

Методы машинного обучения как инструмент диагностических и прогностических исследований при ишемической болезни сердца

Гельцер Б. И., Циванюк М. М., Шахгельдян К. И., Рублев В. Ю.

Методы машинного обучения (МО) являются основным инструментом искусственного интеллекта, использование которых позволяет автоматизировать обработку и анализ больших данных, выявлять на этой основе скрытые или неочевидные закономерности и извлекать новые знания. В обзоре представлен анализ научной литературы по использованию методов МО для диагностики и прогнозирования клинического течения ишемической болезни сердца. Приведены сведения по эталонным базам данных, использование которых позволяет разрабатывать модели и валидировать их (European ST-T Database, Cleveland Heart Disease database, Multi-Ethnic Study of Atherosclerosis и др.). Показаны преимущества и недостатки отдельных методов МО (логистической регрессии, машин опорных векторов, деревьев решений, наивного байесовского классификатора, k-ближайших соседей) для разработки диагностических и прогностических алгоритмов. К наиболее перспективным методам МО относят глубокое обучение, которое реализуется с помощью многослойных искусственных нейронных сетей. Предполагается, что совершенствование моделей на основе методов МО и их внедрение в клиническую практику будет способствовать поддержке принятия врачебных решений, повышению эффективности лечения и оптимизации расходов здравоохранения.

Ключевые слова: методы машинного обучения, ишемическая болезнь сердца, модели диагностики и прогнозирования.

Отношения и деятельность. Работа выполнена при поддержке грантов РФФИ в рамках научных проектов № 18-29-03131, № 19-29-01077.

ФГАОУ ВО Дальневосточный федеральный университет, Владивосток, Россия.

Гельцер Б. И. — д.м.н., профессор, член-корр. РАН, директор департамента клинической медицины школы биомедицины, ORCID: 0000-0002-9250-557X, Циванюк М. М.* — аспирант школы биомедицины, врач отделения рентгено-

хирургических методов диагностики и лечения, кардиолог, ORCID: 0000-0003-3545-3862, Шахгельдян К. И. — д.техн.н., доцент, зав. лабораторией анализа больших данных в биомедицине и здравоохранении школы биомедицины, ORCID: 0000-0002-4539-685X, Рублев В. Ю. — аспирант школы биомедицины, сердечно-сосудистый хирург, ORCID: 0000-0001-7620-4454.

*Автор, ответственный за переписку (Corresponding author):
m_tsivanyuk@list.ru

АД — артериальное давление, ГО — глубокое обучение, ДР — деревья решений, ИБС — ишемическая болезнь сердца, ИИ — искусственный интеллект, ИМ — инфаркт миокарда, ИНС — искусственные нейронные сети, КА — коронарные артерии, ЛЖ — левый желудочек, ЛР — логистическая регрессия, МО — машинное обучение, НБК — наивный байесовский классификатор, СЛ — случайный лес, ССЗ — сердечно-сосудистые заболевания, ССР — сердечно-сосудистый риск, ФВ — фракция выброса, ФР — фактор риска, ХСН — хроническая сердечная недостаточность, ЧСС — частота сердечных сокращений, ЭКГ — электрокардиограмма, ЭхоКГ — эхокардиография, CNN — сверточные нейронные сети, FRS — Framingham Risk Score, kNN — метод k-ближайших соседей, RNN — рекуррентные нейронные сети, SVM — машина опорных векторов.

Рукопись получена 07.07.2020

Рецензия получена 10.08.2020

Принята к публикации 28.08.2020



Для цитирования: Гельцер Б. И., Циванюк М. М., Шахгельдян К. И., Рублев В. Ю. Методы машинного обучения как инструмент диагностических и прогностических исследований при ишемической болезни сердца. *Российский кардиологический журнал*. 2020;25(12):3999. doi:10.15829/1560-4071-2020-3999

Machine learning as a tool for diagnostic and prognostic research in coronary artery disease

Geltser B. I., Tsivanyuk M. M., Shakhgelyan K. I., Rublev V. Yu.

Machine learning (ML) are the central tool of artificial intelligence, the use of which makes it possible to automate the processing and analysis of large data, reveal hidden or non-obvious patterns and learn a new knowledge. The review presents an analysis of literature on the use of ML for diagnosing and predicting the clinical course of coronary artery disease. We provided information on reference databases, the use of which allows to develop models and validate them (European ST-T Database, Cleveland Heart Disease database, Multi-Ethnic Study of Atherosclerosis, etc.). The advantages and disadvantages of individual ML methods (logistic regression, support vector machines, decision trees, naive Bayesian classifier, k-nearest neighbors) for the development of diagnostic and predictive algorithms are shown. The most promising ML methods include deep learning, which is implemented using multilayer artificial neural networks. It is assumed that the improvement of ML-based models and their introduction into clinical practice will help support medical decision-making, increase the effectiveness of treatment and optimize health care costs.

Key words: machine learning, coronary artery disease, diagnostic and predictive models.

Relationships and Activities. This work was supported by grants from the Russian Foundation for Basic Research within the projects № 18-29-03131 and № 19-29-01077.

Far Eastern Federal University, Vladivostok, Russia.

Geltser B. I. ORCID: 0000-0002-9250-557X, Tsivanyuk M. M.* ORCID: 0000-0003-3545-3862, Shakhgelyan K. I. ORCID: 0000-0002-4539-685X, Rublev V. Yu. ORCID: 0000-0001-7620-4454.

*Corresponding author: m_tsivanyuk@list.ru

Received: 07.07.2020 **Revision Received:** 10.08.2020 **Accepted:** 28.08.2020

For citation: Geltser B. I., Tsivanyuk M. M., Shakhgelyan K. I., Rublev V. Yu. Machine learning as a tool for diagnostic and prognostic research in coronary artery disease. *Russian Journal of Cardiology*. 2020;25(12):3999. (In Russ.) doi:10.15829/1560-4071-2020-3999

Сердечно-сосудистые заболевания (ССЗ) являются одной из ведущих причин инвалидизации и смертности населения в большинстве стран мира [1]. По данным 2017г доля ССЗ в структуре смертности от всех причин в Российской Федерации составила 47,2%, из которых больше половины (54%) случаев приходилось на ишемическую болезнь сердца (ИБС) [2]. В последние годы за счет повышения доступности и качества диагностики ИБС удалось добиться значительных успехов в лечении данного заболевания. Вместе с тем совершенствование диагностических технологий и постоянно возрастающий объем научной информации даёт возможность поддержки принятия врачебных решений. Внедрение в клиническую практику технологий искусственного интеллекта (ИИ) позволяет автоматизировать обработку и анализ больших данных, выявлять на этой основе скрытые или неочевидные закономерности и извлекать новые знания [3, 4]. Искусственный интеллект — это область компьютерных наук, которая стремится имитировать мыслительные процессы человека, способность к обучению и хранению знаний, а ее методы применяются в кардиологии для выявления предикторов развития ССЗ, прогнозирования их клинического течения, стратификации риска, повышения эффективности лечения и оптимизации расходов здравоохранения [5]. Методы машинного обучения (МО) являются основным инструментом ИИ и представляют собой обучаемые автоматизированные системы, которые, используя различные виды моделирования и опираясь на доступные данные, принимают необходимые решения для реализации предсказательной функции [6]. Алгоритмы МО используются в смежных областях: когнитивных вычислениях, компьютерном зрении и робототехнике [5]. В отличие от традиционных статистических методов, которые обеспечивают исследования взаимосвязей между ограниченным числом переменных, методы МО предоставляют возможность для обработки и анализа больших сложных разнородных данных. Кроме того, алгоритмы МО основаны на меньшем количестве допущений и имеют более высокую прогностическую точность [3, 7]. В настоящее время методы МО используются в различных шкалах оценки рисков фатальных и не фатальных сердечно-сосудистых событий с горизонтом в 10 лет. К ним относятся Framingham Risk Score (FRS), REYNOLDS, шкала Американской коллегии кардиологов/Американской кардиологической ассоциации (ACC/AHA), а также шкалы SCORE, QRISK, ASSING, PROCAM и их откалиброванные варианты для использования на других популяциях. Эти шкалы оценивают сердечно-сосудистый риск (ССР) на основе методов традиционной статистики и логистической регрессии (ЛР), которые предполагают наличие только линейных связей между предикторами и ис-

ходами, но не учитывают нелинейные взаимосвязи, что ограничивает их прогностический потенциал [8].

Автоматизированные системы и публичные наборы данных

Надежность моделей МО зависит от качества баз данных и алгоритмов отбора предикторов. Для идентификации и классификации ИБС, а также стратификации риска ССЗ сформированы стандартизированные базы данных, находящиеся в свободном электронном доступе (например, электронный ресурс PhysioNet [9] и репозиторий баз данных UCI [10]), которые позволяют исследователям выполнять проверку разработанных моделей. Исследовательский ресурс для сложных физиологических сигналов PhysioNet состоит из трёх ключевых взаимосвязанных компонентов: PhysioBank — архив данных; PhysioToolkit — библиотека программного обеспечения для интеллектуальной обработки и анализа многопараметрических данных; PhysioNet — ресурс распространения и коммуникации. Последний ресурс обеспечивает двустороннюю динамическую связь между архивом и исследовательским сообществом для более производительной работы автоматизированных систем [9]. Репозиторий машинного обучения UCI (University of California, Irvine), созданный в 1987г, представляет собой набор различных баз данных, в т.ч. медицинских, которые широко используются исследователями в качестве основного источника информации для разработки и апробации алгоритмов МО [10].

Для проектирования моделей диагностики ИБС и оценки их прогностической точности используют различные эталонные базы данных. К ним относят Европейский набор данных (European ST-T Database), содержащий записи электрокардиограмм (ЭКГ) продолжительностью 60 мин у 90 пациентов амбулаторного звена с подозрением на ИБС, используемый для анализа преходящих изменений сегмента ST и зубца T [11]. База данных клинической практики США и Европы (The Long-Term ST Database) содержит 86 амбулаторных записей ЭКГ продолжительностью от 21 до 24 ч для демонстрации примеров преходящих изменений сегмента ST, нарушений ритма и проводимости сердца, которые можно использовать для разработки и тестирования алгоритмов, способных автоматически дифференцировать патологические, в т.ч. ишемические, изменения на ЭКГ [12]. Набор данных Санкт-Петербургского института кардиологической техники (St Petersburg INCART 12-lead Arrhythmia Database) включает 75 записей ЭКГ продолжительностью 30 мин, полученных по результатам суточного мониторирования ЭКГ у пациентов с доказанной ИБС, острым инфарктом миокарда (ИМ), нарушениями ритма и проводимости сердца. При создании этой базы учитывались результаты инвазивной коронарографии и уровень кардио-

специфических ферментов [9, 13]. Набор данных Fantasia Database содержит непрерывные 120-минутные записи ЭКГ покоя 40 пациентов с указанием пола, возраста, частоты сердечных сокращений (ЧСС) и дыхания, а у половины исследуемых — результаты неинвазивного измерения артериального давления (АД) [14]. Набор Z-Alizadeh Sani содержит информацию о 303 пациентах, включающую социально-демографические и антропометрические характеристики, результаты лабораторных исследований, ЭКГ, эхокардиографии (ЭхоКГ). Усовершенствованная база данных Extension Z-Alizadeh Sani дополнена результатами инвазивной коронарографии. По мнению авторов, этот набор данных может использоваться не только для идентификации поражений коронарного русла, но и для обнаружения локализации стеноза передней нисходящей, огибающей и правой коронарных артерий (КА) [10, 15]. База данных Кливленда (Cleveland Heart Disease database), опубликованная в 1988г, содержит 76 характеристик 303 пациентов, из которых у 46% была диагностирована ИБС, а у 54% сердечная патология отсутствовала. На веб-странице набора данных указано, что исследователи для разработки прогностических моделей обычно используют 14 параметров из представленных 76, которые включают возраст больных, их гендерную принадлежность, характеристику боли в грудной клетке, уровень АД, ЧСС, концентрацию холестерина и глюкозы в крови натощак, данные ЭКГ, в т.ч. наличие депрессии сегмента ST и ее выраженность на фоне теста с нагрузкой [10]. Набор данных MESA (Multi-Ethnic Study of Atherosclerosis) состоит из 735 показателей 6814 пациентов из медицинских центров США в возрасте 45-84 лет с субклиническими проявлениями атеросклероза. Они включают клинико-демографические и лабораторные показатели, параметры ЭКГ, ЭхоКГ, уровень кальциноза КА по результатам компьютерной томографии и др. [16]. Многопараметрический интеллектуальный мониторинг в отделении интенсивной терапии (MIMIC II и MIMIC III) позволил сформировать базу данных, содержащую информацию двух типов: набор признаков в форме временных рядов (ЭКГ, фотоплетизмография, АД и др.) и клинические характеристики пациентов [17, 18]. В последние годы разработаны продукты облачной инфраструктуры с различными наборами автоматизированных систем МО (Amazon Web Services, Microsoft Azure ML, Google Cloud Auto ML и BigML), предоставляющие возможность использования вычислительных ресурсов, а также услуг по хранению текстовой и графической информации [7]. Эти решения позволяют взаимодействовать с коллегами в реальном времени, что оптимизирует процесс разработки и внедрения алгоритмов в клиническую практику. Необходимо также отметить, что на точность прогностических моделей

помимо размера выборки и качества отбора предикторов влияет локализация источника информации. Показано, в частности, что наилучшую предсказательную ценность разработанные модели демонстрируют в тех популяциях, откуда были получены исходные данные [15, 19]. Именно поэтому наличие и постоянная актуализация региональных регистров больных ИБС является важной задачей здравоохранения.

Базовые или “классические” методы МО в диагностике и прогнозировании ИБС

Методы МО, применяемые в кардиологии с прогностической и диагностической целью, подразделяют на 2 основных класса: обучение без учителя и с учителем. Первый класс включает различные методы кластеризации (K-средних, дискриминантный анализ, сети Кохонена и др.), которые используют для фенотипирования больных с учетом индивидуальных особенностей их генетического, иммунологического или клинико-функционального статуса. МО с учителем используют заранее размеченные наборы данных, обработка которых позволяет выделить факторы, оказывающие влияние на клиническое течение ИБС и ее прогноз [15]. Для реализации этих задач помимо “классической” ЛР используют такие методы МО, как деревья решений (ДР), случайный лес (СЛ), наивный байесовский классификатор (НБК), машины опорных векторов (SVM), k-ближайших соседей (kNN), каждый из которых имеет свои преимущества и недостатки [20]. В настоящее время к наиболее перспективным методам МО относят глубокое обучение (ГО), которое реализуется с помощью многослойных искусственных нейронных сетей (ИНС).

ЛР — это линейный классификатор, относящийся к базовым технологиям МО, заимствованный из традиционной статистики и позволяющий прогнозировать вероятность развития некоторых событий на основе независимых переменных. ЛР наилучшим образом подходит для решения задач, где вероятность развития заболеваний линейно связана с факторами риска (ФР), а они, в свою очередь, должны быть линейно независимы между собой. Необходимость выполнения этих условий существенно ограничивает область применения ЛР в прогностических исследованиях [21]. ЛР часто используется в научных исследованиях для диагностики и прогнозирования развития ССЗ. Так, в работах [22, 23] была представлена прогностическая модель на основе ЛР, которая была способна разделять больных ИБС с обструктивным поражением коронарного русла от пациентов с интактными КА. В работе Verma L, et al. (2016) [24] на клинических данных 335 пациентов была разработана модель ЛР, позволяющая идентифицировать больных ИБС с точностью 88,4%. В японском многоцентровом исследовании с участием 991 пациента

с подозрением на ИБС была создана модель на основе ЛР, в которой традиционные ФР шкалы FRS были дополнены показателем кальциноза коронарного русла. Данная модель показала повышение точности в верификации обструктивного поражения КА по индикаторам С-статистики: у женщин до 0,79 vs 0,66 для FRS; у мужчин до 0,83 vs 0,61 для FRS [25]. В другой работе с помощью ЛР была разработана диагностическая шкала для определения вероятности обнаружения интактных КА при ИМ [26]. Метод ЛР применяли в модели прогнозирования нарушений проводимости сердца, где в качестве предикторов были использованы генотипы транскрипционного фактора TBX5. Было показано, что сочетание генотипов AA и AG является фактором, противостоящим развитию блокады левой ножки пучка Гиса, а наличие генотипа AA у женщин снижает риск развития атриовентрикулярной блокады [27]. Модель ЛР использовалась также для оценки влияния отдельных ФР на развитие ИБС [28]. В работе российских исследователей вероятность развития фатальных и нефатальных сердечно-сосудистых событий у пациентов с хронической ИБС оценивали с помощью многофакторной ЛР, где в качестве предиктора использовали показатель коронарного коллатерального кровотока [29]. При помощи данной модели авторами была установлена взаимосвязь данного фактора с частотой возобновления приступов стенокардии и усугубления ее функционального класса в течение 5 лет наблюдения. Gao Y, et al. (2018) [30] использовали многофакторную ЛР для прогнозирования развития ИБС у пациентов с сахарным диабетом 2 типа. В этом исследовании независимым предиктором ИБС было соотношение уровня моноцитов к концентрации холестерина липопротеидов высокой плотности. При увеличении данного соотношения $>8,2$ чувствительность и специфичность разработанной модели составляла 83,7% и 62%, соответственно, а площадь под ROC-кривой — 0,795.

Метод kNN является одним из наиболее доступных непараметрических методов МО, позволяющий классифицировать данные по степени их близости к заранее размеченной выборке. Полученная таким образом классификация признаков легко интерпретируется [5]. Вместе с тем данный метод не позволяет выделять предикторы и оценить взаимосвязи между анализируемыми факторами и исходами. Кроме того, классификатор на основе kNN может использоваться только в комбинации с обучающей выборкой. В работе Sridhar S, et al. (2016) [31] для автоматизированной идентификации ИБС по данным ЭКГ применяли классификатор kNN, который продемонстрировал максимальную диагностическую точность в 98,67%, а чувствительность и специфичность — в 95% и 99,2%, соответственно. Shi M, et al. (2019) [32] идентификацию ИБС проводили по II стан-

дартному отведению ЭКГ, используя публичные наборы данных Fantasia и St Petersburg INCART 12-lead Arrhythmia Database. Точность, чувствительность и специфичность предложенной модели составила 97,5%, 100% и 95%, соответственно. Tabassian M, et al. (2017) [33] для диагностики ИМ разработали модель на основе kNN с использованием пространственно-временных характеристик кривых деформации миокарда левого желудочка (ЛЖ), полученных по результатам ЭхоКГ. Предложенная модель по точности превосходила заключения экспертов по ультразвуковой диагностике (87% vs 70%).

НБК относится к методам МО, основанным на принципах нечеткой логики, где вместо традиционных значений Истина/Ложь используется более широкий диапазон значений. Основным преимуществом моделей на основе НБК является простота их обучения, а основной недостаток заключается в относительно низком качестве классификации объектов в признаковом пространстве. Прогностическая модель на основе НБК, разработанная Juarez-Orozco LE, et al. (2019) [34], была представлена в клинических рекомендациях Европейского общества кардиологов по диагностике и лечению хронического коронарного синдрома. Данный метод демонстрировал возможность оценки предтестовой вероятности обструктивного поражения КА у лиц с подозрением на ИБС [34, 35]. В исследовании [36] были разработаны модели МО способные предсказывать по результатам внутрисосудистого ультразвукового исследования наличие нестабильных атеросклеротических бляшек в КА. Так, модель на основе НБК для верификации тонкокапсульных фиброатером в когорте из 517 больных ИБС продемонстрировала точность в 80%. В работе Quesada JA, et al. (2019) [8] на испанской когорте из 38527 обследованных приведены результаты оценки прогностической способности моделей на основе 15 методов МО в сравнении со шкалами риска SCORE и REGICOR. В этом исследовании были продемонстрированы преимущества применения для прогнозирования сердечно-сосудистых событий 10 методов МО, к которым относился и НБК. При этом прогностический потенциал модели НБК был существенно выше, чем шкалы SCORE (площадь под ROC-кривой 0,7 vs 0,63).

SVM — один из наиболее популярных методов МО с учителем, который применяется для решения задач классификации и регрессии. Основная идея метода заключается в построении гиперплоскости, разделяющей объекты выборки оптимальным способом. Алгоритм основан на предположении, что чем больше расстояние между гиперплоскостью и объектами разделяемых классов, тем меньше уровень средней ошибки классификатора. SVM может применяться для верификации нелинейных взаимосвязей изучаемых факторов и исходов [21]. В об-

зоре 2020г было показано, что гибридные модели с использованием SVM обладают более высокой точностью в разделении пациентов с истинной ИБС от лиц с кардиалгией [4]. В другой работе точность SVM-модели для диагностики острого коронарного синдрома среди пациентов, госпитализированных в отделение неотложной кардиологии с болью в грудной клетке, составила >99% [37]. В этом исследовании помимо клинико-демографических данных для моделирования были использованы результаты ЭКГ, ЭхоКГ и уровень кардиоспецифических ферментов. Cui S, et al. (2017) [38] на когорте из 400 пациентов показали эффективность модели SVM в прогнозировании рестеноза стентированного сегмента КА с точностью до 90%, используя в качестве предикторов 6 метаболитов сфинголипидов и фосфолипидов плазмы крови.

При построении моделей на основе ДР выполняется расщепление каждого из анализируемых факторов, что позволяет определить границы их референсных значений. Данная процедура позволяет выделить патологические признаки и интерпретировать их [39]. Методологические преимущества ДР состоят в том, что данный метод позволяет структурировать проблему и систематизировать ее компоненты, а итоговое решение принимается на основе логических выводов. Acharya UR, et al. (2017) [40] анализировали записи ЭКГ у 47 пациентов и использовали метод ДР для их дальнейшей классификации. Точность, чувствительность и специфичность идентификации ИБС в этом исследовании составила 99%, 97,7% и 99,4%, соответственно. Juarez-Orozco LE, et al. (2020) [41] была выполнена попытка обработки результатов позитронно-эмиссионной томографии миокарда при помощи ансамблевого метода на основе ДР. В систему интеллектуального анализа были включены 16 клинико-демографических, лабораторных и функциональных признаков 1234 пациентов с подозрением на ИБС. В данной работе в качестве предикторов выступали такие факторы, как ЧСС в покое, систолическое АД, фракция выброса (ФВ) ЛЖ и возраст обследованных. СЛ — это алгоритм МО, применяемый для задач классификации, регрессии и кластеризации, использующий ансамбли ДР. Это универсальный и быстро обучаемый метод интеллектуального анализа для обнаружения линейных и нелинейных взаимосвязей внутри набора данных. СЛ обладает высокой предсказательной ценностью, способен эффективно обрабатывать большое число признаков [21]. В ряде исследований данный метод представлен в качестве одного из наиболее эффективных инструментов моделирования алгоритмов идентификации ИБС. Так, Chicco D и Jurman G (2020) [42], применив в своем исследовании СЛ, определили, что креатинин сыворотки крови и ФВ ЛЖ обладают высоким предиктивным

потенциалом для прогнозирования продолжительности жизни больных ИБС с хронической сердечной недостаточностью (ХСН) (площадь под ROC-кривой 0,800). В данной работе был проанализирован набор данных, включающий 13 показателей 299 пациентов в возрасте от 40 до 95 лет с систолической дисфункцией ЛЖ и ХСН III-IV функциональных классов. В другом исследовании авторы применили метод СЛ с целью улучшения прогностической способности модели для верификации ИБС, дополнив ее результатами генетического тестирования пациентов [43]. Ambale-Venkatesh B, et al. (2017) [44], используя набор данных MESA, продемонстрировали более высокую точность модели на основе СЛ в прогнозировании сердечно-сосудистых событий и исходов (смерть, инсульт, ИМ, декомпенсация ХСН и фибрилляция предсердий) в сравнении со стандартными шкалами ССР АНА/ASCVD, FRS и MESA CHD. Необходимо отметить, что метод СЛ активно используется для автоматизированного анализа изображений ЭхоКГ. Так, в исследовании [45] была разработана ансамблевая модель с тремя различными алгоритмами МО (СЛ, SVM и ИНС), способная дифференцировать физиологические и патологические паттерны гипертрофического ремоделирования миокарда ЛЖ. Для разработки автоматизированной системы оценки морфологических и функциональных характеристик сердечной мышцы были использованы данные ЭхоКГ 62 больных с гипертрофической кардиомиопатией и 77 пациентов с физиологической гипертрофией миокарда ЛЖ. Данная модель демонстрировала диагностическую ценность распознавания отдельных вариантов ремоделирования с чувствительностью 96%, специфичностью 77% и площадью под ROC-кривой 0,795.

ИНС и ГО в диагностике и прогнозировании ИБС

ИНС являются мощным классификатором и регрессором, которые широко используются в различных областях медицинских знаний. Этот метод основан на принципах организации и функционирования биологических нейронных сетей, имитируя работу человеческого мозга [39]. ИНС не программируются, а обучаются, что является одним из главных преимуществ данной технологии перед другими алгоритмами МО. Обучение ИНС заключается в расчете весовых коэффициентов нейронов, в процессе которого она способна выявлять зависимости между входными и выходными данными, а затем обобщать их. Ограничением в использовании ИНС являются высокие требования к вычислительным мощностям и их склонность к переобучению. Arabasadi Z, et al. (2017) [46] предложили гибридный метод на основе ИНС для диагностики ИБС, используя набор данных Z-Alizadeh Sani. Точность, чувствительность и специфичность данного метода составила 93,8%, 97% и 92%, соответственно. Weng

SF, et al. (2017) [47] на когорте из 378256 здоровых лиц Великобритании использовали различные методы МО, включая ИНС, для прогнозирования риска развития сердечно-сосудистых событий. В течение 10-летнего наблюдения у 24907 (6,6%) обследованных были диагностированы ИМ и ишемический инсульт. Авторами было обнаружено, что рассматриваемые в работе прогностические модели на основе ИНС превосходили по точности алгоритм оценки ССР по шкале ACC/АНА на 3,6%. В исследовании Kim JK, et al. (2017) [48] на корейской когорте из 4146 обследованных проводили сравнительную оценку прогностической способности алгоритмов верификации ИБС по “классической” шкале FRS и авторской модели на основе ИНС, предсказательная ценность которой была существенно выше (площадь под ROC-кривой 0,393 vs 0,749). Высокая точность ИНС показана и в работе отечественных авторов [49], в которой с аналогичной целью после предварительного определения ССР по FRS на выборке из 2236 пациентов была использована прогностическая модель. Ее точность составила 78,8% vs 70% по шкале FRS, а специфичность и чувствительность — 67% и 79% vs 83% и 25%, соответственно. Acharya UR, et al. (2018) [50] была разработана модель на основе ИНС для интерпретации ЭКГ у больных ИБС с проявлениями ХСН. По мнению авторов, эта модель требует минимальной предварительной обработки ЭКГ и может применяться как диагностический алгоритм в клинической практике. Ее максимальная точность составила 98,97%, а специфичность и чувствительность — 99%.

ГО — это подкласс ИНС, который использует многослойные нейронные сети. Методы ГО являются базовой платформой для приложений распознавания изображений, которые планируются или уже используются для визуализации: ангиография, ЭхоКГ, компьютерная томография, внутрисосудистое ультразвуковое исследование, оптическая когерентная томография и др. [3, 5, 51]. Наиболее распространенными вариантами ГО являются сверточные (CNN) и рекуррентные нейронные сети (RNN) [51]. В 2016г была впервые опубликована статья, в которой представлены результаты применения CNN для автоматической верификации патологических признаков ЭКГ [52]. В данной работе был продемонстрирован метод обнаружения желудочковых и наджелудочковых экстрасистол по результатам суточного мониторинга ЭКГ, достигавший точности в 99%. По мнению авторов, разработанная интеллектуальная система может быть применима к любому набору данных ЭКГ. В российском исследовании метод CNN применялся к данным 400 больных ИБС для прогнозирования риска летального исхода на горизонте 5 лет после острого коронарного синдрома. Площадь под ROC-кривой данной модели составила

0,74. Авторами были выделены предикторы, влияющие на продолжительность жизни пациентов после острого коронарного синдрома (возраст, ИМ в анамнезе, острое нарушение мозгового кровообращения, фибрилляция предсердий, ХСН, хроническая болезнь почек, ФВ <50%) [53]. В другой работе метод CNN использовали для идентификации кальциноза КА, классифицировав 2175 изображений интракоронарного ультразвукового исследования больных ИБС [54]. Данный метод в сочетании с НБК продемонстрировал 99% точность. Lui HW, et al. (2018) [55] разработали классификатор, позволяющий отличать ЭКГ-сигналы пациентов с ИМ от нормальных или зашумленных сигналов по первому стандартному отведению. В исследовании было проанализировано 549 записей ЭКГ 290 пациентов, из которых 368 кардиограмм принадлежало 148 больным с ИМ. При комбинации CNN и RNN чувствительность и специфичность модели составила 92,4% и 97,7%, соответственно. В работе [56] авторы, используя аналогичные показатели и когорту пациентов, с помощью усовершенствованной комбинации многослойной CNN и RNN продемонстрировали способность модели верифицировать ИМ по ЭКГ с точностью 95,4%, чувствительностью — 98,2% и специфичностью — 86,5%.

Таким образом, методы МО все чаще используются для интеллектуального анализа данных, характеризующих клинический статус пациентов с ИБС, в качестве дополнительного инструмента диагностических и прогностических исследований. К наиболее перспективным направлениям применения методов МО в клинической кардиологии, в т.ч. при ИБС, можно отнести разработку автоматизированных систем диагностики и предсказания сердечно-сосудистых событий. Первая из них позволяет осуществлять автоматический анализ результатов инструментальных и лабораторных исследований, создавать программные продукты, поддерживающие принятие врачебных решений, сокращать время и стоимость диагностических исследований. Модели предсказания событий могут обеспечить более высокую точность прогнозирования ССЗ при эпидемиологических исследованиях, повышать качество оценки рисков развития фатальных и нефатальных осложнений ИБС, в т.ч. в ближайшем и отдаленном периодах после реваскуляризации миокарда. Анализ данных литературы свидетельствует о возрастающем интересе исследователей к совершенствованию технологий ИИ, внедрение которых в клиническую практику является одним из важных факторов становления персонализированной медицины.

Отношения и деятельность. Работа выполнена при поддержке грантов РФФИ в рамках научных проектов № 18-29-03131, № 19-29-01077.

Литература/References

- The World Health Organization. Cardiovascular diseases. 2017. Available at: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- Vaysman DSh, Aleksandrova GA, Leonov SA, et al. The accuracy of indicators and the structure of causes of death from diseases of the circulatory system in the Russian Federation in international comparisons. Current problems of health care and medical statistics. 2019;3:87-102. (In Russ.). Вайсман Д. Ш., Александрова Г. А., Леонов С. А. и др. Достоверность показателей и структуры причин смерти от болезней системы кровообращения в российской Федерации при международных сопоставлениях. Современные проблемы здравоохранения и медицинской статистики. 2019;3:87-102. doi:10.24411/2312-2935-2019-00055.
- Johnson KW, Torres SJ, Glicksberg BS, et al. Artificial Intelligence in Cardiology. *J Am Coll Cardiol*. 2018;71(23):2668-79. doi:10.1016/j.jacc.2018.03.521.
- Geltser BI, Tsvanyuk MM, Shakhgeldyan KI, et al. Machine learning for assessing the pretest probability of obstructive and non-obstructive coronary artery disease. *Russian Journal of Cardiology*. 2020;25(5):3802. (In Russ.). Гельцер Б. И., Циванюк М. М., Шахгельдян К. И. и др. Методы машинного обучения в оценке предстеновой вероятности обструктивных и необструктивных поражений коронарного русла. Российский кардиологический журнал. 2020;25(5):3802. doi:10.15829/1560-4071-2020-3802.
- Krittawanong C, Zhang H, Wang Z, et al. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol*. 2017;69(21):2657-2664. doi:10.1016/j.jacc.2017.03.571.
- Leiner T, Rueckert D, Suinesiaputra A, et al. Machine learning in cardiovascular magnetic resonance: basic concepts and applications. *J Cardiovasc Magn Reson*. 2019;21(1):61. doi:10.1186/s12968-019-0575-y.
- Kagiyama N, Shrestha S, Farjo PD, et al. Artificial Intelligence: Practical Primer for Clinical Research in Cardiovascular Disease. *J Am Heart Assoc*. 2019;8:e012788. doi:10.1161/JAHA.119.012788.
- Quesada JA, Lopez-Pineda A, Gil-Guillen VF, et al. Machine learning to predict cardiovascular risk. *Int J Clin Pract*. 2019;73(10):e13389. doi:10.1111/ijcp.13389.
- Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 2000;101(23):e215-e220. doi:10.1161/01.cir.101.23.e215.
- UCI Machine Learning Repository. Available from: <http://archive.ics.uci.edu/ml>.
- Taddei A, Distante G, Emdin M, et al. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *Eur Heart J*. 1992;13(9):1164-72. doi:10.1093/oxfordjournals.eurheartj.a060332.
- Jager F, Taddei A, Moody GB, et al. Long-term ST database: A reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Med Biol Eng Comput*. 2003;41(2):172-82. doi:10.1007/bf02344885.
- St Petersburg INCART 12-lead Arrhythmia Database. 2008. doi:10.13026/C2V88N.
- Iyengar N, Peng CK, Morin R, et al. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology-Regulatory, Am J Physiol*. 1996;271(4):1078-84. doi:10.1152/ajpregu.1996.271.4.r1078.
- Alizadehsani R, Abdar M, Roshanzamir M, et al. Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Comput Biol Med*. 2019;103346. doi:10.1016/j.cmpbiomed.2019.103346.
- MESA — Multi-Ethnic Study of Atherosclerosis. Available from: <https://www.mesa-nhlbi.org/>.
- Saeed M, Villarroel M, Reiser AT, et al. Multiparameter Intelligent Monitoring in Intensive Care II: A public-access intensive care unit database. *Crit Care Med*. 2011;39(5):952-60. doi:10.1097/ccm.0b013e31820a92c6.
- Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Scientific Data*. 2016;3:160035. doi:10.1038/sdata.2016.35.
- Alizadehsani R, Hosseini MJ, Khosravi A, et al. Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries. *Computer Methods and Programs in Biomedicine*. 2018;162:119-27. doi:10.1016/j.cmpb.2018.05.009.
- Mastoi Q, Wah TY, Gopal Raj R, et al. Automated Diagnosis of Coronary Artery Disease: A Review and Workflow. *Cardiol Res Pract*. 2018;2018:2016282. doi:10.1155/2018/2016282.
- Martin-Isla C, Campello VM, Izquierdo C, et al. Image-based cardiac diagnosis with machine learning: a review. *Front Cardiovasc Med*. 2020;7:1. doi:10.3389/fcvm.2020.00001.
- Yaroslavskaya EI, Kuznetsov VA, Gorbatenko EA, et al. Calculator of non-obstructive coronary atherosclerosis: clinical case of a male patient with suspected coronary artery disease. *The Siberian Medical Journal*. 2018;33(3):93-101. (In Russ.). Ярославская Е. И., Кузнецов В. А., Горбатенко Е. А. и др. Калькулятор неструктурного коронарного атеросклероза: клинический пример использования у мужчины с подозрением на ишемическую болезнь сердца. Сибирский медицинский журнал. 2018;33(3):93-101. doi:10.29001/2073-8552-2018-33-3-93-101.
- Roe MT, Harrington RA, Prosper DM, et al. Clinical and Therapeutic Profile of Patients Presenting with Acute Coronary Syndromes Who Do Not Have Significant Coronary Artery Disease. *Circulation*. 2000;102(10):1101-6. doi:10.1161/01.cir.102.10.1101.
- Verma L, Srivastava S, Negi PC. A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data. *J Med Syst*. 2016;40(7):178. doi:10.1007/s10916-016-0536-z.
- Nakao YM, Miyamoto Y, Higashi M, et al. Sex differences in impact of coronary artery calcification to predict coronary artery disease. *Heart*. 2018;104(13):1118-24. doi:10.1136/heartjnl-2017-312151.
- Ballesteros-Ortega D, Martinez-gonzalez O, Blancas R, et al. Characteristics of patients with myocardial infarction with nonobstructive coronary arteries (MINOCA) from the ARIAM-SEMICYUC registry: development of a score for predicting MINOCA. *Vasc Health Risk Manag*. 2019;15:57-67. doi:10.2147/vhrm.s185082.
- Nikulina SY, Chernova AA, Tretyakova SS, et al. Prediction of cardiac conduction disorders using the methods of mathematical analysis. *Russian Journal of Cardiology*. 2018;(10):53-58. (In Russ.). Никулина С. Ю., Чернова А. А., Третьякова С. С. и др. Прогнозирование нарушений сердечной проводимости с использованием методов математического анализа. Российский кардиологический журнал. 2018;(10):53-58. doi:10.15829/1560-4071-2018-10-53-58.
- Sharma K, Shah K, Brahmabhatt P, et al. Skipping breakfast and the risk of coronary artery disease. *QJM*. 2018;111(10):715-719. doi:10.1093/qjmed/hcy162.
- Kozlova EV, Starostin IV, Bulkina OS, et al. Evaluation of the prevalence of cardiovascular events and mortality in stable coronary heart disease patients depending on baseline coronary collateral blood flow (five-year follow-up). *Russian Journal of Cardiology*. 2018;(3):11-6. (In Russ.). Козлова Е. В., Старостин И. В., Булкина О. С. и др. Оценка частоты сердечно-сосудистых событий и смертности у больных со стабильной ишемической болезнью сердца в зависимости от исходного состояния коронарного коллатерального кровотока (5-летнее наблюдение). Российский кардиологический журнал. 2018;(3):11-6. doi:10.15829/1560-4071-2018-3-11-16.
- Gao Y, Zhang Q, Pan T. Relation of monocyte/high-density lipoprotein cholesterol ratio with coronary artery disease in type 2 diabetes mellitus. *Clin Lab*. 2018;64(6):901-6. doi:10.7754/Clin.Lab.2018.171022.
- Sridhar S, Acharya UR, Bairy GM. Automated diagnosis of Coronary Artery Disease using nonlinear features extracted from ECG signals. 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest. 2016, pp. 000545-000549, doi:10.1109/SMC.2016.7844296.
- Shi M, Zhan C, He H, et al. Renyi distribution entropy analysis of short-term heart rate variability signals and its application in coronary artery disease detection. *Front Physiol*. 2019;10:809. doi:10.3389/fphys.2019.00809.
- Tabassian M, Alessandrini M, Herbots L, et al. Machine learning of the spatio-temporal characteristics of echocardiographic deformation curves for infarct classification. *Int J Cardiovasc Imaging*. 2017;33(8):1159-1167. doi:10.1007/s10554-017-1108-0.
- Juarez-Orozco LE, Saraste A, Capodanno D, et al. Impact of a decreasing pre-test probability on the performance of diagnostic tests for coronary artery disease. *Eur Heart J Cardiovasc Imaging*. 2019;20(11):1198-1207. doi:10.1093/ehjci/jez054.
- Knuuti J, Wijns W, Saraste A, et al. 2019 ESC Guidelines on the diagnosis and management of chronic coronary syndromes: The Task Force for diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). *Eur Heart J*. 2019;41(3):407-77. doi:10.1093/eurheartj/ehz425.
- Bae Y, Kang S-J, Kim G, et al. Prediction of coronary thin-cap fibroatheroma by intravascular ultrasound-based machine learning. *Atherosclerosis*. 2019;288:168-74. doi:10.1016/j.atherosclerosis.2019.04.228.
- Berikol GB, Yildiz O, Ozcan IT. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. *J Med Syst*. 2016;40(4):84. doi:10.1007/s10916-016-0432-6.
- Cui S, Li K, Ang L, et al. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv*. 2017;10(13):1307-1316. doi:10.1016/j.jcin.2017.04.007.
- Ahmadi E, Weckman GR, Masel DT. Decision making model to predict presence of coronary artery disease using neural network and C5.0 decision tree. *J Ambient Intell Human Comput*. 2017;9(4):999-1011. doi:10.1007/s12652-017-0499-z.
- Acharya UR, Sudarshan VK, Koh JEW, et al. Application of higher-order spectra for the characterization of Coronary artery disease using electrocardiogram signals. *Biomedical Signal Processing and Control*. 2017;31:31-43. doi:10.1016/j.bspc.2016.07.003.
- Juarez-Orozco LE, Knol RJ, Sanchez-Catusas CA, et al. Machine learning in the integration of simple variables for identifying patients with myocardial ischemia. *J Nucl Cardiol*. 2020;27:147-55. doi:10.1007/s12350-018-1304-x.
- Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Med Inform Decis Mak*. 2020;20(1):16. doi:10.1186/s12911-020-1023-5.
- Dogan MV, Grumbach IM, Michaelson JJ, et al. Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. *PLOS ONE*. 2018;13(1):e0190549. doi:10.1371/journal.pone.0190549.
- Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. *Circ Res*. 2017;121(9):1092-1101. doi:10.1161/CIRCRESAHA.117.31312.
- Narula S, Shameer K, Salem Omar AM, et al. Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. *J Am Coll Cardiol*. 2016;68:2287-95. doi:10.1016/j.jacc.2016.08.062.

46. Arabasadi Z, Alizadehsani R, Roshanzamir M, et al. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Computer Methods and Programs in Biomedicine*. 2017;141:19-26. doi:10.1016/j.cmpb.2017.01.004.
47. Weng SF, Reys J, Kai J, et al. Can machine learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;12(4):e0174944. doi:10.1371/journal.pone.0174944.
48. Kim JK, Kang S. Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis. *Journal of Healthcare Engineering*. 2017;1-13. doi:10.1155/2017/2780501.
49. Gusev AV, Gavrilov DV, Korsakov IN, et al. Prospects for the use of machine learning methods for predicting cardiovascular disease. *Information technologies for the Physician*. 2019(3):41-47. (In Russ.) 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.
50. Acharya UR, Fujita H, Oh SL, et al. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. *Applied Intelligence*. 2018;49:16-27. doi:10.1007/s10489-018-1179-1.
51. Lih OS, Jahmunah V, San TR, et al. Comprehensive electrocardiographic diagnosis based on deep learning. *Artificial Intelligence in Medicine*. 2020;101789. doi:10.1016/j.artmed.2019.101789.
52. Kiranyaz S, Ince T, Gabbouj M. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Trans Biomed Eng*. 2016;63(3):664-75. doi:10.1109/tbme.2015.2468589.
53. Shvets DA, Karasev AY, Smolyakov MV, et al. Neural network analysis of mortality risk predictors in patients after acute coronary syndrome. *Russian Journal of Cardiology*. 2020;25(3):3645. (In Russ.) 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;25(3):3645. doi:10.15829/1560-4071-2020-3-3645.
54. Sofian H, Chia Ming JT, Noor NM. Calcification Detection Using Deep Structured Learning in Intravascular Ultrasound Image for Coronary Artery Disease, 2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS), Kuching. 2018. pp. 47-52. doi:10.1109/ICBAPS.2018.8527415.
55. Lui HW, Chow KL. Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices. *Informatics in Med*. Unlocked. 2018;13:26-33. doi:10.1016/j.imu.2018.08.002.
56. Feng K, Pi X, Liu H, et al. Myocardial Infarction Classification Based on Convolutional Neural Network and Recurrent Neural Network. *Applied Sciences*. 2019;9(9):1879. doi:10.3390/app9091879.