

ПРОГНОЗИРОВАНИЕ РИСКА РАЗВИТИЯ ЛЕКАРСТВЕННОГО ПОРАЖЕНИЯ ПЕЧЕНИ ПРИ ТЕРАПИИ РЕМДЕСИВИРОМ У ПАЦИЕНТОВ С COVID-19 С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ

Шевчук Ю.В.^{1,2},
 Шамигулов И.И.^{2,3},
 Сычев И.В.^{2,4},
 Крюков А.В.^{1,2},
 Темирбулатов И.И.²,
 Мирзаев К.Б.²,
 Денисенко Н.П.²,
 Абдуллаев Ш.П.²,
 Тучкова С.Н.²,
 Вечорко В.И.^{1,2},
 Аверков О.В.¹,
 Сычев Д.А.²

¹ ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы» (111539, г. Москва, ул. Вешняковская, 23, Россия)

² ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России (123242, г. Москва, ул. Баррикадная, 2/1, стр. 1, Россия)

³ ФГАОУ ВО «Московский физико-технический институт (национальный исследовательский университет)» (117303, г. Москва, ул. Керченская, 1А, корп. 1, Россия)

⁴ ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н.П. Огарева» (430005, г. Саранск, ул. Большевикская, 68, Россия)

Автор, ответственный за переписку:
 Шевчук Юлия Викторовна,
 e-mail: ju-viktorovna@yandex.ru

Статья поступила: 04.07.2024

Статья принята: 15.11.2024

Статья опубликована: 28.12.2024

РЕЗЮМЕ

Обоснование. Противовирусный препарат ремдесивир получил широкое распространение для этиотропного лечения COVID-19. Частота возникновения нежелательных реакций при терапии ремдесивиром достигает 66,2 %, наиболее распространённая нежелательная реакция – повышение уровня печёночных трансаминаз.

Цель исследования. Разработка модели машинного обучения для прогнозирования риска развития лекарственного поражения печени у пациентов с COVID-19 при назначении терапии ремдесивиром.

Методы. Данное проспективное открытое обсервационное исследование проводилось в период с ноября 2021 г. по февраль 2022 г. и включало 154 пациента, получающих терапию ремдесивиром. Пациенты были разделены на две группы: группа 1 (n = 45) – пациенты, у которых при терапии ремдесивиром развились признаки поражения печени; группа 2 (n = 109) – пациенты без данной нежелательной реакции. Всем пациентам были проведены фармакогенетическое исследование и ретроспективный анализ историй болезней, сформирована база данных с результатами проведённых исследований, на основе которой происходило машинное обучение моделей для прогноза риска развития лекарственного поражения печени.

Результаты. Основные прогностические факторы включали индекс массы тела (значимость – 12,03 %) и носительство генотипа AG по полиморфному маркеру rs776746 гена CYP3A5 (значимость – 10,04 %). В дальнейшем по всем полученным факторам на основе категориального бустинга построена модель предсказания развития лекарственного поражения печени, имеющая чувствительность 57,8 % и специфичность 80,7 %.

Заключение. С помощью машинного обучения была сформирована модель риска развития лекарственного поражения печени при терапии ремдесивиром. Индекс массы тела и носительство генотипа AG по полиморфному маркеру rs776746 гена CYP3A5 оказались ключевыми маркерами. Для улучшения точности модели требуется увеличение доли пациентов с нежелательной реакцией в тренировочной выборке. Дальнейшие исследования позволяют улучшить качество модели и интегрировать её в клиническую практику.

Ключевые слова: COVID-19, ремдесивир, нежелательные реакции, гепатотоксичность, фармакогенетическое исследование, предикторы нежелательных реакций, машинное обучение, модель риска

Для цитирования: Шевчук Ю.В., Шамигулов И.И., Сычев И.В., Крюков А.В., Темирбулатов И.И., Мирзаев К.Б., Денисенко Н.П., Абдуллаев Ш.П., Тучкова С.Н., Вечорко В.И., Аверков О.В., Сычев Д.А. Прогнозирование риска развития лекарственного поражения печени при терапии ремдесивиром у пациентов с COVID-19 с помощью машинного обучения. *Acta biomedica scientifica*. 2024; 9(6): 52-62. doi: 10.29413/ABS.2024-9.6.6

PREDICTING THE RISK OF DEVELOPING DRUG-INDUCED LIVER INJURY DURING REMDESIVIR THERAPY IN COVID-19 PATIENTS USING MACHINE LEARNING

Shevchuk Yu.V.^{1,2},
Shamigulov I.I.^{2,3},
Sychev I.V.^{2,4},
Kryukov A.V.^{1,2},
Temirbulatov I.I.²,
Mirzaev K.B.²,
Denisenko N.P.²,
Abdullaev Sh.P.²,
Tuchkova S.N.²,
Vechorko V.I.^{1,2},
Averkov O.V.¹,
Sychev D.A.²

¹ Municipal Clinical Hospital No. 15 named after O.M. Filatov (Veshnyakovskaya str. 23, Moscow 111539, Russian Federation)

² Russian Medical Academy of Continuing Professional Education (Barrikadnaya str. 2/1 build 1, Moscow 125993, Russian Federation)

³ Moscow Institute of Physics and Technology (National Research University) (Kerchenskaya str. 1A, Moscow 117303, Russian Federation)

⁴ National Research Ogarev Mordovia State University (Bolshevistskaya str. 68, Saransk 430005, Russian Federation)

Corresponding author:

Yuliya V. Shevchuk,

e-mail: ju-viktorovna@yandex.ru

ABSTRACT

Background. The antiviral drug Remdesivir has been widely used for etiotropic treatment of COVID-19. The incidence of adverse reactions during Remdesivir therapy reaches 66.2 %, the most common one being an increase in hepatic transaminases.

The aim. To develop a machine learning model for predicting the risk of drug-induced liver damage in patients with COVID-19 when prescribing Remdesivir therapy.

Materials and methods. This prospective open-label observational study was conducted between November 2021 and February 2022, including 154 patients receiving Remdesivir therapy. Patients were divided into two groups: group 1 (n = 45), in which patients developed signs of liver damage during Remdesivir therapy; group 2 (n = 109) – patients without this adverse reaction. All patients underwent pharmacogenetic study and retrospective analysis of medical histories, database with the results of the conducted studies was formed, basing on which machine learning models for predicting the risk of drug-induced liver damage were trained.

Results. The main prognostic factors included body mass index (relevance – 12.03 %) and carriage of AG genotype at polymorphic marker rs776746 of CYP3A5 gene (relevance – 10.04 %). Subsequently, for all obtained factors and based on Categorical boosting a model for predicting the development of drug-induced liver damage with 57.8 % sensitivity and specificity of 80.7 % was developed.

Conclusions. A risk model for the development of drug-induced liver damage during remdesivir therapy was built using machine learning. Body mass index and carriage of AG genotype at polymorphic marker rs776746 of CYP3A5 gene turned out to be key markers. To improve the accuracy of the model, an increase in the proportion of patients with adverse reactions in the training sample is required. Further studies will improve the quality of the model and integrate it into clinical practice.

Key words: COVID-19, remdesivir, adverse drug reactions, hepatotoxicity, pharmacogenetic testing, predictors of adverse reactions, machine learning, risk model

For citation: Shevchuk Yu.V., Shamigulov I.I., Sychev I.V., Kryukov A.V., Temirbulatov I.I., Mirzaev K.B., Denisenko N.P., Abdullaev Sh.P., Tuchkova S.N., Vechorko V.I., Averkov O.V., Sychev D.A. Predicting the risk of developing drug-induced liver injury during remdesivir therapy in COVID-19 patients using machine learning. *Acta biomedica scientifica*. 2024; 9(6): 52-62. doi: 10.29413/ABS.2024-9.6.6

Received: 04.07.2024

Accepted: 15.11.2024

Published: 28.12.2024

ВВЕДЕНИЕ

Пандемия COVID-19 привела к необходимости поиска высокоэффективных противовирусных препаратов. Широкое применение в клинической практике получил противовирусный препарат ремдесивир, одобренный для лечения лёгкой и средней степени тяжести коронавирусной инфекции, однако имеется мало данных о его безопасности, фармакокинетических свойствах и лекарственном взаимодействии при лечении данного заболевания [1].

По данным ряда исследований, распространённость развития нежелательных реакций при применении ремдесивира варьирует от 12 % до 66,2 % [2–5]. При этом распространённость развития тяжёлых нежелательных реакций достигает 13,8 % [4]. Наиболее характерной нежелательной реакцией при терапии ремдесивиром является повышение уровня печёночных трансаминаз, свидетельствующее о повреждении печени; её распространённость составляет от 2 % [2] до 42,9 % [4].

Одним из шагов к повышению безопасности терапии ремдесивиром является выявление предикторов развития нежелательных реакций.

Ранее указанным коллективом авторов было проведено исследование по выявлению предикторов развития лекарственного поражения печени при приёме Ремдесивира у пациентов с COVID-19, в котором были определены конкретные факторы в рамках однофакторного статистического анализа: высокий индекс массы тела; наличие сахарного диабета; высокий уровень ферритина в крови; терапия ингибиторами ангиотензин-превращающего фермента, статинами; наличие гетерозиготы по полиморфному маркеру *rs776746* гена *CYP3A5*. Наличие этих факторов повышает вероятность возникновения лекарственного поражения печени (ЛПП) [6].

Однако для оптимизации процесса выявления пациентов с риском возникновения нежелательных реакций и уменьшения трудозатрат медицинских работников необходима автоматизация, которую может обеспечить использование машинного обучения с построением моделей риска развития нежелательных реакций [7, 8].

ЦЕЛЬ ИССЛЕДОВАНИЯ

Разработка модели машинного обучения для прогнозирования риска развития лекарственного поражения печени у пациентов с COVID-19 при назначении препарата ремдесивир в условиях стационара.

МАТЕРИАЛЫ И МЕТОДЫ

Исследование проводилось в соответствии с этическими принципами Хельсинкской декларации Всемирной медицинской ассоциации и было одобрено Локальным этическим комитетом ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России (Протокол № 15

от 16.10.2021). От всех пациентов или их законных представителей было получено добровольное информированное согласие на участие в данном исследовании.

В период с ноября 2021 г. по февраль 2022 г. было проведено проспективное обсервационное открытое исследование. Исследование было проведено на базе ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы». В исследование были включены мужчины и женщины ($n = 154$) в возрасте 18 лет и старше, госпитализированные с подтверждённой новой коронавирусной инфекцией (COVID-19) (U07.1, U07.2 по Международной классификации болезней 10-го пересмотра (МКБ-10)).

Препарат ремдесивир использовали в стандартной дозе: 200 мг внутривенно в первые сутки, затем – 100 мг 1 раз в сутки в течение 5–10 дней. Исследователь не мог влиять на выбор противовирусного препарата и длительность терапии.

Возраст всех пациентов составил от 44 до 96 лет, средний возраст – $72,1 \pm 13,5$ года. Из 154 пациентов, включённых в исследование, было 58 (38 %) мужчин со средним возрастом $65,3 \pm 13,4$ года и 96 (62 %) женщин со средним возрастом $76,2 \pm 11,8$ года.

Критерии включения в исследование: установленный диагноз новой коронавирусной инфекции (COVID-19) (U07.1, U07.2 по МКБ-10); подписанное добровольное информированное согласие на участие в исследовании; длительность госпитализации > 48 ч; использование ремдесивира в качестве этиотропной терапии.

Критерии невключения: противопоказания к терапии ремдесивиром в соответствии с официальной инструкцией по применению, действующей на момент проведения исследования – скорость клубочковой фильтрации (СКФ) < 30 мл/мин/1,73 м², беременность, период грудного вскармливания, повышение уровня аланинаминотрансферазы (АЛТ) свыше 5 верхних границ нормы; тяжёлая печёночная недостаточность (класс С по Чайлду – Пью); досрочный выход из исследования по желанию пациента. Критерии невключения, за исключением досрочного выхода из исследования по желанию пациента, применялись до начала терапии ремдесивиром.

Демографические данные пациентов представлены в таблице 1.

В дальнейшем, с учётом цели исследования, пациенты были разделены на две группы в зависимости от возникновения нежелательных реакций (НР) в виде поражения печени на фоне терапии ремдесивиром. Группа 1 – пациенты, у которых развилась данная НР; группа 2 – пациенты, у которых не было выявлено поражения печени на фоне терапии ремдесивиром (табл. 2).

Сравниваемые группы были сопоставимы по полу ($\chi^2 = 0,8709$; $df = 1$; $p = 0,3507$) и возрасту ($W = 1894,5$; $p = 0,0267$). Также группы были сопоставимы по тяжести состояния и таким имеющимся нозологиям, как хроническая болезнь почек, наличие активного рака, энцефалопатии и сердечно-сосудистых заболеваний, среди которых в отдельную группу были выделены ишемическая болезнь сердца, хроническая сердечная недоста-

ТАБЛИЦА 1
ДЕМОГРАФИЧЕСКИЕ ДАННЫЕ ПАЦИЕНТОВ,
ВКЛЮЧЁННЫХ В ИССЛЕДОВАНИЕ

Пол	Количество пациентов, <i>n</i>	Доля от общего, %	Средний возраст, лет
Мужчины	58	38	65,3 ± 13,4
Женщины	96	62	76,2 ± 11,8
Все пациенты	154	100	72,1 ± 13,5

TABLE 1
DEMOGRAPHIC DATA OF PATIENTS INCLUDED
IN THE STUDY

ТАБЛИЦА 2
ДЕМОГРАФИЧЕСКИЕ ДАННЫЕ КАЖДОЙ ГРУППЫ
ПАЦИЕНТОВ

Показатели	Группа 1 (<i>n</i> = 45) – 29,22 %	Группа 2 (<i>n</i> = 109) – 70,78 %	$\chi^2; p$
Средний возраст, лет	71,9 ± 13,6	72,1 ± 13,5	–; 0,2672
Мужчины, <i>n</i> (%)	20 (44,44 %)	38 (34,86 %)	0,8709; 0,3507
Женщины, <i>n</i> (%)	25 (55,56 %)	71 (65,14 %)	

TABLE 2
DEMOGRAPHIC DATA OF EACH GROUP OF PATIENTS

точность и артериальная гипертензия. По сопутствующей лекарственной терапии – применение бета-блокаторов, блокаторов кальциевых каналов, блокаторов рецепторов ангиотензина II, диуретиков, нестероидных противовоспалительных средств, салицилатов, антипсихотиков, прокинетики, ингибиторов протонной помпы, бигуанидов, антибактериальных и противогрибковых препаратов, системных глюкокортикостероидов и эноксапарина натрия – сравниваемые группы также были сопоставимы [6].

Всем пациентам в рамках стационарного лечения COVID-19 проводились лабораторные и инструментальные исследования в соответствии с временными методическими рекомендациями по профилактике, диагностике и лечению новой коронавирусной инфекции (COVID-19) от 14.10.2021 и 27.12.2021 и клиническими рекомендациями по другим имеющимся у пациентов нозологиям. Кроме того, пациентам было проведено молекулярно-генетическое исследование для установления генотипа по полиморфному маркеру *rs776746* гена *CYP3A5*, генотипа по полиморфному маркеру *rs35599367* гена *CYP3A4* и генотипа по полиморфному маркеру *rs2244613* гена *CES1*. У пациентов был произведён забор 10 мл венозной крови с помощью вакуумной системы Vacuette® (Greiner Bio-One, Австрия) в пробирки с КЗ-ЭДТА (этилендиаминтетраацетатом). Цельная кровь до выделения ДНК и выделенная ДНК хранились при –80 °С, транспортировку образцов до лаборатории осуществляли при –20 °С. Генотипирование проводилось на базе Научно-исследовательского института молекулярной и персонализированной медицины ФГБОУ ДПО РМАНПО Минздрава России. Выделение геномной ДНК из цельной крови осуществлялось с помощью набора реагентов «S-Сорб» для выделения ДНК на кремниевом сорбенте (ООО «Синтол», Россия). Концентрация экстрагированной ДНК определялась с помо-

щью спектрофотометра NanoDrop 2000 (Thermo Fisher Scientific, США). Определение варианта генотипа по полиморфизму *rs35599367* (*22, *c.522-191C>T*) гена *CYP3A* проводилось методом аллель-специфической полимеразной цепной реакции (ПЦР) в режиме реального времени с использованием коммерческих наборов TaqMan® SNP Genotyping Assays и TaqMan Universal Master Mix II, no UNG (Applied Biosystems, США). Носительство варианта генотипа по полиморфному маркеру *rs776746* (*3, *c.219-237A>G*) гена *CYP3A5* определялось с помощью коммерческого набора реагентов (ООО «Синтол», Россия). Носительство варианта генотипа по полиморфному маркеру *rs2244613* (1165-33C>AA>C) гена *CES1* выявлялось с помощью наборов реагентов «ГенТест *CES1*» (ООО «Номотек», Россия). ПЦР с детекцией в режиме реального времени проводилась на амплификаторе Real-Time CFX96 Touch (Bio-Rad Laboratories Inc., США).

Результаты данного исследования были занесены в таблицу при помощи программы Microsoft Office Excel 2019 (Microsoft Corp., США), где была сформирована база данных исследуемых пациентов. В неё входили 14 лабораторных данных при поступлении пациента в стационар, 19 данных о сопутствующей фармакотерапии, 8 генетических данных, 10 – для шкалы ADRROP, 11 общих данных (демографические, клинические и лабораторные), 11 данных о сопутствующих заболеваниях, 1 целевая переменная (НР в виде ЛПП).

На основании полученных результатов проведённого исследования, а также сформированной базы данных происходило машинное обучение моделей для прогноза риска развития НР в виде поражения печени при терапии препаратом ремдесивир у пациентов с COVID-19 в условиях стационара.

Вся работа с данными, а также обучение искусственного интеллекта проходили в Jupyter notebook в среде conda. Для работы с данными, таблицами и визуализа-

ции диаграмм и графиков были использованы библиотеки следующих версий:

- Pandas 1.3.4;
- Numpy 1.26.4;
- Matplotlib 3.4.3;
- Scikit-learn (sklearn) 1.4.2;
- Catboost 1.1.1;
- XGBoost 1.7.1.

Для поиска наилучшего алгоритма машинного обучения и подбора гиперпараметров к нашим моделям был использован метод поиска по сетке с кросс-валидацией (GridSearchCV) из библиотеки sklearn. Остальные методы, применяемые в работе, за исключением Catboost и XGBoost, были заимствованы из библиотеки sklearn.

В связи с небольшим количеством данных и неравномерно распределёнными классами в целевой переменной обучение проводили с применением метода взвешенных классов, идея которого заключается в присвоении наименее представленному классу (пациенты с НР) большего веса пропорционально отношению количеств представителей при обучении модели. Также обучение проводилось с использованием кросс-валидации для оценки качества модели на комбинациях разных данных, которая хорошо применима для случаев с ограниченным количеством данных. Вместе с тем использовалась стратификация, позволяющая сохранить соотношение количеств представителей разных классов в обучающих и тестовых выборках таким же, как в изначальном наборе данных.

Оценка качества моделей

Для поиска оптимального предсказания конечной точки (НР в виде поражения печени на фоне терапии ремдесивиром) обучались и оценивались модели, основывающиеся на различных алгоритмах машинного обучения для решения задачи классификации: метод k ближайших соседей (KNN, k-nearest neighbors algorithm) [9]; линейный метод – логистическая регрессия [10]; метод, строящий разделяющую гиперплоскость – метод опорных векторов (SVC, support vector classification) [11]; решающие деревья [12]; ансамбль случайных деревьев, ре-

зультатом которого является среднее по деревьям обученных на случайных комбинациях факторов [13]; категориальный бустинг (CatBoost) – градиентный бустинг, последовательные деревья, обученные на ошибках предыдущего по идее градиентного спуска, имеющий преимущества в обучении на категориальных данных [14].

Для оценки эффективности созданной модели использовались такие метрики, как ROC-AUC (Receiver Operating Characteristic Area Under the Curve) и

$$F1score = \frac{2 \times precision \times recall}{precision + recall}$$

где: $precision = \frac{TP}{TP + FP}$;

$$recall = \frac{TP}{TP + FN}$$

$$TPR = \frac{TP}{TP + FN} = \text{чувствительность};$$

$$FRP = \frac{FP}{FP + TN} = (1 - \text{специфичность});$$

TN – true negative; TP – true positive; FP – false positive; FN – false negative [15].

РЕЗУЛЬТАТЫ

Основные этапы работы представлены на рисунке 1. Для составления моделей машинного обучения нами были отобраны 74 параметра, среди которых 53 категориальных и 21 непрерывный, отобранные таким образом, чтобы не было коррелирующих между собой и с конечной точкой (их оценка происходила при поступлении пациента в стационар), и 1 конечная точка (развитие НР), которая оценивалась во время терапии ремдесивиром и после её окончания.

В дальнейшем был проведён разведывательный анализ данных.

Разведывательный анализ данных

Перед проведением обучения моделей искусственного интеллекта (ИИ) предварительно проверены дан-

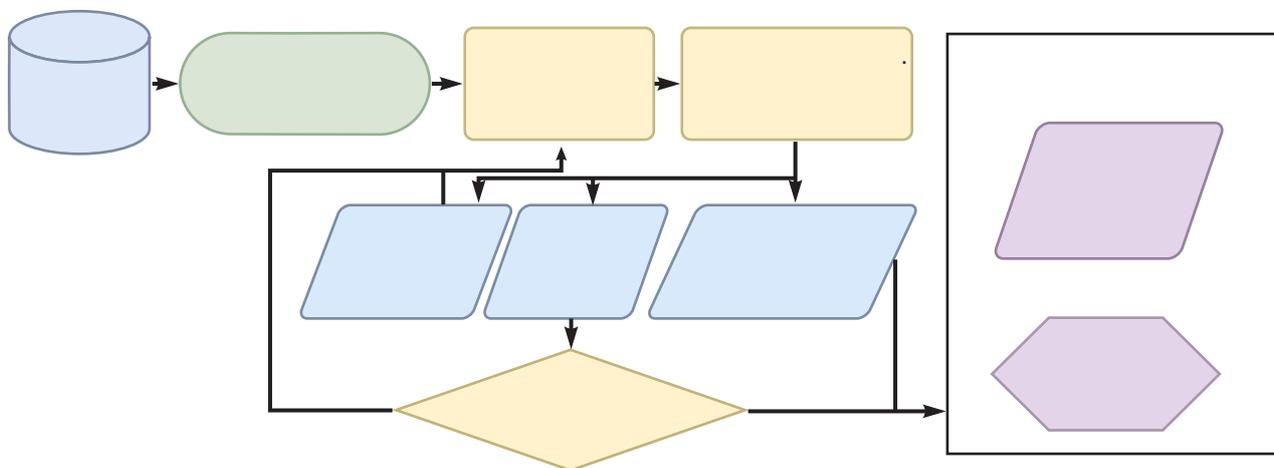


РИС. 1.
Основные этапы обучения моделей

FIG. 1.
The main stages of training models

ные на наличие выбросов, потерянных значений, а также на наличие корреляции между параметрами (рис. 2). Выбросов и пропусков обнаружено не было. Параметры, корреляция которых составляла более 0,7, были признаны взаимоисключающими. В конечном итоге для обучения моделей было использовано 59 параметров, 19 из которых являлись непрерывными и 40 – категориальными.

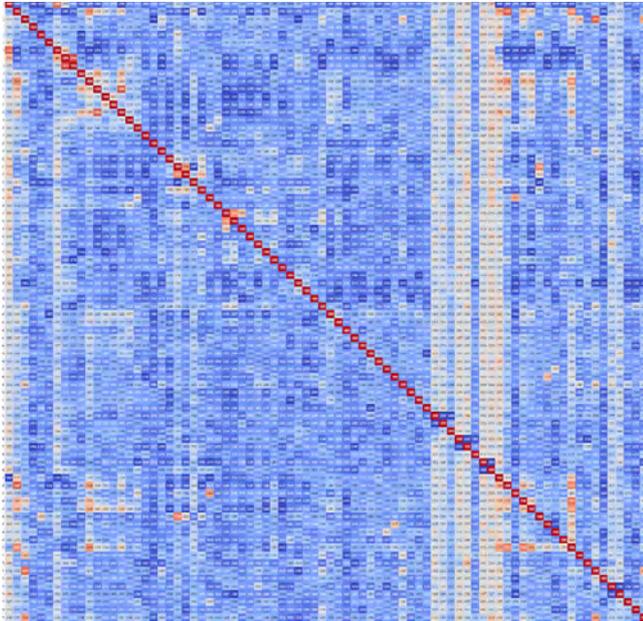


РИС. 2.
Матрица корреляции между параметрами
FIG. 2.
The correlation matrix between the parameters

Поскольку поиск решения происходил на моделях разного типа, нормализация проводилась для непрерывных переменных только для предподготовки к обучению линейных и метрических моделей, так как для них это играет большую роль. К ним относятся метод KNN, логистическая регрессия, метод SVC.

По результатам оценки F1 для указанной конечной точки наиболее подходящей явилась модель ИИ на основе категориального бустинга – CatBoost (рис. 3).

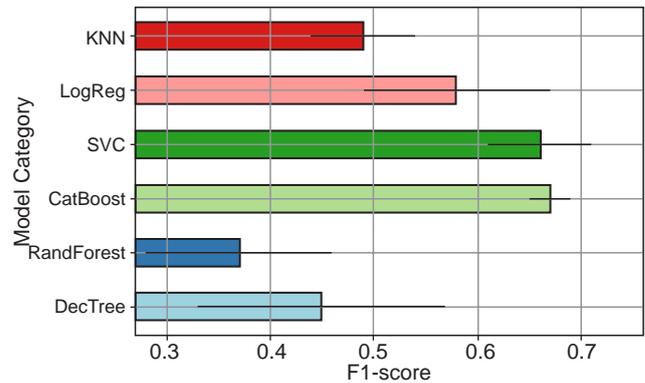


РИС. 3.
Значения F1-метрики для целевой переменной
FIG. 3.
F1 metric values for the target variable

Для предсказания исхода были выделены факторы, имеющие наибольшую прогностическую ценность (рис. 4).

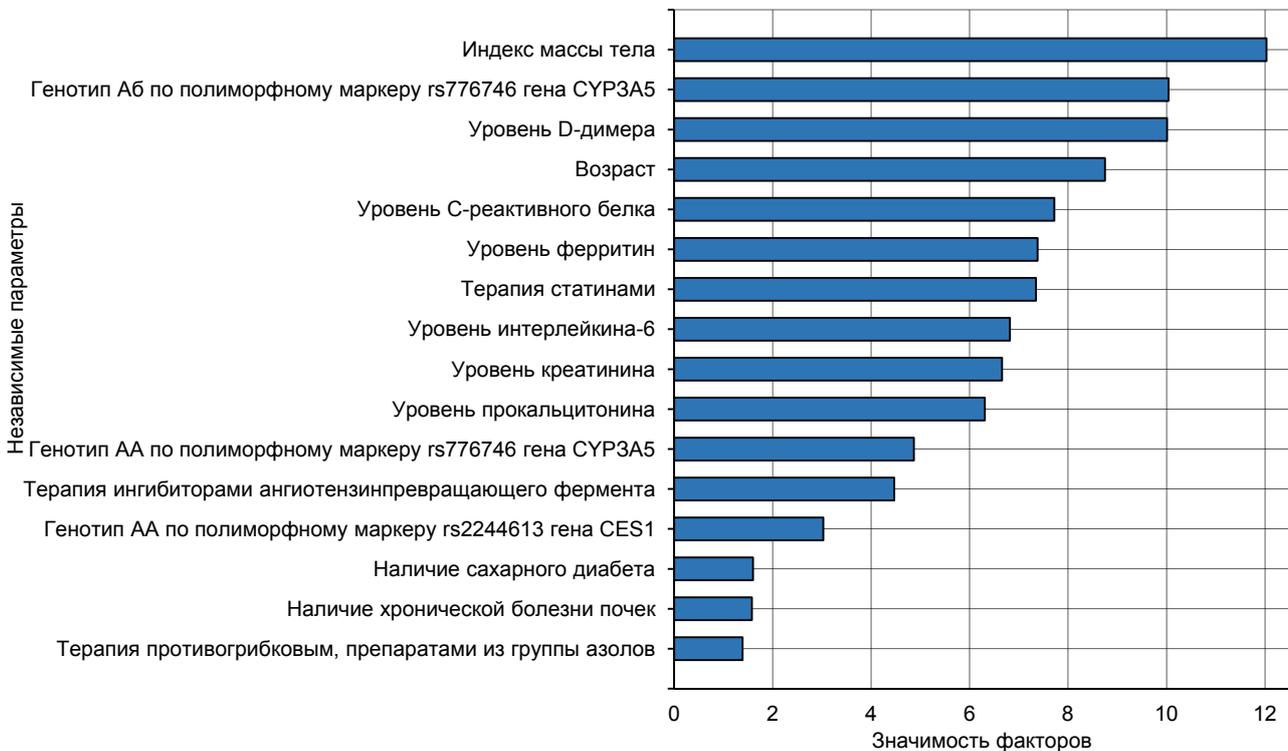


РИС. 4.
Факторы, имеющие наибольшую прогностическую ценность

FIG. 4.
Factors with the highest prognostic value

При обучении модели было выявлено, что в предсказание поражения печени наибольший вклад внесли следующие факторы, представленные в таблице 3: индекс массы тела (ИМТ) (12,03 %); носительство генотипа AG по полиморфному маркеру rs776746 гена CYP3A5 (10,04 %); уровень D-димера (10,01 %); возраст (8,75 %); уровень С-реактивного белка (7,72 %); уровень ферритина (7,38 %); терапия статинами (7,35 %); уровень интерлейкина 6 (6,82 %); уровень креатинина (6,66 %); уровень прокальцитонина (6,31 %); носительство генотипа AA по полиморфному маркеру rs776746 гена CYP3A5 (4,87 %); терапия ингибиторами ангиотензин-превращающего фермента (4,47 %). Остальные факторы также представлены в таблице 3.

По полученным факторам построена окончательная модель, представленная на рисунке 5 в виде матрицы ошибок. Согласно этой модели, вероятность верного предсказания поражения печени соответствует 57,8%, вероятность отсутствия поражения печени – 80,7%. Чувствительность данного метода составляет 57,8%, специфичность – 80,7%.

На рисунке 6 представлено качество прогностической модели, рассчитанное через ROC-AUC.

Методом кросс-валидации оценены среднее значение и погрешность выполнения прогноза. Среднее значение ROC-AUC = 0,68, что превышает показатель случайного классификатора (0,5), в связи с чем модель пригодна для предсказаний. Также отмечено, что погрешность имеет значение 0,02 – это свидетельствует о стабильно-

сти модели, погрешность же может возникать из-за небольших различий в факторах для предсказаний. Следует улучшить чувствительность модели для лучшего предсказания положительных результатов; этого можно достичь, увеличив долю пациентов с НР в обучающей выборке.

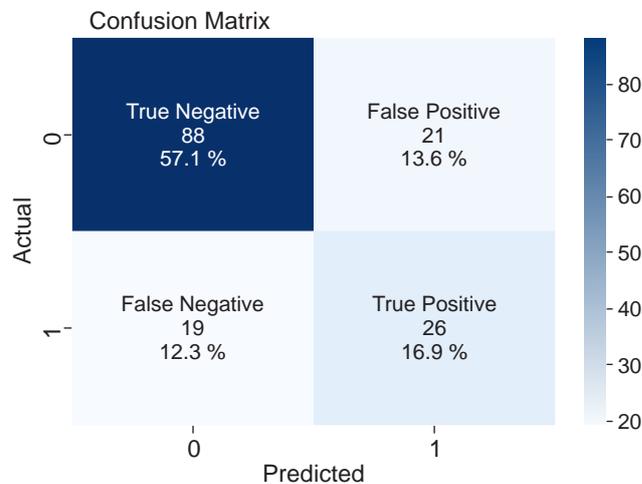


РИС. 5. Матрица ошибок модели предсказания лекарственного поражения печени

FIG. 5. Confusion matrix of the drug-induced liver injury prediction model

ТАБЛИЦА 3
ЗНАЧИМОСТЬ ФАКТОРОВ В ПРЕДСКАЗАНИИ ПОРАЖЕНИЯ ПЕЧЕНИ (%)

Факторы	Важность, %
Терапия противогрибковыми препаратами из группы азолов	1,39
Наличие хронической болезни почек	1,58
Наличие сахарного диабета	1,6
Генотип AA по полиморфному маркеру rs2244613 гена CES1	3,03
Терапия ингибиторами ангиотензинпревращающего фермента	4,47
Генотип AA по полиморфному маркеру rs776746 гена CYP3A5	4,87
Уровень прокальцитонина	6,31
Уровень креатинина	6,66
Уровень интерлейкина 6	6,82
Терапия статинами	7,35
Уровень ферритина	7,38
Уровень С-реактивного белка	7,72
Возраст	8,75
Уровень D-димера	10,01
Генотип AG по полиморфному маркеру rs776746 гена CYP3A5	10,04
Индекс массы тела	12,03

TABLE 3
SIGNIFICANCE OF FACTORS IN PREDICTION OF LIVER DAMAGE (%)

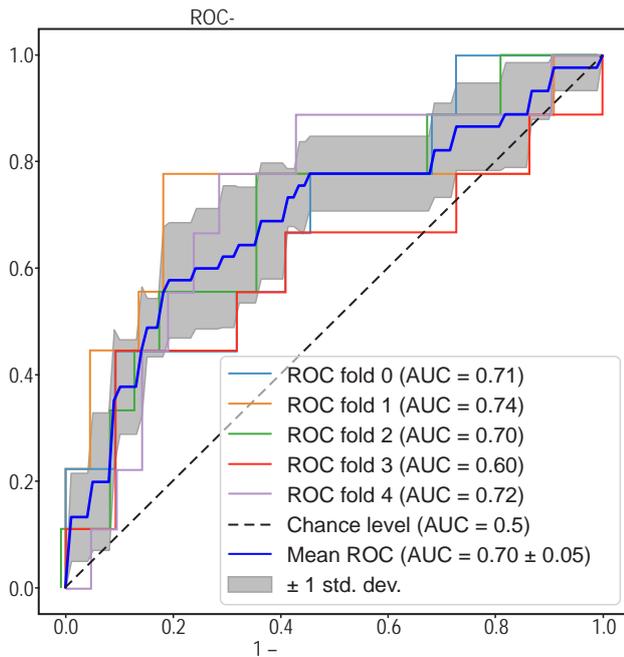


РИС. 6.
 Качество прогностической модели, рассчитанное через ROC-AUC

FIG. 6.
 Quality of the predictive model calculated through ROC-AUC

ОБСУЖДЕНИЕ

В настоящий момент разработаны и внедрены в практику шкалы прогнозирования риска развития НР у пациентов пожилого возраста, такие как ADRROP (Adverse Drug Reaction Risk in Older Persons) и GerontoNet. Эти шкалы имеют балльную оценку и включают в себя факторы риска НР [16, 17]. Также была разработана и валидирована модель прогнозирования риска развития НР у пожилых пациентов PADROI (Prediction of Adverse Drug Reactions in Older Inpatients), имеющая специфичность 90,3 % и чувствительность 87,1 % [18]. Эти шкалы имеют большую прогностическую ценность, позволяя своевременно принять меры для предотвращения развития НР. Также имеются исследования, посвященные прогнозированию развития НР лекарственных препаратов с использованием методов машинного обучения [19–21]; в свою очередь, при анализе литературных источников нами было найдено исследование, в котором проводилось выявление предикторов развития НР и летального исхода при терапии ремдесивиром [22], однако мы не обнаружили данных прогнозирования риска развития НР при терапии этим препаратом. С учётом рисков развития НР при фармакотерапии у госпитализированных пациентов с COVID-19, связанных с наличием у них мультиморбидности и полипрагмазии [23, 24], высокой распространённости НР при терапии ремдесивиром в целом и распространением тяжёлых НР [4], крайне важно использовать инструменты, позволяющие предсказать и благодаря этому снизить риск неблагоприятных исходов терапии ремдесивиром. К тому же предот-

вращение развития НР может снизить продолжительность госпитализации и связанные с этим экономические затраты [25].

В нашем исследовании мы получили модель, позволяющую предсказать развитие ЛПП при терапии ремдесивиром с точностью до 57,8 %. Её эффективность составила 68 % (среднее значение ROC-AUC = 0,68), модель пригодна для предсказаний и является стабильной. Однако для лучшего предсказания положительных результатов можно повысить чувствительность модели с помощью увеличения доли пациентов с НР в обучающей выборке.

Согласно полученной модели, ключевыми маркерами для развития лекарственно индуцированного поражения печени оказались ИМТ (12,03 %) и носительство генотипа AG по полиморфному маркеру *rs776746* гена *CYP3A5* (10,04 %).

Так, ферменты *CYP3A4/5* являются одними из ферментов, участвующих в метаболизме ремдесивира в неактивный метаболит [26]. Вариант *CYP3A5*1/*3*, определяемый интронным вариантом (NM_000777.5: c.219-237A>G; *rs776746*), ассоциирован со сниженной метаболической активностью фермента (исторически известный как неэкспрессорный фенотип) [27], в связи с чем носительство генотипа AG по данному полиморфному маркеру может привести к повышению плазменной концентрации ремдесивира и увеличению риска развития НР. Более высокий ИМТ и, в частности, ожирение являются одним из факторов риска тяжёлого течения COVID-19 [28] и часто сопровождаются другими заболеваниями [29], в том числе неалкогольной жировой болезнью печени [30]. Все эти составляющие могут приводить к повышению риска развития НР при терапии ремдесивиром. В нашем исследовании мы выявили, что более высокий ИМТ прямо пропорционально увеличивает вероятность развития ЛПП при терапии ремдесивиром.

Значения чувствительности и специфичности разработанной модели машинного обучения для прогноза риска развития НР в виде поражения печени у пациентов с COVID-19 на фоне этиотропного лечения препаратом ремдесивир дают основания полагать, что требуется увеличение доли пациентов с НР в тренировочной выборке и указывают на необходимость продолжения исследования с целью улучшения качества модели и её внедрения в клиническую практику в дальнейшем.

ЗАКЛЮЧЕНИЕ

При помощи машинного обучения была сформирована модель прогноза риска развития НР в виде поражения печени у пациентов с COVID-19 на фоне этиотропного лечения препаратом ремдесивир в условиях стационара. По результатам оценки F1 для указанной конечной точки, наиболее подходящей явилась модель ИИ на основе категориального бустинга – CatBoost. Ключевыми маркерами для развития лекарственно индуци-

рованного поражения печени оказались ИМТ (12,03 %) и носительство генотипа AG по полиморфному маркеру rs776746 гена CYP3A5 (10,04 %).

Вид ROC-кривой и значение ROC-AUC дают основания полагать, что модель применима и эффективна, а также не подвержена большим флуктуациям.

Для улучшения чувствительности модели, сформированной при помощи машинного обучения, требуется увеличение доли пациентов с НР в тренировочной выборке, что указывает на необходимость продолжения исследования для улучшения качества модели и её дальнейшего внедрения в клиническую практику.

Финансирование

Данная работа выполнена при финансовой поддержке Министерства здравоохранения Российской Федерации, тематика государственного задания «Разработка системы поддержки принятия врачебных решений для прогнозирования нежелательных лекарственных реакций у пациентов с COVID-19 на основе фармакогенетического тестирования» (ЕГИСУ НИОКТР № 122021800321-2).

Конфликт интересов

Авторы данной статьи заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА / REFERENCES

1. *Временные методические рекомендации по профилактике, диагностике и лечению новой коронавирусной инфекции (COVID-19)*. М.; 2023. [Temporary methodological recommendations for the prevention, diagnosis, and treatment of novel coronavirus infection (COVID-19)]. Moscow; 2023. (In Russ.).
2. Gilead Sciences Biopharmaceutical Companies, Veklury (remdesivir). *U.S. Food and Drug Administration*. 2022. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/214787Orig1s010Lbl.pdf. [date of access: 20.05.2024].
3. Pantazis N, Pechlivanidou E, Antoniadou A, Akinosoglou K, Kalomenidis I, Poulakou G, et al. Remdesivir: Effectiveness and safety in hospitalized patients with COVID-19 (ReEs-COVID-19) – Analysis of data from daily practice. *Microorganisms*. 2023; 11(8): 1998. doi: 10.3390/microorganisms11081998
4. Kang H, Kang CK, Im JH, Cho Y, Kang DY, Lee JY. Adverse drug events associated with remdesivir in real-world hospitalized patients with COVID-19, including vulnerable populations: A retrospective multicenter study. *J Korean Med Sci*. 2023; 38(44): e346. doi: 10.3346/jkms.2023.38.e346
5. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: A randomised, double-blind, placebo-controlled, multicentre trial. *Lancet*. 2020; 395(10236): 1569-1578. doi: 10.1016/S0140-6736(20)31022-9
6. Шевчук Ю.В., Крюков А.В., Темирбулатов И.И., Сычев И.В., Мирзаев К.Б., Денисенко Н.П., и др. Модель прогнозирования риска развития лекарственного поражения печени на фоне терапии ремдесивиром: обсервационное проспективное открытое контролируемое исследование. *Фармация и фармакология*. 2023; 11(3): 228-239. [Shevchuk YuV, Kryukov AV, Temirbulatov II, Sychev IV, Mirzaev KB, Denisenko NP, et al. Model for predicting risk of developing drug-induced liver injury during remdesivir therapy: Observational prospective open case-control study. *Pharmacy & Pharmacology*. 2023; 11(3): 228-239. (In Russ.).] doi: 10.19163/2307-9266-2023-11-3-228-239
7. Falconer N, Barras M, Cottrell N. Systematic review of predictive risk models for adverse drug events in hospitalized patients. *Br J Clin Pharmacol*. 2018; 84: 846-864. doi: 10.1111/bcp.13514
8. Salas M, Petracek J, Yalamanchili P, Aimer O, Kasthuril D, Dhingra S, et al. The use of artificial intelligence in pharmacovigilance: A systematic review of the literature. *Pharm Med*. 2022; 36(5): 295-306. doi: 10.1007/s40290-022-00441-z
9. Goldberger J, Roweis ST, Hinton GE, Salakhutdinov R. *Neighbourhood components analysis*. 2004: 513-520.
10. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*. 2013.
11. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*. 2020; 408: 189-215. doi: 10.1016/j.neucom.2019.10.118
12. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. 2017.
13. Biau G, Scornet E. A random forest guided tour. *TEST*. 2016; 25(1): 197-227. doi: 10.1007/s11749-016-0481-7
14. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulina A. CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*. 2018; 31.
15. Hossain M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. *Int J Data Min Knowl Manag Process*. 2015; 5(2): 1. doi: 10.5121/ijdkp.2015.5201
16. O'Mahony D, O'Connor MN, Eustace J, Byrne S, Petrovic M, Gallagher P. The adverse drug reaction risk in older persons (ADRROP) prediction scale: Derivation and prospective validation of an ADR risk assessment tool in older multi-morbid patients. *Eur Geriatr Med*. 2018; 9(2): 191-199. doi: 10.1007/s41999-018-0030-x
17. Lavan A, Eustace J, Dahly D, Flanagan E, Gallagher P, Cullinane S, et al. Incident adverse drug reactions in geriatric inpatients: A multicentred observational study. *Ther Adv Drug Saf*. 2018; 9(1): 13-23. doi: 10.1177/2042098617736191
18. Yadesa TM, Kitutu FE, Tamukong R, Alele PE. Development and validation of 'Prediction of Adverse Drug Reactions in Older Inpatients (PADROI)' risk assessment tool. *Clin Interv Aging*. 2022; 17: 195-210. doi: 10.2147/CIA.S350500
19. Zhang F, Sun B, Diao X, Zhao W, Shu T. Prediction of adverse drug reactions based on knowledge graph embedding. *BMC Med Inform Decis Mak*. 2021; 21: 1-11. doi: 10.1186/s12911-021-01402-3
20. Galeano D, Li S, Gerstein M, Paccanaro A. Predicting the frequencies of drug side effects. *Nat Commun*. 2020; 11(1): 4575. doi: 10.1038/s41467-020-18305-y
21. Choudhury O, Park Y, Salonidis T, Gkoulalas-Divanis A, Sylla I, Das AK. Predicting adverse drug reactions on distributed health data using federated learning. *AMIA Annu Symp Proc*. 2019; 2019: 313-322.
22. Ayyashi M, Darbashi H, Hakami A, Sharahili F. Evaluation of remdesivir utilization pattern in critically ill patients with COVID-19 in Jazan Province. *Cureus*. 2023; 15(3): e36247. doi: 10.7759/cureus.36247

23. Iloanusu S, Mgbere O, Essien EJ. Polypharmacy among COVID-19 patients: A systematic review. *J Am Pharm Assoc.* 2021; 61(5): e14-e25. doi: 10.1016/j.japh.2021.05.006

24. Lee JY, Ang ASY, Mohd Ali N, Ang LM, Omar A. Incidence of adverse reaction of drugs used in COVID-19 management: A retrospective, observational study. *J Pharm Policy Pract.* 2021; 14: 1-9. doi: 10.1186/s40545-021-00370-3

25. Sendekie AK, Kasahun AE, Limenh LW, Dagnaw AD, Belachew EA. Clinical and economic impact of adverse drug reactions in hospitalised patients: Prospective matched nested case-control study in Ethiopia. *BMJ Open.* 2023; 13: e073777. doi: 10.1136/bmjopen-2023-073777

26. Blair HA. Remdesivir: A review in COVID-19. *Drugs.* 2023; 83(13): 1215-1237. doi: 10.1007/s40265-023-01926-0

27. Pratt VM, Cavallari LH, Fulmer ML, Gaedigk A, Hachad H, Ji Y, et al. CYP3A4 and CYP3A5 genotyping recommendations: A joint consensus recommendation of the association for molecular pathology, clinical pharmacogenetics implementation consor-

tium, College of American Pathologists, Dutch Pharmacogenetics Working Group of the Royal Dutch Pharmacists Association, European Society for Pharmacogenomics and Personalized Therapy, and Pharmacogenomics Knowledgebase. *J Mol Diagn.* 2023; 25(9): 619-629. doi: 10.1016/j.jmoldx.2023.06.008

28. Buscemi S, Corleo D, Randazzo C. Risk factors for COVID-19: Diabetes, hypertension, and obesity. *Coronavirus Therapeutics, Volume II: Clinical Management and Public Health.* 2022; 115-129. doi: 10.1007/978-3-030-85113-2_7

29. Zhang X, Ha S, Lau HCH, Yu J. Excess body weight: Novel insights into its roles in obesity comorbidities. *Semin Cancer Biol.* 2023; 92: 16-27. doi: 10.1016/j.semcancer.2023.03.008

30. Quek J, Chan KE, Wong ZY, Tan C, Tan B, Lim WH, et al. Global prevalence of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in the overweight and obese population: A systematic review and meta-analysis. *Lancet Gastroenterol Hepatol.* 2023; 8(1): 20-30. doi: 10.1016/S2468-1253(22)00317-X

Сведения об авторах

Шевчук Юлия Викторовна – клинический фармаколог, ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы»; аспирант кафедры клинической фармакологии и терапии им. академика Б.Е. Вотчала, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: ju-viktorovna@yandex.ru, <https://orcid.org/0000-0003-2289-1900>

Шамигулов Искандер Ильгамович – инспектор Научно-исследовательского института молекулярной и персонализированной медицины, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России; студент физтех-школы биологической и медицинской физики, ФГАОУ ВО «Московский физико-технический институт (национальный исследовательский университет)», e-mail: shamigulov.ii@phystech.edu, <https://orcid.org/0009-0003-5383-1713>

Сычев Иван Витальевич – младший научный сотрудник Научно-исследовательского института молекулярной и персонализированной медицины, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России; аспирант кафедры факультетской терапии, ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н.П. Огарева», e-mail: sychev_iv@bk.ru, <https://orcid.org/0000-0003-0227-2651>

Крыков Александр Валерьевич – кандидат медицинских наук, заведующий отделом клинической фармакологии, ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы»; доцент кафедры клинической фармакологии и терапии им. академика Б.Е. Вотчала, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: alex.kryukov90@yandex.ru, <https://orcid.org/0000-0001-7903-2977>

Темирбулатов Ильяс Ильдарович – аспирант кафедры клинической фармакологии и терапии им. академика Б.Е. Вотчала, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: temirbulatov.ilyas@gmail.com, <https://orcid.org/0000-0002-1242-0833>

Мирзаев Карин Бадаевич – доктор медицинских наук, доцент, проректор по научной работе и инновациям, профессор кафедры клинической фармакологии и терапии им. академика Б.Е. Вотчала, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: karin05doc@yandex.ru, <https://orcid.org/0000-0002-9307-4994>

Денисенко Наталья Павловна – кандидат медицинских наук, заместитель директора Научно-исследовательского института молекулярной и персонализированной медицины, доцент кафедры клинической фармакологии и терапии им. академика Б.Е. Вотчала, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: natalypilipenko3990@gmail.com, <https://orcid.org/0000-0003-3278-5941>

Абдуллаев Шерзод Пардабоевич – кандидат биологических наук, заведующий отделом предиктивных и прогностических биомаркеров Научно-исследовательского института молекулярной и персонализированной медицины, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: abdullaevsp@gmail.com, <https://orcid.org/0000-0001-9001-1499>

Тучкова Светлана Николаевна – младший научный сотрудник Научно-исследовательского института молекулярной и персонализированной медицины, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: svetlana.tuch1998@gmail.com, <https://orcid.org/0009-0001-2744-2752>

Вечорко Валерий Иванович – доктор медицинских наук, главный врач, ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы»; профессор кафедры организации здравоохранения и общественного здоровья, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: gkb15@zdrav.mos.ru, <https://orcid.org/0000-0003-3568-5065>

Аверков Олег Валерьевич – доктор медицинских наук, профессор, заместитель главного врача, руководитель регионального сосудистого центра, ГБУЗ г. Москвы «Городская клиническая больница № 15 имени О.М. Филатова Департамента здравоохранения города Москвы», e-mail: oleg.averkov@gmail.com, <https://orcid.org/0000-0002-3010-755X>

Сычев Дмитрий Алексеевич – доктор медицинских наук, профессор, Академик РАН, заведующий кафедрой клинической фармакологии и терапии им. академика Б.Е. Вотчала, и. о. ректора, ФГБОУ ДПО «Российская медицинская академия непрерывного профессионального образования» Минздрава России, e-mail: dmitry.alex.sychev@gmail.com, <https://orcid.org/0000-0002-4496-3680>

Information about the authors

Yuliya V. Shevchuk – Clinical Pharmacologist, Municipal Clinical Hospital No. 15 named after O.M. Filatov; Postgraduate at the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Russian Medical Academy of Continuing Professional Education, e-mail: ju-viktorovna@yandex.ru, <https://orcid.org/0000-0003-2289-1900>

Iskander I. Shamigulov – Inspector at the Research Institute of Molecular and Personalized Medicine, Russian Medical Academy of Continuing Professional Education; Student at the School of Biological and Medical Physics, Moscow Institute of Physics and Technology (National Research University), e-mail: shamigulov.ii@phystech.edu, <https://orcid.org/0009-0003-5383-1713>

Ivan V. Sychev – Junior Research Officer at the Research Institute of Molecular and Personalized Medicine, Russian Medical Academy of Continuing Professional Education; Postgraduate at the Department of Intermediate Level Therapy, National Research Ogarev Mordovia State University, e-mail: sychev_iv@bk.ru, <https://orcid.org/0000-0003-0227-2651>

Alexander V. Kryukov – Cand. Sc. (Med.), Head of the Department of Clinical Pharmacology, Municipal Clinical Hospital No. 15 named after O.M. Filatov; Associate Professor at the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Russian Medical Academy of Continuing Professional Education, e-mail: alex.kryukov90@yandex.ru, <https://orcid.org/0000-0001-7903-2977>

Ilyas I. Temirbulatov – Postgraduate at the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Russian Medical Academy of Continuing Professional Education, e-mail: temirbulatov.ilyas@gmail.com, <https://orcid.org/0000-0002-1242-0833>

Karin B. Mirzaev – Dr. Sc. (Med.), Docent, Vice Rector for Research and Innovation, Professor at the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Russian Medical Academy of Continuing Professional Education, e-mail: karin05doc@yandex.ru, <https://orcid.org/0000-0002-9307-4994>

Natalia P. Denisenko – Cand. Sc. (Med.), Deputy Director of the Research Institute of Molecular and Personalized Medicine, Associate Professor at the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Russian Medical Academy of Continuing Professional Education, e-mail: natalypilipenko3990@gmail.com, <https://orcid.org/0000-0003-3278-5941>

Sherzod P. Abdullaev – Cand. Sc. (Biol.), Head of the Department of Predictive and Prognostic Biomarkers, Research Institute of Molecular and Personalized Medicine, Russian Medical Academy of Continuing Professional Education, e-mail: abdullaevsp@gmail.com, <https://orcid.org/0000-0001-9001-1499>

Svetlana N. Tuchkova – Junior Research Officer at the Research Institute of Molecular and Personalized Medicine, Russian Medical Academy of Continuing Professional Education, e-mail: svetlana.tuch1998@gmail.com, <https://orcid.org/0009-0001-2744-2752>

Valery I. Vechorko – Dr. Sc. (Med.), Chief Physician, Municipal Clinical Hospital No. 15 named after O.M. Filatov; Professor at the Department of Healthcare Organization and Public Health, Russian Medical Academy of Continuing Professional Education, e-mail: gkb15@zdrav.mos.ru, <https://orcid.org/0000-0003-3568-5065>

Oleg V. Averkov – Dr. Sc. (Med.), Professor, Deputy Chief Physician, Head of the Regional Vascular Center, Municipal Clinical Hospital No. 15 named after O.M. Filatov, e-mail: oleg.averkov@gmail.com, <https://orcid.org/0000-0002-3010-755X>

Dmitry A. Sychev – Dr. Sc. (Med.), Professor, Member of the RAS, Head of the Department of Clinical Pharmacology and Therapy named after Academician B.E. Votchal, Acting Rector, Russian Medical Academy of Continuing Professional Education, e-mail: dmitry.alex.sychev@gmail.com, <https://orcid.org/0000-0002-4496-3680>