DOI: 10.21294/1814-4861-2025-24-5-27-39
УДК: 616.15-006-08-06:615.273:616.12

CC BY 4.0

Для цитирования: *Эль-Хатиб М.А., Солопов М.В., Склянная Е.В., Попандопуло А.Г.* Сравнительный анализ моделей машинного обучения для прогнозирования антрациклин-индуцированной кардиотоксичности у пациентов с онкогематологическими заболеваниями. Сибирский онкологический журнал. 2025; 24(5): 27–39. – doi: 10.21294/1814-4861-2025-24-5-27-39

For citation: *El-Khatib M.A., Solopov M.V., Sklyannaya E.V., Popandopulo A.G.* Comparative analysis of machine learning models for predicting anthracycline-induced cardiotoxicity in patients with hematologic malignancies. Siberian Journal of Oncology. 2025; 24(5): 27–39. – doi: 10.21294/1814-4861-2025-24-5-27-39

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ АНТРАЦИКЛИН-ИНДУЦИРОВАННОЙ КАРДИОТОКСИЧНОСТИ У ПАЦИЕНТОВ С ОНКОГЕМАТОЛОГИЧЕСКИМИ ЗАБОЛЕВАНИЯМИ

М.А. Эль-Хатиб[1,2], М.В. Солопов[1], Е.В. Склянная[1,2], А.Г. Попандопуло[1]

[1]ФГБУ «Институт неотложной и восстановительной хирургии им. В.К. Гусака» Минздрава России
Россия, Донецкая Народная Республика, 283045, г.о. Донецк, г. Донецк, пр. Ленинский, 47
[2]ФГБОУ ВО «Донецкий государственный медицинский университет им. М. Горького» Минздрава России
Россия, Донецкая Народная Республика, 283003, г.о. Донецк, г. Донецк, пр-т Ильича, 16

Аннотация

Цель исследования – разработка и валидация модели прогнозирования антрациклин-индуцированной кардиотоксичности (АИК) у пациентов с онкогематологическими заболеваниями. Задача решалась в два этапа: сравнительный анализ эффективности различных алгоритмов машинного обучения для выбора наиболее точной и устойчивой модели; идентификация наиболее значимых клинико-инструментальных предикторов на основе лучшей из протестированных моделей. **Материал и методы.** В проспективное исследование включено 155 пациентов (возраст 18–74 года) с онкогематологическими заболеваниями, получавших антрациклины. Анализировались клинические данные, уровни биомаркеров (NT-proBNP, тропонин I) и эхокардиографические параметры диастолической функции (E', E/E', LAVI). Данные проходили предобработку (стандартизация, one-hot encoding), дисбаланс классов устранялся методом SMOTETomek. Модели обучались и оценивались с использованием 5-кратной стратифицированной кросс-валидации по метрикам F1-меры, AUC-ROC, точности и полноты. **Результаты.** Статистически значимыми предикторами АИК явились уровни NT-proBNP ($p<0,001$), TropI ($p=0,004$), эхокардиографические параметры E' ($p<0,001$) и LAVI ($p<0,001$). Включение возраста и соотношения E/E' дополнительно улучшило прогностическую ценность моделей. Логистическая регрессия продемонстрировала наилучшую производительность (F1 0,943 ± 0,070, AUC-ROC 0,963 ± 0,051) при идеальной точности (1,00 ± 0,00) и высокой полноте (0,90 ± 0,12). Линейный дискриминантный анализ показал сопоставимые результаты (F1 0,921 ± 0,066, AUC-ROC 0,963 ± 0,046). Анализ важности признаков в наиболее эффективных моделях (LogReg, LDA) выявил, что наибольший вклад в прогноз вносит эхокардиографический параметр E'. Линейные модели превзошли более сложные алгоритмы (нейронные сети, ансамблевые методы) в данном исследовании. **Заключение.** Линейные модели, в частности логистическая регрессия, показывают высокую точность и надежность в прогнозировании АИК при использовании комбинации биомаркеров и эхокардиографических показателей диастолической функции в качестве предикторов. Данные модели обладают потенциалом для клинического применения с целью стратификации риска и своевременного начала кардиопротективной терапии. Необходима дальнейшая валидация модели на выборках пациентов из разных медицинских центров.

Ключевые слова: антрациклин-индуцированная кардиотоксичность, машинное обучение, прогнозирование, биомаркеры, эхокардиография, онкогематология.

Эль-Хатиб Марьям Аднан Ибрагим, el-khatib.mariam@yandex.ru

COMPARATIVE ANALYSIS OF MACHINE LEARNING MODELS FOR PREDICTING ANTHRACYCLINE-INDUCED CARDIOTOXICITY IN PATIENTS WITH HEMATOLOGIC MALIGNANCIES

M.A. El-Khatib[1,2], M.V. Solopov[1], E.V. Sklyannaya[1,2], A.G. Popandopulo[1]

[1]V.K. Gusak Institute of Emergency and Reconstructive Surgery, Ministry of Health of Russia
47, Leninsky Ave., Donetsk, Donetsk People's Republic, 283045, Russia
[2]M. Gorky Donetsk State Medical University, Ministry of Health of Russia
16, Ilyicha Ave., Donetsk, Donetsk People's Republic, 283003, Russia

Abstract

Objective: a comparative study of various machine learning algorithms (AdaBoost, k-nearest neighbors, linear discriminant analysis, logistic regression, neural networks, random forest, stochastic gradient descent, support vector machines, XGBoost) for predicting anthracycline-induced cardiotoxicity (AIC) in patients with haematological cancer using clinical and instrumental predictors. **Material and Methods.** A prospective study included 155 haematological cancer patients receiving anthracycline-containing therapy. The age of the patients ranged from 18 to 74 years. Clinical data, biomarker levels (NT-proBNP, troponin I), and echocardiographic parameters of diastolic function (E', E/E', LAVI) were analyzed. Data underwent preprocessing (standardization, one-hot encoding), and class imbalance was addressed using SMOTETomek. Models were trained and evaluated via 5-fold stratified cross-validation using F1-score, AUC-ROC, precision, and recall metrics. **Results.** Statistically significant predictors of AIC included NT-proBNP ($p<0.001$), troponin I ($p=0.004$), and echocardiographic parameters E' ($p<0.001$) and LAVI ($p<0.001$). Incorporating age and E/E' ratio further enhanced the model predictive value. Logistic regression demonstrated optimal performance (F1 0.943 ± 0.070, AUC-ROC 0.963 ± 0.051) with perfect precision (1.00 ± 0.00) and high recall (0.90 ± 0.12). Linear discriminant analysis yielded comparable results (F1 0.921 ± 0.066, AUC-ROC 0.963 ± 0.046). Linear models outperformed more complex algorithms (neural networks, ensemble methods). **Conclusion.** Linear models, particularly logistic regression, exhibit high accuracy and reliability in predicting AIC using combined biomarkers and echocardiographic diastolic function parameters. These models show potential for clinical implementation in risk stratification and timely initiation of cardioprotective therapy. Further validation across multi-center patient cohorts is warranted.

Key words: anthracycline-induced cardiotoxicity, machine learning, prediction, biomarkers, echocardiography, oncohaematology.

Введение

Антрациклиновые антибиотики, такие как доксорубицин и даунорубицин, остаются краеугольным камнем в лечении многих онкологических, в том числе гематологических, заболеваний благодаря их высокой эффективности в подавлении репликации и репарации ДНК опухолевых клеток [1, 2]. Однако их клиническое применение существенно ограничивается риском развития антрациклин-индуцированной кардиотоксичности (АИК), которая может манифестировать как в острой, так и в отсроченной форме, приводя к дисфункции левого желудочка (ЛЖ) и сердечной недостаточности. Основными механизмами антрациклин-индуцированного повреждения миокарда считаются окислительный стресс, вызванный генерацией активных форм кислорода [3, 4], ингибирование топоизомеразы II бета в кардиомиоцитах [5] и потенциальная дисрегуляция ренин-ангиотензин-альдостероновой системы [6].

Распространенность АИК варьирует в зависимости от критериев диагностики и исследуемой популяции. В одном исследовании с участием 100 онкологических пациентов кардиотоксичность, определенная как снижение фракции выброса левого желудочка (ФВЛЖ) менее 60 %, была выявлена у 31 % пациентов [7], в то время как другое исследование сообщило о развитии кардиомиопатии у 15 % пациентов [8]. Метаанализ показал, что клинически выраженная кардиотоксичность встречается у 6 % пациентов, получающих антрациклины, а субклиническая – у 18 % [9]. Ключевыми факторами риска являются кумулятивная доза антрациклинов, крайние возрастные группы, а также наличие сопутствующих сердечно-сосудистых заболеваний и факторов риска, таких как артериальная гипертензия и сахарный диабет [7, 9].

Своевременная диагностика и мониторинг АИК имеют решающее значение для предотвращения необратимого повреждения миокарда. Современный подход включает оценку с помощью инструментальных и лабораторных методов. Эхокардиография является основным методом оценки систолической функции ЛЖ, при этом оценка глобальной продольной деформации (GLS) показывает большую чувствительность в выявлении ранних субклинических изменений по сравнению с традиционной оценкой ФВЛЖ [10, 11]. Методики

3D-трекинга спеклов (3D-STI) и тканевой допплерографии могут дополнительно повысить точность диагностики [12, 13]. Магнитно-резонансная томография (МРТ) сердца с использованием Т1- и Т2-картирования и оценки позднего накопления гадолиния позволяет выявлять отек, диффузный и очаговый фиброз миокарда, однако применение этой процедуры ограничено её доступностью и стоимостью [14, 15]. Среди биомаркеров наиболее изучены высокочувствительный тропонин I (hs-cTnI) и N-концевой промозговой натрийуретический пептид (NT-proBNP), повышение которых ассоциировано с повреждением кардиомиоцитов и риском развития сердечной недостаточности соответственно [10, 11, 16, 17]. Комбинированное использование биомаркеров и эхокардиографических параметров, таких как GLS, может улучшить раннее выявление пациентов с высоким риском АИК [18, 19]. Перспективными, но требующими дальнейшего изучения остаются маркеры фиброза (ST2, Galectin-3) и ишемического повреждения (H-FABP) [20, 21]. Тем не менее существующие методы имеют ограничения, включая вариабельность измерений при эхокардиографии, недостаточную специфичность биомаркеров и отсутствие стандартизированных протоколов мониторинга, что затрудняет точное и своевременное прогнозирование риска [22].

Таким образом, в настоящее время с целью раннего выявления антрациклиновых повреждений сердца используются многочисленные неинвазивные методы. Однако ни один из них не является строго специфичным, в связи с чем для исключения возможных диагностических ошибок требуется тщательная интерпретация полученных результатов. Необходимо отметить, что несовершенство диагностики на данный момент не позволяет четко определить ранние доклинические критерии АИК и вовремя предотвратить развитие необратимой кардиальной дисфункции. В связи с вышесказанным необходимость исходной оценки риска развития АИК и выявления категорий пациентов высокого риска среди больных, не имеющих явных, клинически значимых факторов сердечно-сосудистого риска на момент инициации противоопухолевой терапии, включающей в себя кардиотоксичные агенты, не вызывает сомнений.

В последние годы методы машинного обучения (МО) все активнее применяются в кардиоонкологии для решения задач прогнозирования и стратификации риска развития АИК. Модели МО способны интегрировать большие объемы гетерогенных данных (клинических, инструментальных, лабораторных) для выявления сложных нелинейных взаимосвязей и построения точных прогностических моделей. Например, разработаны номограммы на основе регрессионной модели LASSO, интегрирующие уровни NT-proBNP, GLS и клинические переменные и показавшие хорошую прогностическую ценность [23]. В работах с использованием логистической регрессии и регрессии LASSO идентифицированы ключевые предикторы кардиотоксичности [24, 25]. Методы неконтролируемого обучения, такие как кластерный анализ, применялись для идентификации различных фенотипов кардиотоксичности, что потенциально может способствовать персонализации подходов к ведению пациентов [26]. Трансферное обучение использовалось для создания моделей на основе данных ЭКГ [27], а модель случайного леса продемонстрировала высокую точность (AUC-ROC 0,88) в прогнозировании кардиотоксичности [28].

Существует пробел в сравнительных исследованиях, оценивающих эффективность различных алгоритмов МО на одном и том же наборе данных для прогнозирования АИК. Большинство исследований фокусируются на разработке и валидации одной конкретной модели. Кроме того, вопросы интерпретируемости моделей МО и их валидации на независимых внешних когортах остаются критически важными для внедрения этих технологий в клиническую практику [29, 30]. Также необходимо учитывать этические аспекты, связанные с использованием данных пациентов и потенциальной предвзятостью алгоритмов.

Несмотря на растущую популярность сложных моделей, таких как нейронные сети и ансамблевые методы, в медицинской диагностике, их применение на относительно небольших клинических выборках сопряжено с риском переобучения и низкой интерпретируемостью. Возникает важный научный и практический вопрос: оправдано ли усложнение моделей в задаче прогнозирования АИК, или же классические, интерпретируемые подходы могут обеспечить сопоставимую или даже более высокую точность?

Целью исследования является разработка и валидация прогностической модели АИК, основанной на последовательном подходе. Первым шагом являются критическая оценка и сравнительный анализ эффективности серии алгоритмов МО – от классических линейных моделей до сложных ансамблей и нейронных сетей. Вторым шагом, после определения наиболее производительной и надежной модели, является анализ ее структуры для выявления и интерпретации ключевых клинико-инструментальных предикторов, вносящих наибольший вклад в риск развития АИК. Такой двухэтапный дизайн исследования позволяет не только построить точный прогностический инструмент, но и углубить клиническое понимание факторов риска.

Материал и методы

В исходно проведенное проспективное исследование было включено 155 пациентов в возрасте от 18 до 74 лет, проходивших лечение по пово-

ду онкогематологических заболеваний, включая острый миелобластный лейкоз (31 %), острый лимфобластный лейкоз (9 %), лимфогранулематоз (12 %) и неходжкинские лимфомы (48 %). Условиями включения являлись возраст от 18 до 74 лет, установленный диагноз онкогематологического заболевания, по поводу которого планировались антрациклин-содержащие курсы химиотерапии (ХТ), и информированное согласие на участие. Критериями исключения были наличие в анамнезе врожденных и приобретенных пороков сердца, стенокардии напряжения, перенесенного инфаркта миокарда, дилатационной, гипертрофической и рестриктивной кардиомиопатии, гипертонической болезни, сердечной недостаточности, эндокринной патологии, алкогольной и/или наркотической зависимости, ВИЧ-инфекции, а также ожидаемая продолжительность жизни менее 6 мес.

На I этапе исследования пациенты были разделены на 2 группы в зависимости от возраста в соответствии с классификацией возрастов Всемирной организации здравоохранения: 1-ю группу составили 67 пациентов в возрасте 18–44 лет (условно, группа молодого возраста), 2-ю – 88 пациентов в возрасте 45–74 лет (условно, группа среднего и пожилого возраста).

Существенных различий по полу, площади поверхности тела и кумулятивной дозе антрациклинов между сравниваемыми группами не было ($p>0,05$). Далее пациентам проводилось определение уровней кардиальных биомаркеров (тропонин I, NT-proBNP), а также оценивались основные структурно-функциональные параметры сердца. Все обследования выполнялись дважды – до начала антрациклин-содержащих курсов ХТ и после достижения кумулятивной дозы 250 мг/м^2 по доксорубицину.

В ходе исследования проводился сбор клинико-инструментальных данных, включая возраст, сведения о проводимой терапии антрациклинами (доксорубицин, даунорубицин, идарубицин, митоксантрон), уровни кардиомаркеров, а также эхокардиографические параметры. Полученные данные использовались для последующей оценки и анализа развития АИК. Критериями развития КТ считали: изменения сегмента ST; нарушения ритма и проводимости, удлинение интервала QT по данным стандартной ЭКГ и ее суточного мониторирования; снижение ФВЛЖ менее 50 % или более чем на 10 % от исходного значения по данным двухмерной Эхо-КГ.

Второй этап проспективного исследования начинался после достижения пациентами кумулятивной дозы 250 мг/м^2 по доксорубицину. Больные находились под динамическим наблюдением в течение 5 лет. Каждые 6 мес после окончания антрациклин-содержащей химиотерапии проводились сбор жалоб, объективный осмотр, выполнялась трансторакальная Эхо-КГ с последующей оценкой основных структурно-функциональных параметров сердца.

Для выявления клинических и лабораторных параметров, ассоциированных с развитием АИК, проведен статистический анализ данных. Первоначально выполнен корреляционный анализ для оценки взаимосвязи между каждой из исследуемых переменных и развитием АИК. Анализируемые переменные включали возраст, факт проведения лечения, уровень N-терминального промозгового натрийуретического пептида (NT-proBNP), уровень тропонина I (TropI), а также эхокардиографические показатели диастолической функции: скорость раннего диастолического наполнения (E'), соотношение пиковой скорости трансмитрального потока к E' (E/E'), а также индекс объема левого предсердия (LAVI).

Для более детальной оценки различий был проведен сравнительный анализ между группами пациентов с наличием (n=18) и отсутствием (n=137) АИК. Для каждой переменной рассчитаны описательные статистики (среднее значение, медиана, стандартное отклонение) отдельно для каждой группы. Выбор статистического теста для сравнения групп зависел от типа данных и их распределения. Для категориальных переменных использовался критерий χ^2 Пирсона. Для непрерывных переменных предварительно проводилась проверка на нормальность распределения с помощью теста Шапиро–Уилка. При нормальном распределении данных в обеих группах применялся t-критерий Стьюдента. В случае отсутствия нормального распределения хотя бы в одной из групп использовался непараметрический U-критерий Манна–Уитни. Различия считались статистически значимыми при уровне $p<0,05$.

Дополнительно, для оценки мультиколлинеарности между непрерывными предикторами был рассчитан фактор инфляции дисперсии (variance inflation factor, VIF). Статистическая обработка данных выполнялась с использованием языка программирования Python v3.12.8 и библиотек pandas v2.2.3, scipy v1.15.2, statsmodels v0.14.4, numpy v2.2.4, matplotlib v3.10.1 и seaborn v0.13.2.

Для прогнозирования АИК проведено сравнительное исследование серии алгоритмов МО: AdaBoost, K-ближайших соседей (KNN), линейного дискриминантного анализа (LDA), логистической регрессии (LogReg), нейронных сетей (NN), случайного леса (RF), стохастического градиентного спуска (SGD), метода опорных векторов (SVM), XGBoost и нейронных сетей (NN). Обучение алгоритмов было реализовано с использованием библиотек scikit-learn v1.6.1 и PyTorch v2.6.0. Непрерывные клинико-инструментальные признаки подверглись стандартизации методом StandardScaler, а категориальные признаки преобразовывались через one-hot encoding с исключением первого столбца для минимизации мультиколл-

Таблица 1/Table 1

Метрики качества моделей машинного обучения

Machine learning model quality metrics

Метрика/ Metric	Описание/Description	Клиническая интерпретация в контексте прогнозирования АИК/ Clinical interpretation in AIC prediction context
Точность/ Precision	Доля истинно положительных прогнозов среди всех положительных прогнозов/ The proportion of true positive predictions among all positive predictions.	Показывает, какая доля пациентов, отмеченных моделью как группа риска, действительно подвержена АИК. Высокая точность минимизирует ложноположительные диагнозы и ненужные вмешательства/ Shows the proportion of patients identified as at-risk who actually develop AIC. High precision minimizes false positives and unnecessary interventions.
Полнота/ Recall	Доля истинно положительных прогнозов среди всех реальных положительных случаев/ The proportion of actual positives correctly identified.	Показывает, какую долю пациентов с реальным риском АИК модель смогла выявить. Высокая полнота критически важна для минимизации риска пропуска пациентов, нуждающихся в кардиопротективной терапии/ Shows how many actual at-risk patients were identified. High recall is critical to minimize missing patients needing cardioprotective therapy.
F1-мера/ F1-score	Гармоническое среднее между точностью и полнотой/ Harmonic mean of precision and recall.	Комплексный показатель, балансирующий между точностью и полнотой. Особенно полезен при несбалансированных классах/ Balanced metric between precision and recall. Especially useful with imbalanced classes.
AUC-ROC	Площадь под ROC-кривой/ Area under the Receiver Operating Characteristic curve.	Оценивает общую способность модели различать пациентов с риском АИК и без него. Значение ≈1 указывает на превосходную различительную способность/ Measures overall ability to discriminate between at-risk and safe patients. Value ≈1 indicates excellent discriminative power.

Примечание: таблица составлена авторами.

Note: created by the authors.

линеарности. Устранение дисбаланса классов, характерного для исходных данных, выполнено с применением гибридного подхода SMOTETomek, сочетающего синтетическую генерацию примеров миноритарного класса и удаление шумовых точек мажоритарного класса. Для поиска оптимальных гиперпараметров моделей применялся подход, сочетающий GridSearchCV для моделей с ограниченным пространством параметров (LogReg, LDA) и RandomizedSearchCV для моделей с обширным пространством гиперпараметров (KNN, SVM, RF, AdaBoost, XGBoost, SGD).

Исследуемые нейросетевые архитектуры были представлены многослойными перцептронами с вариативным количеством скрытых слоев, функциями активации ReLU и LeakyReLU, а также слоями батч-нормализации (batch normalization) и отсева (dropout) для регуляризации. Процесс оптимизации нейросетей включал кастомную функцию потерь RecallFocusedLoss, направленную на снижение доли ложноотрицательных прогнозов, и адаптивный планировщик ReduceLROnPlateau, автоматически корректирующий скорость обучения в зависимости от динамики валидационных потерь. Наиболее эффективная конфигурация нейросети избиралась на основе максимизации ключевых метрик на валидационных данных для сравнения с другими алгоритмами.

Оценка эффективности моделей проведена методом стратифицированной 5-кратной кросс-валидации с расчетом точности (precision), полноты (recall), F1-меры и AUC-ROC. Интерпретация данных метрик приведена в табл. 1. Порог классификации для каждой модели определялся путем максимизации F1-меры, что позволило балансировать между чувствительностью и специфичностью. Итоговые метрики, усредненные по всем фолдам, были дополнены стандартными отклонениями для оценки стабильности моделей.

Результаты

Проведен корреляционный анализ для оценки взаимосвязи между исследуемыми параметрами и развитием АИК (рис. 1). Наиболее выраженная положительная корреляция с АИК наблюдалась для эхокардиографического показателя E' (r=0,50). Умеренная положительная корреляция выявлена для индекса LAVI (r=0,36) и уровня NT-proBNP (r=0,32). Более слабая положительная корреляция отмечена для уровня TropI (r=0,22), факта проведения лечения (r=0,12) и показателя E/E' (r=0,10). Возраст показал слабую отрицательную корреляцию (r=-0,18) с развитием АИК.

Сравнительный анализ клинико-лабораторных показателей между группами пациентов с АИК и без АИК выявил значимые различия по ключевым параметрам (табл. 2). В группе с АИК зарегистрированы достоверно более высокие уровни биомаркеров кардиального повреждения: медиана NT-proBNP составила 193,5 пг/мл против 134,0 пг/мл в контрольной группе ($p<0,001$, U-критерий Манна–Уитни), а уровень TropI достигал 0,52 нг/мл по сравнению с 0,33 нг/мл у пациентов без АИК ($p=0,004$, U-критерий Манна–Уитни).

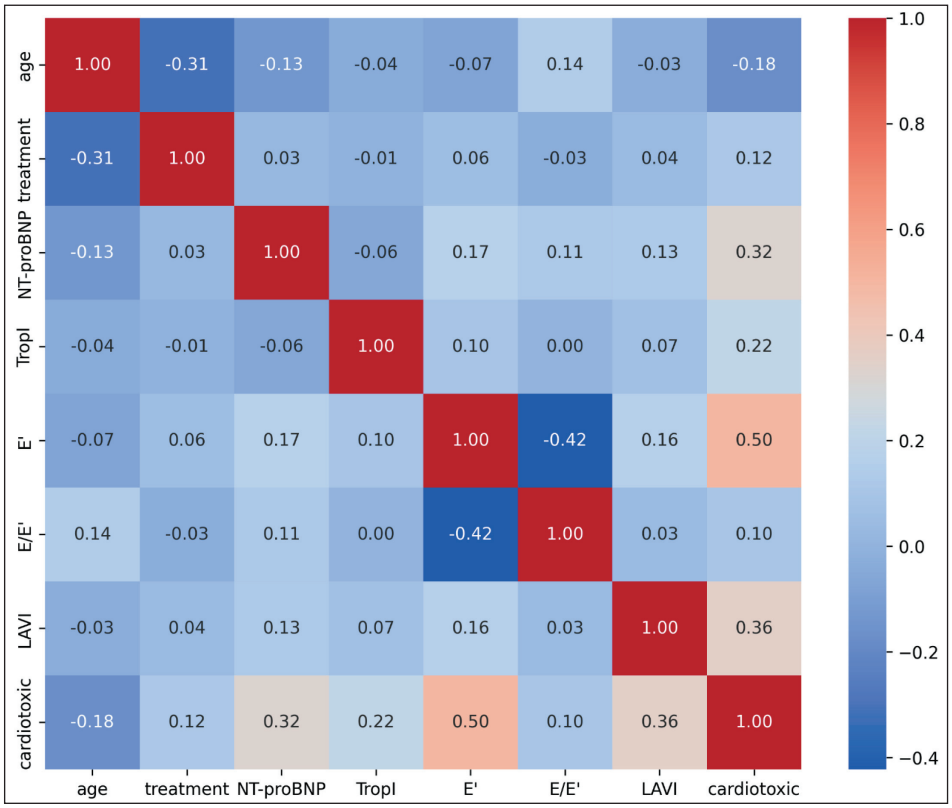

Рис. 1. Тепловая карта коэффициентов корреляции Пирсона между исследуемыми клинико-лабораторными, эхокардиографическими параметрами и развитием антрациклин-индуцированной кардиотоксичности. Примечание: рисунок выполнен авторами
Fig. 1. Heatmap of Pearson correlation coefficients between the investigated clinical, laboratory, and echocardiographic parameters and the development of anthracycline-induced cardiotoxicity. Note: created by the authors

Таблица 2/Table 2

Клинико–лабораторные параметры у пациентов с АИК и без нее
Clinical and laboratory parameters in patients with and without AIC

Параметр/Parameter	АИК/AIC	Среднее ± SD/ Mean ± SD	Медиана/ Median	Критерий/ Criterion	p–value
Возраст/Age	–	0,61 ± 0,49	1,00	χ^2	0,0518
	+	0,33 ± 0,49	0,00		
Лечение/Treatment	–	0,48 ± 0,50	0,00	χ^2	0,2208
	+	0,67 ± 0,49	1,00		
NT-proBNP (пг/мл/pg/mL)	–	133,18 ± 55,67	134,02	Манна–Уитни/ Mann–Whitney	$8{,}93\times10^{-5}$
	+	190,20 ± 36,73	193,53		
Тропонин I (нг/мл)/ Troponin I (ng/mL)	–	0,32 ± 0,17	0,33	Манна–Уитни/ Mann–Whitney	0,00428
	+	0,43 ± 0,18	0,52		
E' (см/с/cm/s)	–	8,34 ± 3,26	8,35	Манна–Уитни/ Mann–Whitney	$5{,}95\times10^{-8}$
	+	14,16 ± 3,52	14,22		
E/E'	–	11,21 ± 5,34	9,87	Манна–Уитни/ Mann–Whitney	0,3553
	+	12,93 ± 6,57	10,88		
LAVI (мл/м²/mL/m²)	–	22,50 ± 3,98	22,22	t-критерий/ t-test	$5{,}69\times10^{-6}$
	+	27,24 ± 4,31	28,35		

Примечание: таблица составлена авторами.

Note: created by the authors.

Эхокардиографическое исследование выявило значимые различия в показателях диастолической функции: медиана E' в группе с АИК составила 14,2 см/с против 8,35 см/с в контрольной группе ($p<0{,}001$, U-критерий Манна–Уитни). Параметр LAVI демонстрировал статистически значимое увеличение до 27,2 ± 4,3 мл/м² у пациентов с АИК по сравнению с 22,5±4,0 мл/м² в группе контроля ($p<0{,}001$, t-критерий Стьюдента).

При этом не выявлено статистически значимых различий между группами по возрасту ($p=0{,}052$, χ^2-критерий), факту проведения терапии ($p=0{,}221$,

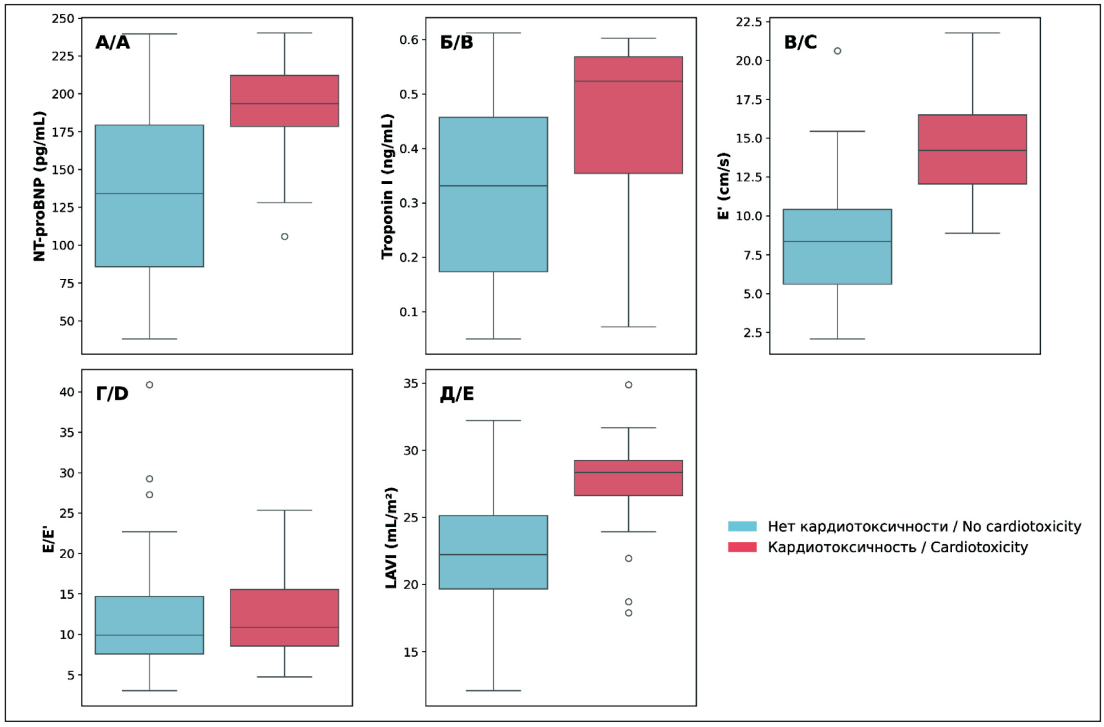

Рис. 2. Сравнение кардиальных биомаркеров между группами пациентов с АИК и без нее. Боксплоты показывают распределение значений: (А) NT-proBNP (пг/мл); (Б) тропонин I (нг/мл); (В) E' (см/с), (Г) E/E', (Д) LAVI (мл/м²). Синие боксплоты соответствуют пациентам без АИК, красные – пациентам с АИК. Боксплоты отображают медиану (центральная линия), межквартильный размах (границы бокса) и выбросы (точки за пределами усов). Примечание: рисунок выполнен авторами

Fig. 2. Comparison of cardiac biomarkers between patients with and without AIC. Box plots show the distribution of values for: (A) NT-proBNP (pg/mL), (B) troponin I (ng/mL), (C) E' (cm/s), (D) E/E', (E) LAVI (mL/m²). Blue box plots represent patients without AIC, red box plots represent patients with AIC. Box plots display the median (center line), interquartile range (box boundaries), and outliers (points beyond whiskers). Note: created by the authors

χ^2-критерий) и показателю соотношения E/E' (p=0,355, U-критерий Манна–Уитни), хотя отмечалась тенденция к более молодому возрасту пациентов в группе с АИК. Полученные данные свидетельствуют о значимом влиянии антрациклиновой терапии на маркеры кардиального повреждения и параметры сердечной функции.

Проверка на мультиколлинеарность с использованием VIF для непрерывных переменных (NT_proBNP, TropI, E', E/E', LAVI) показала низкие значения VIF для всех признаков (максимальный VIF=1,35 для E'). Это свидетельствует об отсутствии значимой линейной зависимости между предикторами, что позволяет включать их совместно в регрессионные модели.

Ведущими предикторами АИК, выявленными в результате статистического анализа, стали NT-proBNP и E', демонстрирующие высокую дискриминативную способность. Параметр NT-proBNP (p<0,001; r=0,32) отражает объемную перегрузку миокарда, а E' (p<0,001; r=0,50) – ранние нарушения диастолической функции, что подтверждается их выраженной межгрупповой вариабельностью (рис. 2). TropI (p=0,004; r=0,22) и LAVI (p<0,001; r=0,36), несмотря на умеренную корреляцию с исходом, дополнили модели за счет интеграции данных о прямом повреждении кардиомиоцитов и внутрисердечной гемодинамике. Эти четыре параметра, включенные во все версии моделей, обеспечили базовую прогностическую мощность, достигая AUC-ROC >0,88.

Включение параметров E/E' и возраста, несмотря на их пограничную статистическую значимость в одномерном анализе, существенно повысило прогностическую эффективность моделей. Например, для логистической регрессии добавление этих признаков привело к увеличению F1-меры с 0,876 до 0,943, AUC-ROC – с 0,955 до 0,963, а полнота возросла с 0,85 до 0,90, что критически важно для минимизации ложноотрицательных прогнозов.

В ходе исследования проведена комплексная оценка эффективности девяти алгоритмов МО для прогнозирования АИК. Ключевые метрики (F1-мера, AUC-ROC) анализировались в контексте клинической релевантности, воспроизводимости и вычислительной эффективности (рис. 3).

Модель логистической регрессии продемонстрировала максимальную прогностическую точность при сохранении высокой полноты (0,90), достигнув F1 0,943 и AUC-ROC 0,963. Нулевое стандартное отклонение точности подтвердило её устойчивость в стратификации риска, что особенно значимо для клинических решений, требующих минимизации ложноположительных диагнозов. Линейный дискриминантный анализ показал сопоставимую эффективность (F1 0,921, AUC-ROC

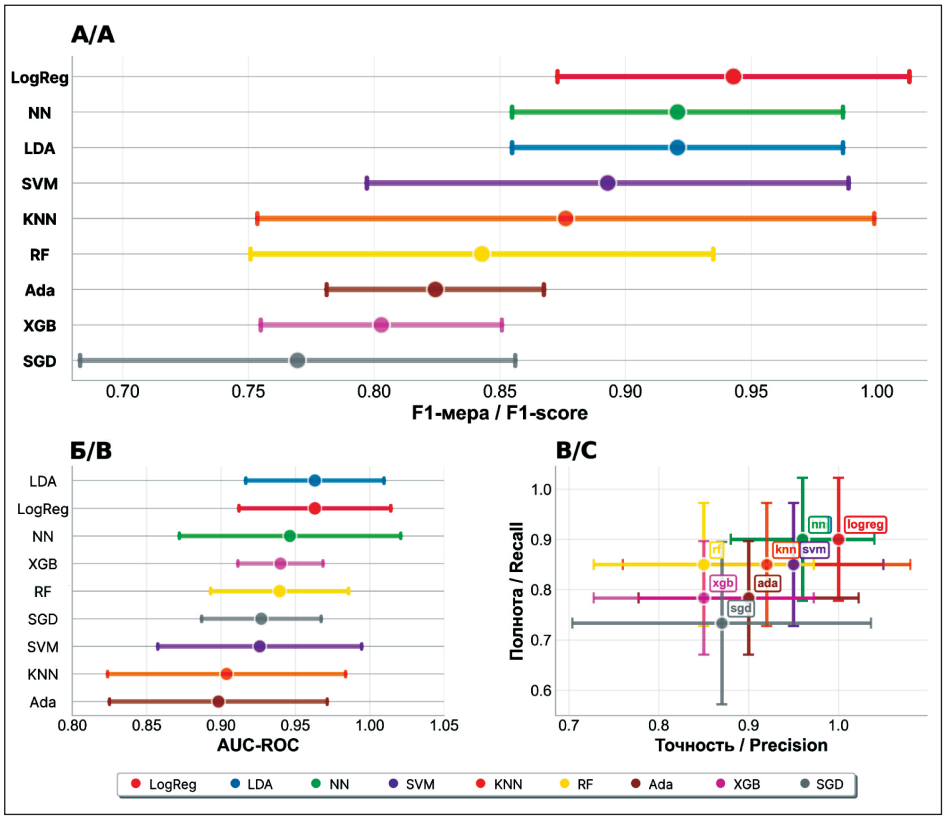

Рис. 3. Комплексный анализ производительности алгоритмов машинного обучения для прогнозирования АИК. Комбинированная диаграмма демонстрирует сравнительную оценку девяти алгоритмов машинного обучения: логистическая регрессия (LogReg), линейный дискриминантный анализ (LDA), метод опорных векторов (SVM), k-ближайших соседей (KNN), случайный лес (RF), адаптивный бустинг (Ada), градиентный бустинг (XGB), стохастический градиентный спуск (SGD) и нейронная сеть (NN). А – сравнение F1-меры; Б – производительность по метрике AUC-ROC; В – диаграмма рассеяния «точность против полноты». Все доверительные интервалы рассчитаны на основе стандартных отклонений, полученных при кросс-валидации. Примечание: рисунок выполнен авторами

Fig. 3. Comprehensive performance analysis of machine learning algorithms for AIC prediction. The composite figure presents a comparative assessment of nine machine learning algorithms: logistic regression (LogReg), linear discriminant Analysis (LDA), support vector machine (SVM), k-nearest neighbors (KNN), random forest (RF), adaptive boosting (Ada), gradient boosting (XGB), stochastic gradient descent (SGD), and neural network (NN). A –F1-score comparison; B – AUC-ROC performance; C – precision vs. recall scatter plot. Confidence intervals were calculated based on standard deviations derived from cross-validation. Note: created by the authors

0,963), однако с более низким показателем точности 0,96.

На основе сравнительного анализа нейросетевых архитектур наиболее эффективным оказался многослойный персептрон с тремя скрытыми слоями размерностью 64, 32 и 16 нейронов с dropout-регуляризацией 0,3, 0,2 и 0,1 соответственно. В качестве функции активации использовалась LeakyReLU. Модель обучалась с размером батча 64, скоростью обучения 0,0001 и коэффициентом регуляризации весов 1e-05. Такая нейронная сеть, несмотря на высокие метрики (F1 0,921, AUC-ROC 0,946), потребовала на два порядка больше вычислительных ресурсов (29,89 с против 0,54 с для LogReg), что ставит под вопрос её практическую применимость.

Нелинейные алгоритмы (SVM, KNN) показали умеренную эффективность (F1 0,89–0,88), при этом KNN выделился максимальной вариабельностью F1-меры, возможно, из-за чувствительности к шуму в гетерогенных клинических данных. Ансамблевые методы (RF, XGBoost) обеспечили высокую AUC-ROC (>0,93), но уступили линейным моделям в точности прогнозирования.

Анализ важности признаков проводился с использованием абсолютных значений коэффициентов для логистической регрессии и коэффициентов линейных дискриминантных функций для LDA (рис. 4). Результаты показали, что наиболее значимым предиктором кардиотоксичности в обеих моделях является показатель E' (LDA: 3.956, LogReg: 2.311), за которым следуют различные комбинации эхокардиографических и биохимических маркеров. LDA продемонстрировала более высокие абсолютные значения важности для большинства признаков по сравнению с логистической регрессией, что указывает на различные механизмы весового распределения между алгоритмами.

Обсуждение

Сравнительная эффективность моделей

Ключевым и, возможно, контринтуитивным результатом нашей работы стало то, что классические линейные модели (LogReg и LDA) превзошли

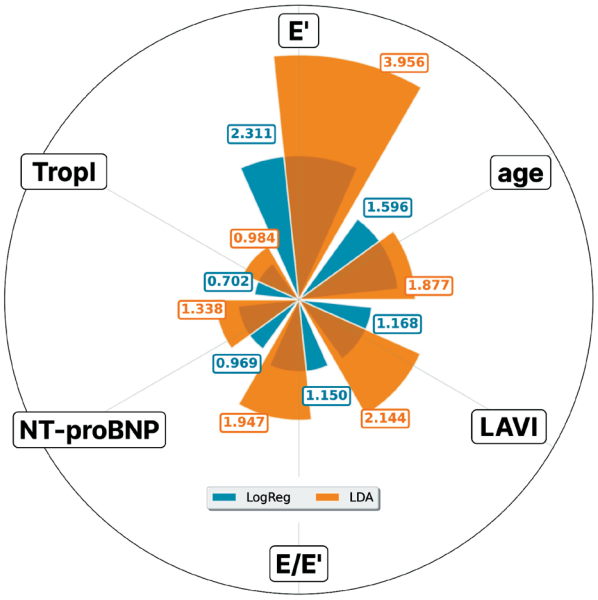

Рис. 4. Сравнение важности прогностических маркеров в моделях логистической регрессии (LogReg) и линейного дискриминантного анализа (LDA).
Примечание: рисунок выполнен авторами
Fig. 4. Comparison of predictive biomarker importance in logistic regression (LogReg) and linear discriminant analysis (LDA) models. Note: created by the authors

более сложные алгоритмы, включая нейронные сети и ансамблевые методы. Этот вывод имеет практическое значение, поскольку он ставит под сомнение мнение о необходимости применения «черных ящиков» для достижения максимальной точности. Наши результаты показывают, что на данных с выраженной, но преимущественно линейной взаимосвязью между предикторами и исходом простые и полностью интерпретируемые модели не только не уступают, но и выигрывают в надежности и устойчивости прогноза.

Сравнение с результатами предыдущих исследований

Сравнение с литературными данными показывает, что достигнутые в представленном исследовании показатели эффективности превышают результаты предыдущих работ. В исследовании, посвященном прогнозированию кардиотоксичности у больных раком молочной железы, сравнивалось несколько алгоритмов МО. Лучшую точность в прогнозировании показала модель многослойного перцептрона с AUC 0,66, а при валидации для прогнозирования сердечной недостаточности со сниженной фракцией выброса та же модель достигла AUC 0,81 [31]. Модель регрессии LASSO для прогнозирования АИК у детей с гематологическими опухолями, включающая такие переменные, как кумулятивная доза антрациклина, фракция выброса и NT-proBNP, достигла значения AUC-ROC 0,773 [23]. Для диагностики субклинической кардиотоксичности, вызванной антрациклином, у пациентов с острым лейкозом с использованием уровней PLT и NT-proBNP модель логистической регрессии достигла AUC-ROC 0,713 [32]. Наши лучшие модели (LogReg, LDA) показали AUC-ROC 0,963, что является самым высоким результатом среди сравниваемых исследований.

Интересно, что в работе W.T. Chang et al. [31] лучшие результаты показала именно нейронная сеть, тогда как в нашем исследовании линейные модели (LogReg, LDA) значительно превзошли нейросетевые и ансамблевые алгоритмы. Это расхождение подчеркивает важность подбора модели под конкретный набор данных, целевую переменную и используемые предикторы. Например, в исследовании W.T. Chang et al. [31] ключевыми детерминантами в модели логистической регрессии были использование трастузумаба, наличие гипертензии и доза антрациклина, в отличие от нашего набора, сфокусированного на биомаркерах (NT-proBNP, TropI) и параметрах диастолической функции (E', LAVI).

Исследуемая нами когорта состояла из взрослых пациентов (возраст 18–74 года) с различными онкогематологическими заболеваниями. Другие исследования фокусировались на специфических группах (например, дети [23], больные раком молочной железы [31]). Различия в возрасте, типе злокачественных новообразований и схемах лечения могут влиять на факторы риска и, следовательно, на прогностическую силу моделей. В нашей работе при первичном статистическом анализе возраст показал слабую отрицательную корреляцию с кардиотоксичностью, однако включение его в список параметров усилило прогностическую способность алгоритмов.

Клиническая интерпретация и значимость предикторов

Несмотря на высокие показатели AUC-ROC, важно отметить, что модель логистической регрессии достигла идеальной точности (1,00) при сохранении высокой полноты (0,90). Это делает ее полезной для клинического применения, где минимизация ложноотрицательных прогнозов (пропуск пациентов с риском) является приоритетом, но и высокая точность (избегание ненужных вмешательств у пациентов без риска) также важна. Нейронные сети, хотя и показали сравнимую эффективность с LogReg и LDA, требовали больше вычислительных ресурсов, что ограничивает их применимость.

В этом исследовании ключевыми предикторами стали NT-proBNP, E', TropI и LAVI, с дополнительным вкладом E/E' и возраста. Этот набор включает как биомаркеры повреждения миокарда и объемной перегрузки, так и эхокардиографические параметры диастолической функции. Сочетание этих параметров, вероятно, позволило моделям идентифицировать различные аспекты развивающейся кардиотоксичности. В других исследованиях использовались иные наборы предикторов (например, LVEF, GLS, кумулятивная доза антрациклина),

что может влиять на итоговую производительность моделей [23, 26, 31, 32].

Полученные результаты анализа важности признаков выявили ключевые различия в интерпретации предикторов АИК между наиболее эффективными алгоритмами (LogReg и LDA). Доминирование показателя E' в обеих моделях согласуется с современными представлениями о ранней диастолической дисфункции как предвестнике кардиотоксических осложнений. Примечательно, что LDA придает существенно большее значение индексу объема левого предсердия (LAVI: 2,144 *vs* 1,168) и отношению E/E' (1,947 *vs* 1,150), что может отражать способность дискриминантного анализа лучше выявлять нелинейные взаимосвязи между показателями диастолической функции. Относительно низкая важность тропонина I в обеих моделях (LDA: 0,984, LogReg: 0,702) предполагает, что эхокардиографические параметры могут быть более чувствительными ранними маркерами кардиотоксичности по сравнению с традиционными биохимическими показателями повреждения миокарда, что имеет важное значение для мониторинга пациентов, получающих антрациклины.

Ограничения исследования

Ключевым ограничением исследования является относительно небольшой размер выборки (155 пациентов), а также малое количество целевых событий (случаев развития АИК). Несмотря на то, что применение техники SMOTETomek для устранения дисбаланса классов способствовало улучшению и стабилизации производительности моделей в отношении миноритарного класса, мы понимаем, что никакие методы ресэмплинга не могут полностью заменить данные, полученные на более крупной когорте. В связи с этим достигнутые показатели эффективности моделей следует интерпретировать с осторожностью. Поэтому полученные результаты стоит рассматривать как предварительные, а разработанные модели – как пилотные.

В дальнейших исследованиях необходима валидация обученных и возможное дообучение моделей на более крупных и независимых внешних когортах для подтверждения их прогностической способности. Кроме того, исследование фокусировалось на прогнозировании АИК на основе данных, собранных в определенные моменты времени; в будущих исследованиях необходимо обратить внимание на динамическое прогнозирование с учетом изменений параметров во времени.

Заключение

В исследовании проведен сравнительный анализ девяти алгоритмов МО для прогнозирования АИК у пациентов с онкогематологическими заболеваниями. Ключевыми предикторами риска в когорте были определены уровень NT-proBNP, эхокардиографические параметры диастолической функции (E', LAVI), уровень тропонина I, а также возраст и соотношение E/E'. Среди протестированных моделей наилучшую общую производительность продемонстрировала логистическая регрессия, достигнув F1-меры 0,943 и AUC-ROC 0,963, при этом показав идеальную точность (1,00) и высокую полноту (0,90). Линейный дискриминантный анализ показал сопоставимые результаты по AUC-ROC и F1-мере, но с немного меньшей точностью. Эти линейные модели превзошли более сложные алгоритмы, включая нейронные сети и ансамблевые методы, в этом конкретном наборе данных демонстрируя высокую эффективность и устойчивость при относительно низких вычислительных затратах.

Полученные результаты подчеркивают потенциал использования относительно простых и интерпретируемых моделей МО для точной стратификации риска АИК в клинической практике. Важно отметить, что разработанные модели, основанные на рутинно используемых биомаркерах и эхокардиографических показателях, являются вычислительно простыми и экономически доступными, что обосновывает их потенциал для широкого клинического применения с целью раннего выявления пациентов группы высокого риска и своевременного начала кардиопротективной терапии. Дальнейшие исследования должны быть направлены на валидацию разработанных моделей на выборках пациентов из разных медицинских центров и изучение возможности их интеграции в системы поддержки принятия врачебных решений.

ЛИТЕРАТУРА/REFERENCES

1. *Васюк Ю.А., Школьник Е.Л., Несветов В.В., Школьник Л.Д., Варлан Г.В.* Кардиоонкология: современные аспекты профилактики антрациклиновой кардиотоксичности. Кардиология. 2016; 56(12): 72–79. [*Vasyuk Y.A., Shkolnik E.L., Nesvetov V.V., Shkolnik L.D., Varlan G.V.* Cardiooncology: Current Aspects of Prevention of Anthracycline Toxicity. Cardiology. 2016; 56(12): 72–79. (in Russian).]. doi: 10.18565/cardio.2016.12.72-79. EDN: XIMODJ.

2. *Morgillo A., Mazzarella M., Randazzo M.F., Mancino N., Talessi M., Marovino E.* Anthracycline Cardiotoxicity: From Mechanisms to Prevention Strategies. J Med – Clin Res Rev. 2023; 7(1): 1–5. doi: 10.33425/2639-944X.1309.

3. *Valcovici M., Andrica F., Serban C., Dragan S.* Cardiotoxicity of anthracycline therapy: current perspectives. Arch Med Sci. 2016; 12(2): 428–35. doi: 10.5114/aoms.2016.59270.

4. *Narezkina A., Narayan H.K., Zemljic-Harpf A.E.* Molecular mechanisms of anthracycline cardiovascular toxicity. Clin Sci (Lond). 2021; 135(10): 1311–32. doi: 10.1042/CS20200301.

5. *Agunbiade T.A., Zaghlol R.Y., Barac A.* Heart Failure in Relation to Anthracyclines and Other Chemotherapies. Methodist DeBakey Cardiovasc J. 2019; 15(4): 243–49. doi: 10.14797/mdcj-15-4-243.

6. *Sobczuk P., Czerwińska M., Kleibert M., Cudnoch-Jędrzejewska A.* Anthracycline-induced cardiotoxicity and renin-angiotensin-aldosterone system-from molecular mechanisms to therapeutic applications. Heart Fail Rev. 2022; 27(1): 295–319. doi: 10.1007/s10741-020-09977-1.

7. *Idrees I., Ahmad R., Nadeem M., Khan A., Gilani A., Khattak R.* Antracycline Induced Early Onset Chronic Cardiotoxicity in Cancer Patients. Pak Armed Forces Med J. 2024; 74(6): 1492–95. doi: 10.51253/pafmj.v74i2.6795.

8. *Tariq M., Ibrahim M., Khan S.R., Baloch S.S., Shahzadi M., Rashid Y.A., Zahid M.J., Khan M.D.* Frequency of Cardiac-Toxicity in Patient Treated with Anthracycline, An Institutional Perspective. Pak J Med Health Sci. 2022; 16(4): 488. doi: 10.53350/pjmhs22164488.

9. *Lotrionte M., Biondi-Zoccai G., Abbate A., Lanzetta G., D'Ascenzo F., Malavasi V., Peruzzi M., Frati G., Palazzoni G.* Review and meta-analysis

of incidence and clinical predictors of anthracycline cardiotoxicity. Am J Cardiol. 2013; 112(12): 1980–84. doi: 10.1016/j.amjcard.2013.08.026.

10. *Sawaya H., Sebag I.A., Plana J.C., Januzzi J.L., Ky B., Tan T.C., Cohen V., Banchs J., Carver J.R., Wiegers S.E., Martin R.P., Picard M.H., Gerszten R.E., Halpern E.F., Passeri J., Kuter I., Scherrer-Crosbie M.* Assessment of echocardiography and biomarkers for the extended prediction of cardiotoxicity in patients treated with anthracyclines, taxanes, and trastuzumab. Circ Cardiovasc Imaging. 2012; 5(5): 596–603. doi: 10.1161/CIRCIMAGING.112.973321.

11. *Bhagat A.A., Kalogeropoulos A.P., Baer L., Lacey M., Kort S., Skopicki H.A., Butler J., Bloom M.W.* Biomarkers and Strain Echocardiography for the Detection of Subclinical Cardiotoxicity in Breast Cancer Patients Receiving Anthracyclines. J Pers Med. 2023; 13(12): 1710. doi: 10.3390/jpm13121710.

12. *Gulati G., Zhang K.W., Scherrer-Crosbie M., Ky B.* Cancer and cardiovascular disease: the use of novel echocardiography measures to predict subsequent cardiotoxicity in breast cancer treated with anthracyclines and trastuzumab. Curr Heart Fail Rep. 2014; 11(4): 366–73. doi: 10.1007/s11897-014-0214-8.

13. *Zhang L., Zhang R., Shuai P., Chen J., Yin L.* A global case meta-analysis of three-dimensional speckle tracking for evaluating the cardiotoxicity of anthracycline chemotherapy in breast cancer. Front Cardiovasc Med. 2022; 9: 942620. doi: 10.3389/fcvm.2022.942620.

14. *Wei X., Lin L., Zhang G., Zhou X.* Cardiovascular Magnetic Resonance Imaging in the Early Detection of Cardiotoxicity Induced by Cancer Therapies. Diagnostics (Basel). 2022; 12(8): 1846. doi: 10.3390/diagnostics12081846.

15. *Musella F., Librera M., Sibilio G., Boccalatte M., Tagliamonte G., Cavaglià E., Ferrara I., Puglia M., Dell'Aversana S., Ducci C.B., Dellegrottaglie S., Savarese G., Scatteia A.* Cardiovascular magnetic resonance parametric techniques to characterize myocardial effects of anthracycline therapy in adults with normal left ventricular ejection fraction: a systematic review and meta-analysis. Curr Probl Cardiol. 2024; 49(7): 102609. doi: 10.1016/j.cpcardiol.2024.102609.

16. *Stevens P.L., Lenihan D.J.* Cardiotoxicity due to Chemotherapy: the Role of Biomarkers. Curr Cardiol Rep. 2015; 17(7): 603. doi: 10.1007/s11886-015-0603-y.

17. *Díaz-Antón B., Madurga R., Zorita B., Wasniewski S., Moreno-Arciniegas A., López-Melgar B., Ramírez Merino N., Martín-Asenjo R., Barrio P., Amado Escañuela M.G., Solís J., Parra Jiménez F.J., Ciruelos E., Castellano J.M., Fernández-Friera L.* Early detection of anthracycline- and trastuzumab-induced cardiotoxicity: value and optimal timing of serum biomarkers and echocardiographic parameters. ESC Heart Fail. 2022; 9(2): 1127–37. doi: 10.1002/ehf2.13782.

18. *Stolojanu C., Steflea R., Micsescu-Olah A.M., Alexandra I., Popoiu A., Doros G.* Combined Utility of Speckle Tracking Echocardiography and Cardiac Biomarkers for Early Detection of Anthracycline-Induced Cardiotoxicity in Pediatric Oncology Patients. Biomedicines. 2024; 12(12): 2849. doi: 10.3390/biomedicines12122849.

19. *Mahjoob M.P., Sheikholeslami S.A., Dadras M., Mansouri H., Haghi M., Naderian M., Sadeghi L., Tabary M., Khaheshi I.* Prognostic Value of Cardiac Biomarkers Assessment in Combination with Myocardial 2D Strain Echocardiography for Early Detection of Anthracycline-Related Cardiac Toxicity. Cardiovasc Hematol Disord Drug Targets. 2020; 20(1): 74–83. doi: 10.2174/1871529X19666190912150942.

20. *Horacek J.M., Vasatova M., Pudil R., Tichy M., Zak P., Jakl M., Jebavy L., Maly J.* Biomarkers for the early detection of anthracycline-induced cardiotoxicity: current status. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2014; 158(4): 511–17. doi: 10.5507/bp.2014.004.

21. *Feng W., Wang Q., Tan Y., Qiao J., Liu Q., Yang B., Yang S., Cui L.* Early detection of anthracycline-induced cardiotoxicity. Clin Chim Acta. 2025; 565: 120000. doi: 10.1016/j.cca.2024.120000.

22. *Murtagh G., Januzzi J.L., Scherrer-Crosbie M., Neilan T.G., Dent S., Ho J.E., Appadurai V., McDermott R., Akhter N.* Circulating Cardiovascular Biomarkers in Cancer Therapeutics-Related Cardiotoxicity: Review of Critical Challenges, Solutions, and Future Directions. J Am Heart Assoc. 2023; 12(21): e029574. doi: 10.1161/JAHA.123.029574.

23. *Liu Z., Wang Y., Huang X., Qi X., Qian C., Zhao S.* Development and Validation of a Diagnostic Nomogram to Predict the Anthracycline-Induced Early Cardiotoxicity in Children with Hematological Tumors. Cardiovasc Toxicol. 2022; 22(9): 802–12. doi: 10.1007/s12012-022-09755-5.

24. *Gómez-Vecino A., Corchado-Cobos R., Blanco-Gómez A., García-Sancha N., Castillo-Lluva S., Martín-García A., Mendiburu-Eliçabe M., Prieto C., Ruiz-Pinto S., Pita G., Velasco-Ruiz A., Patino-Alonso C., Galindo-Villardón P., Vera-Pedrosa M.L., Jalife J., Mao J.-H., Macías de Plasencia G., Castellanos-Martín A., Sáez-Freire M.D.M., Fraile-Martín S., Rodrigues-Teixeira T., García-Macías C., Galvis-Jiménez J.M., García-Sánchez A., Isidoro-García M., Fuentes M., García-Cenador M.B., García-Criado F.J., García-Hernández J.L., Hernández-García M.Á., Cruz-Hernández J.J., Rodríguez-Sánchez C.A., García-Sancho A.M., Pérez-López E., Pérez-Martínez A., Gutiérrez-Larraya F., Cartón A.J., García-Sáenz J.Á., Patiño-García A., Martín M., Alonso-Gordoa T., Vulsteke C., Croes L., Hatse S., Van Brussel T., Lambrechts D., Wildiers H., Hang C., Holgado-Madruga M., González-Neira A., Sánchez P.L., Pérez Losada J.* Intermediate Molecular Phenotypes to Identify Genetic Markers of Anthracycline-Induced Cardiotoxicity Risk. Cells. 2023; 12(15): 1956. doi: 10.3390/cells12151956.

25. *Sharafeldin N., Zhou L., Singh P., Crossman D.K., Wang X., Hageman L., Landier W., Blanco J.G., Burridge P.W., Sapkota Y., Yasui Y., Armstrong G.T., Robison L.L., Hudson M.M., Oeffinger K., Chow E.J., Armenian S.H., Weisdorf D.J., Bhatia S.* Gene-Level Analysis of Anthracycline-Induced Cardiomyopathy in Cancer Survivors: A Report From COG-ALTE03N1, BMTSS, and CCSS. JACC CardioOncol. 2023; 5(6): 807–18. doi: 10.1016/j.jaccao.2023.06.007.

26. *Jacquemyn X., Chinni B.K., Barnes B.T., Rao S., Kutty S., Manlhiot C.* Unsupervised machine learning identifies distinct phenotypes in cardiac complications of pediatric patients treated with anthracyclines. Cardiooncology. 2024; 10(1): 74. doi: 10.1186/s40959-024-00276-4.

27. *Yagi R., Goto S., Himeno Y., Katsumata Y., Hashimoto M., MacRae C.A., Deo R.C.* Artificial intelligence-enabled prediction of chemotherapy-induced cardiotoxicity from baseline electrocardiograms. Nat Commun. 2024; 15(1): 2536. doi: 10.1038/s41467-024-45733-x.

28. *Zhou Y., Hou Y., Hussain M., Brown S., Budd T., Tang W.H.W., Abraham J., Xu B., Shah C., Moudgil R., Popovic Z., Cho L., Kanj M., Watson C., Griffin B., Chung M.K., Kapadia S., Svensson L., Collier P., Cheng F.* Machine Learning-Based Risk Assessment for Cancer Therapy-Related Cardiac Dysfunction in 4300 Longitudinal Oncology Patients. J Am Heart Assoc. 2020; 9(23): e019628. doi: 10.1161/JAHA.120.019628.

29. *Artico J., Abiodun A., Shiwani H., Kurdi H., Chen D., Tyebally S., Moon J.C., Westwood M., Manisty C.H.* Multimodality Imaging for Cardiotoxicity: State of the Art and Future Perspectives. J Cardiovasc Pharmacol. 2022; 80(4): 547–61. doi: 10.1097/FJC.0000000000001281.

30. *Taruna R.N., Wihandono A., Budiarto R.M.* Cardiac Function in Breast Cancer Patients Undergoing Anthracycline Chemotherapy: A Comprehensive Review of Mechanisms, Monitoring, and Management Strategies. Int J Sci Adv. 2024; 5(6): 1565–72. doi: 10.51542/ijscia.v5i6.85.

31. *Chang W.T., Liu C.F., Feng Y.H., Liao C.T., Wang J.J., Chen Z.C., Lee H.C., Shih J.Y.* An artificial intelligence approach for predicting cardiotoxicity in breast cancer patients receiving anthracycline. Arch Toxicol. 2022; 96: 2731–37. doi: 10.1007/s00204-022-03341-y.

32. *Zhou X., Weng Y., Jiang T., Ou W., Zhang N., Dong Q., Tang X.* Influencing factors of anthracycline-induced subclinical cardiotoxicity in acute leukemia patients. BMC Cancer. 2023; 23(1): 976. doi: 10.1186/s12885-023-11060-5.

Поступила/Received 13.07.2025
Одобрена после рецензирования/Revised 06.10.2025
Принята к публикации/Accepted 27.10.2025

СВЕДЕНИЯ ОБ АВТОРАХ

Эль-Хатиб Марьям Аднан Ибрагим, кандидат медицинских наук, гематолог отделения гематологии (гематологии и химиотерапии), ФГБУ «Институт неотложной и восстановительной хирургии им. В.К. Гусака» Минздрава России; доцент кафедры внутренних болезней № 3, ФГБОУ ВО «Донецкий государственный медицинский университет им. М. Горького» Минздрава России (г. Донецк, Россия). SPIN-код: 9885-2438. ORCID: 0000-0003-3160-3974.

Солопов Максим Витальевич, биолог лаборатории клеточного и тканевого культивирования, ФГБУ «Институт неотложной и восстановительной хирургии им. В.К. Гусака» Минздрава России (г. Донецк, Россия). SPIN-код: 3147-0739. Author ID (Scopus): 57196187773. ORCID: 0000-0001-7053-4428.

Склянная Елена Валериевна, кандидат медицинских наук, заведующая отделом гематологии, ФГБУ «Институт неотложной и восстановительной хирургии им. В.К. Гусака» Минздрава России; доцент кафедры внутренних болезней № 3, ФГБОУ ВО

«Донецкий государственный медицинский университет им. М. Горького» Минздрава России (г. Донецк, Россия). SPIN-код: 5608-1287. Author ID (Scopus): 6505508326. ORCID: 0000-0002-0656-7097.
Попандопуло Андрей Геннадиевич, доктор медицинских наук, заведующий лабораторией клеточного и тканевого культивирования, хирург, ФГБУ «Институт неотложной и восстановительной хирургии им. В.К. Гусака» Минздрава России (г. Донецк, Россия). SPIN-код: 4336-2272. Author ID (Scopus): 12782689700. ORCID: 0000-0001-9755-1869.

ВКЛАД АВТОРОВ

Эль-Хатиб Марьям Аднан Ибрагим: сбор и анализ данных проведенного проспективного исследования, статистическая обработка информации, участие в написании черновика статьи, проверка и утверждение текста статьи.
Солопов Максим Витальевич: анализ литературы, разработка алгоритма машинного обучения, написание черновика статьи, подготовка рисунков, проверка и утверждение текста статьи.
Склянная Елена Валериевна: участие в анализе данных проведенного проспективного исследования, статистической обработке информации, проверка и утверждение текста статьи, научное консультирование.
Попандопуло Андрей Геннадиевич: разработка концепции исследования, написание вводной части статьи, научное консультирование, проверка и утверждение текста статьи.
Все авторы одобрили финальную версию статьи перед публикацией, выразили согласие нести ответственность за все аспекты работы, подразумевающую надлежащее изучение и решение вопросов, связанных с точностью и добросовестностью любой части работы.

Финансирование

Результаты получены при грантовом финансировании Фонда поддержки молодых ученых в области биомедицинских наук на основании договора № 03/2025 от 31.01.2025 г.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Соответствие принципам этики

Проведенное исследование соответствует стандартам Хельсинкской декларации, одобрено независимым этическим комитетом ФГБОУ ВО «Донецкий государственный медицинский университет им. М. Горького» Минздрава России (Россия, Донецкая Народная Республика, 283003, г.о. Донецк, г. Донецк, пр-т Ильича, 16), протокол № 36/16-1 от 19.11.21.

Информированное согласие

Все пациенты подписали письменное информированное согласие на публикацию данных в медицинском журнале, включая его электронную версию.

ABOUT THE AUTHORS

Maryam Adnan Ibrahim El-Khatib, MD, PhD, Hematologist, Department of Hematology (Hematology and Chemotherapy), V.K. Gusak Institute of Emergency and Reconstructive Surgery, Ministry of Health of Russia; Associate Professor, Department of Internal Medicine No. 3, M. Gorky Donetsk State Medical University, Ministry of Health of Russia (Donetsk, Russia). ORCID: 0000-0003-3160-3974.
Maxim V. Solopov, Biologist, Laboratory of Cell and Tissue Cultivation, V.K. Gusak Institute of Emergency and Reconstructive Surgery, Ministry of Health of Russia (Donetsk, Russia). Author ID (Scopus): 57196187773. ORCID: 0000-0001-7053-4428.
Elena V. Sklyannaya, MD, PhD, Head of the Hematology Department, V.K. Gusak Institute of Emergency and Reconstructive Surgery, Ministry of Health of Russia; Associate Professor, Department of Internal Medicine No. 3, M. Gorky Donetsk State Medical University, Ministry of Health of Russia (Donetsk, Russia). Author ID (Scopus): 6505508326. ORCID: 0000-0002-0656-7097.
Andrey G. Popandopulo, MD, DSc, Head of the Laboratory of Cell and Tissue Cultivation, Surgeon, V.K. Gusak Institute of Emergency and Reconstructive Surgery, Ministry of Health of Russia (Donetsk, Russia). Author ID (Scopus): 12782689700. ORCID: 0000-0001-9755-1869.

AUTHOR CONTRIBUTIONS

Maryam Adnan Ibrahim El-Khatib: collection and analysis of prospective study data, statistical processing of information, participation in writing the draft manuscript, approval of the manuscript.
Maxim V. Solopov: literature analysis, development of machine learning algorithms, writing a draft manuscript, preparing figures, approval of the manuscript.
Elena V. Sklyannaya: participation in the analysis of prospective study data, statistical processing of information, approval of the manuscript, scientific consulting.
Andrey G. Popandopulo: development of the research concept, scientific consulting, approval of the manuscript.
All authors approved the final version of the manuscript prior to publication and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

Funding

The results were obtained with the grant funding of the Foundation for Support of Young Scientists in Biomedical Sciences on the basis of the contract No. 03/2025 dated 31.01.2025.

Conflict of interests

The authors declare that they have no conflict of interest.

Compliance with Ethical Standards

The study was conducted in accordance with ethical principles outlined in the Declaration of Helsinki approved by Ethics Committee of M. Gorky Donetsk State Medical University (16, Ilyicha Ave., Donetsk, Donetsk People's Republic, 283003, Russia), protocol No. 36/16-1 dated November 19, 2021.

Voluntary informed consent

Written informed voluntaries consents were obtained from the patients for the publication of data in medical journal.