Заключение

Разработанный алгоритм обработки и анализа данных, характеризующих клинико-функциональный статус больных ИБС до КШ, обеспечил высокое качество выделения факторов риска ВГЛ. В отличие от модели EuroSCORE II, структура которой представлена 18 факторами, авторская модель включает 28 предикторов, объединенных в 7 фенотипов, которые формировали признаковое пространство для верификации пациентов, соответствующих тому или иному фенотипу факторов риска. Данный подход позволяет повысить интерпретируемость моделей, объяснимость и точность результатов прогнозирования. Перспективы дальнейших исследований по данной проблеме связаны с совершенствованием технологий ОИИ, использование которых создает условия для разработки информационных систем по управлению рисками, востребованными в повседневной клинической практике.

**Отношения и деятельность.** Работа выполнена при финансовой поддержке проекта FZNS-2023-0010 Госзадания ДВФУ.

### Литература/References

1. Squiers JJ, Schaffer JM, Banwait JK, et al. Long-Term Survival After On-Pump and Off-Pump Coronary Artery Bypass Grafting. *Ann Thorac Surg.* 2022;113(6):1943-52. doi:10.1016/j.athoracsur.2021.07.037.
2. Nashif SA, Roques F, Sharples LD, et al. EuroSCORE II. *European journal of cardiothoracic surgery: official journal of the European Association for Cardio-thoracic Surgery.* 2012;41(4):734-45. doi:10.1093/ejcts/ezs043.
3. Vassileva CM, Aranki S, Brennan JM, et al. Evaluation of The Society of Thoracic Surgeons Online Risk Calculator for Assessment of Risk in Patients Presenting for Aortic Valve Replacement After Prior Coronary Artery Bypass Graft: An Analysis Using the STS Adult Cardiac Surgery Database. *The Annals of thoracic surgery.* 2015;100(6):2109-16. doi:10.1016/j.athoracsur.2015.04.149.
4. 2018 ESC/EACTS guidelines on myocardial revascularization. *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.
5. Valente F, Henriques J, Paredes S, et al. A new approach for interpretability and reliability in clinical risk prediction: Acute coronary syndrome scenario. *Artificial intelligence in medicine.* 2021;117:102-13. doi:10.1016/j.artmed.2021.102113.
6. 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. *Kardiologiya.* 2020;60(10):38-46. (In Russ.) Гельцер Б.И., Шахгельдян К.И., Рублев В.Ю. и др. Методы машинного обучения в прогнозировании

- вании 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.
7. Johnson KW, Torres SJ, Glicksberg BS, et al. Artificial Intelligence in Cardiology. *Journal of the American College of Cardiology*. 2018;71(23):2668-79. doi:10.1016/j.jacc.2018.03.521.
  8. Galderisi M, Cosyns B, Edvardsen T, et al. Standardization of adult transthoracic echocardiography reporting in agreement with recent chamber quantification, diastolic function, and heart valve disease recommendations: an expert consensus document of the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2017;18(12):1301-10. doi:10.1093/ehjci/jex244.
  9. Molnar C. *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. 2022. 328 p. ISBN-13: 979-8411463330. Available at: <https://christophm.github.io/interpretable-ml-book/>.
  10. Shakhgeldyan KI, Rublev VY, Geltser BI, et al. Predictive potential assessment of pre-operative risk factors for atrial fibrillation in patients with coronary artery disease after coronary artery bypass grafting. *The Siberian Journal of Clinical and Experimental Medicine*. 2020;35(4):128-36. (In Russ.) Шахгельдян К. И., Рублев В. Ю., Гельцер Б. И. и др. Оценка предиктивного потенциала дооперационных факторов риска фибрилляции предсердий у больных ишемической болезнью сердца после коронарного шунтирования. *Сибирский журнал клинической и экспериментальной медицины*. 2020;35(4):128-36. doi:10.29001/2073-8552-2020-35-4-128-136.
  11. Geltser BI, Rublev VYu, Tsivanyuk MM, Shakhgeldyan KI. Machine learning in predicting immediate and long-term outcomes of myocardial revascularization: a systematic review. *Russian Journal of Cardiology*. 2021;26(8):4505. (In Russ.) Гельцер Б. И., Рублев В. Ю., Циванюк М. М., Шахгельдян К. И. Машинное обучение в прогнозировании ближайших и отдаленных результатов реваскуляризации миокарда: систематический обзор. *Российский кардиологический журнал*. 2021;26(8):4505. doi:10.15829/1560-4071-2021-4505.
  12. Geltser BI, Shakhgeldyan KI, Rublev VYu, et al. Algorithm for selecting predictors and prognosis of atrial fibrillation in patients with coronary artery disease after coronary artery bypass grafting. *Russian Journal of Cardiology*. 2021;26(7):4522. (In Russ.) Гельцер Б. И., Шахгельдян К. И., Рублев В. Ю. и др. Алгоритм отбора предикторов и прогнозирование фибрилляции предсердий у больных ишемической болезнью сердца после коронарного шунтирования. *Российский кардиологический журнал*. 2021;26(7):4522. doi:10.15829/1560-4071-2021-4522.
  13. Shakhgeldyan K, Geltser B, Rublev V, et al. Feature Selection Strategy for Intrahospital Mortality Prediction after Coronary Artery Bypass Graft Surgery on an Unbalanced Sample. *Proceedings of the 4th International Conference on Computer Science and Application Engineering*. 2020;1-7. doi:10.1145/3424978.3425090.
  14. Jo YY, Cho Y, Lee SY, et al. Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram. *International journal of cardiology*. 2021;328:104-10. doi:10.1016/j.ijcard.2020.11.053.
  15. Taniguchi H, Takata T, Takechi M, et al. Explainable Artificial Intelligence Model for Diagnosis of Atrial Fibrillation Using Holter Electrocardiogram Waveforms. *International heart journal*. 2021;62(3):534-9. doi:10.1536/ihj.21-094.