



Разработка предиктивных моделей для дифференциальной диагностики гипертрофической кардиомиопатии

Зайцев В.В.¹, Сафронов К.С.², Конасов К.С.¹, Бавшин Т.Р.³, Манохин К.А.³, Образцова Л.А.¹, Моисеева О.М.¹

Цель. Разработка предиктивных моделей для дифференциальной диагностики гипертрофического фенотипа у пациентов с сопутствующими заболеваниями, а также их валидация посредством независимой оценки.

Материал и методы. В исследование был включен анализ 1169 медицинских карт из медицинской информационной системы пациентов с выраженной гипертрофией миокарда и предварительным диагнозом гипертрофической кардиомиопатии (ГКМП) (I42.1, I42.2). Пациенты были разделены на 3 группы: пациенты с вероятным диагнозом ГКМП, пациенты с умеренной гипертрофией миокарда (>15 мм) вследствие известного заболевания, а также пациенты с выраженной гипертрофией миокарда, которую сложно объяснить исключительно перегрузкой давлением левого желудочка ("серая зона"). В исходном наборе данных представлено 74 параметра. Построены и оптимизированы модели машинного обучения следующих классов: логистическая регрессия (LR), метод опорных векторов (SVM), дерево принятия решений (DT) и градиентный бустинг на деревьях решений.

Результаты. Все модели обладают достаточной точностью выявления ГКМП, однако точность исключения диагноза довольно низкая. Применение модели машинного обучения с использованием логистической регрессии позволило снизить риск ошибочной диагностики ГКМП в группе сомнительного диагноза до 31%.

Заключение. Разработаны 4 предиктивные модели для дифференциального диагноза при выраженной гипертрофии миокарда левого желудочка с целью улучшения диагностики ГКМП. По результатам валидации слепым методом оптимальной моделью для клинической практики определена логистическая регрессия.

Ключевые слова: гипертрофическая кардиомиопатия, диагностика, машинное обучение.

Отношения и деятельность: нет.

¹ФГБУ Национальный медицинский исследовательский центр им. В.А. Алмазова Минздрава России, Санкт-Петербург; ²ФГБОУ ВО Санкт-Петербургский государственный морской технический университет, Санкт-Петербург; ³ФГАУ ВО Национальный исследовательский университет ИТМО, Санкт-Петербург, Россия.

Зайцев В.В.* — ассистент кафедры кардиологии факультета послевузовского и дополнительного образования Института медицинского образования, ORCID: 0000-0003-1905-2575, Сафронов К.С. — старший преподаватель кафедры прикладной математики и математического моделирования, ORCID: 0000-0003-0879-1225, Конасов К.С. — врач-кардиолог, аспирант кафедры кардиологии факультета послевузовского и дополнительного образования Института медицинского образования, ORCID: 0009-0002-0159-7251, Бавшин Т.Р. — студент, ORCID: 0009-0002-8295-8564, Манохин К.А. — студент, ORCID: 0009-0000-2086-1826, Образцова Л.А. — ординатор кафедры кардиологии факультета послевузовского и дополнительного образования Института медицинского образования, ORCID: 0000-0001-5283-4996, Моисеева О.М. — д.м.н., профессор, г.н.с., руководитель научно-исследовательского отдела некоронарогенных заболеваний сердца Института сердца и сосудов, ORCID: 0000-0002-7817-3847.

*Автор, ответственный за переписку (Corresponding author):

vzaytsev92@yandex.ru

АГ — артериальная гипертензия, ВСС — внезапная сердечная смерть, ВТЛЖ — выносящий тракт левого желудочка, ГКМП — гипертрофическая кардиомиопатия, ЛЖ — левый желудочек, ЛП — левое предсердие, ПЖ — правый желудочек, ДТ — дерево принятия решений, LR — логистическая регрессия, SVM — метод опорных векторов.

Рукопись получена 17.09.2024

Рецензия получена 08.10.2024

Принята к публикации 28.10.2024



Для цитирования: Зайцев В.В., Сафронов К.С., Конасов К.С., Бавшин Т.Р., Манохин К.А., Образцова Л.А., Моисеева О.М. Разработка предиктивных моделей для дифференциальной диагностики гипертрофической кардиомиопатии. *Российский кардиологический журнал*. 2024;29(11):6130. doi: 10.15829/1560-4071-2024-6130. EDN VTNAQV

Development of predictive models for differential diagnosis of hypertrophic cardiomyopathy

Zaitsev V.V.¹, Safronov K.S.², Konasov K.S.¹, Bavshin T.R.³, Manokhin K.A.³, Obratsova L.A.¹, Moiseeva O.M.¹

Aim. To develop predictive models for differential diagnostics of hypertrophic phenotype in patients with concomitant diseases, as well as to validate through independent assessment.

Material and methods. The study included an analysis of 1169 medical records from the medical information system of patients with severe myocardial hypertrophy and a preliminary diagnosis of hypertrophic cardiomyopathy (HCM) (I42.1, I42.2). The patients were divided into 3 following groups: patients with a probable HCM, patients with moderate myocardial hypertrophy (>15 mm) due to a known disease, and patients with severe myocardial hypertrophy that cannot be explained by left ventricular pressure overload ("gray zone"). The original dataset contains 74 parameters. Machine learning models of the following classes were created and optimized: logistic regression, support vector machine, decision tree, and gradient boosting decision trees.

Results. All models have sufficient accuracy in detecting HCM, but the accuracy of ruling out the diagnosis is quite low. The use of a machine learning model using logistic regression reduced the HCM misdiagnosis risk in the group of questionable diagnosis to 31%.

Conclusion. Four predictive models for differential diagnosis of severe left ventricular myocardial hypertrophy were developed to improve the HCM diagnosis. Blinding validation showed that logistic regression is the most optimal model for clinical practice.

Keywords: hypertrophic cardiomyopathy, diagnostics, machine learning.

Relationships and Activities: none.

¹Almazov National Medical Research Center, St. Petersburg; ²Saint Petersburg State Marine Technical University, St. Petersburg; ³National Research University of Information Technologies, Mechanics and Optics, St. Petersburg, Russia.

Zaitsev V.V.* ORCID: 0000-0003-1905-2575, Safronov K.S. ORCID: 0000-0003-0879-1225, Konasov K.S. ORCID: 0009-0002-0159-7251, Bavshin T.R. ORCID: 0009-0002-8295-8564, Manokhin K.A. ORCID: 0009-0000-2086-1826, Obratsova L.A. ORCID: 0000-0001-5283-4996, Moiseeva O.M. ORCID: 0000-0002-7817-3847.

*Corresponding author:
vzaytsev92@yandex.ru

Received: 17.09.2024 Revision Received: 08.10.2024 Accepted: 28.10.2024

For citation: Zaitsev V.V., Safronov K.S., Konasov K.S., Bavshin T.R., Manokhin K.A., Obratsova L.A., Moiseeva O.M. Development of predictive models for differential diagnosis of hypertrophic cardiomyopathy. *Russian Journal of Cardiology*. 2024;29(11):6130. doi: 10.15829/1560-4071-2024-6130. EDN VTNAQV

Ключевые моменты

- Использование методов машинного обучения позволяет улучшить точность диагностики гипертрофической кардиомиопатии, что обеспечит персонализированный подход к выбору тактики ведения пациента.
- Дальнейшее добавление датасетов с информацией о сопутствующих заболеваниях и применяемых методах лечения позволит существенно повысить производительность предлагаемых моделей.

Key messages

- Machine learning methods improve the accuracy of hypertrophic cardiomyopathy diagnosis, which will ensure a personalized approach to choosing patient management tactics.
- Further addition of datasets with information on concomitant diseases and treatment methods will significantly improve the performance of the proposed models.

Гипертрофическая кардиомиопатия (ГКМП) — одно из распространенных наследственных заболеваний, характеризующееся выраженной гипертрофией миокарда желудочков в отсутствие длительной перегрузки давлением и других заболеваний, протекающих с формированием гипертрофического фенотипа. Распространенность ГКМП в зависимости от фенотипа/генотипа составляет 1:200–1:500 в общей популяции [1]. Однако клиническая манифестация заболевания не превышает 1:3000 в общей популяции, что отражает истинное бремя данной патологии. Для ГКМП характерна аутосомно-доминантная природа наследования и распространенность вариантов с неполной пенетрантностью. Возможно вовлечение правого (ПЖ) и левого желудочка (ЛЖ), наличие обструкции в выносящем тракте ЛЖ (ВТЛЖ), внутрижелудочковой обструкции, изолированная апикальная гипертрофия. Многообразие морфологических вариантов гипертрофии с различным вовлечением камер предопределяет своеобразность течения заболевания.

Наиболее часто клинические проявления ГКМП проявляются после достижения возраста 18 лет, когда складывается гипертрофический фенотип. Несмотря на низкую специфичность симптомов, диагностика ГКМП у лиц молодого возраста не представляет значительных трудностей. Отсутствие клинических признаков фенотипов, значимой сопутствующей патологии и врожденных пороков сердца позволяют интерпретировать гипертрофию миокарда как проявление первичной кардиомиопатии. С возрастом увеличивается бремя соматической патологии, растет распространенность артериальной гипертензии (АГ) (74,5% для лиц старше 60 лет [2]),

дегенеративного аортального стеноза, что затрудняет верификацию диагноза ГКМП. При диагностике часто недооценивается длительность и тяжесть течения АГ, приверженность к приему терапии, динамика выраженности гипертрофии по ретроспективным данным. Среди пациентов, отобранных для септальной миоэктомии, в 11% случаев морфологические изменения соответствовали гипертонической болезни [3]. При длительном наблюдении АГ выявляется у 40–70% пациентов с уже подтвержденным диагнозом ГКМП [4]. С учетом неспецифичности клинических проявлений, гипердиагностики заболевания в старшей возрастной группе верификация диагноза ГКМП крайне затруднена. Ситуация еще больше осложняется тем, что ГКМП у возрастных пациентов редко ассоциирована с патологическими вариантами в генах, кодирующих саркомерные белки. Можно предположить, что активное внедрение технологии машинного обучения позволит повысить эффективность раннего выявления ГКМП в реальной клинической практике.

Цель исследования — разработка предиктивных моделей для дифференциальной диагностики гипертрофического фенотипа у пациентов с сопутствующими заболеваниями, а также их валидация посредством независимой оценки.

Материал и методы

В исследовании проведен анализ 1169 медицинских карт из медицинской информационной системы пациентов, обращавшихся за медицинской помощью с 2010 по 2022 г. с диагнозом ГКМП (I42.1, I42.2). Все пациенты имели выраженную гипертрофию миокарда ЛЖ, что потребовало проведения дифференциального диагноза для исключения ГКМП. Из анализа исключены пациенты с подтвержденными

Таблица 1

Общие характеристики групп пациентов

Параметры	Пациент с ГКМП (n=577)	Пациенты без ГКМП (n=134)	p
Возраст, m±sd	53,4 (±13,6)	51,6 (±15,3)	0,110
Мужской пол, n (%)	290 (50%)	60 (45%)	0,260
Семейный анамнез ГКМП, n (%)	36 (7%)	3 (3%)	0,101
SAM-феномен, n (%)	275 (60%)	11 (11%)	<0,001
Митральная недостаточность, n (%)	423 (79%)	80 (62%)	<0,001
Аортальная недостаточность, n (%)	79 (15%)	3 (2%)	0,001
QRS >120 мсек, n (%)	61 (22%)	6 (12%)	0,121
Фибрилляция предсердий, n (%)	130 (23%)	25 (19%)	0,293
Сахарный диабет 2 типа, n (%)	53 (9%)	13 (10%)	0,250
Желудочковые тахикардии, n (%)	115 (23%)	20 (19%)	0,324
Обструкция в выносящем тракте ЛЖ, n (%)	259 (46%)	5 (4%)	<0,001
ИБС, n (%)	184 (33%)	40 (30%)	0,246
Повышение АД в анамнезе, n (%)	385 (69%)	88 (70%)	0,579
Наличие диастолической дисфункции, n (%)	345 (81%)	68 (64%)	0,002
Наличие АВ-блокады, n (%)	86 (20%)	13 (14%)	0,018

Сокращения: АВ — атриовентрикулярный, АД — артериальное давление, ГКМП — гипертрофическая кардиомиопатия, ИБС — ишемическая болезнь сердца, ЛЖ — левый желудочек, SAM — систолическое движение передней митральной створки.

фенокопиями на момент первичного обращения, беременные, лица моложе 18 лет. У всех пациентов проводилась оценка клинических и инструментальных данных, полученных при первичном обращении пациента с неуточненной природой гипертрофии миокарда. В каждом случае были проанализированы следующие данные: жалобы, наличие сопутствующих заболеваний, наследственный анамнез, данные трансторакальной эхокардиографии, суточного мониторингирования электрокардиограммы, магнитно-резонансной томографии сердца с контрастным усилением, лабораторные показатели, а также сопутствующая медикаментозная терапия. На основании полученной информации пациенты были разделены по критериям клинических рекомендаций на 3 группы: пациенты с вероятным диагнозом ГКМП, пациенты с умеренной гипертрофией миокарда (<15 мм) вследствие известного заболевания, а также пациенты с выраженной гипертрофией миокарда, которую сложно объяснить исключительно перегрузкой давлением ЛЖ ("серая зона"). Клиническая характеристика общей выборки представлена в таблице 1.

В исходном наборе данных представлено 74 параметра. Часть признаков, таких как жалобы, объем сопутствующей терапии, содержит значительное количество пропущенных значений (порядка 40-80%), в связи с чем данные были исключены, затем был сформирован итоговый набор данных, содержащий 30 параметров. Полученный набор использовался для обучения нескольких моделей с различной архитектурой с целью найти наилучшую модель для оценки значимости признаков. Построены и оптимизированы модели машинного обучения следую-

щих классов: логистическая регрессия (LR), метод опорных векторов (SVM), дерево принятия решений (DT) и градиентный бустинг на деревьях решений.

Логистическая регрессия является простым и эффективным методом классификации, легко интерпретируемым. Она хорошо работает с линейно разделимыми данными, обладает небольшим количеством гиперпараметров для настройки. Однако она может плохо справляться с нелинейными данными и мультиклассовыми задачами.

Метод опорных векторов хорошо работает с данными, которые не являются линейно разделимыми, благодаря возможности использования ядерных функций для преобразования данных в пространство более высокой размерности. SVM обладает хорошей обобщающей способностью и эффективен в задачах с большим количеством признаков, но может быть чувствителен к выбору гиперпараметров и требует более высоких вычислительных ресурсов.

Деревья решений предоставляют простой способ визуализации принятия решений и понимания важности признаков. Они могут обрабатывать как категориальные, так и числовые данные, не требуют масштабирования признаков и заполнения пропущенных значений. Однако деревья решений склонны к переобучению, особенно на больших данных, и могут быть нестабильными при небольших изменениях в данных. Градиентный бустинг объединяет преимущества слабых моделей, создавая сильную ансамблевую модель. Он обладает высокой точностью, хорошо работает с разнородными данными и способен улавливать сложные взаимосвязи между признаками. Однако градиентный бустинг может быть склонен

Таблица 2

Эхокардиографические показатели пациентов

Параметры	Группа с ГКМП (n=577)	Группа без ГКМП (n=134)	p
Максимальная толщина стенки ЛЖ, мм	22 (20; 24)	12 (11; 14)	<0,001
Переднезадний размер ЛП, мм	46 (43; 51)	42 (37,5; 47)	<0,001
Индексированный объем ЛП, мл/м ²	46 (37; 59)	37,5 (30,75; 50,75)	<0,001
КДР ЛЖ, мм	45 (41; 49)	47 (44; 51)	<0,001
КСР ЛЖ, мм	26 (24; 29)	29 (26; 31)	<0,001
КДО ЛЖ, мл	94 (77,25; 112,75)	99,5 (85,75; 118,25)	0,014
КСО ЛЖ, мл	28 (22; 37)	32 (25; 38)	0,005
Фракция выброса (Симпсон), %	65 (60; 70)	65 (59; 68)	0,061
РСДЛА, мм рт.ст.	35 (30; 41)	33 (28; 37)	0,005
Размер ПЖ (4-камерная позиция), мм	31 (28; 35)	30 (28; 34)	0,085
Размер ПЖ	28 (25; 30)	28 (25; 31,75)	0,186
Толщина стенки ПЖ ≥5, n (%)	175 (56%)	30 (36%)	0,002

Сокращения: ГКМП — гипертрофическая кардиомиопатия, КДО — конечный диастолический объем, КДР — конечный диастолический размер, КСО — конечный систолический объем, КСР — конечный систолический размер, ЛЖ — левый желудочек, ЛП — левое предсердие, ПЖ — правый желудочек, РСДЛА — расчетное систолическое давление в легочной артерии.

к переобучению, требует тщательной настройки гиперпараметров и времени для обучения [5–7].

Программное обеспечение разрабатывалось на персональных компьютерах с операционной системой Windows 11 в среде коллаборативной разработки Google Colab, позволяющей удаленно использовать вычислительные ресурсы на графических ускорителях для значительного увеличения скорости обучения моделей. Непосредственно разработка проводилась на языке Python 3.10.12 с использованием библиотек: NumPy 1.25.2, Pandas 1.5.3, Seaborn 0.13.1 и Matplotlib 3.7.1. Для реализации двух линейных моделей — LR и SVM, — а также для DT использовались классификаторы из библиотеки Scikit-Learn 1.4.1. Для реализации модели градиентного бустинга была выбрана библиотека Yandex Catboost.

В каждом классе моделей выбрана наилучшая с использованием пятиблочной кросс-валидации со стратификацией целевой переменной. Гиперпараметры моделей подбирались эмпирически с использованием метода поиска по сетке (GridSearch) таким образом, чтобы добиться наилучшей чувствительности и специфичности, избежав при этом переобучения моделей.

В моделях логистической регрессии оптимизация проводилась по параметрам: тип решателя, регуляризация (вид и вес), количество итераций и погрешность.

Для метода опорных векторов оценивались следующие параметры: тип ядра, регуляризация (вид и вес), количество итераций и погрешность.

В случае дерева решений оптимизировались параметры: критерий качества разбиения, глубина, баланс классов (class weight), минимальное количество примесей (min impurity), минимальный размер листа (min samples leaf), минимальный размер узла для разделения (min samples split).

Оцениваемые параметры для градиентного бустинга: количество итераций, коэффициент скорости обучения, баланс классов, регуляризация, а также индикатор переобучения.

После построения моделей и выбора наилучшей модели в каждом классе для каждой модели проводилась оценка значимости факторов с помощью векторов Шепли [8] и были отобраны наиболее важные признаки, влияющие на процесс постановки диагноза. Для оценки значимости признаков использовалась библиотека SHAP 0.42.1.

Результаты

При сравнении пациентов с ГКМП и пациентов с вторичной гипертрофией ЛЖ не было получено значимых отличий в отношении возраста, пола и сопутствующей патологии (табл. 1). Распространенность сахарного диабета, ишемической болезни сердца, фибрилляции предсердий и желудочковых нарушений ритма была сопоставима в обеих группах. Пациенты чаще были бессимптомные, частота и характер жалоб не различались, как и анамнез АГ. Повышение артериального давления в анамнезе трактовалось как наличие АГ. Отсутствие данных о степени тяжести и длительности АГ не позволяло объективно оценить её вклад в формирование гипертрофического фенотипа у пациентов с ГКМП.

Пациенты с ГКМП характеризовались большим переднезадним размером и индексированным объемом левого предсердия (ЛП), меньшим размером и объемом полости ЛЖ, более высокими значениями расчетного систолического давления в легочной артерии как проявлением диастолической дисфункции ЛЖ и формирования митральной недостаточности. Для пациентов с ГКМП характерно более частое вовлечение ПЖ. Полученные данные отражают со-

Таблица 3

Метрики моделей машинного обучения

Модель	Общая точность, %	Точность предсказания ГКМП, %	Точность предсказания не ГКМП, %	Чувствительность, %	Специфичность, %
Дерево решений	74	89	38	77	60
Логистическая регрессия	75	94	41	74	80
Метод опорных векторов	73	95	40	71	82
Градиентный бустинг	79	88	44	86	50

Сокращение: ГКМП — гипертрофическая кардиомиопатия.

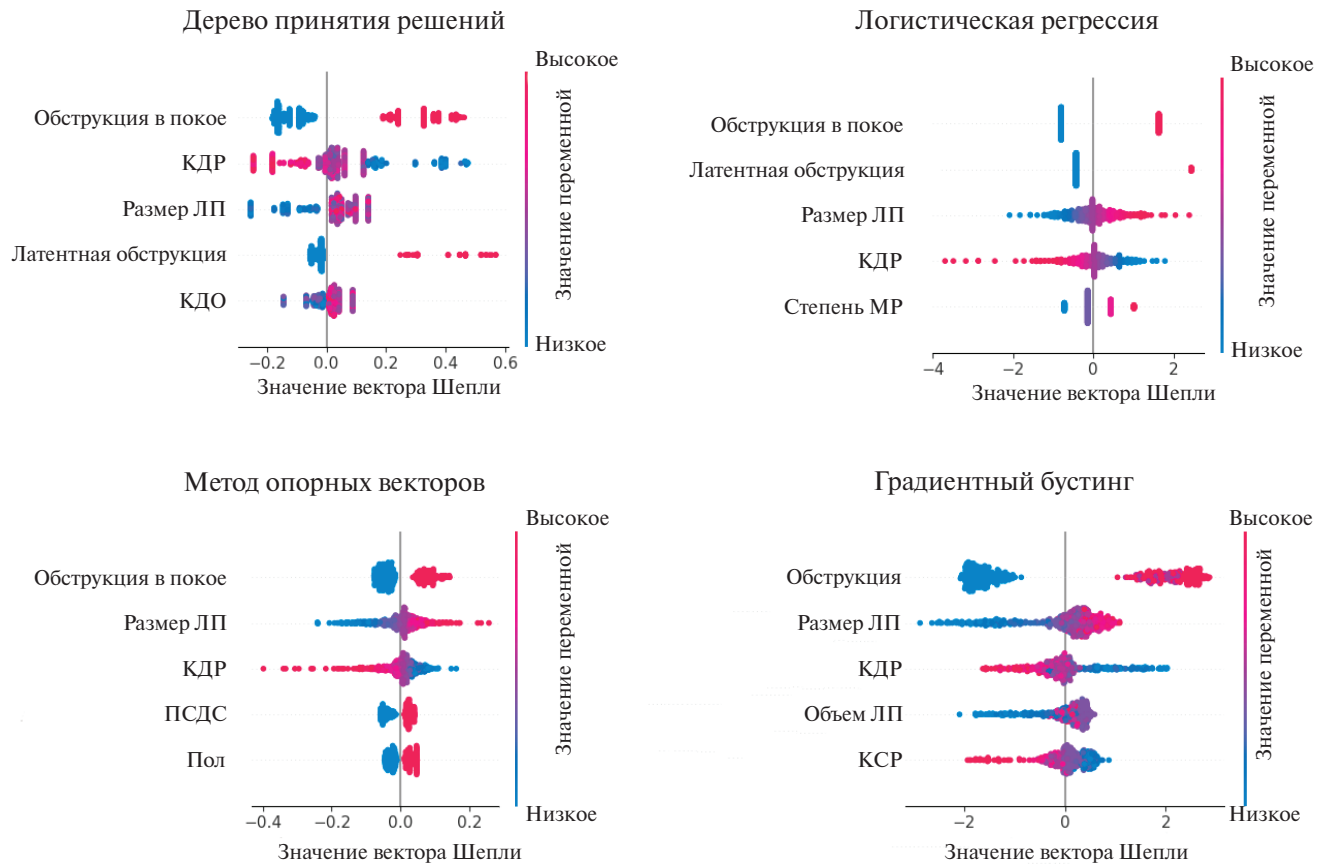


Рис. 1. Графики значений векторов Шепли для каждой модели (признаки упорядочены по степени значимости, для каждой модели выделены пять наиболее важных признаков).

Примечание: цветное изображение доступно в электронной версии журнала.

Сокращения: КДО — конечный диастолический объем, КДР — конечный диастолический размер, ЛП — левое предсердие, МР — митральная регургитация, ПДСД — переднесистолическое движение створки митрального клапана.

временный эхокардиографический фенотип ГКМП. Основные эхокардиографические параметры анализируемой выборки представлены в таблице 2. В клинической практике вероятность наличия ГКМП определяется выраженностью гипертрофии ЛЖ, поэтому и в нашем исследовании выявление пациентов с ГКМП основывалось на оценке толщины стенки ЛЖ. У большей части пациентов с ГКМП толщина стенки ЛЖ составляла не менее 15 мм. Поскольку данный фактор имел сильную корреляцию с наличием ГКМП, принято решение исклю-

чить толщину стенки ЛЖ из рассмотрения при построении моделей, чтобы более объективно оценить влияние других факторов. Основные метрики построенных моделей приведены в таблице 3. Метрики рассчитаны в предположении, что порог вероятности при классификации пациентов составляет 50%. В соответствии с табличными данными все построенные модели обладают достаточной точностью выявления ГКМП (доля истинных случаев ГКМП среди всех определённых как ГКМП), однако точ-

Таблица 4

Процент совпадений в суждениях при использовании моделей машинного обучения и при индивидуальной оценке

Модель	Процент совпадений для ГКМП	Процент совпадений для не ГКМП	Общий процент совпадений
Дерево решений	79	53	66
Логистическая регрессия	72	65	69
Опорные вектора	97	15	57
Градиентный бустинг	87	36	63

Сокращение: ГКМП — гипертрофическая кардиомиопатия.

ность исключения диагноза ГКМП достаточно низкая. Это связано со спецификой данных, поскольку пациентов с определенным диагнозом ГКМП в них в несколько раз меньше, чем пациентов без ГКМП (557 vs 134). Все модели обладают удовлетворительной чувствительностью и специфичностью, кроме градиентного бустинга, чувствительность которого оказалась ниже, чем у других моделей.

Для каждой модели при помощи метода векторов Шепли определены наиболее значимые факторы, влияющие на классификацию. На рисунке 1 приведены значения Шепли для пяти наиболее значимых факторов для каждой модели в порядке уменьшения значимости (сверху вниз). Цвет ассоциирован со значением фактора: красный — большие значения для количественных признаков либо наличие признака для бинарных; синий — меньшие значения для количественных признаков либо отсутствие признака для бинарных. Наблюдения со значением Шепли больше нуля отнесены моделью к ГКМП.

Для модели дерева решений следующие параметры были ассоциированы с наличием ГКМП: наличие обструкции ВТЛЖ, меньший конечный диастолический размер ЛЖ, дилатация ЛП, более молодой возраст.

Для логистической регрессии следующие параметры были ассоциированы с наличием ГКМП: обструкция в ВТЛЖ, дилатация ЛП, относительно небольшой конечный диастолический размер ЛЖ, выраженность митральной и аортальной недостаточности, систолическое движение передней митральной створки (SAM-феномен).

Для метода опорных векторов наиболее значимыми признаками, свидетельствующими о наличии ГКМП, стали: наличие обструкции, дилатация ЛП, меньший конечный диастолический размер ЛЖ, наличие SAM-феномена, утолщение стенки ПЖ, выраженность митральной недостаточности.

Для модели градиентного бустинга наиболее значимыми признаками, свидетельствующими в пользу ГКМП, стали: дилатация ЛП, меньший конечный диастолический размер и конечный систолический объем, больший индекс объема ЛП, более молодой возраст.

Валидация моделей. Затем изучены истории болезни и амбулаторные карты всех пациентов, отнесенных к "серой зоне". Пациенты были разделены на две ка-

тегории в зависимости от клинической вероятности наличия или отсутствия ГКМП. Наиболее значимыми факторами были: возраст, длительность и тяжесть АГ, толщина миокарда ЛЖ, тип гипертрофии миокарда, наличие сопутствующей патологии, не описанной в исходных данных. Выполнены предсказания вероятности наличия ГКМП, произведена оценка совпадения предсказаний между методами машинного обучения и независимой интерпретацией клинических данных. Линейная регрессия продемонстрировала себя как метод с наибольшим числом совпадений с индивидуальной оценкой по данным документации (68%). Результаты представлены в таблице 4.

Обсуждение

Машинное обучение активно внедряется в различные области практической медицины, особенно, в целях улучшения диагностики. Предложены решения для автоматизированной обработки огромных баз данных с изображениями магнитно-резонансной томографии сердца для оценки морфологии камер, определения фенотипа миокарда [9]. Однако данных визуализации часто бывает недостаточно для установления диагноза, определения подходящей тактики ведения.

Использование алгоритмов машинного обучения активно внедряется в диагностику и стратификацию риска внезапной сердечной смерти (ВСС) у пациентов с ГКМП. В одном из исследований пациенты с верифицированной ГКМП подверглись двукратному проведению магнитно-резонансной томографии сердца с автоматическим расчетом толщины стенок ЛЖ с использованием алгоритмов машинного обучения и оценкой выраженности гипертрофии экспертами. Результаты, полученные при автоматическом расчете, оказались более воспроизводимыми, измерения экспертов имели большую вариабельность [10].

В другой работе авторы параллельно оценивали данные электрокардиографии и трансторакальной эхокардиографии с помощью автоматизированных алгоритмов с целью уточнения диагноза ГКМП у пациентов с гипертрофией ЛЖ в сравнении с экспертной оценкой. Машинные методы оказались значительно точнее экспертных в отношении дифференциального диагноза ГКМП [11].

Проблема верификации диагноза с учетом сопутствующих заболеваний сохраняется, несмотря на частую встречаемость ГКМП в популяции. Активное внедрение анализа электронных медицинских баз данных для последующих когортных исследований позволило сократить время верификации диагноза ГКМП, сохранив жизни пациентов, благодаря выбору оптимальной тактики ведения. Однако существующие в настоящее время диагностические алгоритмы имеют ограниченную производительность из-за высокой доли ложноположительных результатов. В ранее опубликованном наблюдательном исследовании треть пациентов с гипертрофическим фенотипом была ошибочно классифицирована в группу ГКМП [12]. Более позднее исследование подтверждает тот факт, что треть пациентов, которым присвоены коды МКБ-10 I42.1, I42.2, не соответствуют общепризнанным диагностическим критериям [13]. По мнению ряда исследователей, только внедрение алгоритмов машинного обучения и искусственного интеллекта позволит существенно улучшить диагностику ГКМП [14, 15]. Результативность такого подхода прежде всего связана с использованием достаточно большого набора клинико-анамнестических и инструментальных данных для идентификации пациентов. В настоящем исследовании при создании предиктивных моделей с использованием машинного обучения проанализированы 30 параметров пациентов с гипертрофическим фенотипом. Применение модели машинного обучения с использованием логистической регрессии позволило снизить риск ошибочной диагностики ГКМП в группе сомнительного диагно-

за до 31%. Дальнейшее добавление датасетов с информацией о сопутствующих заболеваниях и применяемых методах лечения позволит существенно повысить производительность предлагаемых моделей. Несмотря на ценность таких методов машинного обучения, как случайный лес, с медицинской точки зрения они не поддаются интерпретации.

Все представленные в данной работе прогностические модели предлагают оценивать наличие обструкции в ВТЛЖ, переднезадний размер ЛП и более молодой возраст как факторы, которые свидетельствуют в пользу наличия ГКМП у пациента. Разработка программных решений для диагностики и ранней профилактики ВСС является перспективным направлением для нашей работы.

Заключение

На основании регистра были разработаны 4 предиктивные модели для дифференциального диагноза при выраженной гипертрофии миокарда ЛЖ с целью улучшения диагностики ГКМП. Все модели обладали высокой специфичностью, однако недостаточной чувствительностью. По результатам валидации слепым методом, оптимальной моделью для клинической практики определена логистическая регрессия. Использование машинного обучения позволяет улучшить точность диагностики ГКМП, обеспечить своевременную стратификацию риска и первичную профилактику ВСС.

Отношения и деятельность: все авторы заявляют об отсутствии потенциального конфликта интересов, требующего раскрытия в данной статье.

Литература/References

- Maron BJ, Rowin EJ, Maron MS. Global Burden of Hypertrophic Cardiomyopathy. *JACC Heart Fail.* 2018;6(5):376-8. doi:10.1016/j.jchf.2018.03.004.
- Osthega Y, Fryar CD, Nwankwo T, et al. Hypertension Prevalence Among Adults Aged 18 and Over: United States, 2017-2018. *NCHS Data Brief.* 2020;(364):1-8.
- Alashi A, Desai RM, Khullar T, et al. Different Histopathologic Diagnoses in Patients With Clinically Diagnosed Hypertrophic Cardiomyopathy After Surgical Myectomy. *Circulation.* 2019;140(4):344-6. doi:10.1161/CIRCULATIONAHA.119.040129.
- Ezhova AV, Zaitsev VV, Ishmukhametov GI, et al. Association between traditional cardiovascular risk factors and clinical phenotype of hypertrophic cardiomyopathy. "Arterial'naya Gipertenziya" ("Arterial Hypertension"). 2023;29(4):371-9. (In Russ). Ежова А.В., Зайцев В.В., Ишмухаметов Г.И. и др. Значение традиционных факторов риска сердечно-сосудистых заболеваний в фенотипических проявлениях гипертрофической кардиомиопатии. Артериальная гипертензия. 2023;29(4):371-9. doi:10.18705/1607-419X-2023-29-4-371-379.
- Blockeel H, Devos L, Fréney, et al. Decision trees: from efficient prediction to responsible AI. *Front Artif Intell.* 2023;6:1124553. doi:10.3389/rai.2023.1124553.
- Mohri M, Rostamizadeh A, Talwalkar A. *Foundations of Machine Learning.* MIT Press, 2018. ISBN: 9780262018258.
- Deisenroth MP, Faisal AA, Ong CS. *Mathematics for Machine Learning.* Cambridge University Press, 2024. ISBN: 9781108470049.
- Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable.* Lulu.com, 2020. ISBN: 9780244768522.
- Mariscal-Harana J, Asher C, Vergani V, et al. An artificial intelligence tool for automated analysis of large-scale unstructured clinical cine cardiac magnetic resonance databases. *Eur Heart J Digit Health.* 2023;4(5):370-83. doi:10.1093/ehjdh/ztad044.
- Augusto JB, Davies RH, Bhuva AN, et al. Diagnosis and risk stratification in hypertrophic cardiomyopathy using machine learning wall thickness measurement: a comparison with human test-retest performance. *Lancet Digit Health.* 2021;3(1):e20-e28. doi:10.1016/S2589-7500(20)30267-3.
- Soto JT, Weston Hughes J, Sanchez PA, et al. Multimodal deep learning enhances diagnostic precision in left ventricular hypertrophy. *Eur Heart J Digit Health.* 2022;3(3):380-9. doi:10.1093/ehjdh/ztac033.
- Magnusson P, Palm A, Branden E, et al. Misclassification of hypertrophic cardiomyopathy: validation of diagnostic codes. *Clin Epidemiol.* 2017;9:403-10. doi:10.2147/CLEP.S139300.
- Nielsen SK, Rasmussen TB, Hey TM, et al. Frequency of misdiagnosis in hypertrophic cardiomyopathy. *Eur Heart J Qual Care Clin Outcomes.* 2024;qcae031. doi:10.1093/ehjqcc/qcae031.
- Farahani NZ, Arunachalam SP, Sundaram DSB, et al. Explanatory Analysis of a Machine Learning Model to Identify Hypertrophic Cardiomyopathy Patients from EHR Using Diagnostic Codes. *Proceedings (IEEE Int Conf Bioinformatics Biomed).* 2020;2020:1932-7. doi:10.1109/bibm49941.2020.9313231.
- Verma AA, Murray J, Greiner R, et al. Implementing machine learning in medicine. *CMAJ.* 2021;193(34):E1351-E1357. doi:10.1503/cmaj.202434.



Форсига

(Дапаглифлозин)

В классе ИНГЛТ-2

– единственный препарат, показавший снижение риска гСН и СС смерти и входящий в стандарты медицинской помощи пациентам с ХСН¹⁻³

29%

CAP 6,0%
p < 0,001

Всего^a гСН²

ОР: 0,71
95 % ДИ: 0,65–0,78

14% СОР

CAP 1,5%
p = 0,01

Смерть от ССЗ²

ОР: 0,86
95 % ДИ: 0,76–0,97

10% СОР

CAP 1,5%
p = 0,03

Смерть по любой причине²

ОР: 0,90
95 % ДИ: 0,82–0,99



Входит в рекомендации и стандарты по ХСН^{4,5}



Доступна в льготе⁶



**1 таблетка
10 мг¹**



**1 раз
в сутки¹**



**без
титрации¹**



**включен
в ЖНВЛП⁵**

ИНГЛТ-2 — ингибиторы натрий-глюкозного котранспортера 2 типа; СН — сердечная недостаточность; гСН — госпитализация по поводу СН; ХСН — хроническая СН; СС — сердечно-сосудистый; СОР — снижение относительного риска; CAP — снижение абсолютного риска; ЖНВЛП — перечень жизненно необходимых и важнейших лекарственных препаратов.

^a Первичные и повторные.

1. Инструкция по медицинскому применению лекарственного препарата Форсига (таблетки, покрытые пленочной оболочкой, 5 мг, 10 мг). Регистрационное удостоверение ЛП-002596 от 21.08.2014 (переоформлено 16.03.2022) https://grls.rosminzdrav.ru/Grls_View_v2.aspx?routingGuid=b8c830f1-see7-46c5-97d-09f652745a3a.

2. Jhund PS et al. Nat Med. 2022;28(9):1956–1964.

3. Официальный интернет-портал правовой информации. Приказ Министерства здравоохранения Российской Федерации от 20.04.2022 № 272н "Об утверждении стандарта медицинской помощи взрослым при хронической сердечной недостаточности (диагностика, лечение и диспансерное наблюдение)" (Зарегистрирован 02.06.2022 № 68714) <http://publication.pravo.gov.ru/Document/View/0001202206020034> (дата обращения 29.08.2023).

4. Клинические рекомендации по ХСН (проект) 2023 https://scardio.ru/content/Guidelines/project/Project_HSN_2023.pdf (дата обращения 29.08.2023).

5. Перечень жизненно необходимых и важнейших лекарственных препаратов (ЖНВЛП) <http://government.ru/docs/44206/> (дата обращения 31.08.2023).

6. Официальный интернет-портал правовой информации. Приказ Министерства здравоохранения Российской Федерации от 29.09.2022 № 639н "Об утверждении перечня лекарственных препаратов для медицинского применения в целях обеспечения в амбулаторных условиях лиц, находящихся под диспансерным наблюдением, которые перенесли острое нарушение мозгового кровообращения, инфаркт миокарда, а также которым выполнены аортокоронарное шунтирование, ангиопластика коронарных артерий со стентированием и катетерная абляция по поводу сердечно-сосудистых заболеваний, в течение 2 лет даты постановки диагноза и (или) выполнения хирургического вмешательства" (Зарегистрирован 27.10.2022 № 70725) <http://publication.pravo.gov.ru/Document/View/0001202210270019> (дата обращения 29.08.2023).

Материал предназначен для специалистов здравоохранения. Имеются противопоказания. Перед назначением ознакомьтесь, пожалуйста, с полной инструкцией по медицинскому применению препарата Форсига (дапаглифлозин).
ООО «АстраЗенка Фармасьютикалз», 123112, г. Москва, 1-й Красноварварский проезд, д. 21, стр. 1, Башня «ОКО», 30 этаж. Тел.: +7 (495) 799-56-99, факс: +7 (495) 799-56-98. www.astrazeneca.ru, www.az-most.ru
Номер одобрения: FOR- RU-18059. Дата согласования: 01-09-2023. Дата истечения: 31-08-2025.

Инструкция
по применению препарата

