

УДК 614.2

DOI 10.24412/2312-2935-2023-2-804-829

ПРОГНОЗИРОВАНИЕ СЕРДЕЧНО-СОСУДИСТЫХ СОБЫТИЙ С ПОМОЩЬЮ МОДЕЛЕЙ ПРОПОРЦИОНАЛЬНЫХ РИСКОВ И МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ: СИСТЕМАТИЧЕСКИЙ ОБЗОР

И. А. Мишкин^{1,2}, А. В. Концевая¹, А. В. Гусев^{3,4}, 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

Актуальность. Ежегодно во всем мире большое количество людей становятся жертвами сердечно-сосудистых заболеваний. На сегодняшний день основными инструментами прогноза сердечно-сосудистого риска являются шкалы, основанные на моделях пропорциональных рисков (Cox-регрессии). Однако последнее время многие ученые сходятся во мнении, что использование технологий машинного обучения и искусственного интеллекта поможет повысить качество прогноза наступления неблагоприятных сердечно-сосудистых событий.

Цель - провести систематический литературный обзор подходов к формированию прогнозов развития ССЗ, основанных на шкалах пропорциональной оценки рисков и методах ML для обнаружения наиболее эффективных методов анализа данных.

Материалы и методы: Проведен систематический обзор литературы, в который были включены 58 научно-исследовательских работ, использующих методы оценки сердечно-сосудистого риска, основанных на Cox-регрессии, и технологий машинного обучения.

Результаты. Предиктивные возможности машинного обучения превосходят традиционные линейные методы анализа данных. Средние значения AUC 0,82 и 0,75 соответственно, $p=0,003$. Также удалось выделить наиболее частоиспользуемые и эффективные алгоритмы прогноза. Ими оказались random forest, gradient boosting и deep learning. Однако, в отличие от традиционных шкал прогноза, среди представленных алгоритмов машинного обучения 80% не проходили внешнюю валидацию на независимых выборках. Кроме того, использование машинного обучения требует большого объема качественных цифровых данных.

Обсуждение. В результате анализа аналогичных исследований, проведенных отечественными и зарубежными авторами, нам удалось подтвердить, что в среднем прогностические модели, построенные с помощью алгоритмов ML, имеют преимущества по сравнению с традиционными методами анализа данных.

Выводы. Машинное обучение является перспективным методом прогнозирования сердечно-сосудистых событий, одна для его массового использования необходим переход на электронное ведение медицинской документации и агрегация большего количества качественной и структурированной информации.

Ключевые слова: прогноз инсульта, машинное обучение в здравоохранении, прогноз сердечно-сосудистых заболеваний, искусственный интеллект в медицине, прогноз инфаркта миокарда

PREDICTION OF CARDIOVASCULAR EVENTS USING PROPORTIONAL RISK MODELS AND MACHINE LEARNING MODELS: A SYSTEMATIC REVIEW

I. A. Mishkin^{1,2}, A.V. Kontsevaya¹, A.V. Gusev^{3,4}, 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*

Relevance. Every year, a large number of people around the world become victims of cardiovascular diseases. To date, the main tools for predicting cardiovascular risk are scales based on proportional risk models (Cox regression). However, recently many scientists agree that the use of machine learning and artificial intelligence technologies can help to improve the quality of adverse cardiovascular events onset prognosis.

The aim is to conduct a systematic literature review of approaches to the formation of CVD development forecasts based on proportional risk assessment scales and ML methods to identify the most effective methods of data analysis.

Materials and methods: A systematic review of the literature was conducted, which included 58 research papers using methods for assessing cardiovascular risk based on Cox regression and machine learning technologies.

Results. Predictive capabilities of machine learning are superior to traditional linear methods of data analysis. The average AUC values are 0.82 and 0.75, respectively, $p=0.003$. It was also possible to identify the most frequently used and effective prediction algorithms. They turned out to be random forest, gradient boosting and deep learning. However, unlike traditional prediction scales, 80% of the presented machine learning algorithms did not undergo external validation on independent samples. In addition, the use of machine learning requires a large amount of high-quality digital data.

Discussion. As a result of similar studies analysis conducted by domestic and foreign authors, we were able to confirm that, on average, predictive models built using ML algorithms have advantages over traditional data analysis methods.

Conclusions. Machine learning is a promising method of predicting cardiovascular events, but for its mass use it is necessary to switch to electronic medical records management and aggregation of more qualitative and structured information.

Keywords: stroke prognosis, machine learning in healthcare, cardiovascular disease prognosis, artificial intelligence in medicine, myocardial infarction prognosis

Введение. В 21 веке проблема высокого распространения сердечно-сосудистых заболеваний (ССЗ) звучит, как никогда актуально. В 2021 году данная группа болезней унесла жизни 17,9 млн. человек [1]. Увеличение средней продолжительности жизни и тренд на старения населения, несомненно, оказывают влияние и на рост количества ССЗ [2,3]. По прогнозам всемирной организации здравоохранения (ВОЗ), к 2030 г. мировая смертность от данной группы патологий превысит 23,6 млн человек [4].

На сегодняшний день большинство развитых стран мира стремится переформатировать медицинскую отрасль в сторону профилактики развития болезни [5]. Помимо снижения числа хронических неинфекционных заболеваний (ХНИЗ), что, безусловно, должно привести к снижению преждевременной смерти и росту средней продолжительности жизни, а также способствовать снижению затрат на лечение и реабилитацию пациентов.

Важно отметить, что экономический ущерб от ССЗ в РФ в 2016 г. составил 2,7 трлн. рублей, что эквивалентно 3,2% ВВП страны соответствующего года. В структуре ущерба свыше 90% преобладали косвенные потери, обусловленные преждевременной смертью лиц экономически активного возраста. Прямые затраты составили лишь 8,1% [6].

На сегодняшний день в основе концепции индивидуальной профилактики ССЗ лежит оценка сердечно-сосудистого риска (ССР) [8]. Данная технология базируется на использовании шкал. Наиболее известными из них являются SCORE, SCORE2, Framingham и т.д. [9]. Шкалы позволяют оценить риск развития фатальных и нефатальных сердечно-сосудистых событий и наглядно оценить эффект и пользу от профилактической терапии благодаря сравнению конкретного риска у каждого индивида и показателей риска при идеальных параметрах. Однако многие исследователи сходятся во мнении, что шкалы имеют ряд ограничений [1,9-12,14,15]:

- для точной и корректной оценки риска необходима адаптация алгоритмов шкал под конкретную популяционную группу;

- во многих шкалах отсутствуют такие ключевые факторы риска (ФР), как коморбидность;

- многие шкалы, используемые сейчас в клинической практике, были разработаны на данных исследований, проведенных 30-50 лет назад. За прошедший период изменилась не только эпидемиология ССЗ, но также их тактика лечения. Об этом свидетельствуют работы [13], которые демонстрируют эффективность АКШ в лечении ИМ.

- статистические методы обработки информации, основанные на Соx-регрессии, позволяют обрабатывать лишь ограниченное число линейно связанных между собой переменных. При этом такие ФР, как возраст и смертность, индекс массы тела (ИМТ) и заболеваемость, очевидно, связаны нелинейно.

- традиционные шкалы оценивают показатели пациента в определенный момент времени (момент обследования). Данный недостаток не позволяет отслеживать изменения показателей риска в процессе терапии или ее коррекции. Такие ФР, как нелеченая

артериальная гипертензия (АГ) и курения имеют накопительный патологический эффект. Таким образом, если на момент осмотра пациент утверждает, что не курит, но при этом имел длительный стаж курения, нельзя пренебрегать теми патологическими изменениями, которые уже нанесены организму.

- используемые шкалы занижают ССР у молодых людей с выраженными модифицируемыми ФР. Это наглядно продемонстрировано в работе Корягина Н. А. и соавт. [21]. Распространённость таких ФР, как курение и АГ среди лиц в возрасте от 18 до 40 лет превышает 20%. Напротив, у пожилых пациентов при идеальных показателях здоровья риск может быть очень высоким только за счет возраста.

- невысокие показатели ROC-анализа. Для Фрамингемской шкалы показатели AUC варьируются от 0,69 до 0,79. Для шкалы SCORE данные значения составляют от 0,71 до 0,82 (для популяции из России – 0,71).

Традиционные шкалы риска хорошо работают на группе людей, но для конкретного пациента обладают недостаточной точностью. Вышеперечисленные недостатки могут привести к ошибкам в выявлении лиц с высоким ССР. Таким образом, профилактическому воздействию могут подвергаться пациенты, не нуждающиеся в этом, и, наоборот, те кто в ней действительно нуждаются могут не получать необходимые профилактические меры [16].

За последнее десятилетие в медико-биологических исследованиях большую популярность набрали математические методы нелинейного анализа данных [11]. Среди упомянутых методов наиболее активно развиваются технологии искусственного интеллекта (artificial intelligent, AI) и машинного обучения (machine learning, ML) [17,18].

Использование нелинейных моделей позволяет преодолеть барьеры, связанные с линейным анализом данных, и, таким образом, определять ССР, базируясь на индивидуальных особенностях каждого человека. Основа работы данных алгоритмов будет складываться из тех клинических примеров, на которых они были обучены [19].

В связи с актуальностью темы, мы проверили систематический литературный обзор подходов к формированию прогнозов развития ССЗ, основанных на шкалах пропорциональной оценки рисков и методах ML с целью обнаружения наиболее эффективных методов анализа данных.

Материалы и методы. Данный систематический научный обзор был выполнен на основе методологии PRISMA [20]. Обзор содержит результат работ, опубликованных за период 2018 – 2022 гг, а также фундаментальные работы за неограниченный период времени.

Поиск источников производился с использованием баз данных: PubMed, Web of Science, Scopus и Российский индекс научного цитирования (РИНЦ). Для отбора источников литературы использовалась методика PICO по следующим ключевым словам: сердечно-сосудистый риск, машинное обучение, прогноз сердечно-сосудистых заболеваний, искусственный интеллект, stroke prediction, machine learning, cardiovascular diseases, artificial intelligence, asymptomatic subjects, myocardial infarction.

В результате обработки поисковых запросов: фильтрации, удаление дублирующих поисковых выдачей, удалось получить 58 работ (рис.1),

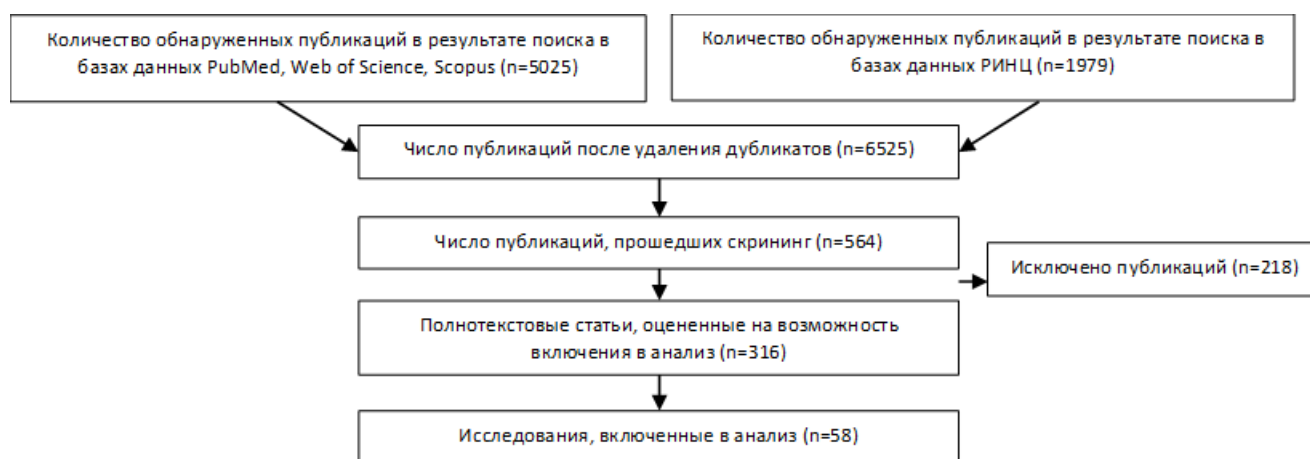


Рисунок 1. Дизайн исследования

удовлетворяющих требованиям данного систематического обзора. Анализ проводили два исследователя которые просматривали резюме публикаций, прошедших скрининг после удаления дубликатов на предмет соответствия целям систематического обзора. При наличии расхождений в мнении экспертов, оно разрешалось путем переговоров. В окончательный анализ включено 18 публикаций по Cox-регрессии и 40 работ по алгоритмам ML.

Результаты. Медианный объем выборки исследований для методов Cox-регрессии составил: $n=7047$, $\min=122$, $\max=7890000$. Меньше 1 тыс. наблюдений – 5,6%, от 1 тыс. до 100 тыс. – 72,2%, от 100 тыс. до 1 млн. – 11,1%, больше 1 млн. наблюдений – 11,1%. Аналогичный показатель для моделей ML: $n=4699$, $\min=155$, $\max=503842$. Меньше 1 тыс. наблюдений – 33,3%, от 1 тыс. до 100 тыс. – 51,3%, от 100 тыс. до 1 млн. – 15,4%, больше 1 млн. наблюдений – 0%.

Среди моделей Cox-регрессии 77,8% прошли внешнюю валидацию на независимых выборках, 16,7% кросс-валидацию. Среди моделей ML – 12,5% внешнюю валидацию на

независимых выборках, 52,5% кросс-валидацию, 27,5% валидация на части от исходной выборки.

Доля исследований, проведенных в США, составила 47,5%, что, очевидно, превалирует над количеством исследований, проведенных в других странах. Также США превалирует по количеству человек в исследуемых и валидационных когортах (рис. 2.). На втором месте находится Китай.



Рисунок 2. Географическое распределение по объему когортных исследований представленных в работе, кол. человек

Для определения алгоритмов ML, продемонстрировавших наилучшие предиктивные возможности, мы сравнивали и выбирали наивысшие показатели AUC среди всех протестированных моделей в каждой работе. Как видно из рис. 3., наилучшими предиктивными возможностями наиболее часто

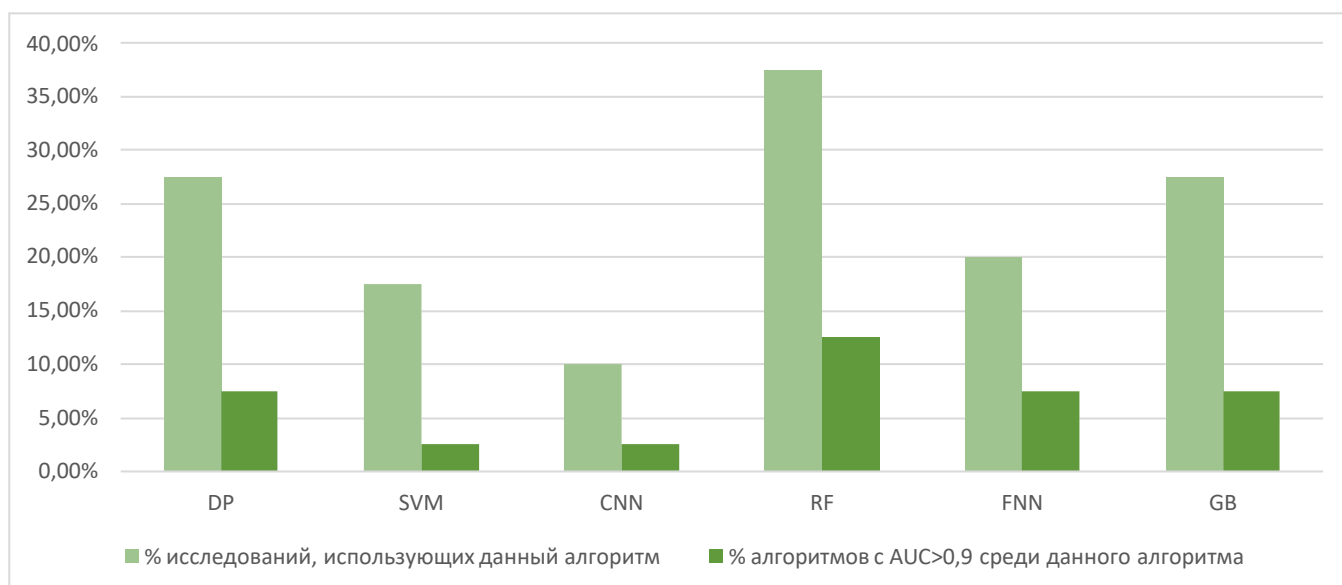


Рисунок 3. Распределение алгоритмов ML по частоте использования и их эффективности в анализируемых работах, %

обладали: RF, GB и DP. Данные алгоритмы наиболее часты использовались в рассмотренных работах, а также имели наибольшую долю алгоритмы с показателями AUC выше 0,90.

Также из данных таблицы 1. видно, что ни у одной модели, основанной на Соx-регрессии, показатель AUC не превышал значения 0,90. При этом у моделей ML - 35% рассмотренных алгоритмов имели $AUC \geq 0,90$. Также в среднем на 0,07 средние показатели AUC у ML ($M=0,82 \pm 0,01$) были выше соответствующих показателей моделей Соx-регрессии ($M=0,75 \pm 0,01$), $p=0,003$.

Между объемами обучающей выборки и показателями AUC у моделей ML статистически значимых корреляций не обнаружено. Также не было обнаружено значимых различий между средними показателями AUC среди трех методов валидации (кросс-валидация, часть исходной выборки и внешняя независимая выборка).

Таблица 1

Характеристики моделей прогноза ССС, основанных на модели пропорциональных рисков

Название шкалы	Прогнозируемое событие	Предикторы	Объём выборки/регион проведения исследования, чел.	Объём валидационной выборки/регион валидации, чел.	Ограничения использования	AUC
Прогноз ССС у здоровых людей						
SCORE	10-ти летняя сердечно-сосудистая смерть	Возраст, сАД, ОХС, пол, курение	205,178 тыс., Европейский регион	Европейский регион	Возраст 45-64 гг., отсутствие коморбидности, прогноз фатальных исходов	0,72 - 0,81
SCORE2	10-ти летний риск развития фатальных и нефатальных ССЗ	Возраст, сАД, ОХС, пол, курение	677,684 тыс., Европейский регион	1,13 млн., Европейский регион	Возраст 40-69 лет, отсутствие коморбидности	0,67 - 0,81
SCORE2-OP	5-ти и 10-ти летний риск сердечно-сосудистой смерти, ИМ и инсульта	Возраст, сАД, ОХС, пол, курение	28,503 тыс., Европейский регион	340 тыс., Европейский региона	Возраст от 70 лет, отсутствие коморбидности	0,63 - 0,67

Научно-практический рецензируемый журнал
 "Современные проблемы здравоохранения и медицинской статистики" 2023 г., № 2
 Scientific journal "Current problems of health care and medical statistics" 2023 г., № 2
 ISSN 2312-2935

PCE ASCVD Risk Estimator +	10-ти летний риск сердечно-сосудистой смерти, ИМ и инсульта	Возраст, раса, сАД, ОХС, ЛПВП, ЛПНП, пол, курение, СД, АГПТ, терапия статинами, прием аспирина	Данные Фрамингемского исследования (≈5,348 тыс.)	Северная Америка	Возраст 20-79 лет	0,66 - 0,81
QRISK3	10-ти летний риск развития ИМ, ИБС, инсульта, ТИА	Возраст, раса, сАД, ОХС, ЛПВП, ЛПНП, пол, курение, СД, ХБП, сАД, и т.д.	7,89 млн., Соединенное королевство	2,67 млн., Соединенное королевство	Возраст 25-84 лет	0,88
Astro-CHARM	10-ти летний риск развития атеросклеротических заболеваний сердца	Возраст, раса, сАД, ОХС, ЛПВП, пол, курение, СД, АГПТ и т.д.	7,382 тыс., данные исследований: MESA	FHS Offspring и Third Generation cohorts	Возраст 40-65 лет	0,78
Прогноз ССС у коморбидных пациентов						
SMART	10-ти летний риск сердечно-сосудистой смерти, атеросклеротической болезнью сердца ИМ и инсульта у пациентов с	Возраст, сАД, ОХС, ЛПВП, пол, курение, СД, уровень С-реактивного белка, СКФ, длительность заболевания и т.д.	5,788 тыс., Нидерланды	18,436 тыс, Европа, Израиль, США, Канада, Мексика, Южная Африка, Австралия и Новая Зеландия	Возраст 30-90 лет	0,62 - 0,68
CADC2	Развитие обструктивной формы у пациентов с ИБС	Возраст, пол, характер боли в груди, сАД, курение, дислипидемия, наличие АГ, СД, кальциевый индекс	5,677 тыс., Европа и США	Кросс-валидация	-	0,69-0,79

ADVANCE	4-х и 10-ти летний риск развития сердечно-сосудистой смерти, ИМ и инсульта у пациентов с СД 2 типа	Возраст, пол, длительность СД, курение, ФП, ретинопатия, АГПТ, сАД, дАД и т.д.	11,140 тыс., 20 стран всех регионов мира	?	Возраст 55-90 лет	?
Инфаркт миокарда						
EDACS-ADP	ССС в течение 30 суток после перенесенного ИМ	Возраст, пол, наличие ИБС, потливость, иррадиация болей в плечо/руку, усиление боли при дыхании и т.д	1,974 тыс. пациентов с низким ССР Австралии и Новой Зеландии	608, Австралия и Новая Зеландии	Пациенты с изначально низким ССР	0,87
HEART	Повторный ИМ, реваскуляризация, смерть в течение 30 суток, 6 недель, года у пациентов, перенесших ИМ	Возраст, данные анамнеза, результаты ЭКГ, уровень тропонина	122, Нидерланды	?	-	0,83-0,86
GRACE	Повторная госпитальная смерть, смерть в течение 6-ти месяцев у пациентов, перенесших ИМ	Возраст, сАД, ЧСС, уровень креатинина, класс по Killip, биомаркеры, изменение сегмента ST на ЭКГ, остановка сердца	21, 688 тыс. и 94, госпитали 14-ти стран Европы, Северной и Южной Америки, Австралии и Новой Зеландии	22, 122 тыс. независимая выборка	-	0,81-0,84
Фибрилляция предсердий						
FRS	10-ти летний появления ФП	Возраст, ИМТ, сАД, наличие антигипертензивной терапии, интервал PR на ЭКГ и т.д.	4,764 тыс., когорта Фрамингемского исследования	49,599 тыс., когорта Фрамингемского исследования	Возраст 45-95 лет	0,71-0,78

RED-AF	30-дневный риск обращения за неотложной помощью, повторные госпитализации с ФП, сердечно-сосудистой смерти	Возраст, пол, одышка, сердцебиение, курение, периферические отеки, ЧСС, прием β -блокаторов и т.д.	1,870 тыс., Чикаго США	mixed-effects regression models	-	0,65-0,70
Сердечная недостаточность						
SHFM	Смерть в течении 1, 2, 5-ти лет у пациентов с СН	Возраст, пол, ишемическая этиология, класс NYHA, ФВЛЖ, сАД, уровень гемоглобина и т.д.	1,125 тыс., США и Канада	9,942 тыс., США и Канады	Для пациентов с СН с ФВЛЖ \leq 40%	0,69-0,79
MAGGIC	Смерть в течение 1-го и 3-х лет у пациентов с ХСН	Возраст, пол, ИМТ, класс NYHA, ФВЛЖ, сАД, уровень креатинина, наличие СД, прием β -блокаторов и т.д.	39,372 тыс., Великобритания	Кросс-валидация и 382 пациента из Шведского регистра	-	0,74
Инсульт						
QStroke	Вероятность развития инсульта в течение 10-ти лет	Раса, возраст, пол, статус курения, сАД, индекс атерогенности, ИМТ, семейный анамнез и т.д.	3,5 млн., Великобритания	1,9 млн., Великобритания	Возраст 25-84 года	0,78-0,87
R-FSRS	Вероятность развития инсульта в течение 10-ти лет	Возраст, пол, сАД, антигипертензивная терапия, СД, наличие ССЗ, ФП	6,712 тыс., когорта Фрамингемского исследования	Когорта Фрамингемского исследования	-	0,72

Примечание. АГТП – антигипертензивная терапия

Таблица 2

Характеристики моделей прогноза ССЗ, основанных на технологии машинного обучения

Авторы исследования	Алгоритм ИИ	Прогнозируемое событие	Предикторы	Объем выборки/регион проведения исследования, чел.	Объем валидационной выборки/регион валидации, чел.	Ограничения использования	AUC
Прогноз ССЗ у здоровых людей							
Kim et al. (2021)	10 видов машинного обучения (extreme GB*, GB*, RF и т.д.)	ССЗ	Пол, возраст, лабораторные данные, физическая активность, анамнез, наследственность, ИМТ и т.д.	4,699 тыс., National Health Insurance Corporation (NHIC)	4-fold cross-validation	Старше 45 лет	0,79 - 0,82
Ward et al. (2020)	LR, RF, GB*, extreme GB	Атеросклеротические заболевания	1,175 тыс. предикторов	262,923 тыс., Northern California	5-fold cross-validation	-	0,85
Плехова и др. (2020)	FNN 5 слоев	Исходы и риски ССЗ	ИМТ, вес, обхват талии, обхват бедер, возраст, пол, среднее ЧСС, триглицериды и т.д.	467, г. РФ	187 человек из исходной выборки	Возраст 25-64 года	0,97
Jamthikar et al. (2020)	SVM	Прогноз 10-ти летнего ССР	возраст, пол, сАД, дАД, АГ, гликированный гемоглобин (HbA1c), и т.д.	202, Ohashi Medical Center, Toho University Japan	10-fold cross-validation	-	0,88

Соболева и др. (2020)	Oracle Data Mining	Прогноз ССЗ	возраст, пол, ИМТ, АД, наличие или отсутствие СД и т.д.	10 тыс. Фрамингеское исследование, «Heart Disease» - 303 записи, Набор данных хакатона AgeHack - 100 000 записей (искусственно созданные)	нет	Обучение на искусственно-созданных данных	0,73 - 0,87
Quesada et al. (2019)	Cox regression, NB*, DA*, LR, K-nearest neighbours, SVM, FNN*, classification trees, bagging, AdaBoost, RF	Оценка ССР	возраст, пол, ОХС, сАД и курение	38,527 тыс..Spanish ESCARVAL RISK clinical practice cohort	30% от исходной выборки	Пациенты старше 40-ка лет	0,70
Гусев и др. (2019)	FNN 5 слоев	Прогноз ССС в течение 10-ти лет	возраст, пол, ИМТ, АД, наличие или отсутствие СД и т.д.	2,236 тыс., РФ	25% от исходной выборки	Возраст 41-84 года	0,84
Poplin et al. (2018)	DL	Прогноз ССС в течение 5-ти лет	Изображения сетчатки глаза	284,335 тыс., UK Biobank и EyePACS	12,026 тыс. и 999 человек из независимой выборки	-	0,66-0,73
Nakanishi et al. (2021)	LogitBoost	Прогноз 10-ти летней смертности от ИБС и других ССЗ	Оценка риска ССЗ, возраст, пол, раса, ИМТ, гипертония, СД, гиперлипидемия и т.д.	66,636 тыс. США	10-fold cross validation	-	0,82-0,86

Eisenberg et al. (2020)	DL	Прогноз ССС	Данные антропометрии, анамнеза, данные лабораторных обследований, прием медикаментов и т.д.	2,068 тыс., EISNER trial	Кросс-валидация	Возраст 45–80 года	0,76
Commandeur (2020)	Extreme GB	Прогноз долгосрочного риска развития ИИ и сердечно-сосудистой смерти	Данные антропометрии, анамнеза, данные лабораторных обследований и т.д. (22 предиктора)	1,912 тыс. EISNER trial	10-fold cross validation	-	0,82
Мишкин и др. (2020)	RF, GB*	Оценка риска развития ССЗ	пол, возраст, ИМТ, ЧСС, сАД, сАД и т.д.	800, РФ	5-fold cross validation	-	0,82 - 0,97
Jamthikar et al. (2019)	RF	Оценка ССР	Данные антропометрии, анамнеза, данные лабораторных обследований и т.д. (47 предиктора)	202, Toho University, Japan	10-fold cross-validation		0,80
Прогноз ССС у коморбидных пациентов							
Невзорова и др. (2020)	FNN 5 слоев	Прогнозирование исходов и рисков ССС у пациентов с АГ	ИМТ, окружность талии, ЧСС, курение и т.д.	733, РФ	245 человек от исходной выборки	-	0,95
de Gonzalo-Calvo (2020)	regression tree models	Прогноз ССР у пациентов с почечной недостаточностью	возраст, пол, регион, ИМТ, КТ/В	2,776 тыс., AURORA trial	Кросс-валидация	Возраст 50-80 лет	0,71

Cheng et al. (2022)	FNN	риск стеноза коронарных артерий у пациентов с ИБС	ИМТ, объем тали, данные анализов крови (13-предикторов)	155, Third People's Hospital of Hefei	10-fold cross validation	-	0,84 - 0,95
Chu et al. (2021)	DL	Прогноз ССЗ у пациентов с СД	пол, возраст, уровень образования, HbA1c, уровень глюкозы натощак и т.д.	834, США	10-fold cross validation	Возраст 25–78 лет	0,91
Tran et al. (2021)	LR, RF*, extra trees and GB*	Прогноз летальных исходов у пациентов с ССЗ	демографические данные пациентов, показатели здоровья, биомедицинские изображения или клинические заметки	346,201 тыс., Австралия	5-fold cross-validation	-	0,95 - 0,97
Hong et al. (2021)	LASSO regularization method	5-ти летний ССР у пациентов с СД 2 типа	возраст, раса/этническая принадлежность, пол, лабораторные анализы, вес, рост и т.д.	6,245 тыс., T2DM cohort in Ochsner Health System	Исходная выборка	-	0,85
Инфаркт миокарда							
Швец и др. (2020)	FNN, CNN*	Прогнозирование рисков летальных исходов у пациентов перенесших ОКС	Данные анамнеза, данные антропометрии, показатели анализов крови (28 ФР)	400, РФ	60 человек от исходной выборки	-	0,63 - 0,74
Zack et al. (2019)	RF	30-ти дневная повторная госпитализация и смерть в течение 180-ти дней после перенесенного ИМ	Данные анамнеза, данные антропометрии, показатели анализов крови (508 предикторов)	11,709 тыс., Mayo Clinic, US	8-fold cross-validation	-	0,88 - 0,92

Xue et al. (2021)	k-means, hierarchical, partitioning around medoids, and k- medoids algorithms	Стратификация риска пациентов с ИМ	Данные липидограммы	1,355 тыс., STEMI	?	-	?
Макарихин и др. (2021)	GB	Прогноз ИМ	130 тыс. показателей	7,557 тыс., Тюменская область, РФ	Независимая выборка г. Тюмень	-	F1 score – 0,83 -0,93
Саостин и др. (2019)	decision trees, k- ближайших соседей*, logistic regression, RF*, SVM*	Определение ИМ по данным ЭКГ	Данные ЭКГ	535 записей из открытой база записей ЭКГ The PTB Diagnostic ECG Database	?	-	F1- score - 0,85- 0,98
Фибрилляция предсердий							
Новиков и др. (2019)	RF, дерево классифика ции, FNN из 30 узлов, FNN (3 слоя из 17 нейронов)*, DA*, SVP*, метод k- ближайших соседей	Прогнозирование исходов и осложнений у пациентов с ФП	Данные анамнеза, данные антропометрии, показатели анализов крови (25 ФР)	234, РФ	50 человек от исходной выборки	-	0,75- 0,76

Poh et al. (2018)	DL	Диагностика ФП	фотоплетизмографические импульсные сигналы	149,048 тыс., PPG database	3,039 тыс. независимая выборка	-	0,99
Мазаев и др. (2020)	CNN	Выявление скрытой ФП, сниженной фракции выброса левого желудочка и нарушения обмена калия	Данные ЭКГ	180,922 тыс., Клиника Майо США	52,870 тыс. и часть исходной выборки	-	0,85-0,96
Xu et al. (2019)	CNN 12 слоев	Диагностика ФП	Данные ЭКГ	MIT-BIH Atrial Fibrillation Database	5-fold cross validation	-	0,88
Сердечная недостаточность							
Blecker et al. (2018)	Logistic-linear regression, machine learning	Острая декомпенсированная сердечная недостаточность	демографические данные, результаты лабораторных исследований и тд.	37,229 тыс., NYU Langone Health Нью-Йорк, США	25% от исходной выборки	-	0,96-0,99
Golas et al. (2018)	LR, GB, CNN*	30-ти дневная повторная госпитализация у пациентов с СН	Записи электронных медицинских карт	11,510 тыс., Enterprise Data Warehouse (EDW) и Research Patient Data Repository (RPDR)	10-fold cross-validation	-	0,65-0,72
Frizzell et al (2017)	RF, GB*	30-ти дневная повторная госпитализация у пациентов с СН	Записи электронных медицинских карт	56,477 тыс. American Heart Association	30% от исходной выборки	Пациенты старше 65 лет	0,59-0,62
Kobayashi et al. (2022)	cluster analysis, decision tree	Прогноз частоты СН у бессимптомных пациентов	возраст, пол, рост, вес, статус курения и т.д.	827, STANISLAS cohort	1,792 тыс., Malmö Preventive Project cohort, 10-fold cross-validation	-	?

Au-Yeung et al. (2018)	RF*, SVM	5-ти секундное и 10-ти минутное предупреждение желудочковой тахикардии у пациентов с СН	Вариабельность сердечного ритма	11,764 тыс.	Кросс-валидация	-	0,81-0,87
Инсульт							
Xie et al. (2019)	GB	Восстановление в течение 90 дней после перенесенного инсульта	возраст, пол, исходные и 24-часовые показатели NIHSS, сторона инсульта и 90-дневный показатель mRS	512, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland	5-fold cross validation	-	0,74-0,78
Kurtz et al. (2022)	7 видов моделей машинного обучения (RF*)	Продолжительность пребывания в больнице и прогнозирование 30-дневной смертности при инсульте	Записи электронных медицинских карт	17,115 тыс. 43 Brazilian hospitals	Кросс-валидация	-	0,73-0,90
Alanazi et al. (2021)	NB, BayesNet, J48 (Java implementation of C4.5 algorithm, RF*)	Риск инсульта	Данные лабораторных исследований	15,714 тыс., National Health and Nutrition Examination Survey data sets	10-fold cross-validation	-	0,96-0,99

Lineback et al. (2021)	LR, NB, SVM, RF, GB*, extreme GB*	Прогноз осложнений после перенесенного инсульта в течение 30-ти дней	возраст, пол, раса, этническая принадлежность, статус страхования, статус брака, статус курения	2,305., Cerner, Kansas City, MO	5-fold cross validation	-	0,62 - 0,64
Wu et al. (2021)	RF	Риск ишемического и геморрагического инсульта у пациентов с концентрацией ЛПНП <70 мг/дл	Данные анамнеза, данные антропометрии, показатели анализов крови (20 предикторов)	9,327 тыс., Китай	1,753 тыс., независимая выборка	-	0,76 - 0,79
Ramos et al. (2019)	SVM, RF, logistic-linear regression, CNN	Развитие субарахноидального кровоизлияния	Данные компьютерной томографии	317	Кросс-валидация	-	0,68 - 0,74
Chun et al. (2021)	RF, logistic regression, SVM, GB tree*, FNN	Риск развития инсульта в течение 9-ти лет	143 фактора риска	503,842 тыс., Китай	12,75% от исходной выборки	-	0,83

Примечание. * - наиболее эффективный алгоритм; - модели, с наилучшими прогностическими возможностями ($AUC \geq 0,9$); GB-gradient busting; RF-random forest; FNN-feedforward neural network; NB-naïve Bayes; DL-deep learning; CNN-convolutional neural network; SVM-support vector machine, LR – logistics regression, DA – discriminant analysis.

В таблицах 1 и 2 представлены характеристики основных исследований, включенных в систематический обзор. Наилучшие предиктивные возможности в прогнозе ССС у здоровых людей среди алгоритмов ML продемонстрировали разработки отечественных авторов. В обеих работах AUC превышает 0,90. Из зарубежных работ наилучшие прогностические возможности продемонстрировал модель, разработанная Jamthikar и соавт. Исследователи использовали метод опорных векторов и такие классические ФР для обучения, как возраст, пол, липидограмму, вредные привычки и т.д. Несмотря на то, что AUC шкал пропорциональных рисков не превышал 0,88, данные инструменты прогнозы прошли внешнюю валидацию на крупных независимых выборках, что делает результаты из работы более достоверными.

Наиболее точные модели ML, прогнозирующие ССС у коморбидных пациентов использовали в своей структуре FNN, RF и GB, однако тестирование большинства из них проводилось с использованием кросс-валидации. В качестве предикторов также использовались традиционные ФР. Для моделей Cox-регрессии, прошедших внешнюю валидацию, AUC не превышало 0,70.

ИМ также лучше прогнозировали модели, разработанные отечественными учеными. Методы ML, продемонстрировавшие наилучших результат: RF, SVM и GB. В модель, как правило включались данные регистрации ЭКГ. Однако, также по сравнению с традиционными шкалами, единичные модели ML прошли внешнюю валидацию.

Среди всех моделей прогноза среднее количество предикторов превалировало у алгоритмов ML. Данный факт обусловлен тем, что многие алгоритмы были обучены на данных записей ЭМК, содержащих большое количество как структурированной, так и не структурированной информации. Также данный факт свидетельствует о том, что итоговый алгоритм оценки может быть представлен в качестве компьютерной программы и может быть внедрен в МИС. Это позволяет не сокращать число предикторов, для создания оценочных шкалах в виде таблиц. Также это то в значительной степени облегчает использование данной технологии, поскольку анализ ФР может быть произведен автоматически.

У большинства шкал Cox-регрессии, прогнозирующих ССС у коморбидных и бессимптомных пациентов, временной лаг прогноза составил 10 лет. У шкал, прогнозирующих исключительно ФП, ИМ, инсульт, СН, временной лаг прогноза варьируется от 30 суток, до 5 лет. У алгоритмов ML временной лаг прогноза часто отсутствовал поскольку данный тип анализа не требует цензурируемых переменных, благодаря чему алгоритмы

работают, как классификатор, выдавая значения процентной (0-100%) вероятности принадлежности элемента к той или иной группе (больной/здоровый).

Если сравнивать возрастные ограничения использования двух методов прогноза ССР, то значительных различий обнаружить не удалось. Минимальный возраст для прогноза среди алгоритмов Соx-регрессии составил 20 лет, максимальный возраст 90 лет. Для прогноза ССС у бессимптомных пациентов в среднем возрастной диапазон составлял 40-70 лет – возраст наиболее частого развития ССЗ [16]. Для алгоритмов ML часто возрастные ограничения не учитывались.

Обсуждение. В настоящей статье представлен первый систематический обзор, выполненный в РФ в котором наглядно продемонстрированы сравнительные характеристики двух методов прогноза ССС. Также данная работа является первой, где рассматривается эффективность алгоритмов, разработанных отечественными учеными.

В работе Швеца и соавт.[1] также производится сравнение традиционных методов прогноза ССР и методов ML. Средние показатели AUC для AI также превалировали над аналогичными значениями моделей Соx-регрессии. Однако авторы отмечают, что в отдельных случаях преимуществ в прогнозе у ML выявлено не было. Также авторы подчеркивают, что внедрение технологий AI в систему поддержки принятия врачебных решений (СППВР) и обработки электронных медицинских карт (ЭМК) может привести к серьезным ошибкам, связанными с порой неструктурированными и низкокачественными исходными данными. В итоге авторы сходятся во мнении, что значимых преимуществ в использовании ML выявлено не было.

В обзоре отечественных ученых Голубкова и соавт. [29] напротив продемонстрирована высокая эффективность работы алгоритмов ML. Представленные работы показывают точность диагностики и прогноза развития атеросклеротических заболеваний сердца и ишемической болезни сердца (ИБС) не ниже 90%, СН – не ниже 80%. Также в обзоре наглядно продемонстрирована эффективность работы ML в профилактике других заболеваний.

Chowdhury и соавт. был проведен обзор 52 работы прогноза ССР, в 34 из которых были использованы традиционные алгоритмы. Среднее число ФР, включенных в модели, не превышало 7. Данные обзора аналогичны данным нашего исследования. Средние показатели AUC для ML превосходили таковые у моделей Соx-регрессии (0,75 и 0,76 соответственно).

В работе Vazoukis и соавт. [34] авторы сравнивают предиктивные возможности традиционных методов прогноза сердечно-недостаточности (СН) и методов ML. В

исследование были включены 122 работы. Более, чем в 90% случаев использование методов AI позволило улучшить прогноз болезни и превзойти традиционные методы анализа данных. Полученные результаты косвенно подтверждает результаты нашего исследования. Средние показатели AUC для ML в прогнозе СН были в среднем на 0,04 выше таковых у шкал пропорциональных рисков, однако различия оказались недостоверными.

В работе Ванерjee и соавт. [35] также сравнивают предиктивные возможности ML и традиционных методов обработки данных в прогнозе СН, острого коронарного синдрома (ОКС) и ФП. Авторы рассмотрели 97 работ: 57 с использованием и 40 работ без использования AI. Исследователи заключают, что большинство исследований в данной сфере основаны на когорте людей, проживающих на территории США. Также авторы подчеркивают такие недостатки, как недостаточный объем выборки, отсутствие внешней валидации для сконструированных моделей ML. Данные недостатки также были наглядно продемонстрированы в нашем исследовании.

В систематическом обзоре, посвященном изучению использования ML в прогнозе осложнений после проведения реваскуляризации миокарда, авторы подчеркивают, что разработка и совершенствования технологий поддержки принятия врачебных решений является одной из наиболее перспективных сфер цифрового развития здравоохранения. Исследователи рассматривают возможность применения методов ML в прогнозировании рисков неблагоприятных ССС у больных с ИБС после реваскуляризации миокарда [24]. Средний показатель AUC для моделей ML составил 0,8. Авторы сходятся во мнении, что повышения точности прогностических моделей необходимо проводить валидацию на различных популяциях с учетом клинических течений ИБС и ресурсного обеспечения кардиологических служб. Для этого в первую очередь необходимо формирование и постоянное обновление региональных и федеральных электронных баз данных, хранящих информацию о больных с ИБС.

В работе Гусева и соавт. [26] также говорится, что использование ретроспективных данных с возможностью включения неограниченного числа предикторов в качестве субстрата для ML дает возможность лучше выявлять скрытые нелинейные закономерности и причинно-следственные связи между изучаемыми переменными. Данная особенность алгоритмов AI позволяет строить более качественные модели прогноза, что, несомненно, является ключевым в ранней диагностике и профилактике заболеваний. Результаты обзора демонстрируют, что использование алгоритмов ML позволяет строить прогностические модели более высокой

точности, по сравнению с традиционными шкалами оценки ССР. Также в аналогичных работах данного автора [30], продемонстрирована эффективность внедрения технологий ML в СППВР и систему ведения ЭМК.

В работе Baashar и соавт. [36] в анализ были включены 17 статей. Было продемонстрировано, что алгоритмы ML имели явное превосходство. GB хорошо зарекомендовал себя в диагностике и прогнозе СН. Средние показатели ROC-анализа составили 0,84 с точностью 91,1%, что в среднем на 4,4% было выше среди остальных моделей. Метод SVM продемонстрировал хорошие показатели в прогнозировании инсульта.

В большом метаанализе Krittanawong и соавт. [31] для оценки эффективности использования ML для прогноза различных ССС были собраны и проанализированы средние показатели AUC. В прогнозе атеросклеротических заболеваний средние значения составили 0,87 для CNN, 0,88 для GB и 0,93 для других алгоритмов. В прогнозе инсульта показатели AUC составили 0,90, 0,92 и 0,91 для CNN, SVM и GB соответственно. Наилучшая эффективность использования данных методов ML также была продемонстрирована среди работ, рассмотренных в нашем обзоре.

В обзоре, посвященном анализу методов прогноза инсульта с помощью ML, средний AUC составил 0,84, что также было в среднем выше, чем у традиционных шкал прогноза [32].

Исследовали Naseri Jahfari и соавт. [33] проанализировали 55 работ, посвященных использованию ML в качестве прогноза ССС. Авторы подчеркивают, что ни одна из представленных работ не была внедрена в клиническую практику. Также у представленных исследований отсутствовали 2 и 3 фаза клинических исследований, уровень технической готовности не превышал 8.

В работе Wang и соавт. [37] проанализированы сконструированные модели ML, оценивающие исходы инсульта, за период 1990 – 2019 гг., всего 18 работ. Смертельный исход являлся наиболее частой конечной точкой в анализируемых моделях. Наиболее часто используемые алгоритмы ML: RF, SVM, DT, и ANN. Средний объем выборок составил 475, среднее количество анализируемых предикторов – 22. Авторы сходятся во мнении для дальнейшего развития необходимы точные и качественные наборы данных, новые клинические исследования и больше протестированных моделей.

В работе Lee и соавт. [22] авторами проанализировано 102 работы, использующие алгоритмы AI для прогноза различных ССС. 26 работ были посвящены диагностике и прогнозу ФП. Средние показатели чувствительности и специфичности 94,8% и 96,9%

соответственно. Средние показатели AUC для DP – 0,98, CNN – 0,96. Также различия в точности прогнозов зависели от валидационной выборки. В нашем исследовании связи между методом валидации и точности выборки выявлено не было.

Таким образом в данном систематическом обзоре нам удалось доказать, что, несмотря на многие недостатки, прогностические возможности технологии ML превышают таковые у традиционных методов прогноза ССР. Об этом свидетельствуют как данные нашего исследования, так и результаты множества других аналогичных работ.

Выводы.

По данным проведенного систематического обзора и сравнения его результатов с результатами аналогичных исследований, можно сказать, что предиктивные возможности технологий ML превосходят традиционные методы прогноза ССР. Стоит отметить, что на сегодняшний день данные технологии наиболее активно развиваются на территории США и странах Европейского региона. Также большинство алгоритмов не прошли внешнюю валидацию на независимых выборках, что также не дает достоверного подтверждения эффективности использования моделей на других популяциях. Кроме того, ограничением их применения в реальной практике является необходимость полноценного электронного ведения медицинской документации и агрегация большего количества качественной и структурированной информации.

Однако несмотря на очевидные ограничения за технологиями оценки риска на основе ML будущее, поскольку они позволят оценивать риск на основании гораздо большего количества параметров по сравнению с традиционными шкалами и станут одним из компонентов персонализированной медицины.

Список литературы

1. Швец Д. А., Поветкин С. В. Сравнительный обзор использования методов машинного обучения для прогнозирования сердечно-сосудистого риска. Вестник новых медицинских технологий. Электронное издание. 2020;5:74-82. doi 10.24411/2075-4094-2020-16711
2. Roth GA, Forouzanfar MH, Moran AE, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. N Engl J Med. 2015 Apr 2;372(14):1333-41. doi: 10.1056/NEJMoa1406656.

3. Мишкин И. А., Сахаров А. А. Изучение использования алгоритмов машинного обучения в оценке риска развития сердечно-сосудистых заболеваний. *Инновации. Наука. Образование.* 2020;17:187-202.
4. WHO. About cardiovascular diseases. <https://www.who.int/home/cms-decommissioning>
5. 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.
6. Концевая А. В., Драпкина О. М., Баланова Ю. А. и др. Экономический ущерб сердечно-сосудистых заболеваний в Российской Федерации в 2016 году. *Рациональная фармакотерапия в кардиологии.* 2018;14(2):156-166. doi: 10.20996/1819-6446-201814-2-156-166
7. Weintraub WS. Reduce Acute Care Costs, and All Other Healthcare Costs Too. *J Am Heart Assoc.* 2019;8(8):e012604. doi:10.1161/JAHA.119.012604
8. Организация проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения. Методические рекомендации по практической реализации приказа Минздрава России от 13 марта 2019 г. N 124н "Об утверждении порядка проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения". М. 2019 - 165 с. Утверждены главным внештатным специалистом по терапии и общей врачебной практике Минздрава России О.М. Драпкиной и главным внештатным специалистом по медицинской профилактике Минздрава России Л.Ю. Дроздовой 22 октября 2019 г.
9. Беялов Ф.И. Прогнозирование и шкалы в медицине. 3-е изд., перераб. и доп. М.: МЕДпресс-информ, 2020. 248 с.: ил. ISBN 978-5-00030-761-8.
10. Беялов Ф.И. Прогнозирование заболеваний с помощью шкал. *Комплексные проблемы сердечно-сосудистых заболеваний.* 2018;7(1):84-93. doi: 10.17802/2306-1278-2018-7-1-84-93
11. 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.
12. Benjamins JW, Hendriks T, Knuuti J, et al. A primer in artificial intelligence in cardiovascular medicine. *Neth Heart J.* 2019;27(9):392-402. doi:10.1007/s12471-019-1286-6

13. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization [published correction appears in *Eur Heart J*. 2019 Oct 1;40(37):3096]. *Eur Heart J*. 2019;40(2):87-165. doi:10.1093/eurheartj/ehy394
14. Johnson KW, Shameer K, Glicksberg BS, et al. Enabling Precision Cardiology Through Multiscale Biology and Systems Medicine. *JACC Basic Transl Sci*. 2017;2(3):311-327. Published 2017 Jun 26. doi:10.1016/j.jacbts.2016.11.010
15. Shah RU, Rumsfeld JS. Big Data in Cardiology. *Eur Heart J*. 2017;38(24):1865-1867. doi:10.1093/eurheartj/ehx284
16. Гундаров И. А. Профилактическая медицина на рубеже веков: от факторов риска - к резервам здоровья и социальной профилактике. М.: Общество с ограниченной ответственностью Издательская группа "ГЭОТАР-Медиа". 2016. с. 256. ISBN 978-5-9704-3871-8.
17. Zhang Y, Guo SL, Han LN, et al. Application and Exploration of Big Data Mining in Clinical Medicine. *Chin Med J (Engl)*. 2016;129(6):731-738. doi:10.4103/0366-6999.178019
18. Brink H. Real-world machine learning. Publisher: Shelter Island, NY: Manning Publications Co. 2017. p. 264. ISBN 9781617291920.
19. Han H, Liu W. The coming era of artificial intelligence in biological data science. *BMC Bioinformatics*. 2019;20(Suppl 22):712. Published 2019 Dec 30. doi:10.1186/s12859-019-3225-3
20. Серeda А. П. Рекомендации по оформлению дизайна исследования // *Травматология и ортопедия России*. 2019;25(3):165-184. doi: 10.21823/2311-2905-2019-25-3-165-184
21. Корягина Н. А., Рямзина И. Н., Шапошникова А. И. и др. Основные факторы риска сердечно-сосудистых заболеваний у молодого работающего населения. *КВТиП*. 2013;3:40-42. doi.org/10.15829/1728-8800-2013-3-40-42
22. Lee S, Chu Y, Ryu J, et al. Artificial Intelligence for Detection of Cardiovascular-Related Diseases from Wearable Devices: A Systematic Review and Meta-Analysis. *Yonsei Med J*. 2022;63(Suppl):S93-S107. doi:10.3349/ymj.2022.63.S93
23. Gleeson S, Liao YW, Dugo C, et al. ECG-derived spatial QRS-T angle is associated with ICD implantation, mortality and heart failure admissions in patients with LV systolic dysfunction. *PLoS One*. 2017;12(3):e0171069. Published 2017 Mar 30. doi:10.1371/journal.pone.0171069
24. Гельцер Б. И., Рублев В. Ю., Циванюк М. М. И др. Машинное обучение в прогнозировании ближайших и отдаленных результатов реваскуляризации миокарда:

систематический обзор. Российский кардиологический журнал. 2021;26(8):115-124. doi:10.15829/1560-4071-2021-4505.

25. Roederer A, Holmes JH, Smith MJ, et al. Prediction of significant vasospasm in aneurysmal subarachnoid hemorrhage using automated data. *Neurocrit Care*. 2014;21(3):444-450. doi:10.1007/s12028-014-9976-9

26. 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;26(12):171-180. doi:10.15829/1560-4071-2021-4618

27. Мишкин И. А., Сахаров А. А. Изучение использования алгоритмов машинного обучения в оценке риска развития сердечно-сосудистых заболеваний. *Инновации. Наука. Образование*. 2020;17:187-202

28. Ward A, Sarraju A, Chung S, et al. Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. *NPJ Digit Med*. 2020;3:125. Published 2020 Sep 23. doi:10.1038/s41746-020-00331-1

29. Голубков А. В., Гаврилова М. П. Применение искусственных нейронных сетей в профилактической и клинической медицине (научный обзор). *Профилактическая и клиническая медицина*. 2020;4(77):30-39.

30. 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.

31. Krittanawong C, Virk HUH, Bangalore S, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep*. 2020;10(1):16057. Published 2020 Sep 29. doi:10.1038/s41598-020-72685-1

32. Mainali S, Darsie ME, Smetana KS. Machine Learning in Action: Stroke Diagnosis and Outcome Prediction. *Front Neurol*. 2021;12:734345. Published 2021 Dec 6. doi:10.3389/fneur.2021.734345

33. Naseri Jahfari A, Tax D, Reinders M, et al. Machine Learning for Cardiovascular Outcomes From Wearable Data: Systematic Review From a Technology Readiness Level Point of View. *JMIR Med Inform*. 2022;10(1):e29434. Published 2022 Jan 19. doi:10.2196/29434

34. Bazoukis G, Stavrakis S, Zhou J, et al. Machine learning versus conventional clinical methods in guiding management of heart failure patients-a systematic review. *Heart Fail Rev.* 2021;26(1):23-34. doi:10.1007/s10741-020-10007-3
35. Banerjee A, Chen S, Fatemifar G, et al. Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility. *BMC Med.* 2021;19(1):85. Published 2021 Apr 6. doi:10.1186/s12916-021-01940-7
36. Baashar Y, Alkawsy G, Alhussian H, et al. Effectiveness of Artificial Intelligence Models for Cardiovascular Disease Prediction: Network Meta-Analysis. *Comput Intell Neurosci.* 2022;2022:5849995. Published 2022 Feb 24. doi:10.1155/2022/5849995
37. Wang W, Kiik M, Peek N, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS One.* 2020;15(6):e0234722. Published 2020 Jun 12. doi:10.1371/journal.pone.0234722

References

1. Shvets D.A., Povetkin S.V. Sravnitel'nyj obzor ispol'zovaniya metodov mashinnogo obucheniya dlya prognozirovaniya serdechno-sosudistogo riska. [A comparative review of the use of machine learning methods for predicting cardiovascular risk]. *Vestnik novyh medicinskih tekhnologij. Elektronnoe izdanie.* [Bulletin of New Medical Technologies. Electronic edition]. 2020;14(5):74-82. (in Russian) doi 10.24411/2075-4094-2020-16711.
2. Roth GA, Forouzanfar MH, Moran AE, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. *N Engl J Med.* 2015 Apr 2;372(14):1333-41. doi: 10.1056/NEJMoa1406656.
3. Mishkin I.A., Sakharov A.A. Izuchenie ispol'zovaniya algoritmov mashinnogo obucheniya v ocenke riska razvitiya serdechno-sosudistykh zabolevanij. *Innovacii* [The study of the use of machine learning algorithms in assessing the risk of developing cardiovascular diseases]. *Nauka. Obrazovanie* [Innovation. The science. Education]. 2020;17:187-202. (in Russian)
4. WHO. About cardiovascular diseases. <https://www.who.int/home/cms-decommissioning>
5. Gusev A.V., Kuznetsova T.Yu., Korsakov I.N. Iskusstvennyj intellekt v ocenke riskov razvitiya serdechno-sosudistykh zabolevanij [Artificial intelligence in assessing the risks of developing cardiovascular diseases]. *ZHurnal telemeditsiny i elektronnoho zdravoohraneniya* [Journal of Telemedicine and E-Health]. 2018;3(8):85-90. (in Russian)

6. Kontsevaya A.V., Drapkina O.M., Balanova Y.uA. et al. Ekonomicheskij ushcherb serdechno-sosudistyh zabolevanij v Rossijskoj Federacii v 2016 godu [Economic damage of cardiovascular diseases in the Russian Federation in 2016]. Racional'naya farmakoterapiya v kardiologii [Rational pharmacotherapy in cardiology]. 2018;14(2):156-166. (in Russian) doi: 10.20996/1819-6446-201814-2-156-166
7. Weintraub WS. Reduce Acute Care Costs, and All Other Healthcare Costs Too. J Am Heart Assoc. 2019;8(8):e012604. doi:10.1161/JAHA.119.012604
8. Organizaciya provedeniya profilakticheskogo medicinskogo osmotra i dispanserizacii opredelennyh grupp vzroslogo naseleniya. Metodicheskie rekomendacii po prakticheskoj realizacii prikaza Minzdrava Rossii ot 13 marta 2019 g. N 124n "Ob utverzhdenii poryadka provedeniya profilakticheskogo medicinskogo osmotra i dispanserizacii opredelennyh grupp vzroslogo naseleniya". M. 2019 - 165 s. Utverzhdeny glavnym vneshtatnym specialistom po terapii i obshchej vrachebnoj praktike Minzdrava Rossii O.M. Drapkinoj i glavnym vneshtatnym specialistom po medicinskoj profilaktike Minzdrava Rossii L.YU. Drozdova October 22, 2019 [Organization of preventive medical examination and medical examination of certain groups of the adult population. Methodological recommendations for the practical implementation of the order of the Ministry of Health of the Russian Federation dated March 13, 2019 N 124n "On approval of the procedure for preventive medical examination and medical examination of certain groups of the adult population". M. 2019 - 165 p. Approved by the chief freelance specialist in therapy and General Medical practice of the Ministry of Health of Russia O.M. Drapkina and the chief freelance specialist in medical prevention of the Ministry of Health of Russia L.Y. Drozdova]. (in Russian)
9. Belyalov F.I. Prognozirovanie i shkaly v medicine. 3-e izd., pererab. i dop. M.: MEDpress-inform, 2020. p.248. [Forecasting and scales in medicine. 3rd ed., reprint. and additional M.: MEDpress-inform] (in Russian) ISBN 978-5-00030-761-8
10. Belyalov F.I. Prognozirovanie zabolevanij s pomoshch'yu shkal [Disease prediction using scales]. Kompleksnye problemy serdechno-sosudistyh zabolevanij [Complex problems of cardiovascular diseases]. 2018;7(1):84-93. (in Russian)
11. 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)

12. Benjamins JW, Hendriks T, Knuuti J, et al. A primer in artificial intelligence in cardiovascular medicine. *Neth Heart J*. 2019;27(9):392-402. doi:10.1007/s12471-019-1286-6
13. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization [published correction appears in *Eur Heart J*. 2019 Oct 1;40(37):3096]. *Eur Heart J*. 2019;40(2):87-165. doi:10.1093/eurheartj/ehy394
14. Johnson KW, Shameer K, Glicksberg BS, et al. Enabling Precision Cardiology Through Multiscale Biology and Systems Medicine. *JACC Basic Transl Sci*. 2017;2(3):311-327. Published 2017 Jun 26. doi:10.1016/j.jacbts.2016.11.010
15. Shah RU, Rumsfeld JS. Big Data in Cardiology. *Eur Heart J*. 2017;38(24):1865-1867. doi:10.1093/eurheartj/ehx284
16. Gundarov I.A. Profilakticheskaya medicina na rubezhe vekov: ot faktorov riska - k rezervam zdorov'ya i social'noj profilaktike. M.: Obshchestvo s ogranichennoj otvetstvennost'yu Izdatel'skaya gruppa "GEOTAR-Media" [Preventive medicine at the turn of the century: from risk factors to health reserves and social prevention. Moscow: Limited Liability Company Publishing Group "GEOTAR-Media"]. 2016. p. 256. (in Russian)
17. Zhang Y, Guo SL, Han LN, et al. Application and Exploration of Big Data Mining in Clinical Medicine. *Chin Med J (Engl)*. 2016;129(6):731-738. doi:10.4103/0366-6999.178019
18. Brink H. Real-world machine learning. Publisher: Shelter Island, NY: Manning Publications Co. 2017. p. 264. ISBN 9781617291920.
19. Han H, Liu W. The coming era of artificial intelligence in biological data science. *BMC Bioinformatics*. 2019;20(Suppl 22):712. Published 2019 Dec 30. doi:10.1186/s12859-019-3225-3
20. Sereda A.P. Rekomendacii po oformleniyu dizajna issledovaniya [Recommendations for the design of the study]. *Travmatologiya i ortopediya Rossii* [Traumatology and orthopedics of Russia]. 2019;25(3):165-184. (in Russian) doi: 10.21823/2311-2905-2019-25-3-165-184
21. Koryagina N.A., Ryamzina I.N., Shaposhnikova A.I., et al. Osnovnye faktory riska serdechno-sosudistyh zabolevanij u molodogo rabotayushchego naseleniya [The main risk factors for cardiovascular diseases in the young working population]. *KVTiP [CVTandP]*. 2013;3:40-42. (in Russian) doi.org/10.15829/1728-8800-2013-3-40-42
22. Lee S, Chu Y, Ryu J, et al. Artificial Intelligence for Detection of Cardiovascular-Related Diseases from Wearable Devices: A Systematic Review and Meta-Analysis. *Yonsei Med J*. 2022;63(Suppl):S93-S107. doi:10.3349/ymj.2022.63.S93

23. Gleeson S, Liao YW, Dugo C, et al. ECG-derived spatial QRS-T angle is associated with ICD implantation, mortality and heart failure admissions in patients with LV systolic dysfunction. *PLoS One*. 2017;12(3):e0171069. Published 2017 Mar 30. doi:10.1371/journal.pone.0171069
24. Geltser B.I., Rublev V.Yu., Tsivanyuk M.M., et al. Mashinnoe obuchenie v prognozirovanii blizhajshih i otdalennyh rezul'tatov revaskulyarizacii miokarda: sistematiceskij obzor [Machine learning in predicting the immediate and long-term results of myocardial revascularization: a systematic review]. *Rossijskij kardiologicheskij zhurnal [Russian Journal of Cardiology]*. 2021;26(8):115-124 (in Russian) doi:10.15829/1560-4071-2021-4505
25. Roederer A, Holmes JH, Smith MJ, et al. Prediction of significant vasospasm in aneurysmal subarachnoid hemorrhage using automated data. *Neurocrit Care*. 2014;21(3):444-450. doi:10.1007/s12028-014-9976-9
26. 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
27. Mishkin I.A., Sakharov A.A. Izuchenie ispol'zovaniya algoritmov mashinnogo obucheniya v ocenke riska razvitiya serdechno-sosudistykh zabolevanij [The study of the use of machine learning algorithms in assessing the risk of developing cardiovascular diseases]. *Innovacii. Nauka. Obrazovanie [Innovation. The science. Education]*. 2020;17:187-202. (in Russian)
28. Ward A, Sarraju A, Chung S, et al. Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. *NPJ Digit Med*. 2020;3:125. Published 2020 Sep 23. doi:10.1038/s41746-020-00331-1
29. Golubkov A.V., Gavrilova M.P. Primenenie iskusstvennykh nejronnykh setej v profilakticheskoy i klinicheskoy medicine (nauchnyj obzor) [Application of artificial neural networks in preventive and clinical medicine (scientific review)]. *Profilakticheskaya i klinicheskaya medicina [Preventive and clinical medicine]*. 2020;4(77):30-39. (in Russian)
30. Gusev A.V., Kuznetsova T.Yu., Korsakov I.N. Iskusstvennyj intellekt v ocenke riskov razvitiya serdechno-sosudistykh zabolevanij [Artificial intelligence in assessing the risks of developing cardiovascular diseases]. *ZHurnal telemeditsiny i elektronnoho zdavoohraneniya [Journal of Telemedicine and E-Health]*. 2018;3(8):85-90. (in Russian)

31. Krittanawong C, Virk HUH, Bangalore S, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep.* 2020;10(1):16057. Published 2020 Sep 29. doi:10.1038/s41598-020-72685-1
32. Mainali S, Darsie ME, Smetana KS. Machine Learning in Action: Stroke Diagnosis and Outcome Prediction. *Front Neurol.* 2021;12:734345. Published 2021 Dec 6. doi:10.3389/fneur.2021.734345
33. Naseri Jahfari A, Tax D, Reinders M, et al. Machine Learning for Cardiovascular Outcomes From Wearable Data: Systematic Review From a Technology Readiness Level Point of View. *JMIR Med Inform.* 2022;10(1):e29434. Published 2022 Jan 19. doi:10.2196/29434
34. Bazoukis G, Stavrakis S, Zhou J, et al. Machine learning versus conventional clinical methods in guiding management of heart failure patients-a systematic review. *Heart Fail Rev.* 2021;26(1):23-34. doi:10.1007/s10741-020-10007-3
35. Banerjee A, Chen S, Fatemifar G, et al. Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility. *BMC Med.* 2021;19(1):85. Published 2021 Apr 6. doi:10.1186/s12916-021-01940-7
36. Baashar Y, Alkaws G, Alhussian H, et al. Effectiveness of Artificial Intelligence Models for Cardiovascular Disease Prediction: Network Meta-Analysis. *Comput Intell Neurosci.* 2022;2022:5849995. Published 2022 Feb 24. doi:10.1155/2022/5849995
37. Wang W, Kiik M, Peek N, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS One.* 2020;15(6):e0234722. Published 2020 Jun 12. doi:10.1371/journal.pone.0234722

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: Ilya.mischkin@yandex.ru, ORCID 0000-0003-4850-0648; SPIN: 4086-1507

Концевая Анна Васильевна - д.м.н., 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 - к.т.н. 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

Драпкина Оксана Михайловна - д.м.н. директор ФГБУ "Национального медицинского исследовательского центра терапии и профилактической медицины" Министерства здравоохранения Российской Федерации, профессор, академик РАН, главный внештатный специалист по терапии и общей врачебной практике Минздрава России, заслуженный врач РФ, 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, ORCID 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

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: 04.03.2023 г.
Принята к публикации: 28.06.2023 г.