

СОВРЕМЕННЫЕ ПОДХОДЫ К ПРОГНОЗИРОВАНИЮ РИСКА РАЗВИТИЯ НЕБЛАГОПРИЯТНЫХ СЕРДЕЧНО-СОСУДИСТЫХ СОБЫТИЙ У ПАЦИЕНТОВ С ОСТРЫМ КОРОНАРНЫМ СИНДРОМОМ

© Снимщикова И.А., Ревякина М.О., Строев С.П.

Орловский государственный университет имени И.С. Тургенева (ОГУ им. И.С. Тургенева)

Россия, 302026, Орловская область, г. Орёл, ул. Комсомольская, д. 95

Сердечно-сосудистые заболевания представляют собой основную причину заболеваемости и смертности населения всего мира. Несмотря на достигнутый прогресс в прогнозировании риска развития и неблагоприятных сердечно-сосудистых исходов, остается большое количество пациентов, которые не распознаются при использовании стандартных прогнозных моделей.

Цель – оценка возможности применения машинного обучения для построения мультимаркерной модели прогноза исхода острого коронарного синдрома в виде острого инфаркта миокарда или нестабильной стенокардии.

Материал и методы. Разработка прогностических моделей проводилась на языке программирования Python (библиотеки Pandas, Seaborn и Scikit-learn) с использованием базы данных, составленной на основании результатов клинико-лабораторных исследований 60 пациентов в возрасте от 18 до 59 лет с острым коронарным синдромом.

Результаты. Из общего перечня, включающего 181 показатель (10860 значений), были отобраны 3 наиболее информативных биомаркера, используемых для идентификации установленных конечных точек (диагноза пациента). Классификатор «случайный лес» превзошел остальные модели в отношении точности, чувствительности и специфичности, что позволило с 96% точностью определять наступление неблагоприятного исхода ОКС.

Заключение. На основе полученных данных была разработана программа для ЭВМ «Калькулятор расчета вероятностного исхода острого коронарного синдрома» для практического использования в учреждениях здравоохранения.

Ключевые слова: острый инфаркт миокарда; нестабильная стенокардия; машинное обучение; прогностическая модель; острый коронарный синдром.

Снимщикова Ирина Анатольевна – д-р мед. наук, профессор, директор Медицинского института, зав. кафедрой иммунологии и специализированных клинических дисциплин, вед. науч. с. лаборатории новых медицинских технологий, ОГУ им. И.С. Тургенева, г. Орёл. ORCID iD: 0000-0002-4258-963X. E-mail: snimshikova@mail.ru

Ревякина Мария Олеговна – канд. мед. наук, доцент кафедры иммунологии и специализированных клинических дисциплин, ст. науч. с. лаборатории новых медицинских технологий, ОГУ им. И.С. Тургенева, г. Орёл. ORCID iD: 0000-0003-1593-5290. E-mail: revyakina_masha@mail.ru (автор, ответственный за переписку)

Строев Сергей Павлович – канд. экон. наук, доцент, зав. кафедрой алгебры и математических методов в экономике, ОГУ им. И.С. Тургенева, г. Орёл. ORCID iD: 0000-0002-2271-5264. E-mail: stroewsp@mail.ru

В Российской Федерации, как и во всем мире, сердечно-сосудистые заболевания (ССЗ) являются основной причиной заболеваемости и смертности населения, представляя собой не только острую медицинскую, но и социально-экономическую проблему [1].

По данным Росстата, Орловская область занимает одну из лидирующих позиций по смертности от ССЗ, показатель которой в 2023 году составил 901,4 на 100 тыс. населения и превысил его значение по РФ в целом более чем в 1,5 раза. При этом в структуре кардиоваскулярной патологии в Орловской области лидирует смертность от ишемической болезни сердца: 569,9 человек на 100 тыс. населения [2, 3].

За последние годы в мире благодаря более широкому использованию инвазивных технологий и терапевтических стратегий, включающих применение антиагрегантных и антикоагулянтных препаратов, вторичной профилактики, направленной на снижение скорости про-

грессирования атеросклероза и особенно улучшению стратификации риска, достигнут определенный прогресс в снижении смертности от острых форм ИБС [4, 5].

Вместе с тем, несмотря на наличие различных подходов для прогнозирования риска ССЗ и неблагоприятных сердечно-сосудистых событий (GRACE – Global Registry of Acute Coronary Events, Framingham Risk Score и др.), остается большое количество пациентов с высоким риском ССЗ, которые данными инструментами не распознаются [6-8].

Использование машинного обучения в комбинации с биомаркерами, отражающими различные патофизиологические процессы, такие как повреждение миокарда, воспаление, нестабильность атеросклеротической бляшки и др., позволяет создавать более точные персонализированные прогностические модели для диагностики и стратификации риска при ОКС [4, 9–11].

По мере совершенствования понимания патогенеза ОКС и открытия новых биомаркеров, мультимаркерная парадигма все шире используется в построении прогностических моделей, в том числе в комбинации со стандартными клинико-лабораторными показателями и факторами риска ССЗ, что позволяет существенно улучшить результаты лечения пациентов с ОКС [4, 9, 10, 12].

Учитывая это, возникает насущная необходимость разработки и создания новых алгоритмов диагностики на основе мультимаркерных прогностических моделей, включающих патогенетически значимые биомаркеры оценки состояния и прогнозных исходов острых форм ИБС.

Использование данного подхода с расширением набора исследуемых признаков и современных технологий анализа данных эмпирических исследований, в частности методов машинного обучения, может быть эффективным для оценки риска неблагоприятного исхода ССЗ в разные периоды заболевания.

К настоящему моменту времени имеется достаточно много публикаций, описывающих применение алгоритмов классификации к диагностированию сердечно-сосудистых заболеваний [11, 13, 14]. В зависимости от располагаемых данных и постановки задачи для построения подобных моделей наиболее часто применяются методы множественного регрессионного анализа, методы классификации (метод логистической регрессии, метод опорных векторов, случайный лес, градиентный бустинг) и нейронные сети [15-17].

При этом отмечается, что использование линейных моделей сопряжено с определенными недостатками. Во-первых, в этом случае предполагается наличие линейной зависимости между заболеванием и совокупным воздействием исследуемых признаков, а во-вторых, признаки должны быть линейно независимы между собой.

Применение нейросетевых технологий предполагает наличие значительных объемов данных. В условиях эмпирических исследований, характеризующихся, как правило, небольшими объемами данных, применение нейронных сетей к построению прогностических моделей развития сердечно-сосудистых заболеваний существенно ограничено [7, 17-19]. В связи с чем в данной работе для построения классификационных моделей были использованы нелинейные модели.

Целью работы явилась оценка возможности применения машинного обучения для построения мультимаркерной модели прогноза исхода

ОКС (острый инфаркт миокарда или нестабильная стенокардия).

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

В проводимое исследование вошли 60 пациентов в возрасте от 18 до 59 лет, госпитализированных в кардиологические отделения регионального сосудистого центра БУЗ Орловской области «Орловская областная клиническая больница» с предварительным диагнозом ОКС.

С учетом клинико-лабораторного и инструментального обследования, проведенного в приемном покое, пациенты были разделены на 2 группы: I группа – больные острым инфарктом миокарда (31 человек, возраст $51,2 \pm 6,3$ года), II группа – пациенты с нестабильной стенокардией (29 человек, возраст $52,2 \pm 5,9$ года), в которых преобладали мужчины, составив 80% общей когорты.

Обследование и лечение пациентов в стационаре осуществлялось в рамках действующих клинических рекомендаций в соответствии с диагнозом.

На каждого пациента заполнялась унифицированная электронная карта и присваивался уникальный цифровой код, дальнейшая интерпретация и обработка результатов проводилась в обезличенной форме. В работу не включались пациенты с неполными результатами обследования.

Во время госпитализации больных осуществлялось построение классификационных моделей для прогнозирования исхода ОКС в соответствии с поставленной целью на основе использования алгоритмов машинного обучения.

Все данные, касающиеся исходных характеристик, собирались как первое значение в течение 24 часов после поступления пациента в клинику. Исходная выборочная совокупность, ассоциированная с известным конечным диагнозом, состояла из 181 результата исследования на каждого пациента. Анализируемые переменные включали основные демографические данные (возраст, пол, факторы риска ССЗ), результаты лабораторных показателей периферической крови, данных электрокардиографии, эхокардиографии, диагностической коронароангиографии.

Для улучшения традиционного способа прогнозирования вероятностного исхода ОКС и на основе наших предыдущих исследований морфогенных белков WNT-сигнального пути и оксидативного статуса, показавших их существенную прогностическую ценность при оценке течения и прогноза ССЗ, нами была выбрана

панель биомаркеров, включавшая β -катенин, склеростин, WIF-1, DVL-1, GSK-3 β , GSK-3 α , оксид азота, супероксиддисмутазу [20, 21]. Уровень основных белков WNT-сигнального пути и супероксиддисмутазы определяли методом иммуноферментного анализа (ELISA). Сывороточный уровень оксида азота оценивался непрямой метод по стандартной методике с использованием реактива Грисса.

Электронная база данных пациентов (с известными конечными точками – клиническим диагнозом) была сформирована из результатов обследования больных ОКС с использованием программы для работы с электронными таблицами – Microsoft Excel.

Предварительная обработка данных, выявление статистически значимых различий в значениях показателей и последующая разработка прогностических моделей проводились с использованием языка программирования Python (библиотек Pandas, Seaborn и Scikit-learn).

Общий объем выборки включал 10860 анализируемых значений. Формирование обучающего и тестирующего массивов для обучения классификаторов производилось случайным образом в соотношении 80/20, подбор наилучших гиперпараметров осуществлялся с помощью компоненты Gridsearch (рис. 1).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Острый коронарный синдром, связанный с полной или частичной окклюзией коронарного кровотока, приводит к развитию нестабильной стенокардии (НС), инфаркта миокарда без подъема сегмента ST или инфаркта миокарда с подъемом сегмента ST, являясь одной из основных причин смерти во всем мире [22, 23].

В настоящее время в мире для диагностики ОКС используется целый ряд критериев, включающих факторы риска (курение, артериальная гипертензия и др.), наличие семейного анамнеза по ОКС, клиническую картину, данные ЭКГ в 12 отведениях и повышение уровня hscTn тропонина, позиционируемого как «золотой стандарт» дифференциальной диагностики острого инфаркта миокарда [23-25].

Однако эти методы связаны со многими ограничениями, что приводит как к недостаточной, так и гипердиагностике острого инфаркта миокарда. Так, показатели ЭКГ часто неспецифичны и могут отсутствовать примерно у 35-40% пациентов, а повышение уровня тропонина может регистрироваться при ряде других сердечно-сосудистых заболеваний, что усугубляет диагностическую дилемму [22-26].

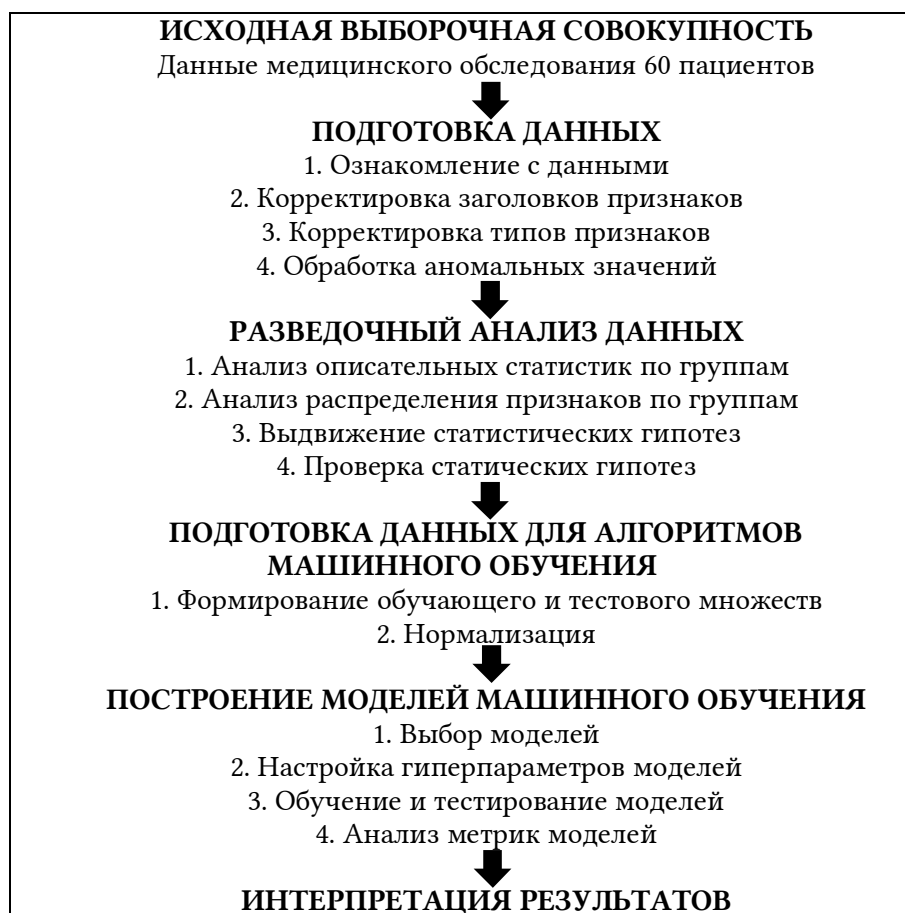


Рис. 1. Блок-схема, описывающая общую концепцию исследования.

Fig. 1. Flow chart describing the general concept of the study.

Машинное обучение является мощным инструментом, охватывающим широкий спектр методов, посредством которых искусственный интеллект учится выполнять задачи по анализу больших объемов данных (Big Data) для составления прогнозов [13, 18, 19].

Одним из ключевых направлений применения машинного обучения в диагностике ОКС является разработка прогностических моделей, которые могут помочь специалистам как первичного звена здравоохранения, так и специализированного кардиологического профиля, в оценке риска развития осложнений и принятии клинических решений. Такие модели, как правило, включают в себя комбинацию различных клинических, лабораторных и инструментальных данных, в том числе биомаркеры [4, 9, 4, 13].

Хотя во многих исследованиях, посвященных оценкам течения и исходов ОКС, предложены к использованию различные алгоритмы машинного обучения и представлены прогностические модели более точной идентификации ОКС [4, 11, 13, 15], все еще остаются актуальными научные изыскания, оценивающие возможности использования машинного обучения для разработки оптимальной панели биомаркеров сыворотки крови, способных независимо диагностировать вероятностные исходы ОКС в виде нестабильной стенокардии или острого инфаркта миокарда.

Первоначально с целью подготовки исходных данных для разведочного анализа проводилась их предварительная обработка: проверка на наличие дублирующихся записей, аномальных значений и пропусков по каждому признаку. С помощью методов библиотеки Pandas было установлено, что в исходных данных отсутствуют дубликаты и пропущенные значения.

Анализ аномальных значений признаков осуществлялся посредством изучения описательных статистик (минимальное и максимальное значения, первого и третьего квартилей, медианы) с их последующей визуализацией в виде

«Box-plot». Записи, содержащие аномальные значения признаков, удалялись. В результате объем исследуемой выборки сократился до 49 записей, из которых 25 относились к пациентам I группы (острый инфаркт миокарда) и 24 – к больным II группы (нестабильная стенокардия).

Следующим этапом проводился эксплораторный факторный анализ данных с использованием библиотек Pandas и Seaborn языка программирования Python для формирования единого признакового пространства, элементы которого в дальнейшем использовались для построения классификационных моделей. В нашем исследовании были использованы алгоритмы машинного обучения, включающие 5 классификаторов (стохастический градиент, метод опорных векторов, логистическая регрессия, случайный лес и градиентный бустинг), которые были применены к выборке из 10860 анализируемых показателей (181 результат исследования на каждого из 60 пациентов).

С помощью методов корреляционного анализа были отобраны наиболее значимые показатели, характеризующиеся высокой степенью линейной зависимости с целевым показателем – конечным диагнозом пациента.

Проверка гипотез о статистической значимости полученных коэффициентов корреляции с использованием критерия Стьюдента ($\alpha=0.05$) позволила сократить объем исходной выборочной совокупности до 17 показателей, вошедших в конечную выборочную совокупность.

По каждому из исследуемых показателей статистические гипотезы формулировались следующим образом:

H0: различия между значениями показателя по нозологиям статистически незначимы;

H1: различия между значениями показателя по нозологиям статистически значимы.

Принцип принятия решений представлен в таблице 1.

Таблица 1

Table 1

Принятие гипотез для генеральной совокупности

Acceptance of the hypothesis for the general design

Для генеральной совокупности For the general population	Решение на основе критерия Criterion Based Decision	
	Принять H0 Accept H0	Отвергнуть H0 Reject H0
H0 верна H0 is correct	Верное решение с вероятностью $1-\alpha$ Correct solution with probability $1-\alpha$	Ошибочное решение с вероятностью α Wrong decision with probability α
H0 неверна H0 is not true	Ошибочное решение с вероятностью β Wrong decision with probability β	Верное решение с вероятностью $1-\beta$ Correct solution with probability $1-\beta$

Для выявления статистически значимых отличий показателей у пациентов с инфарктом миокарда и нестабильной стенокардией использовалась статистическая визуализация функций плотности методом Distplot библиотеки Seaborn языка программирования Python для отображения дистрибутива с различными его вариациями, который представляет собой одномерное распределение данных (рис. 2).

Как видно на рисунке 2А, графики плотностей одного показателя по различным нозологиям смещены относительно друг друга по оси абсцисс. Это означает, что значения центральных положений данных плотностей отличаются друг от друга, соответствуя различиям между группами пациентов по данному показателю.

Далее для отобранных таким образом 17 показателей проводилась статистическая проверка значимостей наблюдаемых различий. С помощью критерия Колмогорова-Смирнова проверялась гипотеза о нормальном распределении показателя в разрезе целевого признака. Для всех признаков гипотезу H_0 об их нормальности отвергли. Далее по критерию Вилконсона-Манна-Уитни проверялась гипотеза о равенстве выборочных средних признака, рассчитанных для больных острым инфарктом миокарда и нестабильной стенокардией. Для рассматриваемых признаков при уровне значимости $\alpha=0.05$ получены следующие результаты: количество лейкоцитов ($p\text{-value} = 0.02$), общий белок ($p\text{-value} = 0.43$), холестерин общий ($p\text{-value} = 0.33$), коэффициент атерогенности ($p\text{-value} = 0.28$), липо-

протеины низкой плотности ($p\text{-value} = 0.46$), липопротеины очень низкой плотности ($p\text{-value} = 0.22$), конечно-диастолический объем сердца ($p\text{-value} = 0.1$), конечный систолический объем сердца ($p\text{-value} = 0.39$), индекс массы тела ($p\text{-value} = 0.19$), глюкоза ($p\text{-value} = 0.24$), катенин ($p\text{-value} = 0.44$), склеростин ($p\text{-value} = 0.12$), WIF-1 ($p\text{-value} = 0.03$), DVL-1 ($p\text{-value} = 0.27$), GSK-3 β ($p\text{-value} = 0.29$), GSK-3 α ($p\text{-value} = 0.33$), оксид азота ($p\text{-value} = 0.01$).

Проведенная проверка позволила отобрать 3 биомаркера (число лейкоцитов, уровень WIF-1 и оксида азота периферической крови) и сформировать новое признаковое пространство для построения прогностической модели.

Дополнительно для установления связи между отобранными биомаркерами и целевым признаком (диагноз) проводился корреляционный анализ, который также показал тесную связь между ними. Так, например, парный коэффициент корреляции для количества лейкоцитов с целевой переменной составил $r=0,69$. Кроме того, были получены данные о том, что рост значений показателей (количества лейкоцитов, WIF-1) сопряжен с более высокой вероятностью развития ОИМ у пациента с ОКС.

Построение моделей дифференциальной диагностики нозологий в виде острого инфаркта миокарда и нестабильной стенокардии проводилось с использованием актуальных библиотек для машинного обучения, таких как Scikit-learn, Scipy и статистического пакета Statsmodels.

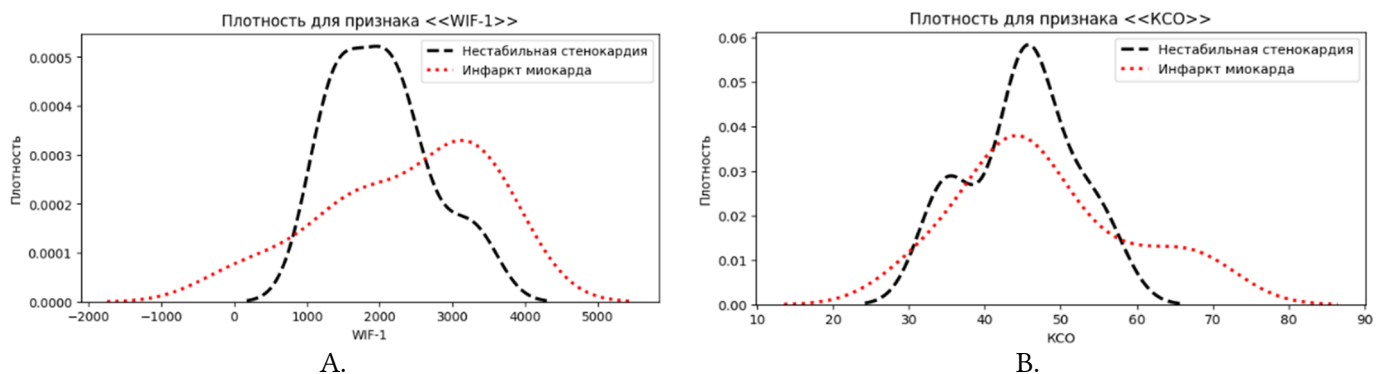


Рис. 2. Графики плотностей показателей: А – наличие смещения параметров положения; В – отсутствие смещения параметров положения.

Fig. 2. Graphs of the density of the indicators: A – the presence of a shift in the position parameters; B – the absence of a shift in the position parameters.

Примечание: КСО – конечный систолический объем сердца.

Note: the red (dotted) line on the graph corresponds to the values of patients with acute myocardial infarction, the black (dashed) line – to patients with unstable angina. ESV – cardiac end systolic volume.

Построение прогнозных моделей проводилось поэтапно:

1 этап. Выделение массивов объясняющих признаков и целевой переменной.

Листинг 1 – Разделение на независимый и зависимый наборы

```
X = df_full[var]
y = df_full['болезнь']
```

2 этап. Разделение обучающей и тестовой выборки в соотношении 80/20.

Листинг 2 – Обучающее и тестирующее множества. Стандартизация данных.

```
X_train, X_test, y_train, y_test = train_test_split(X, y,
                                                    test_size = 0.2,
                                                    random_state = 42)

scaler = StandardScaler()
X_train_st = scaler.fit_transform(X_train)
X_test_st = scaler.transform(X_test)
```

3 этап. Гиперпараметрический поиск.

Подбор наилучших гиперпараметров каждого из используемых алгоритмов на сетке с помощью компоненты Gridsearch () со значением кросс-валидации, равным 5 (cv=5), и метрикой ошибки, показывающей процент правильных прогнозов (точность, scoring='accuracy').

Листинг 3 – Подбор оптимальных параметров на сетке

```
grid_searcher = GridSearchCV(
    RandomForestClassifier(),
    param_grid={
        'n_estimators': [1, 50, 100, 150, 200],
        'max_depth': [1, 2, 3, 4, 5, 6]
    },
    cv=5,
    scoring='accuracy')
```

```
grid_searcher.fit(X_train_st, y_train);
print ('best_score --
      {}'.format(grid_searcher.best_score_))
print ('best_params --
      {}'.format(grid_searcher.best_params_))
```

4 этап. Оптимизация гиперпараметров в Python.

Листинг 4 – Обучение модели

```
rnd_clf = RandomForestClassifier(n_estimators = 50,
                                max_depth = 5, random_state=42)
rnd_clf.fit(X_train_st, y_train)
rnd_imp_var =
pd.DataFrame({'feature':X_train.columns,
              'coeff':
rnd_clf.feature_importances_}, )
rnd_imp_var['coeff'] =
rnd_imp_var['coeff'].apply(abs)
rnd_imp_var.sort_values(by = 'coeff', ascending =
False, inplace = True)
```

5 этап. Тестирование моделей машинного обучения на проверочном наборе данных.

Листинг 5 – Прогнозирование

```
pred = rnd_clf.predict(X_test_st)
```

6 этап. Сравнение результатов работы алгоритмов машинного обучения по выбранным метрикам.

Листинг 6 – оценка качества моделей

```
Print ('Метрики качества для
RandomForestClassifier', '\n',
classification_report(y_test, pred))
```

При построении классификационной модели пациентов с острым коронарным синдромом использовались следующие алгоритмы машинного обучения: логистическая регрессия (logistic regression, LR), метод опорных векторов (support vector machine , SVM), стохастический градиент (stochastic gradient descent, SG), случайный лес (Random Forest Classifier, RF) и градиентный бустинг (eXtreme Gradient Boosting, GB). В качестве критерия эффективности работы алгоритмов применялась метрика точности (accuracy).

Исходный массив данных был разделен на обучающую и тестирующую выборки. Результаты работы применяемых алгоритмов на тестирующей выборке представлены в таблице 2.

Сравнение различных алгоритмов обучения по достигнутому показателю точности показывает, что классификатор случайный лес (RF) превосходит все остальные используемые методы и имеет точность 96%.

Полученные ранги значимости показателей в итоговой прогнозной модели представлены на рис. 3 с возрастанием их значимости в следующем порядке: число лейкоцитов – WIF-1 – оксид азота.

В целях верификации модели отобранный признаковый набор расширился за счет поочередного введения в прогнозную модель дополнительных биомаркеров (последовательно 1 из 14, не вошедших в модель). Однако это привело к снижению значений установленных метрик. Пример оценки качества построенных классификационных моделей с использованием липопротеинов низкой плотности в качестве четвертого биомаркера представлен в таблице 3.

Таким образом, из общего количества показателей были отобраны 3 биомаркера (включая ассоциированные с воспалительной реакцией при остром ишемическом процессе миокарда) для идентификации НС и ОИМ с использованием классификационной модели случайного леса, которая превосходила все остальные используемые методы (в отношении точности, чувствительности и специфичности) и позволяла с высокой точностью (96%) определять

Таблица 2

Table 2

Матрицы производительности, основанные на соотношении данных, для обучения и тестирования 80/20

Performance matrices based on 80/20 training/test data ratio

Алгоритм классификации Classification algorithm	Точность Precision		Полнота Recall		f1-мера f1-score	
	0	1	0	1	0	1
SG	0.97	0.85	0.65	0.99	0.78	0.91
SVM	0.77	0.68	0.12	0.98	0.20	0.80
LR	0.91	0.96	0.92	0.95	0.92	0.96
RF	0.91	0.96	0.93	0.95	0.92	0.96
GB	0.71	0.90	0.82	0.83	0.76	0.86

Примечание. Здесь и табл. 3: LR – logistic regression, SVM – support vector machine, SG – stochastic gradient, RF-Random Forest Classifier, GB – Gradient Boosting.

Note: Here and in Tab. 3: LR - logistic regression, SVM- support vector machine, SG - stochastic gradient, RF-Random Forest Classifier, GB - Gradient Boosting.

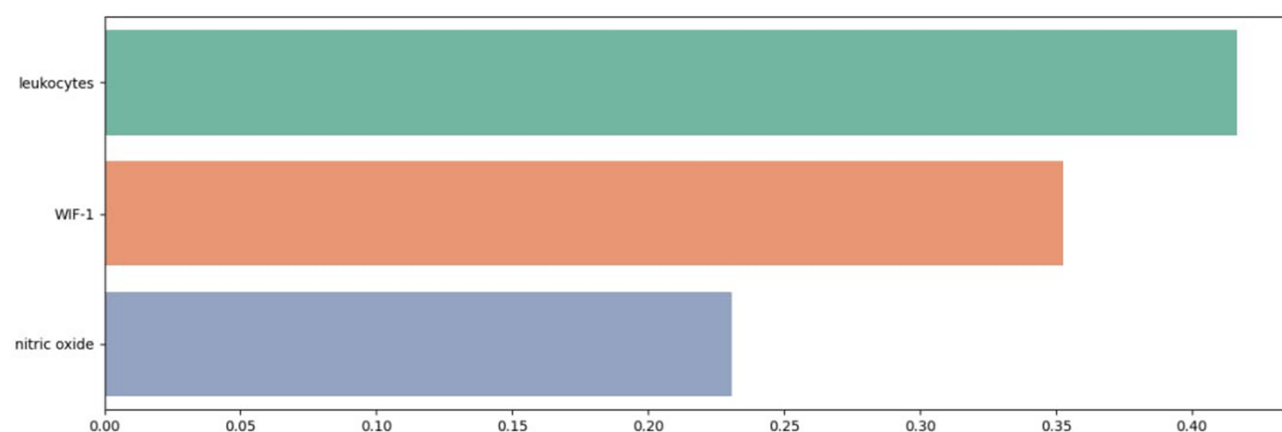


Рис. 2. Распределение значимости признаков в модели случайного леса.

Fig. 2. Distribution of feature importance in the random forest model.

наступление неблагоприятного исхода ОКС (развитие острого инфаркта миокарда).

На основании проведенной работы по построению классификационной модели мы разработали программу для ЭВМ «Калькулятор расчета вероятностного исхода острого коронарного синдрома», позволяющую использовать полученную модель в практическом здравоохранении, в медицинских и образовательных организациях, предназначенную для любых устройств с возможностью выхода в Интернет, и установленным браузером [27]. Программа базируется на модели клиент-серверной архитектуры с функциональными возможностями корректировки параметров моделей машинного обучения и оценки исхода заболевания. Полученные данные показали, что калькулятор может значительно повысить уровень диагностики в стационарах кардиологического профиля.

Приводим пример работы программы для ЭВМ «Калькулятор расчета вероятностного ис-

хода острого коронарного синдрома» в виде клинического случая.

В приемный покой регионального сосудистого центра БУЗ Орловской области «Орловская областная клиническая больница» скорой медицинской помощью доставлен мужчина 47 лет с предварительным диагнозом «острый коронарный синдром». Жалобы на момент поступления: внезапно возникшая боль в груди, сохраняющаяся в состоянии покоя и не связанная с травмой или физической нагрузкой. Боль сохранялась более 3 часов, не купировалась нестероидными противовоспалительными средствами (пациент принимал самостоятельно Ibuprofen 200 mg), через 1 час с момента начала болевого синдрома появилась иррадиация в шею и нижнюю челюсть.

Со слов пациента, 1 год назад проходил диспансеризацию по месту работы, все клинико-лабораторные показатели были в пределах нормы, за исключением повышенного артериального давления (из назначенных кардиологом

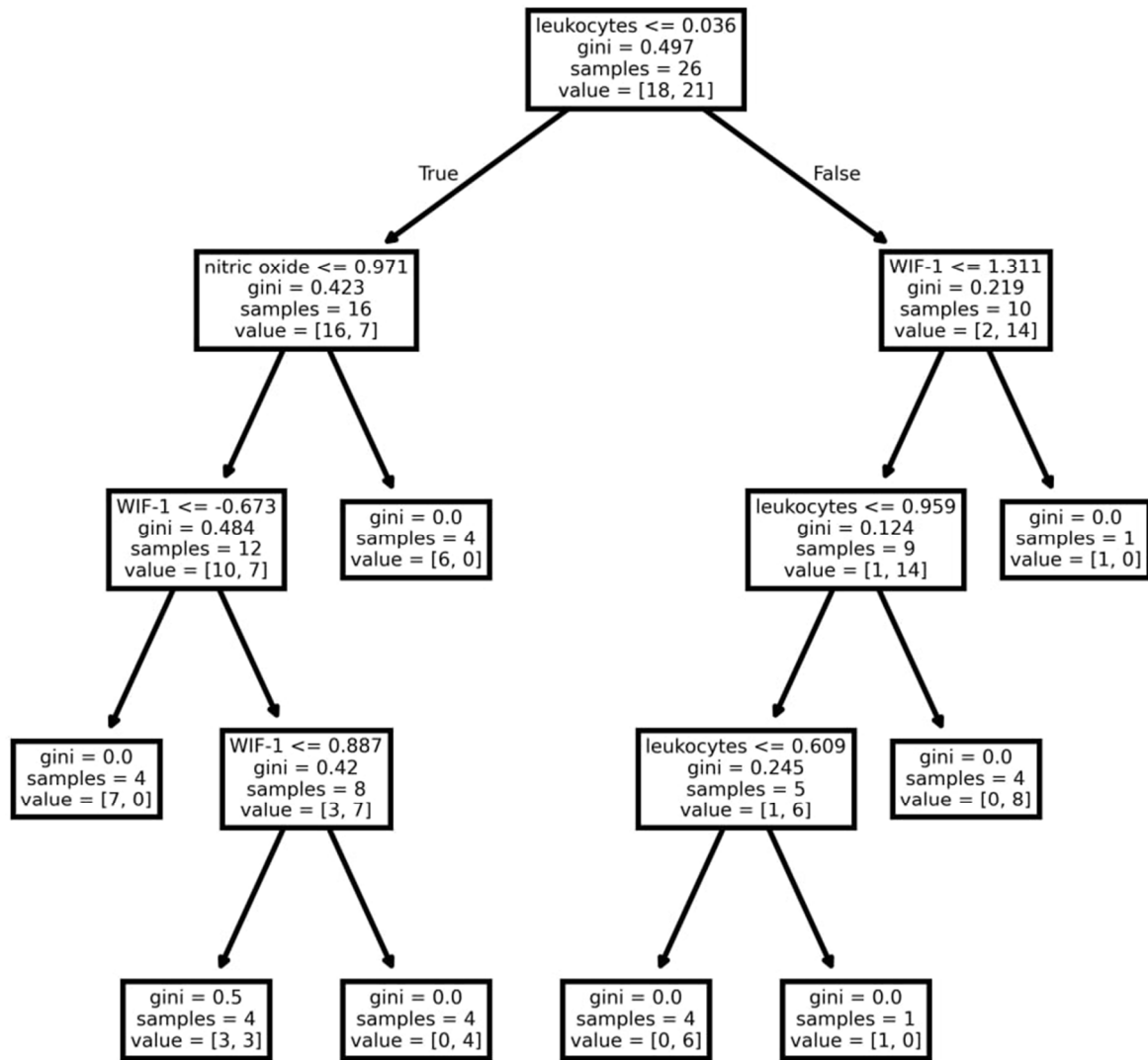


Рис. 3. Алгоритм работы классификационной модели случайного леса.

Fig. 3. Algorithm of the random forest classification model.

Таблица 3

Table 3

Оценки качества построенных классификационных моделей

Quality assessments of the constructed classification models

Алгоритм классификации Classification algorithm	Точность Precision		Полнота Recall		f1-мера f1-score	
	0	1	0	1	0	1
SG	0.92	0.85	0.65	0.86	0.78	0.91
SVM	0.77	0.69	0.59	0.59	0.25	0.81
LR	0.88	0.79	0.89	0.88	0.91	0.90
RF	0.90	0.95	0.91	0.85	0.93	0.91
GB	0.79	0.88	0.81	0.73	0.79	0.76

лекарственных препаратов принимал только Captopril 25 mg при повышении давления выше 140 мм рт. ст.)

Семейный анамнез отягощен по ССЗ: отец пациента умер в возрасте 42 лет от инфаркта миокарда, дедушка умер от «сердечного приступа» в молодости (со слов пациента).

Физикальный осмотр: общее состояние относительно удовлетворительное. Телосложение

нормостеническое (ИМТ 22,8 кг/м²). Периферических отеков нет. В легких ослабленное везикулярное дыхание, хрипов нет. Тоны сердца ясные, ритмичные, слегка приглушены. Артериальное давление – 134/72 мм рт. ст., пульс 86 уд/мин., ритмичный.

Результаты лабораторного обследования, полученные в первый час с момента поступления в приемный покой (табл. 4), характеризовались

незначительными изменениями: в общем анализе крови умеренный лейкоцитоз ($10,0 \times 10^9/\text{л}$); в липидограмме – умеренная гиперхолестеринемия (общий холестерин 6,6 ммоль/л).

На 12-канальной ЭКГ покоя при поступлении в стационар – депрессия сегмента ST $\geq 0,05$ мВ в отведениях II, III, aVF, V1, V2.

Анализ на кардиоспецифичные формы тропонина (сTn-I и сTn-T) при поступлении отрицательный.

На основании осмотра, клинико-анамнестических данных и полученных результатов лабораторно-инструментального обследования пациенту выставлен диагноз: ОКС без подъема сегмента ST.

В дополнении к стандартному обследованию пациенту были определены уровни оксида азота и WIF-1, концентрация которых составила 9,0 мкмоль/л и 1600,0 пг/мл соответственно. Полученные данные были внесены в программу для ЭВМ «Калькулятор расчета вероятностного исхода острого коронарного синдрома», соглас-

но которой наиболее вероятным диагнозом у пациента является нестабильная стенокардия (рис. 4).

Дальнейшее наблюдение за пациентом в стационаре позволило подтвердить полное совпадение диагнозов (установленного лечащим врачом и полученным в программе для ЭВМ), что демонстрирует высокий потенциал дальнейшего изучения возможностей использования методов машинного обучения с применением различных биомаркеров.

Использование методов машинного обучения в диагностике острого коронарного синдрома представляет собой перспективное и многообещающее направление, которое может значительно улучшить уровень диагностики и, как следствие, повысить качество оказываемой медицинской помощи. Интеграция прогностических моделей с учетом биомаркеров и других клинических данных позволит более точно оценивать риск и своевременно диагностировать критические ситуации.

Таблица 4

Table 4

Результаты стандартного лабораторного обследования, полученные в первый час с момента поступления в приемный покой

Results of standard laboratory examination obtained in the first hour from the moment of admission to the emergency room

Показатель Parameter	Результаты Result	Показатель Parameter	Результаты Result
Глюкоза Glucose	5.31 ммоль/л ($\mu\text{mol/L}$)	Общий белок Total protein	65.5 г/л (g/L)
Гемоглобин Hemoglobin	140.5 г/л (g/L)	ЩФ ALP	64.2 МЕ/л (IU/L)
Эритроциты RBC	4.7×10^{12} /л (L)	Креатинин Creatinine	72.1 мкмоль/л ($\mu\text{mol/L}$)
Лейкоциты WBC	$10,0 \times 10^9$ /л (L)	ТГ TG	1.6 ммоль/л (mmol/L)
СОЭ ESR	18.0 мм/ч (mm/h)	ЛПВП HDL	1.4 ммоль/л (mmol/L)
Общий холестерин Cholesterol	6.6 ммоль/л (mmol/L)	ЛПНП LDL	3.6 ммоль/л (mmol/L)
Общий билирубин Total bilirubin	17.7 мкмоль/л ($\mu\text{mol/L}$)	ЛПОНП VLDL	1.8 ммоль/л (mmol/L)
АЛТ ALT	31.8 МЕ/л (IU/L)	Натрий Sodium	118.7 ммоль/л (mmol/L)
АСТ AST	30.6 МЕ/л (IU/L)	Калий Potassium	4.7 ммоль/л (mmol/L)

Примечание. СОЭ – скорость оседания эритроцитов, АЛТ – аланинаминотрансфераза, АСТ – аспаратаминотрансфераза, ЛПВП – липопротеиды высокой плотности, ЛПНП – липопротеиды низкой плотности, ЛПОНП – липопротеиды очень низкой плотности, ТГ – триглицериды, ЩФ – щелочная фосфатаза.

Note. RBC – red blood cells, WBC – white blood cells, ESR – erythrocyte sedimentation rate, ALT – alanine aminotransferase, AST – aspartate aminotransferase, HDL – high-density lipoproteins, LDL – low-density lipoproteins, VLDL – very low-density lipoproteins, TG – triglycerides, ALP – alkaline phosphatase.

Рис. 4. Скриншот интерфейса программы «Калькулятор расчета вероятностного исхода острого коронарного синдрома».

Fig. 4. Screenshot of the interface of the program "Calculator for calculating the probable outcome of acute coronary syndrome".

КОНФЛИКТ ИНТЕРЕСОВ

Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Работа выполнена в ОГУ имени И.С. Тургенева в рамках государственного задания № 075-00196-24-08 на 2024 год и на плановый период 2025 и 2026 г. от 23.08.2024 г., проект № FSGN-2024-0014 (1024041900023-6-1.2.1;2.6.2;3.1.3;3.2.12;3.2.4).

СООТВЕТСТВИЕ ПРИНЦИПАМ ЭТИКИ

Исследование было одобрено Локальным Этическим комитетом ФГБОУ ВО «Орловский государственный университет имени И.С. Тургенева» (Протокол №31 от 27.06.2024 г.), от каждого пациента получено письменное информированное согласие.

ЛИЧНЫЙ ВКЛАД АВТОРОВ

Снимщикова И.А. – разработка концепции и дизайна исследования, редактирование, окончательное утверждение для публикации рукописи; Ревякина М.О. – разработка дизайна исследования, сбор материала, редактирование, подготовка текста; Строев С.П. – математический анализ полученных данных.

ЛИТЕРАТУРА / REFERENCES

1. *World Heart Report 2023: Confronting the World's Number One Killer*. Geneva, Switzerland. World Heart Federation. 2023
2. *Российский статистический ежегодник. 2023*. Москва: Росстат, 2023. 701 с. [Russian Statistical Yearbook. 2023. Moscow: Rosstat, 2023. 701 p. (in Russ.)]
3. *Здравоохранение в России. 2023*. Москва: Росстат, 2023. 179 с. [Healthcare in Russia. 2023. Moscow: Rosstat, 2023. 179 p. (in Russ.)]
4. Montone R.A., Ford T.J., Galli M., Rinaldi R., Bland A., Morrow A., Angiolillo D.J., Berry C. et al. Stratified medicine for acute and chronic coronary syndromes: A patient-tailored approach. *Prog Cardiovasc Dis*. 2024;85:2–13. DOI: 10.1016/j.pcad.2024.06.003.
5. Jensen R.V., Hjortbak M.V., Bøtker H.E. Ischemic Heart Disease: An Update. *Semin Nucl Med*. 2020;50(3):195–207. DOI: 10.1053/j.semnuclmed.2020.02.007.
6. Shuvy M., Beerl G., Klein E., Cohen T., Shlomo N., Minha S., Pereg D. Accuracy of the Global Registry of Acute Coronary Events (GRACE) Risk Score in Contemporary Treatment of Patients With Acute Coronary Syndrome. *Can J Cardiol*. 2018;34(12): 1613–1617. DOI: 10.1016/j.cjca.2018.09.015.
7. Zhang Q., Gao J., Yin X., Zhang S., Wang Y., Ji H., Zhang X., Song D. et al. Risk Prediction Models for Ischemic Cardiovascular Outcomes in Patients with Acute Coronary Syndrome. *Rev Cardiovasc Med*. 2023;24(4):106. DOI: 10.31083/j.rcm2404106.
8. Wong Y.K., Cheung C.Y.Y., Tang C.S., Au K.W., Hai J.S.H., Lee C.H., Lau K.K. et al. Age-Biomarkers-Clinical Risk Factors for Prediction of Cardiovascular Events in Patients With Coronary Artery Disease. *Arterioscler Thromb Vasc Biol*. 2018;38(10):2519–2527. DOI: 10.1161/ATVBAHA.118.311726.
9. Expert Group on Biomarkers. Biomarkers in Cardiology – Part 2: In Coronary Heart Disease,

- Valve Disease and Special Situations. *Arq Bras Cardiol.* 2015;104(5):337–346. DOI: 10.5935/abc.20150061.
10. Expert Group on Biomarkers. Biomarkers in cardiology – part 1 – in heart failure and specific cardiomyopathies. *Arq Bras Cardiol.* 2014;103(6):451–459. DOI: 10.5935/abc.20140184.
11. Severino P., D'Amato A., Pucci M., Infusino F., Adamo F., Birtolo L.I., Netti L., Montefusco G. et al. Ischemic Heart Disease Pathophysiology Paradigms Overview: From Plaque Activation to Microvascular Dysfunction. *Int J Mol Sci.* 2020;21(21):8118. DOI: 10.3390/ijms21218118.
12. Saeedbakhsh S., Sattari M., Mohammadi M., Najafian J., Mohammadi F. Diagnosis of Coronary Artery Disease based on Machine Learning algorithms Support Vector Machine, Artificial Neural Network, and Random Forest. *Adv Biomed Res.* 2023;12:51. DOI: 10.4103/abr.abr_383_21.
13. Song L., Li Y., Nie S., Feng Z., Liu Y., Ding F., Gong L., Liu L., et al. Using machine learning to predict adverse events in acute coronary syndrome: A retrospective study. *Clin Cardiol.* 2023;46(12):1594–1602. DOI: 10.1002/clc.24127.
14. D'Ascenzo F., De Filippo O., Gallone G., Mittone G., Deriu M.A., Iannaccone M., Ariza-Solé A., Liebetrau C. et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. *Lancet.* 2021;397(10270):199–207. DOI: 10.1016/S0140-6736(20)32519-8.
15. Collazo R.A., Pessôa L.A.M., Bahiense L., Pereira B. de B., Reis A.F. dos, Silva N.S. et al. A Comparative study between artificial neural network and support vector machine for acute coronary syndrome prognosis. *Pesqui Oper.* 2016;36(2):321–343. DOI: 10.1590/0101-7438.2016.036.02.0321
16. Zhang X., Wang X., Xu L., Liu J., Ren P., Wu H. The predictive value of machine learning for mortality risk in patients with acute coronary syndromes: a systematic review and meta-analysis. *Eur J Med Res.* 2023;28(1):451. DOI: 10.1186/s40001-023-01027-4.
17. Na L., Lin J., Kuiwu Y. Risk prediction model for major adverse cardiovascular events (MACE) during hospitalization in patients with coronary heart disease based on myocardial energy metabolic substrate. *Front Cardiovasc Med.* 2023;10:1137778. DOI: 10.3389/fcvm.2023.1137778.
18. Tao K., Li J., Li J., Shan W., Yan H., Lu Y. Estimation of Heart Rate Using Regression Models and Artificial Neural Network in Middle-Aged Adults. *Front Physiol.* 2021;12:742754. DOI: 10.3389/fphys.2021.742754.
19. Li R., Shen L., Ma W., Yan B., Chen W., Zhu J., Li L., Yuan J. et al. Use of machine learning models to predict in-hospital mortality in patients with acute coronary syndrome. *Clin Cardiol.* 2023;46(2):184–194. DOI: 10.1002/clc.23957.
20. Snimshchikova I.A., Plotnikova M.O. Role of morphogenic proteins of the WNT signaling pathway in coronary artery disease. *Medical Immunology (Russia).* 2023;25(4):985–990. DOI: 10.15789/1563-0625-ROM-2835.
21. Плотникова М.О., Снимщикова И.А., Шкодкина С.А. Диагностическая значимость показателей окислительного стресса у пациентов с ишемической болезнью сердца. *Медико-фармацевтический журнал «Пulse».* 2022;24(5):63–67 [Plotnikova M.O., Snimshchikova I. A., Shkodkina S.A. Diagnostic significance of oxidative stress in patients with coronary heart disease. *Medical & pharmaceutical journal «Pulse».* 2022;24(5):63–67 (in Russ.)]. DOI: 10.26787/nydha-2686-6838-2022-24-5-63-67. EDN: LZZWAZ.
22. Nohria R., Viera A.J. Acute Coronary Syndrome: Diagnosis and Initial Management. *Am Fam Physician.* 2024;109(1):34–42.
23. Byrne R.A., Rossello X., Coughlan J.J., Barbato E., Berry C., Chieffo A., Claeys M.J., Dan G.A. et al. 2023 ESC Guidelines for the management of acute coronary syndromes. *Eur Heart J.* 2023;44(38):3720–3826. DOI: 10.1093/eurheartj/ehad191.
24. Garg P., Morris P., Fazlanie A.L., Vijayan S., Dancso B., Dastidar A.G., Plein S., Mueller C. et al. Cardiac biomarkers of acute coronary syndrome: from history to high-sensitivity cardiac troponin. *Intern Emerg Med.* 2017;12(2):147–155. DOI: 10.1007/s11739-017-1612-1.
25. Celik S., Giannitsis E., Wollert K.C., Schwöbel K., Lossnitzer D., Hilbel T., Lehrke S. et al. Cardiac troponin T concentrations above the 99th percentile value as measured by a new high-sensitivity assay predict long-term prognosis in patients with acute coronary syndromes undergoing routine early invasive strategy. *Clin Res Cardiol.* 2011;100(12):1077–1085. DOI: 10.1007/s00392-011-0344-x.
26. Wu A.H. Early detection of acute coronary syndromes and risk stratification by multimarker analysis. *Biomark Med.* 2007;1(1):45–57. DOI: 10.2217/17520363.1.145.
27. Программа «Калькулятор расчета вероятностного исхода острого коронарного синдрома». Плотникова М.О., Строев С.П., Снимщикова И.А., Вишнеvский В.И., Дорофеева В.И., Киселёв Д.П., Честнихина А.Д. № 2022668140 Свидетельство о государственной регистрации для ЭВМ РФ. 4 октября 2022 г. [The program "Calculator for calculating the probabilistic outcome of acute coronary syndrome". Plotnikova M.O., Stroeve S.P., Snimshchikova I.A., Vishnevsky V.I., Dorofeeva V.I., Kiselyov D.P., Chestnykhina A.D. No. 2022668140 Certificate of state registration of a computer program of the Russian Federation. Oct 4, 2022 (in Russ.)]

Поступила в редакцию 18.10.2024
Подписана в печать 25.11.2024

Для цитирования: Снимщикова И.А., Ревякина М.О., Строев С.П. Современные подходы к прогнозированию риска развития неблагоприятных сердечно-сосудистых событий у пациентов с острым коронарным синдромом. *Человек и его здоровье.* 2024;27(3):39–50. DOI: 10.21626/vestnik/2024-3/04. EDN: LFSICQ.

MODERN APPROACHES TO PREDICTING THE RISK OF ADVERSE CARDIOVASCULAR EVENTS IN PATIENTS WITH ACUTE CORONARY SYNDROME

© Snimshchikova I.A., Revyakina M.O., Stroyev S.P.

Orel State University named after I.S. Turgenev (I.S. Turhenev OSU)

95, Komsomolskaya Str., Orel, Orel region, 302026, Russian Federation

Cardiovascular diseases are the leading cause of morbidity and mortality worldwide. Despite the progress achieved in predicting the risk of development and adverse cardiovascular outcomes, there remains a large number of patients who are not recognized using standard predictive models.

Objective: to assess the possibility of using machine learning to build a multi-marker model for predicting the outcome of acute coronary syndrome in the form of acute myocardial infarction or unstable angina.

Materials and methods. The development of predictive models was carried out in the Python programming language (Pandas, Seaborn and Scikit-learn libraries) using a database compiled from the results of clinical and laboratory studies of 60 patients aged 18 to 59 years with acute coronary syndrome.

Results. From a total list of 181 indicators (10,860 values), 3 most informative biomarkers were selected to identify the established endpoints (patient diagnosis). The random forest classifier outperformed other models in terms of accuracy, sensitivity and specificity, which allowed us to determine the occurrence of an unfavorable outcome of ACS with 96% accuracy.

Conclusion. Based on the obtained data, a computer program "Calculator for calculating the probabilistic outcome of acute coronary syndrome" was developed for practical use in healthcare institutions.

Keywords: acute myocardial infarction; unstable angina; machine learning; prognostic model; acute coronary syndrome.

Snimshchikova Irina A. – Dr. Sci. (Med.), Professor, Director of the Medical Institute, Head of the Department of Immunology and Specialized Clinical Disciplines, Leading Researcher at the Laboratory of New Medical Technologies, I.S. Turhenev OSU, Orel, Russian Federation. ORCID iD: 0000-0002-4258-963X. E-mail: snimshikova@mail.ru

Revyakina Maria O. – Cand. Sci. (Med.), Associate Professor at the Department of Immunology and Specialized Clinical Disciplines, Senior Researcher at the Laboratory of New Medical Technologies, I.S. Turhenev OSU, Orel, Russian Federation. ORCID iD: 0000-0003-1593-5290. E-mail: revyakina_masha@mail.ru (corresponding author).

Stroev Sergey P. – Cand. Sci. (Econ.), Associate Professor, Head of the Department of Algebra and Mathematical Methods in Economics, I.S. Turhenev OSU, Orel, Russian Federation. ORCID iD: 0000-0002-2271-5264. E-mail: stroewsp@mail.ru

CONFLICT OF INTEREST

The author declares the absence of obvious and potential conflicts of interest related to the publication of this article.

SOURCE OF FINANCING

The work was performed at the I.S. Turgenev OSU within the framework of state assignment No. 075-00196-24-08 for 2024 and for the planning period 2025 and 2026 dated 08/23/2024, project No. FSGN-2024-0014 (1024041900023-6-1.2.1;2.6.2;3.1.3;3.2.12;3.2.4).

COMPLIANCE WITH PRINCIPLES OF ETHICS

The study was approved by the Local Ethics Committee of the I.S. Turgenev Oryol State University (Protocol No. 31 dated 06/27/2024), written informed consent was received from each patient.

AUTHORS CONTRIBUTION

Snimshchikova I.A. – development of the concept and design of the study, revision, final approval for the publication of the manuscript; Revyakina M.O. – development of the design of the study, collection of material, editing, preparation of the text; Stroyev S.P. – mathematical analysis of the data obtained.

Received 18.10.2024

Accepted 25.11.2024

For citation: Snimshchikova I.A., Revyakina M.O., Stroyev S.P. Modern approaches to predicting the risk of adverse cardiovascular events in patients with acute coronary syndrome. *Humans and their health*. 2024;27(3):39–50. DOI: 10.21626/vestnik/2024-3/04. EDN: LFSICQ.