



Разработка прогностической модели исхода течения вирусного поражения лёгких с использованием алгоритмов машинного обучения

© Е.А. Бородулина^{1*}, Е.С. Вдоушкина¹, К.И. Шахгельдян^{2,3}, Е.А. Курдюкова^{3,4},
Б.Е. Бородулин¹, С.Ю. Пушкин^{1,5}, К.В. Жилинская¹

¹ Самарский государственный медицинский университет, Самара, Россия

² Дальневосточный федеральный университет, Школа медицины и наук о жизни, Владивосток, Россия

³ Владивостокский государственный университет, НОЦ Искусственный интеллект, Владивосток, Россия

⁴ Институт автоматизации и процессов управления Дальневосточного отделения РАН, Владивосток, Россия

⁵ Самарская областная клиническая больница им. В.Д. Середавина, Самара, Россия

* Е.А. Бородулина, Самарский государственный медицинский университет, 443099, Самара, ул. Чапаевская, 89, borodulinbe@yandex.ru

Поступила в редакцию 1 июля 2025 г. Исправлена 29 августа 2025 г. Принята к печати 23 октября 2025 г.

Резюме

Актуальность: В периоды роста респираторных инфекций актуальным является стратификация помощи пациентам с вирус-ассоциированными поражениями лёгких.

Цель: Разработка прогностической модели исхода течения вирусного поражения лёгких с использованием алгоритмов машинного обучения.

Материалы и методы: Исследование включало 295 пациентов с ковид-ассоциированной пневмонией. Сформированы группы: 1-я группа – умершие (n=78), 2-я группа – выжившие (n=217). Обработка данных осуществлялась с помощью IDE R Studio (Version 4.3.1). Прогностическая модель разработана с использованием потенциальных предикторов и алгоритмов машинного обучения: многофакторная логистическая регрессия, случайный лес и стохастический градиентный бустинг. Использован метод 5-кратной перекрёстной проверки (KFold). Модели прогнозирования смертности были реализованы с помощью среды Google Colaboratory для работы с Python. Качество моделей оценивали по метрикам: площадь под ROC-кривой (AUC) и матрицы неточностей (confusion matrix), позволяющей определить точность (Precision), полноту (Recall) или чувствительность (Sensitivity), специфичность (Specificity), F1-score.

Результаты: На основании предикторов летальности построена прогностическая модель исхода с высоким уровнем статистической значимости отличия в предсказываемых группах: наличие двух и более сопутствующих заболеваний, дыхательной недостаточностью выше 2-й степени, выраженной миалгии и одышки, отношения нейтрофилов к лимфоцитам >5,1, критическим снижением уровня эозинофилов на фоне повышенного уровня СОЭ, глюкозы, АСАТ, мочевины, С-реактивного белка и снижения уровня гемоглобина ($p < 0,05$).

Заключение: Прогностическая модель риска смерти, построенная методами машинного обучения с наилучшей производительностью у алгоритма случайного леса (AUC=0,99) с высокой точностью, позволяет на этапе поступления в стационар оценивать риск летального исхода, оптимизировать выбор тактики ведения.

Ключевые слова: факторы риска, системный анализ, пандемия, пневмония, логистическая регрессия

Цитировать: Бородулина Е.А., Вдоушкина Е.С., Шахгельдян К.И. и др. Разработка прогностической модели исхода течения вирусного поражения лёгких с использованием алгоритмов машинного обучения. *Инновационная медицина Кубани*. 2026;11(1):7–15. <https://doi.org/10.35401/2541-9897-2026-11-1-7-15>



Development of a Prognostic Model for the Outcome of Viral Lung Disease Using Machine Learning Algorithms

©Elena A. Borodulina^{1*}, Elizaveta S. Vdoushkina¹, Karina I. Shakhgeldyan^{2,3}, Elena A. Kurdyukova^{3,4}, Boris E. Borodulin¹, Sergey Y. Pushkin^{1,5}, Kristina V. Zhilinskaya¹

¹Samara State Medical University, Samara, Russian Federation

²Far Eastern Federal University, School of Medicine and Life Sciences, Vladivostok, Russian Federation

³Vladivostok State University, Artificial Intelligence Research Center, Vladivostok, Russian Federation

⁴Vladivostok Institute of Automation and Control Processes, Vladivostok, Russian Federation

⁵Samara Regional Clinical Hospital named after V.D. Seredavin, Samara, Russian Federation

* Elena A. Borodulina, Samara State Medical University, 89 Chapaevskaya St., Samara, 443099, Russian Federation, borodulinbe@yandex.ru

Received: July 1, 2025. Received in revised form: August 29, 2025. Accepted: October 23, 2025.

Abstract

Background: During periods of increasing respiratory infections, stratifying patient care for virus-associated lung injury is crucial.

Objective: To develop a prognostic model for the outcome of viral lung injury using machine learning algorithms.

Materials and methods: The study included 295 patients with COVID-associated pneumonia, divided into two groups: 1) deceased (n=78) and 2) survivors (n=217). Data processing was performed using IDE R Studio (Version 4.3.1). The prognostic model was created using potential predictors and machine learning algorithms: multivariate logistic regression, random forest, and stochastic gradient boosting. A 5-fold cross-validation method (KFold) was used. Prediction models for mortality were implemented in Google Colaboratory with Python. Model quality was assessed by metrics such as ROC-AUC and confusion matrix, which provided measures like precision, recall, sensitivity, specificity, and F1-score.

Results: The prognostic outcome model was based on mortality predictors, demonstrating a high level of statistical significance between the predicted groups, with significant predictors including the presence of two or more comorbidities, respiratory failure above grade 2, pronounced myalgia and dyspnea, neutrophil-to-lymphocyte ratio >5.1, a critical decrease in eosinophil count accompanied by elevated ESR, glucose, AST, CRP, urea levels, and decreased hemoglobin ($p < 0.05$).

Conclusion: The mortality risk prediction model developed using machine learning methods, particularly the random forest algorithm (AUC=0.99), demonstrates high accuracy and enables the assessment of mortality risk at hospital admission, guiding management decisions and optimizing treatment strategies.

Keywords: risk factors, systemic analysis, pandemic, pneumonia, logistic regression

Cite this article as: Borodulina EA, Vdoushkina ES, Shakhgeldyan KI, et al. Development of a prognostic model for the outcome of viral lung disease using machine learning algorithms. *Innovative Medicine of Kuban*. 2026;11(1):7–15. <https://doi.org/10.35401/2541-9897-2026-11-1-7-15>

Введение

В первые десятилетия XXI века практически ежегодно фиксировались вспышки вирусных инфекций. Благодаря современному развитию биотехнологий появляются новые опасные виды вирусов [1]. Остро протекающие респираторно-вирусные заболевания плохо контролируются ввиду разнообразия возбудителей, высокой заразности и отсутствия эффективных методов вакцинации. Количество заражённых лиц в периоды эпидемий существенно превосходит число официально зарегистрированных случаев [2]. Ярким примером является вспышка новой корона-вирусной инфекции, известная также как тяжёлый острый респираторный синдром коронавируса второго типа (SARS-CoV-2). Риск возникновения новых инфекционных вспышек сохраняется каждый год [3].

Искусственный интеллект обладает большим потенциалом в противодействии пандемиям. Тем не менее, опыт показывает, что лишь отдельные модели способны демонстрировать высокую точность прогнозирования тяжёлых последствий, включая смертельные случаи при COVID-19 [4, 5]. Появление

новых инфекций позволяет искусственному интеллекту эффективно выявлять ранние признаки вспышек заболеваний и предсказывать развитие эпидемиологических процессов.

Цель исследования

Разработка прогностической модели исхода течения вирусного поражения лёгких с использованием алгоритмов машинного обучения.

Материалы и методы

Настоящее исследование проводилось в рамках выполнения НИОКТР № 121051700033–3 «Поражение лёгких инфекционной этиологии. Совершенствование методов выявления, диагностики и лечения» (14.05.2021 г.). Исследование выполнено в соответствии с принципами Хельсинкской декларации Всемирной медицинской ассоциации и правилами надлежащей клинической практики Международной конференции по гармонизации. Протокол исследования был одобрен комитетом по биоэтике (протокол № 211 от 07.10.2020 г.). Исследование выполнялось по результатам ретроспективного

анализа историй болезни 295 больных с диагнозом COVID-19 (116 мужчин и 179 женщин), госпитализированных в «ковид»-стационар с мая 2020 по май 2022 г. В зависимости от исхода выделено две группы: 1-я группы – 78 больных с летальным исходом, вторая – 217 больных с благоприятным исходом. Конечная точка исследования представлена бинарным признаком: фатальным событием или его отсутствием.

Дизайн исследования включал несколько этапов. На первом этапе были проанализированы 84 показателя клинико-функционального статуса больных при сборе жалоб, анамнеза, физикальных, лабораторных, инструментальных данных в день поступления. Использовались статистические критерии: критерий Шапиро-Уилка для проверки на нормальность распределения, рассчитывались t-критерий Стьюдента – для нормально распределенных переменных, или критерий Манна-Уитни – для ненормально распределенных, критерии хи-квадрат и Фишера – для категориальных (номинальных) переменных и отношение шансов (ОШ) – для оценки связи между определённым исходом и фактором риска. Статистически значимыми считали при $p < 0,05$. Реализация анализа и обработка данных осуществлялись с помощью IDE R Studio (Version 4.3.1).

С помощью логистической регрессии на втором этапе исследований по стандартизованным данным были построены однофакторные модели. Для дальнейшего анализа были отобраны наиболее значимые признаки с весовыми коэффициентами модели выше 1 и сформирован пул потенциальных предикторов. На третьем этапе исследования с помощью методов многофакторной логистической регрессии (МЛР), случайного леса (СЛ) и стохастического градиентного бустинга (СГБ) были разработаны прогностические модели летального исхода.

Подбор предикторов и настройка гиперпараметров моделей выполнялись по метрике площади под ROC-кривой (AUC). Из пула потенциальных предикторов методом перебора были выбраны такие признаки, которые формировали модель с максимальной AUC. Оценку эффективности разработанных классификаторов выполняли с помощью метода 5-кратной перекрестной проверки (KFold). Итоговое тестирование разработанных моделей выполняли на 20% данных. Качество моделей оценивали также по следующим метрикам: чувствительность (Sensitivity), специфичность (Specificity) и F1-score, для вычисления которых определялся порог отсечения по минимизации разницы между чувствительностью и специфичностью на перекрестной проверке. При расчёте 95%-й доверительного интервала для метрик качества использовалось процентильное распределение результатов бутстрэп-пинга с сохранением пропорций классов в каждой

бутстрэп-выборке такими же, как в исходной. Отбор гиперпараметров проводился методом перебора с помощью GridSearchCV. Для оценки степени влияния предикторов прогностической модели на конечную точку использовали метод аддитивного объяснения Шепли. Исследование выполнялось с помощью среды Google Colaboratory для работы с Python.

Результаты исследования

Возраст больных был от 20 до 92 лет (медиана – 58 лет), 55,5 лет – среди мужчин, 59 лет – среди женщин. Доля пациентов старше 65 лет составила 28%.

В результате анализа 84 факторов, полученных при комплексном обследовании в день поступления, установлена достоверность различий по 46 из них ($p < 0,05$), при этом максимальный уровень статистической значимости ($p < 0,0001$) отмечен по 24 показателям (табл. 1, 2). Подтвердили высокую значимость уровень сатурации периферической крови, наличие сахарного диабета 2-го типа, температура тела выше 38 °C (фебрильная: $p < 0,0001$, пиретическая: $p = 0,02$). Из одиннадцати клинических симптомов в 50,8% были общая слабость (98,6%), одышка (78,3%), кашель (74,9%). Наиболее частыми сопутствующими заболеваниями являлись гипертоническая болезнь (55,9%), хронические заболевания ЖКТ (28,1%) и ИБС (стенокардия) (27,1%). Летальность составила от 36,4% у больных с гипертонической болезнью ($n = 165$) до 75,3% у больных с ожирением ($n = 73$) (табл. 1).

При моделировании в качестве потенциальных предикторов для модели выбирались наиболее значимые: предшествующее лечение жаропонижающими, наличие повышенной температуры тела при поступлении ($p < 0,05$) и показатель, характеризующий соотношение нейтрофилы/лимфоциты (во избежание мультиколлинеарности в процессе обучения модели между показателями нейтрофилов и лимфоцитов ($r = -0,9$), использовался показатель, характеризующий соотношение нейтрофилы/лимфоциты и имеющий значимые различия в группах исхода).

Многофакторные прогностические модели МЛР, СЛ и СГБ имели идентичный набор предикторов. Их набор определялся максимальным уровнем метрики AUC при перекрестной проверке. Весовые коэффициенты модели МЛР, разработанной на стандартизованных предикторах, демонстрируют доминирующее влияние ожирения на летальность при COVID (табл. 3).

Прогностические модели на основе СЛ и СГБ имели схожие с МЛР метрики качества не только при перекрестной проверке, но и на 20% данных, выделенных под итоговое тестирование (табл. 4).

Таблица 1
Показатели клинико-функционального статуса больных, госпитализированных с COVID-19 (n=295), p<0,05
Table 1
Clinical and functional status parameters of patients hospitalized with COVID-19 (n=295)

Показатели	Выжившие, n=217	Умершие, n=78	ОШ [ДИ95%]	p-value
Предшествующее лечение, абс. (%)	177 (81,6)	75 (96,2)	5,6 [1,7; 29,3]	0,003
• Жаропонижающие	102 (47)	65 (83,3)		<0,0001
• Глюкокортикостероиды	0 (0)	10 (12,8)	5,6 [2,9; 11,8]	<0,0001
Тяжёлое состояние, абс. (%)	18 (8,3)	43 (55,1)	13,4 [6,7; 27,8]	<0,0001
Количество симптомов, Ме (Q1-Q3)	4 (3; 5)	5 (4; 6)		<0,0001
Жалобы:				
Наличие повышенной температуры, абс. (%)	132 (60,8)	70 (89,7)	5,6 [2,5; 14,2]	<0,0001
• Фебрильная, абс. (%)	24 (11,1)	33 (42,3)	5,9 [3,03; 11,4]	<0,0001
• Пиретическая, абс. (%)	3 (1,4)	6 (7,7)	5,9 [1,2; 37,4]	0,02
Наличие одышки, абс. (%)	161 (74,2)	70 (89,7)	3,03 [1,3; 7,8]	0,007
Боли в мышцах, абс. (%)	16 (7,4)	25 (32,1)	5,9 [2,8; 12,7]	<0,0001
Наличие сопутствующей патологии, абс. (%)	168 (77,4)	75 (96,2)	7,3 [2,2; 37,6]	0,0005
Количество сопутствующих патологий, Ме (Q1-Q3)	1 (1; 2)	5 (3; 6)		<0,0001
Инфаркт миокарда, абс. (%)	4 (1,8)	10 (12,8)	7,8 [2,2; 35,0]	0,0003
Сахарный диабет 2-го типа, абс. (%)	31 (14,3)	27 (34,6)	3,2 [1,7; 6,04]	0,0003
ХОБЛ, абс. (%)	2 (0,9)	6 (7,7)	8,9 [1,5; 91,9]	0,006
Ожирение, абс. (%)	18 (8,3)	55 (70,5)	25,9 [12,7; 5,8]	<0,0001
Гипотериоз, абс. (%)	3 (1,4)	6 (7,7)	5,9 [1,2; 37,4]	0,02
Операция в анамнезе, абс. (%)	42 (19,4)	32 (41)	2,9 [1,6; 5,3]	0,0003
ИБС, стенокардия, абс. (%)	35 (16,1)	45 (57,7)	7,03 [3,8; 13,1]	<0,0001
ГБ, абс. (%)	105 (48,4)	60 (76,9)	3,5 [1,9; 6,8]	<0,0001
ОНМК абс. (%)	8 (3,7)	9 (11,5)	3,4 [1,1; 10,5]	0,02
ХБП, абс. (%)	18 (8,3)	33 (42,3)	8,03 [4,0; 16,6]	<0,0001
Заболевания ЖКТ, абс. (%)	30 (13,8)	53 (67,9)	13,04 [6,9; 5,5]	<0,0001
Онкология, абс. (%)	7 (3,2)	8 (10,3)	3,4 [1,04; 11,5]	0,0337
Сатурация, Ме (Q1-Q3)	95 (94; 97)	92,5 (86; 94)		<0,0001
Сатурация меньше нормы, абс. (%)	84 (38,7)	59 (75,6)	4,9 [2,7; 9,3]	<0,0001
ДН 0 степени абс. (%)	45 (20,7)	2 (2,6)	0,1 [0,01; 0,4]	0,0003
ДН 3 степени абс. (%)	5 (2,3)	21 (26,9)	15,4 [5,4; 54,7]	<0,0001
Поражение лёгких на КТ абс. (%)				
• 1-я степень (КТ-1)	92 (42,4)	18 (23,1)	0,4 [0,2; 0,8]	0,004
• 4-я степень (КТ-4)	4 (1,8)	10 (12,8)	7,8 [2,2; 35]	0,0003

Прим.: ДН – дыхательная недостаточность; ХОБЛ – хроническая обструктивная болезнь лёгких; ИБС – ишемическая болезнь сердца; ГБ – гипертоническая болезнь; ОНМК – острое нарушение мозгового кровообращения; ХБП – хроническая болезнь почек; ЖКБ – заболевания желудочно-кишечного тракта

Note: ДН, respiratory failure; ХОБЛ, chronic obstructive pulmonary disease; ИБС, ischemic heart disease; ГБ, hypertension; ОНМК, acute cerebrovascular accident; ХБП, chronic kidney disease; ЖКБ, gastrointestinal tract disease

Таблица 2
Показатели общего и биохимического анализов крови больных, госпитализированных с COVID-19 (n=295), p<0,05

Table 2
Parameters of complete blood count and biochemical blood tests of patients hospitalized with COVID-19 (n=295)

Показатели	Выжившие, n=217	Умершие, n=78	p-value
Общий анализ крови			
Лейкоциты 4,0–8,8×10 ⁹ /л, Me (Q1–Q3)	7 (4,3; 8,1)	7,5 (4,9; 11,3)	0,02
СОЭ 1–15 мм/ч, Me (Q1–Q3)	20 (12; 32)	25,5(12,8; 37,8)	0,05
Сегментоядерные нейтрофилы 47–67%, Me (Q1–Q3)	67 (60; 74)	75 (65,5; 79)	<0,0001
Эозинофилы 0–1%, Me (Q1–Q3)	1 (0; 2)	0 (0; 0)	<0,0001
Лимфоциты 25–35%, Me (Q1–Q3)	22 (14–29)	14,5 (11–21)	<0,0001
Моноциты 2–6%, Me (Q1–Q3)	5 (4; 7)	5 (3; 6)	<0,0001
Соотношение нейтрофилы/лимфоциты, Me (Q1–Q3)	2,96 (2,1; 5,3)	5,1 (3,2; 6,8)	<0,0001
Биохимический анализ крови			
Глюкоза 3,3–6,0 ммоль/л, Me (Q1–Q3)	6,2 (5,2; 7,9)	8,8 (6,9; 12,1)	<0,0001
Общий белок 64–84 г/л, Me (Q1–Q3)	67 (62; 72)	63,9 (60,2;67,6)	0,001
АсАТ до 45 ед./л, Me (Q1–Q3)	37,8 (23,2;47,5)	45,6 (31,5;68,6)	<0,0001
Билирубин 5–20 ммоль/л, Me (Q1–Q3)	11,8 (10,2;13)	12,75 (11;14,6)	0,002
СРБ менее 5 мг/л, Me (Q1–Q3)	51,8 (18,1; 118)	121 (40,4; 178,2)	<0,0001
Мочевина 2,5–8,3 ммоль/л, Me (Q1–Q3)	5,5 (4,3; 7)	7,1 (5,2; 9,3)	<0,0001
Фибриноген 2–4 г/л, Me (Q1–Q3)	4,5 (4,2; 4,9)	4,9 (4,4; 5,3)	0,004

Прим.: СОЭ – скорость оседания эритроцитов; АсАТ – аспаратаминотрансфераза; СРБ – С-реактивный белок

Note: СОЭ, erythrocyte sedimentation rate; АсАТ, aspartate aminotransferase; СРБ, C-reactive protein

Таблица 3
Предикторы и их весовые коэффициенты прогностической модели многофакторной логистической регрессии

Table 3
Predictors and their weighting coefficients in the multivariable logistic regression predictive model

Предикторы	Весовые коэффициенты
Ожирение	2,46
Температура тела	1,86
Дыхательная недостаточность 4-й степени	1,73
Заболевания ЖКТ	1,61
ХОБЛ	1,3
Боли в мышцах	1,08
ИБС, стенокардия	1,07
СРБ	0,96
Глюкоза	0,74
ХБП	0,67
Сатурация периферической крови	-0,67
АСАТ	0,47
Соотношение нейтрофилы/лимфоциты	0,33
Свободный член	-4,16

Значения AUC (0.96-0.97) при итоговом тестировании разработанных моделей прогнозирования, согласно экспертной шкале для значений AUC, по которым можно судить о качестве моделей, соответствуют отличному уровню (рис.1).

Метод аддитивного объяснения Шепли позволил оценить степень влияния предикторов на фатальное событие. Высокие значения shap-value ассоциируются с высокими рисками летальности. Уровень значений признаков поставлен в соответствие с цветовой шкалой от синего (минимальное значение признака) до красного (максимальное значение признака). Согласно Шепли, как и в МЛР, в модели СГБ наибольшее влияние оказывает ожирение (агрегированный индекс влияния – 1.83) (рис. 2). Наличие заболеваний ЖКТ (1.79), высокой температуры тела (1.51) и значений АСАТ (1.46) также существенно влияет на летальность пациентов с COVID-19.

Метод Шепли позволяет объяснить и отдельные клинические случаи (рис. 3). Так, в приведённом примере смерть пациента обусловлена прежде всего ожирением, заболеванием ЖКТ и гипергликемией.

Таблица 4
Метрики качества прогностических моделей летальности при итоговом тестировании

Table 4
Quality metrics of mortality predictive models in the final testing

Метод	AUC	Sensitivity	Specificity	F1-score
МЛР	0.97 [0.96; 0.98]	0.89 [0.83; 0.96]	0.9 [0.87; 0.93]	0.83 [0.78; 0.88]
СГБ	0.96 [0.94; 0.97]	0.88 [0.83; 0.94]	0.89 [0.87; 0.92]	0.81 [0.79; 0.84]
СЛ	0.96 [0.95; 0.98]	0.91 [0.87; 0.95]	0.9 [0.878 0.93]	0.84 [0.81; 0.87]

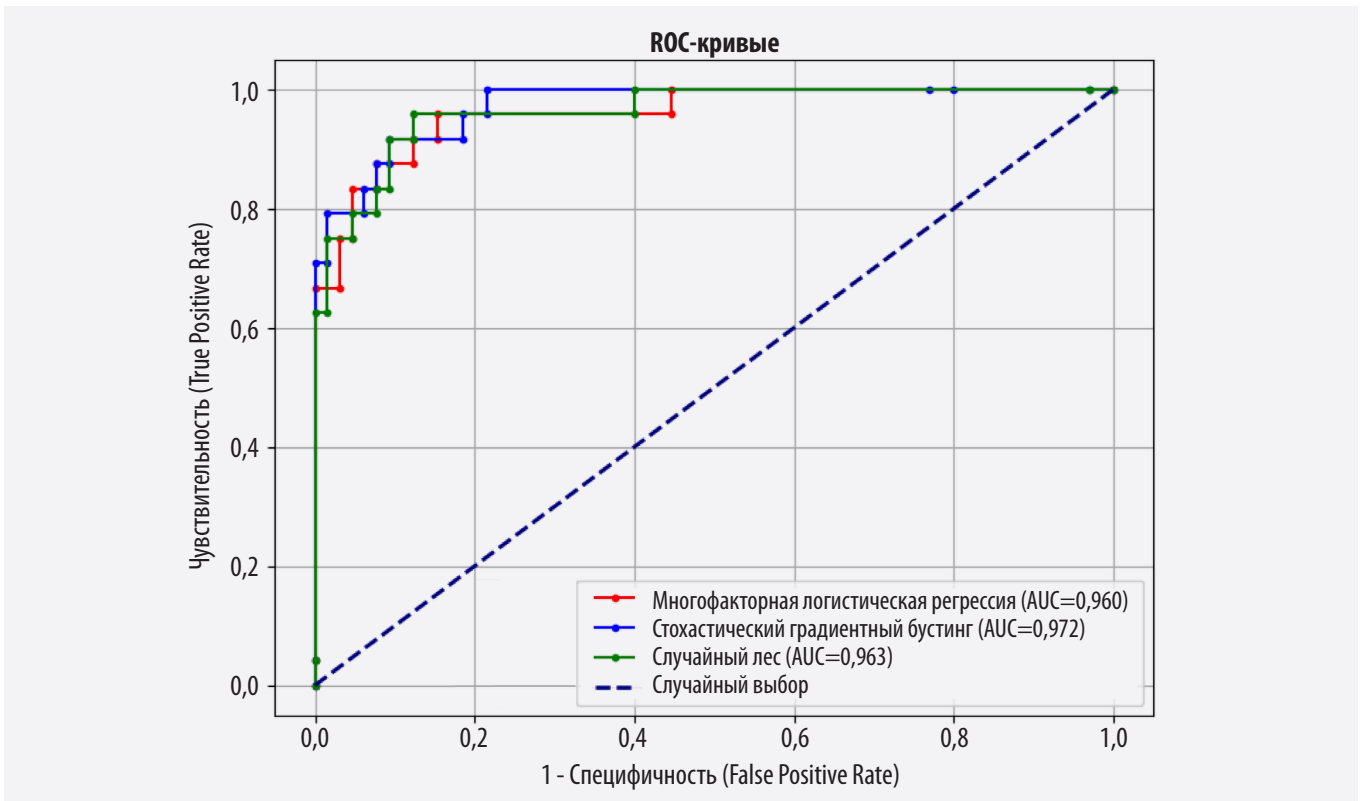


Рисунок 1. Графики ROC-кривой для прогностических моделей летальности при COVID-19
 Figure 1. ROC curve plots for mortality predictive models in COVID-19

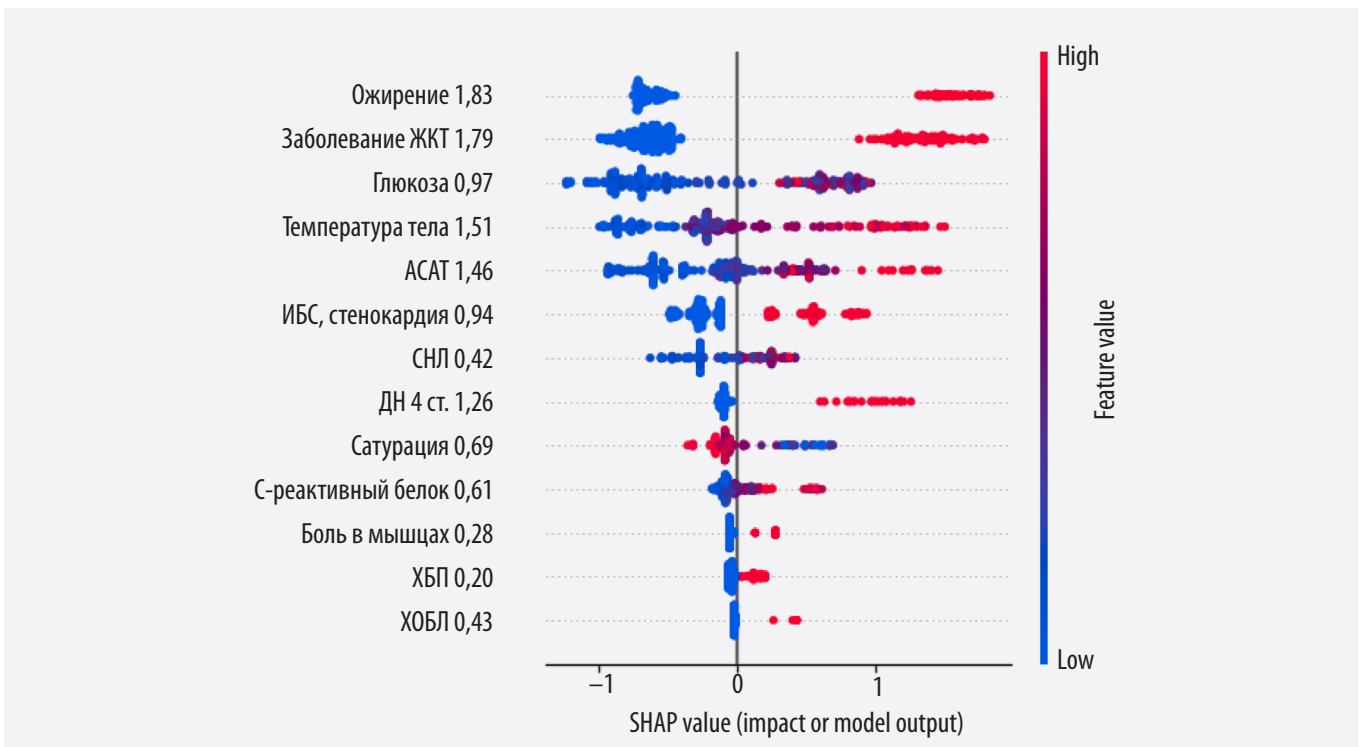


Рисунок 2. Важность признаков, библиотека Шепли (SHAP): влияние признаков на результат прогноза
 Figure 2. Feature importance using the SHAP library: impact or features on the prediction outcome

Прим.: SHAP-value (impact on model output) – SHAP-значения (влияние признаков на результат модели); Feature value – значения признаков; СНЛ – соотношение нейтрофилов к лимфоцитам; ДН – дыхательная недостаточность; ЖКТ – желудочно-кишечный тракт; ХОБЛ – хроническая обструктивная болезнь лёгких; ХБП – хроническая болезнь почек
 Note: SHAP-value (impact on model output); Feature value; СНЛ, neutrophil-to-lymphocyte ratio; ДН, respiratory failure; ЖКТ, gastrointestinal tract; ХОБЛ, chronic obstructive pulmonary disease; ХБП, chronic kidney disease



Рисунок 3. Объяснение причины фатального события по результатам оценки Шепли
 Figure 3. Explanation of the cause of fatal outcomes based on SHAP analysis

Дополнительными факторами риска были высокий уровень АСАТ (68,9), низкий уровень сатурации (80) и боль в мышцах.

Обсуждение

Исход заболевания зависит от множества факторов [6]; среди них изначально выделялись возраст старше 65 лет [7], срок обращения за медицинской помощью и контакт с больными COVID-19 [8]. У пациентов старше 80 лет вероятность смерти увеличивается примерно в шесть раз по сравнению с лицами моложе 65 лет; большое значение имеют сопутствующие патологии, особенно сахарный диабет [9]. Наше исследование показало отсутствие статистически значимой разницы в возрасте между группами с разными исходами болезни; средний возраст больных с неблагоприятным течением составил около $56,7 \pm 10,9$ лет. Среди всех исследуемых групп была отмечена высокая частота сопутствующих заболеваний, причём ожирение ($\text{ИМТ} > 30 \text{ кг/м}^2$) стало важным предиктором ухудшения состояния пациента. Доказано, что увеличенные и функционально нарушенные жировые клетки выделяют избыточное количество воспалительных цитокинов, снижают подвижность диафрагмы, уменьшают резерв дыхательной системы, повышают склонность к тромбозам, вызывают инсулинорезистентность и нарушение функций иммунной системы, способствуя повреждению лёгких [10, 11]. Уже с начала развития COVID-19 заболевание характеризуется феноменом «цитокиновой бури» [12, 13], который сопровождается высоким риском тромбообразования и эмболии, приводящих к мультиорганным повреждениям [14]. Многие авторы подчеркивают важность степени поражения легочной ткани, определяемой компьютерной томографией, как ключевого фактора негативно прогноза [9, 15].

Выявленные в нашем исследовании значимые предикторы неблагоприятного исхода коррелируют с данными, полученными другими авторами (наличие одышки, дыхательной недостаточности 3-й степени, низкой сатурации, степени поражения лёгких КТ-4, выраженность лихорадки, уровень СРБ, эозинофи-

лов, нейтрофилов, лимфоцитов, фибриноген) [16]. По данным литературы, соотношение нейтрофилы/лимфоциты $> 5,1$ достоверно коррелирует с повышенной смертностью и тяжёлыми осложнениями у госпитализированных с COVID-19 [17, 18].

Процесс трансформации системы здравоохранения в рамках реализации Стратегии научно-технического развития Российской Федерации включает обязательное внедрение технологий искусственного интеллекта, таких как методы машинного обучения и создание прогностических моделей, используемых в условиях пандемии [19].

Настоящее исследование было направлено на разработку прогностической модели, основанной на факторах риска, позволяющих определить возможный исход заболевания уже при первом контакте с пациентом.

Заключение

Использование общедоступных и экономически эффективных показателей обеспечивает быструю стратификацию пациентов по группам риска, что способствует своевременному распределению ресурсов: пациенты с высоким прогнозируемым риском осложнений направляются на углублённое обследование или раннюю госпитализацию, а лица с низким риском – на амбулаторное наблюдение. Интеграция моделей машинного обучения в клиническую практику повысит точность принятия решений, снизит нагрузку на стационары и персонал, а также персонализирует подходы к терапии.

Вклад авторов

Разработка концепции и дизайна: Е.А. Бородулина, К.И. Шахгельдян, Е.С. Вдоушкина

Разработка модели, оценка важности признаков:

Е.А. Курдюкова

Сбор, анализ и интерпретация данных: Е.С. Вдоушкина, К.В. Жилинская

Проведение статистического анализа: Е.А. Курдюкова, К.И. Шахгельдян

Подготовка текста: Е.С. Вдоушкина, Е.А. Курдюкова, К.И. Шахгельдян

Редактирование текста: Б.Е. Бородулин, С.Ю. Пушкин, Е.А. Бородулина

Author contributions

Research concept and design: Borodulina, Shakhgelyan, Vdoushkina

Model development and feature importance assessment: Kurdyukova

Acquisition, analysis, or interpretation of data: Vdoushkina, Zhilinskaya

Statistical analysis: Kurdyukova, Shakhgelyan

Manuscript drafting: Vdoushkina, Kurdyukova, Shakhgelyan

Manuscript revising: Borodulin, Pushkin, Borodulina

Литература/References

1. Ершов Ф.И. Почему XXI век может стать “веком пандемий”? Вопрос для дискуссии. *Вопросы вирусологии.* 2024;69(1):88-90. <https://doi.org/10.36233/0507-4088-227>

Ershov FI. Why may the 21st century become the «century of pandemics»? A question for discussion. *Problems of Virology.* 2024;69(1):88–90. (In Russ.). <https://doi.org/10.36233/0507-4088-227>

2. Мартянова А.Е., Ажмухамедов И.М. SEIRD-модель динамики распространения вирусных инфекций с учетом возникновения новых штаммов. *Прикаспийский журнал: управление и высокие технологии.* 2022;4(60):38-46. https://doi.org/10.54398/20741707_2022_4_38

Martyanova AE, Azhmukhamedov IM. SEIRD model describing the dynamics of spread viral infections considering the appearance of new strains. *Caspian journal: Control and High Technologies.* 2022;4(60):38-46. (In Russ.). https://doi.org/10.54398/20741707_2022_4_38

3. Азимова Н.Н., Бедойдзе М.В., Холодова С.Н. и др. Статистическая оценка биогенного риска для человеческой популяции со стороны новых вирусных инфекций на примере COVID-19. *Безопасность техногенных и природных систем.* 2023;1:4-15. <https://doi.org/10.23947/2541-9129-2023-1-4-15>

Azimova NN, Bedoidze MV, Kholodova SN, et al. Statistical assessment of biogenic risk to the human population from new viral infections based on COVID-19. *Safety of technogenic and natural systems.* 2023;1:4-15. (In Russ.). <https://doi.org/10.23947/2541-9129-2023-1-4-15>

4. Ipekci AM, Buitrago-Garcia D, Meili KW, et al. Outbreaks of publications about emerging infectious diseases: the case of SARS-CoV-2 and Zika virus. *BMC Med Res Methodol.* 2021;21(1):50. PMID: 33706715. PMID: PMC7948668. <https://doi.org/10.1186/s12874-021-01244-7>

5. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ.* 2020;369:m1328. PMID: 32265220. PMID: PMC7222643. <https://doi.org/10.1136/bmj.m1328>

6. Мареев В.Ю., Беграмбекова Ю.Л., Мареев Ю.В. Как оценивать результаты лечения больных с новой коронавирусной инфекцией (COVID-19)? Шкала Оценки Клинического Состояния (ШОКС–КОВИД). *Кардиология.* 2020;60(11):35–41. <https://doi.org/10.18087/cardio.2020.11.n1439>

Mareev VYu, Begrambekova YuL, Mareev YuV. How evaluate results of treatment in patients with COVID-19? Symptomatic Hospital and Outpatient Clinical Scale for COVID-19 (SHOCS-COVID). *Kardiologiya.* 2020;60(11):35-41. (In Russ.). <https://doi.org/10.18087/cardio.2020.11.n1439>

7. Дворецкий Л.И., Комарова И.С., Мухина Н.В., Черкасова Н.А., Дятлов Н.В. Новая коронавирусная инфекция (COVID-19) у больных пожилого и старческого возраста. *Медицинский вестник Северного Кавказа.* 2022;3(17):335-341. <https://doi.org/10.14300/mnnc.2022.17082>

Dvoretzky LI, Komarova IS, Mukhina NV, Cherkasova NA, Dyatlov MV. New coronaviral infection (COVID-19) in patients of elderly and senile age. *Medical Bulletin of the North Caucasus.* 2022;3(17):335-341. (In Russ.). <https://doi.org/10.14300/mnnc.2022.17082>

8. Вдоушкина Е.С., Бородулина Е.А., Поваляева Л.В., Суханова А.В., Жилинская К.В., Сутягин А.В. Сроки обращения и тяжесть состояния пациентов с поражением легких и подозрением на новую коронавирусную инфекцию при поступлении в стационар в период начала пандемии. *Врач.* 2020;31(11):60-63. <https://doi.org/10.29296/25877305-2020-11-12>

Vdoushkina ES, Borodulina EA, Povalyaeva LV, Sukhanova AV, Zhilinskaya KV, Sutyagin AV. The timing of referral and the severity of condition in patients with lung damage and suspected novel coronavirus infection on admission to hospital during the beginning of the pandemic. *Vrach.* 2020;31(11):60-63. (In Russ.). <https://doi.org/10.29296/25877305-2020-11-12>

9. Бахитов В.В., Алиев С.Р., Марцинкевич В.М., Дмитриева К.В., Масленников Р.В., Васильева Е.В. Структура смертности пациентов амбулаторного центра в период пандемии новой коронавирусной инфекции (COVID-19) *Современные проблемы здравоохранения и медицинской статистики.* 2022;S2:67-76. <https://doi.org/10.24412/2312-2935-2022-2-67-76>

Bakhitov VV, Aliev SR, Marcinkevich VM, Dmitrieva KV, Maslennikov RV, Vasilieva EV. Structure of mortality among patients of an outpatient center during the pandemic of the new coronavirus infection (COVID-19) *Current problems of health care and medical statistics.* 2022;S2:67-76. (In Russ.). <https://doi.org/10.24412/2312-2935-2022-2-67-76>

10. Korakas E, Ikonomidis I, Kousathana F, et al. Obesity and COVID-19: immune and metabolic derangement as a possible link to adverse clinical outcomes. *Am J Physiol Endocrinol Metab.* 2020;319(1):E105-E109. PMID: 32459524. PMID: PMC7322508. <https://doi.org/10.1152/ajpendo.00198.2020>

11. Giamarellos-Bourboulis EJ, Netea MG, Rovina N, et al. Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. *Cell Host Microbe.* 2020;27(6):992-1000.e3. PMID: 32320677. PMID: PMC7172841. <https://doi.org/10.1016/j.chom.2020.04.009>

12. Йокота Ш., Куройва Е., Нишиока К. Новая коронавирусная болезнь (COVID-19) и “цитокиновый шторм”. Перспективы эффективного лечения с точки зрения патофизиологии воспалительного процесса. *Инфекционные болезни: новости, мнения, обучение.* 2020;9(4):13-25. <https://doi.org/10.33029/2305-3496-2020-9-4-13-25>

Yokota SH, Kuroiwa E, Nishioka K. Novel coronavirus disease (COVID-19) and cytokine storms. For more effective treatment from the viewpoints of an inflammatory pathophysiology perspective. *Infectious diseases: news, opinions, training.* 2020;9(4):13-25. (In Russ.). <https://doi.org/10.33029/2305-3496-2020-9-4-13-25>

13. Zhang C, Wu Z, Li JW, Zhao H, Wang GQ. Cytokine release syndrome in severe COVID-19: interleukin-6 receptor antagonist tocilizumab may be the key to reduce mortality. *Int J Antimicrob Agents.* 2020;55(5):105954. PMID: 32234467. PMID: PMC7118634. <https://doi.org/10.1016/j.ijantimicag.2020.105954>

14. Жумабаева Т.Т., Турсунбаева А.Т., Кадырбаева А.А. Изменение биохимических показателей крови больных при COVID-19 со смертельным исходом. *Бюллетень науки и практики.* 2023;9(4.):256-264. <https://doi.org/10.33619/2414-2948/89/29>

Zhumabaeva TT, Tursunbaeva AT, Kadyrbaeva AA. Changes in biochemical indicators of the blood of patients with fatal COVID-19. *Bulletin of science and practice.* 2023;9(4.):256-264. (In Russ.). <https://doi.org/10.33619/2414-2948/89/29>

15. Биличенко Т.Н. Факторы риска, иммунологические механизмы и биологические маркеры тяжелого течения COVID-19 (обзор исследований). *PMЖ. Медицинское обозрение*. 2021;5(5):237-244. <https://doi.org/10.32364/2587-6821-2021-5-5-237-244>

Bilichenko TN. Risk factors, immunologic mechanisms and biological markers of severe course of COVID-19 course (study overview). *Russian Medical Journal*. 2021;5(5):237-244. (In Russ.). <https://doi.org/10.32364/2587-6821-2021-5-5-237-244>

16. Tian S, Hu W, Niu L, Liu H, Xu H, Xiao SY. Pulmonary Pathology of Early-Phase 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer. *J Thorac Oncol*. 2020;15(5):700-704. PMID: 32114094. PMCID: PMC7128866. <https://doi.org/10.1016/j.jtho.2020.02.010>

17. Lagunas-Rangel FA. Neutrophil-to-lymphocyte ratio and lymphocyte-to-C-reactive protein ratio in patients with severe coronavirus disease 2019 (COVID-19): A meta-analysis. *J Med Virol*. 2020;92(10):1733-1734. PMID: 32242950. PMCID: PMC7228336. <https://doi.org/10.1002/jmv.25819>

18. Liu Y, Du X, Chen J, et al. Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. *J Infect*. 2020;81(1):e6-e12. PMID: 32283162. PMCID: PMC7195072. <https://doi.org/10.1016/j.jinf.2020.04.002>

19. Борукаева И.Х., Абазова З.Х., Темиржанова Ф.Х., Юсупова М.М. COVID-19: Некоторые соображения о стандартных алгоритмах лечения. *Медицинская иммунология*. 2021;23(4):909-914. <https://doi.org/10.15789/1563-0625-COO-2265>

Borukaeva I.H., Abazova Z.H., Temirzhanova F.H., Yusupova M.M. COVID-19: Observations on standard treatment algorithms. *Medical Immunology (Russia)*. 2021;23(4):909-914. (In Russ.). <https://doi.org/10.15789/1563-0625-COO-2265>

Сведения об авторах

Бородулина Елена Александровна, д. м. н., профессор, заведующий кафедрой фтизиатрии и пульмонологии, Самарский государственный медицинский университет (Самара, Россия). <https://orcid.org/0000-0002-3063-1538>

Влоушкина Елизавета Сергеевна, к. м. н., доцент кафедры фтизиатрии и пульмонологии, Самарский государственный медицинский университет (Самара, Россия). <https://orcid.org/0000-0003-0039-6829>

Шахгельдян Карина Иосифовна, д. т. н., заведующий научной лабораторией Школа медицины и наук о жизни Дальневосточного федерального университета; профессор НОЦ «Искусственный интеллект» Владивостокского государственного университета (Владивосток, Россия). <https://orcid.org/0000-0002-4539-685X>

Курдюкова Елена Александровна, к. б. н., главный специалист лаборатории искусственного интеллекта, Институт автоматизации и процессов управления Дальневосточного отделения РАН; магистрант 2-го курса НОЦ «Искусственный интеллект» Владивостокского государственного университета (Владивосток, Россия). <https://orcid.org/0009-0004-2563-2845>

Бородулин Борис Евгеньевич, д. м. н., профессор кафедры фтизиатрии и пульмонологии, Самарский государственный

медицинский университет (Самара, Россия). <https://orcid.org/0000-0002-8847-9831>

Пушкин Сергей Юрьевич, д. м. н., заведующий кафедрой хирургических болезней детей и взрослых, Самарский государственный медицинский университет; главный врач, Самарская областная клиническая больница им. В.Д. Середавина, главный внештатный специалист-торакальный хирург Самарской области и ПФО (Самара, Россия). <https://orcid.org/0000-0003-2206-6679>

Жилинская Кристина Васильевна, аспирант кафедры фтизиатрии и пульмонологии, Самарский государственный медицинский университет (Владивосток, Россия). <https://orcid.org/0000-0003-4209-3025>

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Финансирование

Работа выполнена при финансовой поддержке проекта FZNS-2023-0010 Госзадания ДВФУ.

Author credentials

Elena A. Borodulina, Dr. Sci. (Med.), Professor, Head of the Department of Phthisiology and Pulmonology, Samara State Medical University (Samara, Russian Federation). <https://orcid.org/0000-0002-3063-1538>

Elizaveta S. Vdoushkina, Cand. Sci. (Med.), Associate Professor, Department of Phthisiology and Pulmonology, Samara State Medical University (Samara, Russian Federation). <https://orcid.org/0000-0003-0039-6829>

Karina I. Shakhgelydyan, Dr. Sci. (Tech.), Head of Scientific Laboratory, School of Medicine and Life Sciences, Far Eastern Federal University; Professor, Artificial Intelligence Research Center, Vladivostok State University (Vladivostok, Russian Federation) <https://orcid.org/0000-0002-4539-685X>

Elena A. Kurdyukova, Cand. Sci. (Biol.), Chief Specialist of Artificial Intelligence Laboratory, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences; 2nd year Master's Student of Artificial Intelligence Research Center, Vladivostok State University (Vladivostok, Russian Federation). <https://orcid.org/0009-0004-2563-2845>

Boris E. Borodulin, Dr. Sci. (Med.), Professor, Department of Phthisiology and Pulmonology, Samara State Medical University (Samara, Russian Federation). <https://orcid.org/0000-0002-8847-9831>

Sergey Y. Pushkin, Dr. Sci. (Med.), Head of the Department of Surgical Diseases of Children and Adults, Samara State Medical University; Chief Physician, Samara Regional Clinical Hospital named after V.D. Seredavin; Chief Freelance Thoracic Surgeon of the Samara Region and the Volga Federal District (Samara, Russian Federation). <https://orcid.org/0000-0003-2206-6679>

Kristina V. Zhilinskaya, Resident, Department of Phthisiology and Pulmonology, Samara State Medical University (Samara, Russian Federation). <https://orcid.org/0000-0003-4209-3025>

Conflict of interest: none declared.

Funding:

This work was funded by FZNS-2023-0010 project under the State Assignment of the Far Eastern Federal University.