

УДК 614.2

DOI 10.24412/2312-2935-2023-4-936-955

РАЗРАБОТКА И ТЕСТИРОВАНИЕ НОВЫХ МЕТОДИЧЕСКИХ ПОДХОДОВ ПРОГНОЗИРОВАНИЯ 6-ТИ МЕСЯЧНОЙ ПОСТГОСПИТАЛЬНОЙ ЛЕТАЛЬНОСТИ У ПАЦИЕНТОВ, ПЕРЕНЕСШИХ ИНФАРКТ МИОКАРДА, С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ НА БАЗЕ МЕЖДУНАРОДНОГО ИССЛЕДОВАНИЯ «ОСТРЫЙ ИНФАРКТ МИОКАРДА В РОССИЙСКОЙ ФЕДЕРАЦИИ: СУЩЕСТВУЮЩАЯ ПРАКТИКА И ПРЕПЯТСТВИЯ НА ПУТИ ЭФФЕКТИВНОГО ЛЕЧЕНИЯ НА РАЗНЫХ УРОВНЯХ СИСТЕМЫ ЗДРАВООХРАНЕНИЯ»

И. А. Мишкин^{1,2}, А. В. Концевая¹, А. В. Гусев^{3,4}, А. А. Сахаров⁵, О. М. Дранкина¹

¹ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Министерства здравоохранения Российской Федерации, г. Москва

²ГУЗ ТО «Киреевской ЦРБ», г. Киреевск

³ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» Министерства здравоохранения Российской Федерации, г. Москва

⁴Компания «К-Скай», г. Петрозаводск

⁵ПАО «Группа Ренессанс Страхование», г. Тула

Актуальность. На сегодняшний день это на долю инфаркта миокарда (ИМ) приходится 38,9 смертельных исходов на 100 тыс. человек. Чтобы снизить данный показатель важно сократить число постгоспитальных осложнений у пациентов после выписки. Для таких целей сегодня активно используются шкалы прогноза: EDACS-ADP, HEART, GRACE и др. Однако, последнее время многие ученые высказываются о возможности применения технологии искусственного интеллекта с целью повышения качества прогноза.

Цель - разработать и протестировать новые методические подходы прогнозирования смертельных исходов на протяжении 6-ти месяцев после выписки из стационара у пациентов, перенесших ИМ, с использованием технологии машинного обучения.

Материалы и методы: работа была выполнена на данных международного исследования «Острый инфаркт миокарда в Российской Федерации: существующая практика и препятствия на пути эффективного лечения на разных уровнях системы здравоохранения», n=1,128 тыс.: мужчины n=845 (74,9%); женщины n=248 (21,9%). Обе когорты были разделена на 2 группы: 1-ая группа n= 980 (86,9%) – пациенты, выжившие за 6-ти месячный период наблюдения после выписки из стационара. 2-ая группа n=65 (5,8%) - пациенты, умершие за 6-ти месячный период наблюдения после выписки из стационара, включая смерти в стационаре. Для построения модели прогноза мы использовали шесть алгоритмов классификации. Предикторы, включенные в анализ, n=193. Для оценки эффективности моделей прогноза мы использовали показатели ROC-анализа.

Результаты. Наилучшим алгоритмом оказался LGBMClassifier AUC-0,84, наихудшие результаты показала логистическая регрессия AUC-0,79. Наиболее значимыми факторами являлись наблюдение терапевта и кардиолога в течение года, занятость пациентов и показатели скорости клубочковой фильтрации. Также важную роль играли такие факторы риска, как прохождение диспансеризации, наличие хронической сердечной недостаточности в анамнезе и коморбидность.

Обсуждение. В результате обсуждения удалось сравнить наиболее важные факторы риска, влияющие на постгоспитальную летальность и сравнить их с полученными в результате эксперимента. Также мы удостоверились, что среднее значение AUC, полученное в аналогичных исследованиях – 0,82 соответствует полученным нами результатам.

Выводы. В результате исследования нам удалось сконструировать модель классификации с достаточно хорошим качеством детерминации. Дальнейшие изучения данной темы поможет лучше понять природу развития патологических изменений и заранее предупредить неблагоприятные исходы.

Ключевые слова: постгоспитальная летальность, инфаркт миокарда, прогнозирование, искусственный интеллект, машинного обучение, факторы риска

DEVELOPMENT AND TESTING OF NEW METHODOLOGICAL APPROACHES FOR PREDICTING 6-MONTH POST-HOSPITAL MORTALITY IN PATIENTS WITH MYOCARDIAL INFARCTION USING MACHINE LEARNING TECHNOLOGY BASED ON THE INTERNATIONAL STUDY "ACUTE MYOCARDIAL INFARCTION IN THE RUSSIAN FEDERATION: EXISTING PRACTICE AND OBSTACLES TO EFFECTIVE TREATMENT AT DIFFERENT LEVELS OF THE HEALTHCARE SYSTEM"

I. A. Mishkin^{1,2}, A.V. Kontsevaya¹, A.V. Gusev^{3,4}, A. A. Saharov⁵, O. M. Drapkina¹

¹*Federal State Budgetary Institution National Medical Research Center for Therapy and Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Moscow*

²*Tula State Healthcare Institution district Kireevskaya central district hospital, Kireevsk*

³*Russian Research Institute of Health, Moscow*

⁴*"K-Sky" company, Petrozavodsk*

⁵*PJSC "Renaissance Insurance Group", Tula*

Relevance. To date, the spread of myocardial infarction (MI) accounts for 38.9 deaths per 100 thousand people. To reduce this indicator, it is important to reduce the number of post-hospital complications in patients after discharge. For such purposes, forecast scales are actively used today: EDACS-ADP, HEART, GRACE, etc. However, recently many scientists have been talking about the possibility of using artificial intelligence technology in order to improve the quality of the forecast.

The goal is to develop and test new methodological approaches for predicting deaths for 6 months after discharge from the hospital in patients who have undergone MI using machine learning technology.

Materials and methods: the work was carried out on the data of the international study "Acute myocardial infarction in the Russian Federation: current practice and obstacles to effective treatment at different levels of the healthcare system", n=1,128: men n=845 (74.9%); women n=248 (21.9%). Both cohorts were divided into 2 groups: First group n= 980 (86.9%) is patients who survived for a 6-month follow-up period after discharge from the hospital. Second group n=65 (5.8%) is patients who died during the 6-month follow-up period after discharge from the hospital, including deaths in the hospital. To build the prediction model, we used six classification algorithms. Predictors included in the analysis are n=193. To evaluate the effectiveness of forecast models, we used ROC analysis indicators.

Results. The best algorithm turned out to be LGBMClassifier AUC-0.84, the worst results were shown by logistic regression AUC-0.79. The most significant factors were the observation of a

therapist and a cardiologist during the year, the employment of patients and glomerular filtration rate indicators. Risk factors such as medical examination, a history of chronic heart failure and comorbidity also played an important role.

Discussion. As a result of the discussion, it was possible to compare the most important risk factors affecting post-hospital mortality and compare them with those obtained as a result of the experiment. We also made sure that the average AUC value obtained in similar studies – 0.82 corresponds to the results we obtained.

Conclusions. As a result of the study, we were able to construct a classification model with a fairly good quality of determination. Further studies of this topic will help to better understand the nature of pathological changes development and to prevent adverse outcomes in advance.

Keywords: post-hospital mortality, myocardial infarction, prediction, artificial intelligence, machine learning, risk factors

Введение. По данным Росстата на 2021 год заболеваемость болезнями системы кровообращения составила 3050 случаев на 100 тыс. человек и до сих пор продолжает занимать первое место по летальности среди всех нозологий. Смертность от инфаркта миокарда (ИМ) – одного из наиболее тяжелых осложнений ишемической болезни сердца составила 38,9 случаев на 100 тыс. человек. Таким образом, почти каждый 80-ый заболевший умирает. При этом среди мужчин заболеваемость в среднем выше на 3,7%, чем у женщин и увеличивается с возрастом [1].

По данным исследования Сайгитова и соавт [2] процент 15-ти месячной постгоспитальной летальности после перенесённого ИМ в РФ составил 12,2% на 2006 г. Шишкина и соавт. [3] сообщают, что процент 12-ти месячной постгоспитальной летальности на 2020 г уже составлял 5,2%. Шальнова и соав. [4] приводит аналогичные показатели. Средний процент постгоспитальной летальности в их работе составил 5,8% на 2022 г.

Несмотря на то, что очевидно наблюдается значительное снижение смертности благодаря существенным успехам в области эндоваскулярного лечения больным с острым коронарным синдромом (ОКС), необходимо улучшать методы профилактического воздействия для снижения данного показателя.

На сегодняшний день основными инструментами прогноза летальности от ИМ являются шкалы оценки риска. Наиболее популярными из них являются EDACS-ADP, оценивающая развитие неблагоприятных сердечно-сосудистых событий (ССС) в течение 30 суток после перенесенного ИМ, HEART, оценивающая риск развития повторного ИМ, повторной реваскуляризации и смерть в течение 30 суток, 6 недель и года у пациентов, перенесших ИМ, и шкала GRACE, оценивающая риск постгоспитальной смерти в течение 6-ти месяцев у пациентов, перенесших ИМ [5-7]. Данные шкалы хорошо зарекомендовали себя,

так как имеют высокие показатели C-статистики от 0,81 до 0,87. Данный факт позволяет врачам во всем мире активно их использовать.

Однако, несмотря на многие преимущества, последнее время многие исследователи склоняются к использованию технологий машинного обучения (machine learning, ML) для прогноза ССС, поскольку они позволяют избежать множества ошибок связанных с линейным анализом данных [8-11].

Применение искусственного интеллекта позволяет в автоматическом режиме проводить анализ большого количества данных, а основанные на нем программные средства могут быть интегрированы в системы поддержки принятия врачебных решений и мобильные приложения (mHealth). В сумме это даст возможность как врачам, так и пациентом в реальном времени отслеживать состояние здоровья и вносить необходимые корректировки при его ухудшении [12,13]

Цель работы - разработать и протестировать новые методические подходы прогнозирования смертельных исходов на протяжении 6-ти месяцев после выписки из стационара у пациентов, перенесших ИМ, с использованием технологии ML.

Материалы и методы. Работа была выполнена на данных международного исследования «Острый инфаркт миокарда в Российской Федерации: существующая практика и препятствия на пути эффективного лечения на разных уровнях системы здравоохранения», проведенного в 2015-2016 гг., на территории РФ в 15-ти регионах при сотрудничестве ФГБУ «НМИЦ ТПМ» Минздрава России, Лондонской школой гигиены и тропической медицины Великобритания), и Арктического университета Норвегии Университета Тромсё [14-17]. Материалами для работы послужили «Экспертная карта пациента 1-ый этап», «Контрольная карта пациента (6 месяцев)». Всего выборка содержала информацию о 1,128 тыс. пациентах: мужчины n=845 (74,9%), сред. возраст - 57,7; женщины n=248 (21,9%), сред. возраст - 62,3 лет. В среднем женщины были старше на 4,6 лет, $p < 0,001$. 15,8% значений гендерной принадлежности были пропущены в исходной выборке. Пациенты поступили в лечебные учреждения с диагнозом ОКС и впоследствии при выписке были переквалифицированы в следующие диагнозы МКБ10: I20 – 6 (0,6%); I21 – 352 (32,4%); I22 – 38 (3,7%); I23 – 1 (0,1%); I24 – 1 (0,1%); I48 – 0,1%); 63% информации о диагнозах отсутствовала. 41 (3,6%) человек умерли в стационаре.

Социально-демографические характеристики выборки представлены на рис. 1.

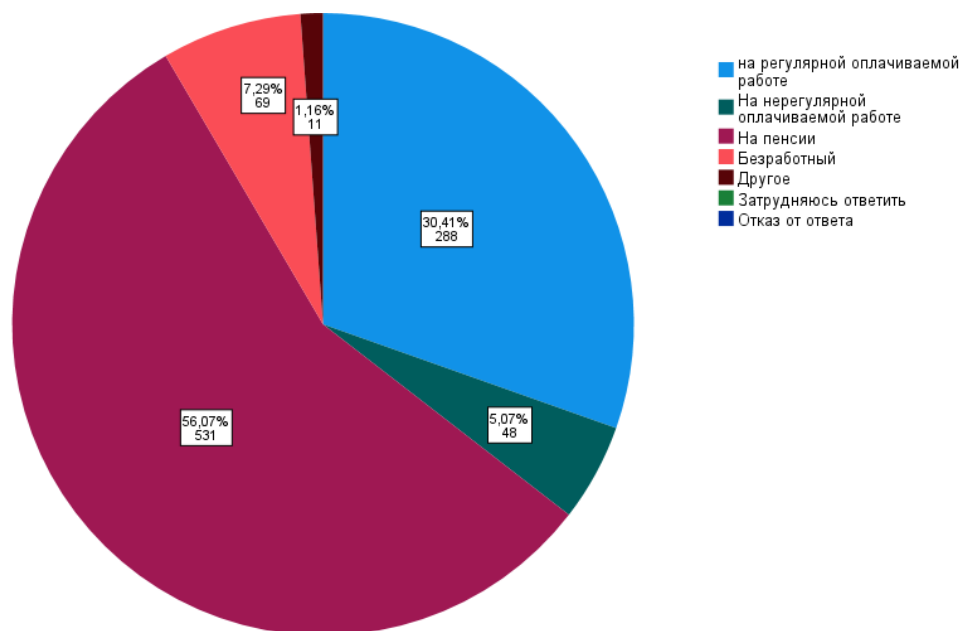


Рисунок 1. Занятость участников исследования, n, %

Обе когорты были разделены на 2 группы: 1-ая группа n= 980 (86,9%), сред. возраст - 58,6±0,24 лет – пациенты, выжившие за 6-ти месячный период наблюдения после выписки из стационара. 2-ая группа n=65 (5,8%), сред. возраст 61,5±0,81 лет (люди из второй группы в среднем были старше на 2,8 лет, p=0,003) пациенты, умершие за 6-ти месячный период наблюдения после выписки из стационара, включая смерти в стационаре, от следующих диагнозов МКБ10: I20 – 1 (0,1%); I21 – 23 (2,3%); I22 – 7 (0,7%); I23 – 1 (0,1%); R57.0 – 9 (0,9%); 95,9% информации о диагнозах отсутствовала.

Неизвестные исходы в процессе наблюдения в исследование не включались. Конечная точка для первой группы была обозначена «1» - наступление события, для второй группы – «0» - событие не наступило. Для расчета бинарных значений прогноза за 1 принимали значения вероятностей $\geq 0,5$.

Для построения модели прогноза мы использовали шесть наиболее подходящих алгоритмов классификации на среде программирования Python версия: 3.11.1.:

- RandomForestClassifier из пакета sklearn.ensemble - реализация алгоритма случайного леса. Это непараметрический контролируемый метод машинного обучения, используемый для классификации. Цель состоит в том, чтобы создать модель, которая предсказывает значение целевой независимой переменной, изучая простые правила принятия решений, выведенные из характеристик данных;

- GradientBoostingClassifier из пакета sklearn.ensemble - реализация алгоритма градиентного бустинга. Суть метода заключается в построении ансамбля слабых моделей

(например, деревьев принятия решений), в которых (в отличие от бэггинга) модели строятся не независимо (параллельно), а последовательно. Это означает, что следующее дерево учится на ошибках предыдущего, затем этот процесс повторяется, наращивая количество слабых моделей;

- `ExtraTreesClassifier` из пакета `sklearn.ensemble` - реализация алгоритма дополнительных деревьев;

- `XGBClassifier` из пакета `xgboost` - реализация алгоритма градиентного бустинга;

- `LGBMClassifier` из пакета `lightgbm` - реализация алгоритма градиентного бустинга;

- `LogisticRegression` из пакета `sklearn.linear_model` - реализация алгоритма логистической регрессии. Статистическая модель, используемая для прогнозирования вероятности возникновения некоторого события путём его сравнения с логистической кривой. Эта регрессия выдаёт ответ в виде вероятности бинарного события (1 или 0).

Мы использовали данные алгоритмы поскольку `gradient boosting` и `random forest` вместе с глубоким обучением искусственных нейронных сетей (ИНС) входили в тройку наиболее часто используемых алгоритмов для создания прогнозов на табличных данных, а также имели наиболее высокую долю результатов с AUC выше 90% в ранее проведенном нами исследовании [13]. Также предпочтительным использование деревьев решений является их эффективность в выборе глобальных характеристик с наибольшим статистическим объемом информации [18-21]. Было принято решение не использовать ИНС, поскольку архитектуры глубокого обучения показывают хорошую производительность в задачах, связанных с «неструктурированными данными», такими как изображения, звук и текст в свободной форме [22].

Предикторы, включенные в анализ, составили $n=193$, охватывающие информацию об анамнезе, особенностях образа жизни, данных лабораторных и инструментальных исследований.

Все данные были подвергнуты очистке. Мы исключили переменные с количеством пропущенных значений $\geq 50\%$ - 86 переменных (44,6%). Данная операция была произведена с целью минимизации возникновения потенциальной ошибки при обучении.

Аномальные значения составили не более 0,01%. Поскольку процент оказался небольшим, данные значения были удалены.

Затем мы произвели операцию восстановления среди переменных с количеством пропущенных значений $<50\%$ - 107 переменных (55,4%).

В качестве алгоритма восстановления пропущенных значений был выбран способ IterativeImputer, включенный в пакет sklearn.impute языка программирования Python. Данный метод восстанавливает пропущенные данные, представляя каждый пропущенный признак как функцию других признаков с использованием алгоритма Round-robin.

Данный метод восстановления пропущенных значений был выбран в связи с особенностями датасета и, в частности, специфики пропущенных значений-признаков, поскольку поле, пропущенное в одной строке, необязательно коррелирует со значением поля другой строки, сходной по другим параметрам.

Благодаря использованию данного алгоритма нам удалось восстановить до 7,2% пропущенных значений благодаря чему удалось улучшить качество модели в среднем на 5%.

Поскольку датасет содержал относительно небольшое количество записей, для тестирования модели мы использовали алгоритм GridSearchCV из пакета sklearn.model_selection, который реализует метод кросс-валидации на 5 фолдах (5-fold cross-validation). Метрика качества - CV-AUC.

Для оценки эффективности моделей прогноза мы использовали показатели ROC-анализа (ROC - Receiver Operating Characteristic (рабочая характеристика приёмника). Результатом данного анализа служит метрика AUC – Area Under the Curve (площадь под кривой), которая позволяет оценить качество классификации. Данный показатель высчитывался на cross-validation (CV) выборке. ROC-кривую можно вычислить по формуле:

$$ROC = \text{sensitivity} / 1 - \text{specificity}, \text{ где}$$

Sensitivity (чувствительность) – показатель, демонстрирующий на сколько тест способен идентифицировать инстиноположительный результат, как таковой;

$\text{Sensitivity} = TP / P$, где TP – количество правильно распознанных положительных ответов, как таковых; P – количество положительных ответов;

Specificity (специфичность) – показатель, демонстрирующий на сколько тест способен идентифицировать инстиноотрицательный результат, как таковой;

$\text{Specificity} = TN / N$, где TN - количество правильно распознанных отрицательных ответов, как таковых; N – количество отрицательных ответов;

Accuracy (точность) – показатель, демонстрирующий сумму правильно распознанных как положительных, так и отрицательных ответов относительно всех случаев;

$$\text{Accuracy} = TP + TN / P + N.$$

Для оценки мультиколлинеарности использовали показатель VIF (Variance Inflation Factor, коэффициент инфляции дисперсии). За неудовлетворительный показатель принимали значения $VIF > 10$.

Для оценки влияния каждого предиктора на результат детерминации модели использовали два способа анализа: permutation importance и feature importance.

Для оценки распределения мы использовали тест Колмогорова-Смирнова для количественных переменных, одновыборочный критерий хи-квадрат для категориальных небинарных переменных и одновыборочный бинаминальный критерий для категориальных бинарных переменных.

Статистическая обработка данных производилась с использованием прикладных программ Microsoft Excel 2021 и IBM SPSS Statistics 27.

Результаты. В результате анализа все 19 количественных переменных (100%) имели нормальное распределение, 77 категориальных переменных (97,5%) имели распределение с равными вероятностями и 24 бинарные переменные (100%) были распределены с вероятностями 0,5 и 0,5.

Таблица 1.

Показатели ROC-анализа для вероятностных значений шести алгоритмов

Алгоритм	AUC	Стандартная ошибка	p	Асимптотический 95% доверительный интервал	
				Нижняя граница	Верхняя граница
Самарская когорта					
RandomForestClassifier	0,831	0,03	<0,001	0,771	0,89
GradientBoostingClassifier	0,821	0,035	<0,001	0,752	0,889
ExtraTreesClassifier	0,831	0,031	<0,001	0,77	0,891
XGBClassifier	0,8	0,033	<0,001	0,735	0,864
LGBMClassifier*	0,842	0,031	<0,001	0,781	0,903
LogisticRegression	0,796	0,031	<0,001	0,736	0,857

Примечание: *-лучший алгоритм

Как видно из таблицы 1, наилучшим алгоритмом оказался LGBMClassifier, наилучшие результаты показала логистическая регрессия. В среднем все алгоритмы ML продемонстрировали результаты выше 0,8.

Таблица 2.

Показатели значений C-статистики для шести алгоритмов при бинарных значениях классификации

Метрики	Алгоритмы					
	<i>RandomForestClassifier</i>	<i>GradientBoostingClassifier</i>	<i>ExtraTreesClassifier</i>	<i>XGBClassifier*</i>	<i>LGBMClassifier</i>	<i>LogisticRegression</i>
Чувствительность	0,992	0,996	0,994	0	0,997	0,983
Специфичность	0,4	0,215	0,323	0	0,308	0,231
Точность	0,961	0,95	0,957	0,938	0,956	0,951
AUC	0,698	0,606	0,658	0,5	0,652	0,607

Примечание: $p < 0,05$, $*-p > 0,05$

Как видно из таблицы 2 в среднем все алгоритмы, за исключением XGBClassifier, продемонстрировали высокие показатели чувствительности и точности, и достаточно низкие показатели специфичности. Такие результаты можно объяснить неравным соотношением исходов конечных точек (1:15). Таким образом алгоритмам не хватило информации для более точного обучения.

Для оценки наличия мультиколлинеарности датасета мы рассчитали VIF (variance inflation factor) с помощью `statsmodels.stats.outliers_influence import variance_inflation_factor`. VIF > 10 наблюдался у 9 (5,2%) переменных (максимальное значение – 22,66, средний показатель – 3,09).

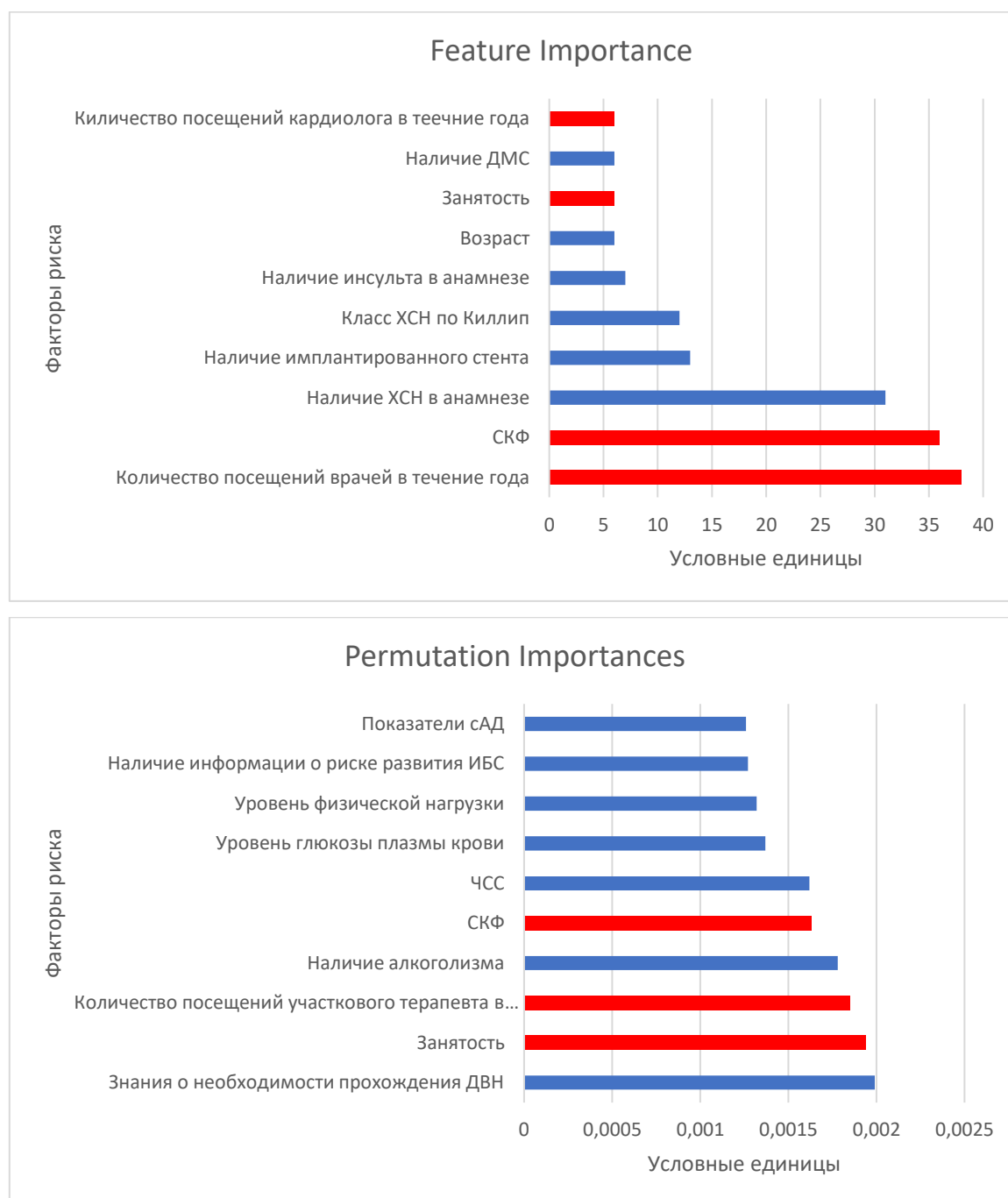


Рисунок 2. Ранжирование 10 наиболее значимых факторов риска для LGBMClassifier

Примечание: ■ - общие факторы риска для двух методов оценки

Для наилучшего алгоритма мы провели ранжирования ФР по степени их влияния на результат с помощью двух методов permutation importance и feature importance. Как видно из рисунка 2 для двух методов оценки наиболее значимыми факторами являлись наблюдение терапевта и кардиолога в течение года, занятость пациентов и показатели скорости

клубочковой фильтрации (СКФ). Также важную роль играли такие ФР, как прохождение диспансеризации, наличие сердечной недостаточности в анамнезе и коморбидность.

Обсуждение. Ссылаясь на результаты исследования, проведенного на той же выборке [4], с применением многофакторной модели Кокса, были достоверно выделены следующие факторы риска (ФР), оказывающих вклад в развитие неблагоприятных сердечно-сосудистых событий в постгоспитальный период: свойственные для ИМ изменения на ЭКГ при поступлении, мужской пол, возраст, курение и наличие среднего или ниже образования. Данные результаты отдаленно коррелируют с полученными нами результатами.

Шишкина и соавт. [3] на данных когорты, состоящих из пациентов, перенесших ИМ (n=155), провели собственное исследование с использованием однофакторного и многофакторного регрессионного анализа для прогноза 12-ти месячной летальности у лиц трудоспособного возраста. AUC модели составил 0,85, достоверно значимые ФР, влияющие на прогноз – конечный систолический объем при выписке $\geq 67,5$ мл, креатинин при поступлении ≥ 91 мкмоль/л, острая сердечная недостаточность при поступлении (ОСН) \geq III класса по Killip, фракция выброса левого желудочка (ФВЛЖ) при выписке $\leq 50\%$, частота сердечных сокращений при поступлении (ЧСС) ≥ 100 уд. в мин., тромбоцитарно-лимфоцитарный индекс при поступлении ≥ 122 , наличие ИМ в анамнезе.

Также Шишкина и соавт. [23] провели аналогичное исследование на выборке, содержащей данные о 424 пациентах трудоспособного возраста, перенесших ИМ. Период наблюдения составлял 12 месяцев. С использованием однофакторного и многофакторного анализов для прогностической модели удалось достичь AOC – 0,88, достоверно значимые предикторы: наличие постинфарктного кардиосклероза; высокая коморбидность, кардиогенный шок; ОСН II-IV класса по Killip при поступлении; ЧСС более 100 уд. в мин. при поступлении; ФВЛЖ $\leq 40\%$ при выписке; митральная регургитация при поступлении; нейтрофильно-лимфоцитарный индекс (НЛИ) $\geq 4,52$; уровень гемоглобина менее 130 г/л.

Сайгитов и соавт. [2] сделали прогноз постгоспитальной летальности, наблюдающихся по поводу ОКС. Когорта содержала данные о 567 пациентах, срок наблюдения 443 дня (14,8 мес.) Дискриминантная мощность (AUC) моделей многофакторного анализа составила от 0,72 до 0,92. Наиболее значимые ФР: возраст; повторный ИМ; наличие инсульта; наличие хронических неспецифических заболеваний легких; ЧСС при поступлении ≥ 110 уд. в мин.; депрессия сегмента ST; наличие патологического зубца Q; уровень гематокрита $< 40\%$;

уровень калия; хронические заболевания почек в анамнезе; ОН II-IV класса по Killip при поступлении; уровень креатинина более 110 мкмоль\л и уровень глюкозы более 6 ммоль\л.

Макарихин и соавт. [24] в своей работе по прогнозу развития ИМ использовали градиентный бустинг и деревья решений. Датасетом послужили данные ретроспективного исследования с количеством участников 7,557 тыс. человек. F1-score модели составил 0,74. Удалось выделить наиболее значимые ФР такие, как пол, возраст, уровень общего холестерина плазмы крови, уровень липопротеиды низкой плотности, индекс массы тела (ИМТ) и уровень гематокрита.

Швец и соавт. [25] провели аналогичное исследование, где использовали сверточную ИНС и ИНС прямого распространения. Когорта содержала данные о 400 пациентах, прошедших лечение по поводу ИМ. Результаты AUC сверточной и ИНС прямого распространения составили 0,74 и 0,63 соответственно. Для выделения значимых предикторов авторы использовали однофакторный анализ. Достоверно удалось выделить влияние возраста, функционального класса сердечной недостаточности, СКФ, ФВЛЖ, ИМ в анамнезе, ИМТ, коморбидности и ЧСС.

Chad J Zack и соавт. [26] для прогноза рисков 30-ти дневной повторной госпитализации и смерти в течение 180-ти дней после перенесенного ИМ использовали алгоритм случайного леса. Выборка насчитывала 11,709 тыс. пациентов клиники Майо. Для тестирования проводилось методом кросс-валидации на 8-ми фолдах. AUC модели составил 0,93. Наиболее значимые ФР – возраст, кардиогенный шок при поступлении и застойная сердечная недостаточность.

Rashmee U Shah и соавт. [27] использовали модель пропорциональных рисков (Сох-регрессия) на выборке, состоящей из 112,668 тыс. пациентов старше 65 лет из ACTION Registry-GWTG, перенесших ИМ. ФР, ассоциируемые с одногодичной постгоспитальной летальностью – возраст, уровень креатинина, проведение ЧКВ, количество госпитализаций за год, уровень гемоглобина, коморбидность, пол, кардиогенный шок, уровень тропонина, наличие ИМ в анамнезе, статус курения, дислипидемия.

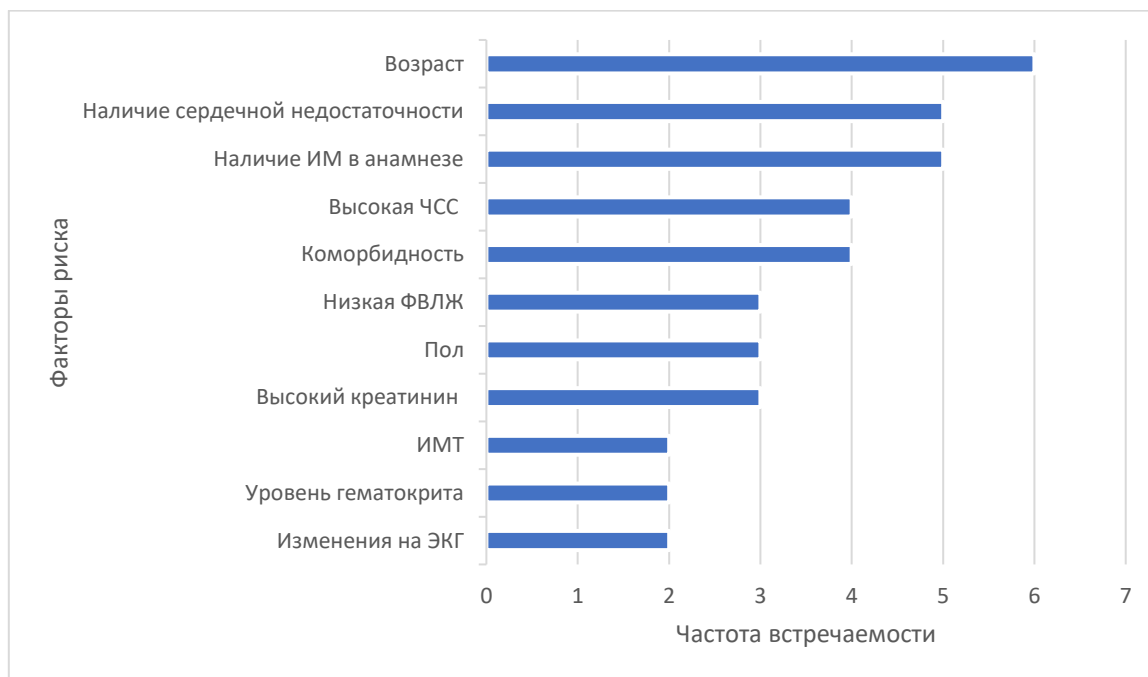


Рисунок 3. Частота встречаемости наиболее значимых ФР, оказывающих значительный вклад в прогноз постгоспитальной летальности от ИМ среди 8-и анализируемых работ

Таким образом, в результате анализа аналогичных работ нам удалось соотнести полученные нами результаты с результатами проведенных исследований и выявить наиболее значимые ФР.

Выводы. В результате исследования нам удалось сконструировать модель классификации с достаточно хорошим качеством детерминации. Также нам удалось выделить наиболее значимые ФР, оказывающие влияние на развитие 6-ти месячных осложнений у пациентов, перенесших ИМ. Дальнейшие изучения данной темы поможет лучше понять природу развития патологических изменений и заранее предупреждать неблагоприятные исходы.

Список литературы

1. Федеральная служба государственной статистики. Здравоохранение. Заболеваемость. 2023. <https://rosstat.gov.ru/folder/13721>
2. Сайгитов Р. Т., Глезер М. Г., Семенцов Д. П., и др. Прогнозирование постгоспитальной летальности у мужчин и женщин, наблюдавшихся по поводу острого коронарного синдрома. Российский кардиологический журнал. 2006;(3):41-48.

3. Шишкина Е.А., Агафонов А.В., Хлынова О.В. и др. Возможности прогнозирования постгоспитальной летальности у больных трудоспособного возраста с инфарктом миокарда без стойкого подъема сегмента ST. *Врач.* 2020;31(10):40–45. doi:10.29296/25877305-2020-10-07
4. Шальнова С. А., Драпкина О. М., Куценко В. А., и др. Инфаркт миокарда в популяции некоторых регионов России и его прогностическое значение. *Российский кардиологический журнал.* 2022;27(6):9-19. doi:10.15829/1560-4071-2022-4952
5. Than M, Flaws D, Sanders S, et al. Development and validation of the Emergency Department Assessment of Chest pain Score and 2 h accelerated diagnostic protocol. *Emerg Med Australas.* 2014 Feb;26(1):34-44. doi: 10.1111/1742-6723.12164. Epub 2014 Jan 15. PMID: 24428678
6. Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Heart J.* 2008;16(6):191-196. doi:10.1007/BF03086144
7. Fox KA, Dabbous OH, Goldberg RJ, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: prospective multinational observational study (GRACE). *BMJ (Clinical research ed.).* 2006;333(7578):1091. doi:10.1136/bmj.38985.646481.55
8. Гусев А. В., Кузнецова Т. Ю., Корсаков И. Н. Искусственный интеллект в оценке рисков развития сердечно-сосудистых заболеваний. *Журнал телемедицины и электронного здравоохранения.* 2018;3(8):85-90.
9. Гусев А.В., Гаврилов Д.В., Корсаков И.Н. и др. Перспективы использования методов машинного обучения для предсказания сердечно-сосудистых заболеваний. *Врач и информационные технологии.* 2019;(3):41-47.
10. Гусев А. В., Гаврилов Д. В., Новицкий Р. Э., и др. Совершенствование возможностей оценки сердечно-сосудистого риска при помощи методов машинного обучения. *Российский кардиологический журнал.* 2021;(12):171-180.
11. Гусев А.В., Владзимирский А.В., Шарова Д.Е. и др. Развитие исследований и разработок в сфере технологий искусственного интеллекта для здравоохранения в Российской Федерации: итоги 2021 года. *Digital Diagnostics.* 2022;3(3):178–194. doi:10.17816/DD107367
12. Мишкин И. А., Концевая А. В., Гусев А. В. и др. Прогнозирование сердечно-сосудистых событий с помощью моделей пропорциональных рисков и моделей машинного

обучения: систематический обзор. Современные проблемы здравоохранения и медицинской статистики. 2023;(2):804-829 doi:10.24412/2312-2935-2023-2-804-829.

13. Мишкин И.А., Гусев А.В., Концевая А.В. и др. Эффективность использования mHealth в качестве инструмента профилактики сердечно-сосудистых заболеваний. Систематический обзор. Врач и информационные технологии. 2022;4:12-27. doi:10.25881/18110193_2022_4_12.

14. Концевая А. В., Бейтс К., Горячкин Е. А., и др. Госпитальный этап лечения инфаркта миокарда в 13 регионах Российской Федерации по результатам международного исследования. Рациональная фармакотерапия в кардиологии. 2018;14(4):474-487.

15. Kontsevaya A, Bobrova N, Barbarash O, et al. The management of acute myocardial infarction in the Russian Federation: protocol for a study of patient pathways. Wellcome Open Res. 2017;2:89. Published 2017 Sep 25. doi:10.12688/wellcomeopenres.12478.2

16. Kontsevaya AV, Bates K, Schirmer H, et al. Management of patients with acute ST-segment elevation myocardial infarction in Russian hospitals adheres to international guidelines. Open Heart. 2020;7(1):e001134. Published 2020 Jan 23. doi:10.1136/openhrt-2019-001134

17. Kontsevaya A, Sabgaida T, Ivanova A, et al. How has the management of acute coronary syndrome changed in the Russian Federation during the last 10 years?. Health Policy. 2017;121(12):1274-1279. doi:10.1016/j.healthpol.2017.09.018

18. Grabczewski K, Jankowski N. Feature selection with decision tree criterion. Fifth International Conference on Hybrid Intelligent Systems (HIS'05). 2005:6.

19. Tin Kam H. The random subspace method for constructing decision forests. PAMI 20. 1998;8:832–844.

20. Tianqi C, Carlos G. XGBoost: A Scalable Tree Boosting System. In KDD. 2016.

21. Guolin K, Qi M, Thomas F, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In NIPS. 2017.

22. Что такое глубокое обучение? <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/deep-learning-overview>

23. Шишкина Е. А., Хлынова О. В., Черемных А. Б. Прогнозирование постгоспитальной летальности у больных инфарктом миокарде молодого и среднего возраста. Доктор.Ру. 2020;19(5):24-29. doi: 10.31550/1727-2378-2020-19-5-24-29 Г

24. Макарихин А.В., Немков А.Г., Рейтблат О.М., и др. Разработка прогностической модели наступления инфаркта миокарда на основе технологий машинного обучения. Вестник новых медицинских технологий. 2021;28(4):103-106. doi:10.24412/1609-2163-2021-4-103-106
25. Швец Д. А., Карасёв А. Ю., Смоляков М. В., и др. Нейросетевой анализ предикторов летального риска у больных после перенесенного острого коронарного синдрома. Российский кардиологический журнал. 2020;(3):68-75. doi:10.15829/1560-4071-2020-3-3645
26. Zack CJ, Senecal C, Kinar Y, et al. Leveraging Machine Learning Techniques to Forecast Patient Prognosis After Percutaneous Coronary Intervention. JACC Cardiovasc Interv. 2019;12(14):1304-1311. doi:10.1016/j.jcin.2019.02.035
27. Shah RU, de Lemos JA, Wang TY, et al. Post-Hospital Outcomes of Patients With Acute Myocardial Infarction With Cardiogenic Shock: Findings From the NCDR. J Am Coll Cardiol. 2016;67(7):739-747. doi:10.1016/j.jacc.2015.11.048

References

1. Federal'naya sluzhba gosudarstvennoj statistiki. Zdravooхранenie. Zaboлеваemost'. 2023. [Federal State Statistics Service. Healthcare. Morbidity. 2023.] (in Russian) <https://rosstat.gov.ru/folder/13721>
2. Sajgitov R. T., Glezer M. G., Semencov D. P., i dr. Prognozirovanie postgospital'noj letal'nosti u muzhchin i zhenshchin, nablyudavshihsy po povodu ostrogo koronarnogo sindroma. [Prediction of posthospital mortality in men and women who were observed for acute coronary syndrome] Rossijskij kardiologicheskij zhurnal. [Russian Journal of Cardiology]. 2006;(3):41-48. (in Russian)
3. SHishkina E.A., Agafonov A.V., Hlynova O.V. i dr. Vozmozhnosti prognozirovaniya postgospital'noj letal'nosti u bol'nyh trudosposobnogo vozrasta s infarktomiokarda bez stojkogo pod"ema segmenta ST. [Possibilities of predicting posthospital mortality in working-age patients with myocardial infarction without persistent ST segment elevation]. Vrach. [Doctor]. 2020;31(10):40–45. (in Russian) doi:10.29296/25877305-2020-10-07
4. SHal'nova S. A., Drapkina O. M., Kucenko V. A., i dr. Infarkt miokarda v populyacii nekotoryh regionov Rossii i ego prognosticheskoe znachenie. [Myocardial infarction in the population of some regions of Russia and its prognostic value] Rossijskij kardiologicheskij zhurnal. [Russian Journal of Cardiology]. 2022;27(6):9-19. (in Russian) doi:10.15829/1560-4071-2022-4952

5. Than M, Flaws D, Sanders S, et al. Development and validation of the Emergency Department Assessment of Chest pain Score and 2 h accelerated diagnostic protocol. *Emerg Med Australas.* 2014 Feb;26(1):34-44. doi: 10.1111/1742-6723.12164. Epub 2014 Jan 15. PMID: 24428678
6. Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Heart J.* 2008;16(6):191-196. doi:10.1007/BF03086144
7. Fox KA, Dabbous OH, Goldberg RJ, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: prospective multinational observational study (GRACE). *BMJ (Clinical research ed.).* 2006;333(7578):1091. doi:10.1136/bmj.38985.646481.55
8. Gusev A. V., Kuznecova T. YU., Korsakov I. N. Iskusstvennyj intellekt v ocenke riskov razvitiya serdechno-sosudistyh zabolevanij. [Artificial intelligence in risk assessment of cardiovascular diseases] *ZHurnal telemeditsiny i elektronogo zdravooohraneniya.* [Journal of Telemedicine and E-Health]. 2018;3(8):85-90. (in Russian)
9. Gusev A.V., Gavrilov D.V., Korsakov I.N. et al. Perspektivy ispol'zovaniya metodov mashinnogo obucheniya dlya predskazaniya serdechno-sosudistyh zabolevanij [Prospects of using machine learning methods to predict cardiovascular diseases]. *Vrach i informacionnye tekhnologii* [Doctor and Information Technologies]. 2019;(3):41-47. (in Russian)
10. Gusev A.V., Gavrilov D.V., Novitsky R.E., et al. Sovershenstvovanie vozmozhnostej ocenki serdechno-sosudistogo riska pri pomoshchi metodov mashinnogo obucheniya [Improving the possibilities of assessing cardiovascular risk using machine learning methods]. *Rossijskij kardiologicheskij zhurnal* [Russian Journal of Cardiology]. 2021;26(12):171-180. (in Russian) doi:10.15829/1560-4071-2021-4618
11. Gusev A.V., Vladzimirskij A.V., SHarova D.E. i dr. Razvitie issledovanij i razrabotok v sfere tekhnologij iskusstvennogo intellekta dlya zdravooohraneniya v Rossijskoj Federacii: itogi 2021 goda. [Development of research and development in the field of artificial intelligence technologies for healthcare in the Russian Federation: results of 2021] *Digital Diagnostics.* [Digital Diagnostics]. 2022;3(3):178–194. (in Russian) doi:10.17816/DD107367
12. Mishkin I. A., Koncevaya A. V., Gusev A. V. i dr. Prognozirovanie serdechno-sosudistyh sobytij s pomoshch'yu modelej proporcional'nyh riskov i modelej mashinnogo obucheniya: sistemicheskij obzor. [Prediction of cardiovascular events using proportional risk models and machine learning models: a systematic review] *Sovremennye problemy zdravooohraneniya i*

medicinskoj statistiki. [Modern problems of healthcare and medical statistics]. 2023;(2):804-829. (in Russian) doi:10.24412/2312-2935-2023-2-804-829.

13. Mishkin I.A., Gusev A.V., Koncevaya A.V. i dr. Effektivnost' ispol'zovaniya mHealth v kachestve instrumenta profilaktiki serdechno-sosudistyh zabolevanij. Sistematicheskij obzor. [The effectiveness of using mHealth as a tool for the prevention of cardiovascular diseases. Systematic review] Vrach i informacionnye tekhnologii. [Doctor and information technology]. 2022;4:12-27. (in Russian) doi:10.25881/18110193_2022_4_12.

14. Koncevaya A. V., Bejts K., Goryachkin E. A., i dr. Gospital'nyj etap lecheniya infarkta miokarda v 13 regionah Rossijskoj Federacii po rezul'tatam mezhdunarodnogo issledovaniya. [The hospital stage of treatment of myocardial infarction in 13 regions of the Russian Federation according to the results of an international study] Racional'naya farmakoterapiya v kardiologii. [Rational pharmacotherapy in cardiology]. 2018;14(4):474-487. (in Russian)

15. Kontsevaya A, Bobrova N, Barbarash O, et al. The management of acute myocardial infarction in the Russian Federation: protocol for a study of patient pathways. Wellcome Open Res. 2017;2:89. Published 2017 Sep 25. doi:10.12688/wellcomeopenres.12478.2

16. Kontsevaya AV, Bates K, Schirmer H, et al. Management of patients with acute ST-segment elevation myocardial infarction in Russian hospitals adheres to international guidelines. Open Heart. 2020;7(1):e001134. Published 2020 Jan 23. doi:10.1136/openhrt-2019-001134

17. Kontsevaya A, Sabgaida T, Ivanova A, et al. How has the management of acute coronary syndrome changed in the Russian Federation during the last 10 years?. Health Policy. 2017;121(12):1274-1279. doi:10.1016/j.healthpol.2017.09.018

18. Grabczewski K, Jankowski N. Feature selection with decision tree criterion. Fifth International Conference on Hybrid Intelligent Systems (HIS'05). 2005:6.

19. Tin Kam H. The random subspace method for constructing decision forests. PAMI 20. 1998;8:832–844.

20. Tianqi C, Carlos G. XGBoost: A Scalable Tree Boosting System. In KDD. 2016.

21. Guolin K, Qi M, Thomas F, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In NIPS. 2017.

22. CHto takoe glubokoe obuchenie? [What is deep learning?]. (in Russian)<https://learn.microsoft.com/ru-ru/dotnet/machine-learning/deep-learning-overview>

23. SHishkina E. A., Hlynova O. V., CHERemnyh A. B. Prognozirovanie postgospital'noj letal'nosti u bol'nyh infarktom miokarde molodogo i srednego vozrasta. [Prediction of posthospital

mortality in patients with myocardial infarction of young and middle age] *Doktor.Ru.* [Doktor.Ru.]. 2020;19(5):24-29. (in Russian) doi: 10.31550/1727-2378-2020-19-5-24-29 Г

24. Makarihin A.V., Nemkov A.G., Rejtlat O.M., i dr. Razrabotka prognosticheskoy modeli nastupleniya infarkta miokarda na osnove tekhnologij mashinnogo obucheniya. [Development of a predictive model of the onset of myocardial infarction based on machine learning technologies] *Vestnik novyh medicinskih tekhnologij* [Bulletin of New Medical Technologies]. 2021;28(4):103-106. (in Russian) doi:10.24412/1609-2163-2021-4-103-106

25. SHvec D. A., Karasyov A. YU., Smolyakov M. V., i dr. Nejrosetevoj analiz prediktorov letal'nogo riska u bol'nyh posle perenesennogo ostrogo koronarnogo sindroma. [Neural network analysis of predictors of lethal risk in patients after acute coronary syndrome] *Rossijskij kardiologicheskij zhurnal.* [Russian Journal of Cardiology]. 2020;(3):68-75. (in Russian) doi:10.15829/1560-4071-2020-3-3645

26. Zack CJ, Senecal C, Kinar Y, et al. Leveraging Machine Learning Techniques to Forecast Patient Prognosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv.* 2019;12(14):1304-1311. doi:10.1016/j.jcin.2019.02.035

27. Shah RU, de Lemos JA, Wang TY, et al. Post-Hospital Outcomes of Patients With Acute Myocardial Infarction With Cardiogenic Shock: Findings From the NCDR. *J Am Coll Cardiol.* 2016;67(7):739-747. doi:10.1016/j.jacc.2015.11.048

Финансирование. Исследование не имело спонсорской поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Acknowledgments. The study did not have sponsorship.

Conflict of interests. The authors declare no conflict of interest.

Сведения об авторах

Мишкин Илья Александрович – аспирант ФГБУ "Национального медицинского исследовательского центра терапии и профилактической медицины" Министерства здравоохранения Российской Федерации, 101990, Россия, Москва, Петроверигский пер., 10, стр. 3; врач-терапевт ГУЗ ТО «Киреевской ЦРБ», 301260, Россия, Киреевск, ул. Ленина 44. e-mail: Iya.mischkin@yandex.ru, ORCID 0000-0003-4850-0648; SPIN: 4086-1507

Концевая Анна Васильевна – доктор медицинских наук, заместитель директора по научной и аналитической работе ФГБУ "Национального медицинского исследовательского центра терапии и профилактической медицины" Министерства здравоохранения Российской Федерации 101990, Россия, Москва, Петроверигский пер., 10, стр. 3. e-mail: koncanna@yandex.ru, ORCID [0000-0003-2062-1536](https://orcid.org/0000-0003-2062-1536); SPIN: 6787-2500

Гусев Александр Владимирович – кандидат технических наук, член экспертного совета Минздрава по вопросам использования ИКТ, старший научный сотрудник ФГБУ «Центрального научно-исследовательского института организации и информатизации здравоохранения» Министерства здравоохранения Российской Федерации, 127254, Москва, ул. Добролюбова 11; директор по развитию компании «К-Скай», 185031, Россия, Петрозаводск, наб. Варкауса, 17. e-mail: agusev@webiomed.ai, ORCID 0000-0002-7380-8460; SPIN: 9160-7024

Сахаров Андрей Алексеевич – программист, ведущий разработчик ПАО «Группа Ренессанс Страхование», 300035, Россия, Тула, ул. Софьи Перовской 4, e-mail: andrei.saharoff3011@gmail.com

Драпкина Оксана Михайловна – доктор медицинских наук, профессор, академик РАН, директор, ФГБУ "Национального медицинского исследовательского центра терапии и профилактической медицины" Министерства здравоохранения Российской Федерации, 101990, Россия, Москва, Петроверигский пер., 10, стр. 3. e-mail: drapkina@bk.ru, ORCID 0000-0002-4453-8430; SPIN: 4456-1297

About the authors

Mishkin Ilya Alexandrovich - Ph.D. student of the Federal State Budgetary Institution National Medical Research Center for Therapy and Preventive Medicine of the Ministry of Healthcare of the Russian Federation, 101990, Russia, Moscow, Petroverigsky lane, 10, p. 3; physician of Tula State Healthcare Institution district Kireevskaya central district hospital, 301260, Russia, Kireevsk, 44 Lenin Street. e-mail: Ilya.mischckin@yandex.ru, ORCHID 0000-0003-4850-0648; SPIN: 4086-1507

Kontsevaya Anna Vasilievna - Ph.D., Deputy Director for Scientific and Analytical Work of the Federal State Budgetary Institution National Medical Research Center for Therapy and Preventive Medicine of the Ministry of Healthcare of the Russian Federation, 101990, Russia, Moscow, Petroverigsky lane, 10, p. 3. e-mail: koncanna@yandex.ru, ORCID 0000-0003-2062-1536; SPIN: 6787-2500

Gusev Alexandr Vladimirovich - Ph.D. of Engineering Sciences Candidate, member of the Russian Federation Expert Healthcare Ministry Council on the use of information computer technologies, senior Researcher at the Russian Research Institute of Health, 127254, Russia, Moscow, Dobrolyubova str. 11; development director of the "K-Sky" company, 185031, Russia, Petrozavodsk, nab. Varkaus, 17. e-mail: agusev@webiomed.ai, ORCID 0000-0002-7380-8460; SPIN: 9160-7024

Sakharov Andrey Alekseevich – Programmer, Lead Developer of Renaissance Insurance Group PJSC, 300035, Russia, Tula, Sofya Perovskaya Str. 4, e-mail: andrei.saharoff3011@gmail.com

Drapkina Oksana Mihaylovna - Ph.D., Director of the Federal State Budgetary Institution National Medical Research Center for Therapy and Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Professor, Academician of the Sciences Russian Academy, Chief Freelance Specialist in Therapy and General Medical Practice of the Russia Healthcare Ministry, Honored Doctor of the Russian Federation, e-mail: drapkina@bk.ru, ORCID 0000-0002-4453-8430; SPIN: 4456-1297

Статья получена: 10.09.2023 г.
Принята к публикации: 25.12.2023 г.