

11. Pearce N., Rhodes S., Stocking K., Pembrey L., van Veldhoven K., Brickley E.B., Robertson S., Davoren D., Nafilyan V., Windsor-Shellard B., Fletcher T., van Tongeren M. Occupational differences in COVID-19 incidence, severity, and mortality in the United Kingdom: Available data and framework for analyses. *Wellcome Open Research*. 2021; 6: 102 (version 2 published 15 Dec 2023). DOI: 10.12688/wellcomeopenres.16845.2. <https://wellcomeopenresearch.org/articles/6-102/v2> Date of access: 15.06.2025
12. Saraghi I.D., Tonapa S.I., Saraghi I.S., Advani S., Abdurahim S.A., Lin C.-J., Fisher M.L. Global prevalence of mental health problems among healthcare workers during the COVID-19 pandemic: A systematic review and meta-analysis. *International Journal of Nursing Studies*. 2021; 121: 104002. DOI: 10.1016/j.ijnurstu.2021.104002. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9701545/> Date of access: 15.06.2025
13. Платонова Т.А., Голубкова А.А., Смирнова С.С. COVID-19 у сотрудников медицинских организаций: характеристика клинических проявлений в острый период и в период реконвалесценции Журнал эпидемиологии и инфекционных болезней. 2021; 11 (4): 25–30. DOI: <https://dx.doi.org/10.18565/epidem.2021.11.4.25-30>
- Platonova T.A., Golubkova A.A., Smirnova S.S. COVID-19 in healthcare workers: clinical characteristics during acute phase and convalescence. *Journal of Epidemiology and Infectious Diseases*. 2021; 11 (4): 25–30. (In Russ.). DOI: <https://dx.doi.org/10.18565/epidem.2021.11.4.25-30>

Статья поступила / Received 09.07.2025
Получена после рецензирования / Revised 14.07.2025
Принята в печать / Accepted 22.07.2025

Сведения об авторах

Устарханова Амира Кюраевна, аспирант 1 года обучения¹, аналитик².
E-mail: ustarhanova95@gmail.com ORCID: 0009-0002-1882-5038
Стрижаков Леонид Александрович, д.м.н., профессор, заместитель директора по научной и лечебной работе³, профессор кафедры медицины труда, авиационной, космической и водолазной медицины, руководитель центра профессиональной патологии². E-mail: strijzhakov@irioh.ru. ORCID: 0000-0002-2291-6453
Бухтияров Игорь Валентинович, д.м.н., профессор, академик РАН, директор³, заведующий кафедрой медицины труда, авиационной, космической и водолазной медицины². E-mail: info@irioh.ru. ORCID: 0000-0002-8317-2718
Костенко Наталья Алексеевна, начальник управления медицинской деятельностью⁴. E-mail: kostenkona@dmo.org.rzd.ru
Акимкин Василий Геннадьевич, д.м.н., профессор, академик РАН, директор¹. E-mail: vgakimkin@yandex.ru, SPIN-код: 4038-7455. ORCID: 0000-0003-4228-9044

About authors

Ustarhanova Amira K., 1st year postgraduate student¹, analyst².
E-mail: ustarhanova95@gmail.com ORCID: 0009-0002-1882-5038
Strizhakov Leonid A., DM Sci (habil.), professor, deputy director for Research and Treatment³, professor at Dept of Occupational Medicine, Aviation, Space and Diving Medicine, head of Center for Occupational Pathology². E-mail: strijzhakov@irioh.ru. ORCID: 0000-0002-2291-6453.
Bukhtiyarov Igor V., DM Sci (habil.), professor, RAS academician, director³, head of Dept of Occupational Medicine, Aviation, Space and Diving Medicine². E-mail: info@irioh.ru. ORCID: 0000-0002-8317-2718
Kostenko Natalia A., head of Medical Activities Dept⁴. E-mail: kostenkona@dmo.org.rzd.ru
Akimkin Vasily G., DM Sci (habil.), professor, RAS academician, director¹. E-mail: vgakimkin@yandex.ru, SPIN-code: 4038-7455. ORCID: 0000-0003-4228-9044

¹ ФБУН «Центральный научно-исследовательский институт эпидемиологии» Роспотребнадзора, Москва, Россия
² ФГАОУ ВО «Первый Московский государственный медицинский университет имени И.М. Сеченова» Минздрава России (Сеченовский университет), Москва, Россия
³ ФГБУН «Научно-исследовательский институт медицины труда им. акад. Н.Ф. Измерова» Минобрнауки России, Москва, Россия
⁴ Центральная дирекция здравоохранения – филиал ОАО «Российские железные дороги», Москва, Россия

¹ Central Research Institute of Epidemiology, Moscow, Russia
² I.M. Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia
³ Izmerov Research Institute of Occupational Health, Moscow, Russia
⁴ Central Healthcare Directorate – Branch of JSC “Russian Railways”, Moscow, Russia

Автор для переписки: Устарханова Амира Кюраевна.
E-mail: ustarhanova95@gmail.com

Corresponding author: Ustarhanova Amira K. E-mail: ustarhanova95@gmail.com

Для цитирования: Устарханова А.К., Стрижаков Л.А., Бухтияров И.В., Костенко Н.А., Акимкин В.Г. COVID-19 и медицинские работники: заболеваемость, летальность и отдаленные последствия. *Медицинский алфавит*. 2025; (16): 7–12. <https://doi.org/10.33667/2078-5631-2025-16-7-12>

For citation: Ustarhanova A.K., Strizhakov L.A., Bukhtiyarov I.V., Kostenko N.A., Akimkin V.G. COVID-19 and healthcare workers: morbidity, mortality, and long-term consequences. *Medical alphabet*. 2025; (16): 7–12. <https://doi.org/10.33667/2078-5631-2025-16-7-12>



DOI: 10.33667/2078-5631-2025-16-12-20

Оценка риска развития средней и тяжелой степени тяжести пневмонии при COVID-19 методами машинного обучения

О.В. Сенько^{1,4}, А.В. Кузнецова³, И.А. Демина^{1,2}, Е.М. Воронин¹, А.А. Плоскирева¹,
Ю.Р. Посынкина¹, В.Г. Акимкин¹

- ¹ ФБУН «Центральный научно-исследовательский институт эпидемиологии» Роспотребнадзора, Москва, Россия
- ² ГБУЗ «Московский многопрофильный научно-клинический центр имени С.П. Боткина» Департамента здравоохранения города Москвы, Россия
- ³ ФГБУН «Институт биохимической физики имени Н.М. Эмануэля» Российской академии наук, Москва, Россия
- ⁴ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия

РЕЗЮМЕ

Работа посвящена разработке компьютерного метода оценки степени тяжести течения пневмонии, являющейся осложнением при COVID-19, по наборам клинико-лабораторных показателей методами машинного обучения (МО). Работа состояла в исследовании возможности предсказания с использованием МО градации степени тяжести течения пневмонии, основанной на результатах компьютерной томографии (КТ). Сравнивались группы из 31 пациента со средней или тяжелой степенью пневмонии (КТ2 – КТ4) и 113 пациентов без пневмонии или с легкой степенью тяжести пневмонии (КТ0 – КТ1). База данных включала 105 клинико-лабораторных показателей. Для сравнения групп применили стандартные непараметрические критерии χ^2 и критерий Манна-Уитни (U-тест) с коррекцией на множественное тестирование по Холма–Бонферрони. Наряду с традиционными статистическими методами использовался также оригинальный метод интеллектуального анализа данных, позволяющий статистически обоснованно выделять

информативные интервалы значений показателей. Для прогнозирования градаций степени тяжести течения пневмонии использовался набор методов МО, включавший наряду с широко известными методами также и оригинальную разработку. В исследованиях использовался программный комплекс Data Master Azforus. Исследование подтвердило возможность прогнозирования результатов оценок степени тяжести течения на основе КТ по клиничко-лабораторным показателям с использованием методов машинного обучения. При этом эффективность прогноза по ROC AUC составила около 0,9. Внедрение модели в практику будет способствовать улучшению точности и оперативности диагностики тяжелого течения пневмонии.

КЛЮЧЕВЫЕ СЛОВА: COVID-19, пневмония, тяжесть течения, компьютерная томография, машинное обучение, интеллектуальный анализ данных.

КОНФЛИКТ ИНТЕРЕСОВ. Авторы заявляют об отсутствии конфликта интересов.

Assessing the risk of developing moderate and severe pneumonia in COVID-19 patients using machine learning methods

O. V. Senko^{1,4}, A. V. Kuznetsova³, I. A. Demina^{1,2}, E. M. Voronin¹, A. A. Ploskireva¹,
I. R. Posynkina¹, V. G. Akimkin¹

¹ Central Research Institute of Epidemiology, Moscow, Russia

² Moscow Multidisciplinary Scientific and Clinical Center n.a. S.P. Botkin (Botkin hospital), Moscow, Russia

³ Emanuel Institute of Biochemical Physics of Russian Academy of Sciences, Moscow, Russia

⁴ Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

SUMMARY

The work is devoted to the development of a computer method for assessing the severity of pneumonia, which is a complication of COVID-19, based on sets of clinical and laboratory indicators using machine learning (ML) methods. The work consisted of investigating the possibility of predicting using ML the computed tomography (CT) grade of severity of pneumonia. Groups of 31 patients with moderate or severe pneumonia (CT2–CT4) and 113 patients without pneumonia or with mild pneumonia (CT0–CT1) were compared. The database included 105 clinical and laboratory parameters. To compare groups, standard nonparametric χ^2 tests and the Mann-Whitney test (U-test) with Bonferroni-Holm correction for multiple testing were used. Along with traditional statistical methods, an original method of data mining was also used, allowing statistically sound identification of informative intervals of indicator values. To predict severity gradations, a set of ML methods was used, which included, along with widely known methods, also an original development. The research used the Data Master Azforus software package. The study confirmed the possibility of predicting the results of CT based severity grades by clinical and laboratory indicators using machine learning methods. The forecast efficiency according to ROC AUC was about 0.9. The introduction of the model into practice will help improve the accuracy and efficiency of diagnosing severe pneumonia.

KEYWORDS: COVID-19, pneumonia, severity, computed tomography, machine learning, data mining.

CONFLICT OF INTEREST. The authors declare no conflict of interest.

Введение

Одним из наиболее тяжелых осложнений при заболевании COVID-19, вызываемого вирусом SARS-CoV-2, является пневмония различной тяжести течения, которая может привести к неблагоприятным исходам. Анализ факторов тяжелого течения пневмонии, а также показателей, характеризующих тяжесть течения, необходим как для правильного выбора способа лечения, так и для понимания механизмов развития тяжелых форм заболевания.

Согласно обзору [1] факторами тяжелого течения заболевания являются пожилой возраст, мужской пол, наличие коморбидности (сахарный диабет, сердечно-сосудистые заболевания, болезни органов дыхания, почек, аутоиммунные заболевания и иммунодефицитные состояния). Также прогноз тяжелого течения пневмонии связан с лимфопенией, повышенными уровнями в сыворотке амилоидного белка, С-реактивного белка, прокальцитонина, интерлейкина-6.

В рамках существующих Временных клинических рекомендаций «Профилактика, диагностика и лечение новой коронавирусной инфекции COVID-19» (версия 16 от 18.08.2022) наличие и тяжесть течения пневмонии подтверждается таким объективным диагностическим исследованием, как компьютерная томография (КТ), играющим решающую роль в оценке тяжести течения и степени

поражения легких. Среди различных шкал оценки тяжести течения можно выделить получившую широкое распространение в России «эмпирическую» визуальную шкалу КТ 0–4: КТ-0 – отсутствие признаков пневмонии; КТ-1 – легкая степень тяжести пневмонии с участками «матового стекла», выраженность патологических изменений менее 25%; КТ-2 – умеренная пневмония, поражено 25–50% легких; КТ-3 – среднетяжелая пневмония, поражено 50–75% легких; КТ-4 – тяжелая степень тяжести пневмонии, поражено >75% легких. Особого внимания требуют пациенты со среднетяжелыми и тяжелыми пневмониями, при которых поражения легких составляет от 25 до 100% (КТ2 – КТ4). Сравнение различных шкал показывает их значительную взаимосвязь [2, 4, 5].

Важность компьютерной томографии подтверждается исследованиями связи различных шкал оценки тяжести течения по КТ как с выживаемостью [2, 3, 6, 7], так и с клиническими оценками течения заболевания [3, 4, 5, 6]. В клинической практике оценка тяжести течения с использованием КТ обычно дополняется анализом клиничко-лабораторных показателей. В связи с этим представляет интерес исследование связи совокупности клиничко-лабораторных показателей с градацией тяжести

течения на основе КТ. В работе [8] была выявлена связь оценок тяжести течения по шкале КТ 0–4 с выживаемостью, а также с возрастом, ИМТ и рядом сопутствующих заболеваний. Была отмечена связь оценок тяжести течения по КТ с уровнем С-реактивного белка, D-димера, ферритина, лимфопений [9].

В настоящее время в медицинских исследованиях наряду со статистическими критериями и методами регрессионного анализа все большее распространение получают методы машинного обучения (МО), позволяющие эффективно строить мультифакторные нелинейные модели компьютерной диагностики или прогнозирования. Преимущество подобного подхода состоит в существовании арсенала современных методов МО, обеспечивающих высокую точность автоматической диагностики, а также способов корректной оценки эффективности получаемых алгоритмов. Использование коллективных решений по ансамблям алгоритмов позволяет повысить точность диагностики.

Методы МО ранее использовались для оценивания тяжести течения и прогнозирования исхода пневмонии.

В работе [10] была показана возможность точного прогноза степени тяжести состояния пациента с внебольничной пневмонией (ВП), который позволяет улучшить уход за пациентами и управление больницей. Индекс тяжести течения состояния пациента пневмонии (Pneumonia Severity Index, PSI) был разработан в 1997 году в качестве инструмента для руководства клинической практикой путем стратификации форм тяжести состояния пациентов с ВП. В то время как PSI оценивался по сравнению с другими инструментами клинической стратификации, он не оценивался по нескольким классическим классификаторам машинного обучения по различным показателям при большом объеме выборки. В этой статье сравнили эффективность прогнозирования девяти классических классификаторов машинного обучения с PSI по данным 34 720 пациентов для взрослых (возраст 18+), собранных из 749 больниц в период с 2009 по 2018 год в Соединенных Штатах. Классификаторы машинного обучения, такие как Random Forest, обеспечили статистически значимое улучшение ($p < 0,001$) ($\approx 33\%$ в PR AUC и $\approx 6\%$ в ROC AUC) по сравнению с PSI и потребовали всего 7 входных значений (по сравнению с 20 параметрами, используемыми в PSI).

Целью работы является исследование возможности создания компьютерного метода диагностики тяжести течения состояния пациента пневмонии при COVID-19 по набору клиничко-лабораторных показателей, наиболее соответствующего объективным градациям тяжести течения на основе КТ.

Для достижения указанной цели на первом этапе с использованием стандартных статистических методов и оригинального метода интеллектуального анализа данных производится поиск статистически значимых связей между объективными градациями степени тяжести течения на основе КТ и клиничко-лабораторных показателями. Метод интеллектуального анализа данных используется для статистически обоснованного поиска интервалов значений показателей, а также областей значений парных сочетаний показателей, соответствующих средней и тяжелой формами

пневмонии. Целью первого этапа является поиск информативных наборов показателей, которые далее используются в методах мультифакторной автоматической диагностики, основанной на машинном обучении. Другой целью первого этапа является достижение большей ясности в понимании биологических эффектов изучаемых патологических процессов на основании данных, полученных с использованием алгоритмов, основанных на машинном обучении.

Целью второго этапа является оценка точности предсказания градаций тяжести течения на основе КТ по различным наборам клиничко-лабораторных показателей с использованием ряда методов машинного обучения.

Материалы и методы

Материалы

Было проведено ретроспективное исследование. Группа пациентов состояла из 144 взрослых, госпитализированных в 2021 г. в Московский многопрофильный научно-клинический центр имени С. П. Боткина. У всех госпитализированных был положительный результат на SARS-CoV-2, полученный методом ПЦР, и у большей части госпитализированных пневмония была подтверждена компьютерной томографией грудной клетки. Госпитализированные пациенты были разделены на две группы: группа 1 включает 31 пациента со средней и тяжелой степенью тяжести течения COVID-19-ассоциированной пневмонии, подтвержденными на КТ, группа 2 включает 113 пациентов без пневмонии или с легкой степенью тяжести течения пневмонии при COVID-19.

Клиничко-лабораторная база данных содержала 105 показателей. В первичной базе данных вместе с показателями Суточного Мониторинга Артериального Давления (СМАД) было 144 показателя. В том числе анамнез, симптомы, сопутствующие заболевания, антропометрия, физикальные данные, общий анализ крови, биохимия, коагулограмма, гормоны, динамика Ig G.

Методы анализа данных

Наряду с традиционными статистическими методами, включая критерий Манна-Уитни и метод Холма – Бонферрони [11], в работе использовались альтернативные методы интеллектуального анализа данных (ИАД) и методы машинного обучения (МО).

Методы ИАД. Из методов ИАД был использован метод оптимально достоверных разбиений (ОДР) [12], который представляет собой оригинальную технологию анализа данных, основанную на построении оптимальных разбиений пространства признаков. При этом разбиение считается оптимальным, если оно позволяет наилучшим образом разделить сравниваемые непересекающиеся группы. Поиск оптимальных разбиений производится внутри множеств (семейств) разбиений с числом элементов, не превышающем некоторое фиксированное число. В настоящем исследовании рассматривались два семейства: 1) семейство разбиений интервалов значений отдельного признака на два интервала с помощью одной граничной точки; 2) семейство разбиений двумерной области совместных значений пар признаков на четыре подобласти с помощью границ, параллельных координатным осям. Примеры одномерного и двумерного ОДР представлены на рисунках 1 и 2.

Каждое найденное оптимальное достоверное разбиение фактически описывает различия между группами и по соответствующим показателям или сочетаниям показателей.

Необходимым критерием эффективности анализа является оценка статистической значимости выявленных различий. Для оценки значимости в методе ОДР используется известный непараметрический перестановочный тест, не требующий априорных предположений о характере распределений и не использующий асимптотических оценок.

Метод ОДР ранее использовался при решении многих медицинских задач, включая исследования в эпидемиологии, и может рассматриваться как вспомогательный инструмент для достижения интерпретируемости и прозрачности методов машинного обучения.

Методы машинного обучения. В работе были использованы традиционные статистические методы автоматической классификации, включая линейный дискриминант Фишера, и основанные на машинном обучении методы автоматической классификации, включая логистическую регрессию [14], метод опорных векторов [15,16], вариант градиентного бустинга XGBoost [16,17,18]. Также использовался оригинальный метод автоматической классификации статистически взвешенные синдромы (СВС) [19,20].

Метод СВС является ансамблевым методом автоматического отнесения объектов к одной из групп. Преимуществом метода СВС является простота и естественный способ работы с данными, содержащими пропущенные значения. Метод демонстрирует высокую эффективность при работе с медицинскими данными относительно небольшого объема. При этом он нередко превосходит по точности альтернативные подходы. Однако, недостатком метода является проведение большого объема вычислений при поиске решающего правила.

Оценка статистической значимости найденных закономерностей производилась с использованием метода кроссвалидации (LeaveOneOut). Для оценки значимости использовались показатели чувствительности, специфичности

и общей точности, а также ROC анализ [21]. Использовались реализации ОДР и перечисленных методов в программном комплексе Data Master Azforus.

Результаты и обсуждение

На первом этапе исследования сравниваемые группы анализировали с помощью стандартных статистических методов: χ^2 и Манна–Уиттн. Показатели объединены в наборы по периоду заболевания: острый период и период реконвалесценции.

Вследствие большого числа показателей, сравнимого с числом объектов в обучающей выборке, проведена проверка на множественное тестирование Холма–Бонферрони. Метод Холма (поправка Холма–Бонферрони) равномерно более мощный, чем поправка Бонферрони, и, в определенной степени, решает проблему падения мощности при росте числа гипотез. Значимыми с учетом поправки Холма–Бонферрони стали различия по 12 показателям, представленные в таблице 1. В столбцах «Среднее в группе 1» и «Среднее в группе 2» даны средние значения показателей в группах 1 и 2. Для бинарных показателей, обозначающих наличие или отсутствие симптома, в указанных столбцах даны доли пациентов с наличием этого симптома соответственно в группах 1 и 2. Наиболее значимыми стали различия, касающиеся дыхательной недостаточности «1-я степень дыхательной недостаточности» и «Отсутствие дыхательной недостаточности». Даже после коррекции по Холма–Бонферрони значимость у них сохраняется на уровне $p < 0,000001$. Значимость показателей «Изменение веса», «Астенизация», «Нарушение сна», «Выпадение волос» после коррекции оценивалась на уровне $p < 0,001$. Остальные показатели оказались значимыми на уровне $p < 0,02$.

Оценка информативности клинико-лабораторных показателей с помощью метода ОДР

Метод ОДР использовался для изучения закономерностей связи номера группы (1 или 2) с клинико-лабораторными показателями.

Таблица 1
Сравнение исследуемых групп пациентов с применением традиционных статистических методов с проверкой Бонферрони–Холма

№	Показатель	Среднее или частота в группе 1	Среднее или частота в группе 2	χ^2 р	U-тест р	U-тест с коррекцией Холма–Бонферрони (р)
Острый период						
1	1-я степень дыхательной недостаточности	0,613	0,1239	<0,000001	<0,000001	<0,000001
2	Отсутствие дыхательной недостаточности	0,129	0,8407	<0,000001	<0,000001	<0,000001
3	Длительность пиретической лихорадки (дни)	2,355	0,354	0,000188	0,00002	0,001442
4	Индекс сатурация	98,19	98,65	0,000122	0,00002	0,001408
5	2 степень дыхательной недостаточности	0,2258	0,0177	0,000862	0,000024	0,001748
6	Пиретическая лихорадка до 41С	0,3871	0,0885	0,000844	0,000046	0,003284
7	Частота дыхательных движений (ЧДД)	16,39	15,85	0,000400	0,000173	0,012378
Период реконвалесценции						
1	Снижение веса	0,677	0,2035	0,000006	<0,000001	0,000031
2	Отсутствие снижения веса	0,29	0,7788	0,000004	<0,000001	0,000022
3	Длительность астенизации, дни	91,58	28,34	0,00009	<0,000001	0,000004
4	Нарушения сна	0,7097	0,2566	0,000026	0,000003	0,000234
5	Выпадение волос	0,5161	0,1416	0,000163	0,00001	0,000691

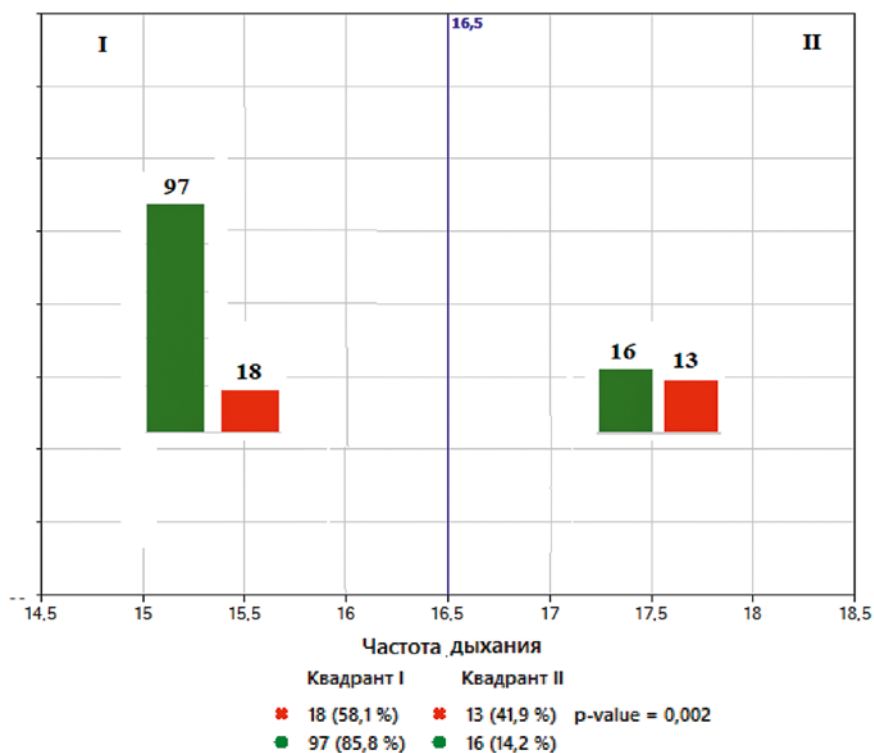


Рисунок 1. Различия между случаями из группы 1 (красный цвет) и группы 2 (зелёный цвет) по показателю «ЧДД», выявленные методом ОДР

Эффективность метода ОДР иллюстрируется результатами для показателя «ЧДД», представленным на *рисунке 1*. На одномерной диаграмме случаи из группы 1 со средним и тяжелым течением пневмонии обозначены красным цветом, а случаи из группы 2 с легким течением пневмонии или без пневмонии – зеленым (*рис. 1*).

Как видно на *рисунке 1*, слева от границы «16,5», оптимально выставленной алгоритмом ОДР для показателя «ЧДД», отображаются 97 случаев (85,8%) из группы 2 (легкая степень тяжести течения пневмонии) и 18 случаев (58,1%) из группы 1 (средняя и тяжелая степень тяжести течения пневмонии). При этом справа от границы отображаются 13 случаев из группы 1 (41,9%), и только 16 случаев из группы 2 (14,2%). Значимость такого разделения оценена с помощью перестановочного теста на уровне $p < 0.002$.

Таблица 2
Значимые показатели для распознавания групп 1 и 2 при одномерном разбиении методом ОДР

№	Показатель	Граница	Ниже границы или отсутствие симптомов		Выше границы или наличие симптомов		p-значение ПТ**
			Группа 1, абс. и (%)	Группа 2, абс. и (%)	Группа 1, абс. и (%)	Группа 2, абс. и (%)	
Острое состояние							
1	Отсутствие дыхательной недостаточности	-	27 (87,1)	18 (15,9)	4 (12,9)	95 (84,1)	<0,0005
2	1-я степень дыхательной недостаточность	-	12 (38,7)	99 (87,6)	19 (61,3)	14 (12,4)	<0,0005
3	Длительность пиретической лихорадки (дни)	3,5	20 (64,5)	109 (96,5)	11 (35,5)	4 (3,5)	<0,0005
4	Индекс сатурация	98,5	24 (77,4)	39 (34,5)	7 (22,6)	74 (65,5)	0,0005
5	2 степень дыхательной недостаточность	-	24 (77,4)	111 (98,2)	7 (22,6)	2 (1,8)	0,0005
6	Пиретическая лихорадка до 41С	-	19 (61,3)	103 (91,2)	12 (38,7)	10 (8,8)	0,0005
7	Головная боль		6 (19,4)	58 (51,3)	25 (80,6)	55 (48,7)	0,001
8	ИМТ	34	20 (64,5)	105 (92,9)	11 (35,5)	8 (7,1)	0,0015
9	Частота дыхательных движений (ЧДД)	16,5	18 (58,1)	97 (85,8)	13 (41,9)	16 (14,2)	0,0015
10	Длительность фебрильной лихорадка (дни)	6,5	22 (71)	106 (93,8)	9 (29)	7 (6,2)	0,0055
11	Масса тела (кг)	116	27 (87,1)	113 (100)	4 (12,9)	0 (0)	0,0055
12	Повышенная потливость	0,5	27 (87,1)	112 (99,1)	4 (12,9)	1 (0,9)	0,006
13	Сухость слизистых	0,5	22 (71)	101 (89,4)	9 (29)	12 (10,6)	0,017
14	Длительность субфебрильной лихорадки (дни)	17,5	26 (83,9)	111 (98,2)	5 (16,1)	2 (1,8)	0,0195
15	Лихорадки нет		30 (96,8)	90 (79,6)	1 (3,2)	23 (20,4)	0,023
16	Артралгия		12 (38,7)	71 (62,8)	19 (61,3)	42 (37,2)	0,017
Период реконвалесценции							
17	Длительность астенизации (дни)	25,5	6 (19,4)	85 (75,2)	25 (80,6)	28 (24,8)	<0,0005
18	Отсутствие снижения веса		22 (71)	25 (22,1)	9 (29)	88 (77,9)	<0,0005
19	Снижение веса		10 (32,3)	90 (79,6)	21 (67,7)	23 (20,4)	<0,0005
20	Нарушения сна		9 (29)	84 (74,3)	22 (71)	29 (25,7)	<0,0005
21	Выпадение волос		15 (48,4)	97 (85,8)	16 (51,6)	16 (14,2)	<0,0005
22	Астенизация		2 (6,5)	54 (47,8)	29 (93,5)	59 (52,2)	0,0005
23	Повышение веса		23 (74,2)	106 (93,8)	8 (25,8)	7 (6,2)	0,0045
24	Тромбоцит (РСТ)	0,16	5 (17,9)	2 (1,9)	23 (82)	104 (98)	0,0135
25	МНО (коагулограмма)	1,3	28 (90,3)	113 (100)	3 (9,7)	0 (0)	0,0305
26	Миалгия	0,5	10 (32,3)	60 (53,1)	21 (67,7)	53 (46,9)	0,0435

Выявленная взаимосвязь может быть интерпретирована как худший прогноз для пациентов с ЧДД выше 16,5 вдохов в минуту. Этот результат может приниматься во внимание для оценки перспектив терапии.

Определенные методом ОДР границы для каждого значимого показателя для пациентов из групп 1 и 2 позволяют оценивать количественное (или – доленое) распределение пациентов в исследуемых группах справа или слева от границы оптимального разбиения и прогнозировать степени тяжести течения заболевания для каждого конкретного пациента.

Для показателей, имеющих граничные значения, в *таблице 2* приведены числовые значения оптимальных границ показателей, полученных методом ОДР при одномерном разбиении. В *таблице* представлено распределение пациентов из полных групп 1 и 2 по подгруппам (ниже или выше границы для пациентов из групп 1 и 2) в зависимости от того, какое значение имеет исследуемый показатель конкретного пациента по отношению к границе оптимального достоверного разбиения (ниже или выше границы разбиения).

Для бинарных показателей (отсутствие или наличие соответствующего симптома) указаны абсолютные количества пациентов и их доли (%) из групп 1 и 2. В скобках даны процентные доли подгрупп выше или ниже границы в полных группах 1 и 2.

В *таблице 2* представлены значимые показатели как для периода острого состояния (16 показателей), так и для периода реконвалесценции (10 показателей).

В графическом виде информативность сочетаний показателей представлена на рисунке двумерного разбиения методом ОДР – по осям двумерной диаграммы отложены два значимых показателя. Результаты для двумерной модели, описывающей связь степени тяжести течения пневмонии с сочетанием длительности пиретической лихорадки и индекса сатурации крови SO_2 представлены на *рисунке 2*.

Из *рисунка 2* видно, что при длительности пиретической лихорадки выше 3 дней и уровне сатурации ниже 98,5 % (квадрант III) абсолютно преобладают случаи с тяжелым

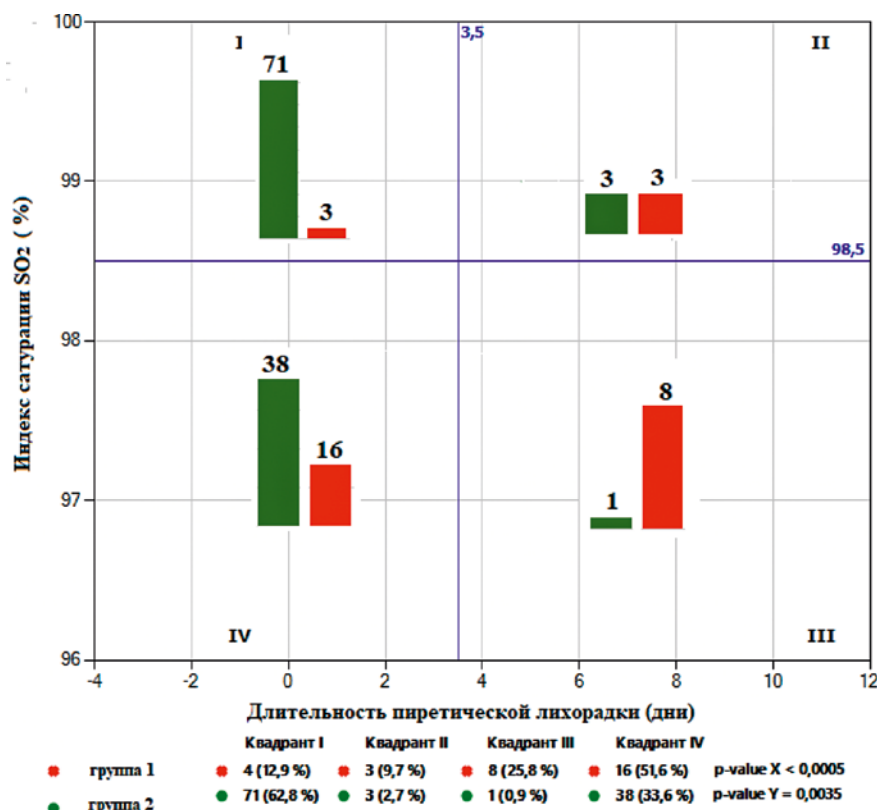


Рисунок 2. Различия между случаями из группы 1 и группы 2 по сочетанию длительности пиретической лихорадки (ось X) и сатурации крови O_2 (ось Y)

течением пневмонии: из 9 случаев, удовлетворяющих указанным условиям, тяжелому течению пневмонии соответствуют 8. Наоборот, в квадранте I (уровень сатурации выше 98,5 и длительности пиретической лихорадки менее 4 дней), расположены 71 случай без пневмонии или с легким течением пневмонии и только 3 случая с тяжелым течением пневмонии. Значимость различий была оценена на уровне $p < 0,0005$ для длительности пиретической лихорадки и $p < 0,0035$ для сатурации крови O_2 . Следует отметить, что практически все показатели, приведенные в *таблице 2*, отмечались в научной периодике или клинической практике как информативные для оценки степени тяжести течения пневмония в острый период или в период реконвалесценции.

Многофакторные модели для оценивания степени тяжести течения по клиническим показателям, основанные на технологиях машинного обучения

Методы машинного обучения использовались для компьютерной оценки степени тяжести течения пневмонии по 4 наборам клиничко-лабораторных показателей:

- 16 показателей из *таблицы 2*, соответствующих острой стадии заболевания (набор 1);
- 10 показатели из *таблицы 2*, соответствующих периоду реконвалесценции (набор 2);
- набор из 26 показателей, являющийся объединение наборов 1 и 2 (набор 3);
- набор из 7 показателей в остром периоде и 5 показателей в период реконвалесценции из *таблицы 1*, различия по которым сохранили значимость после проведения коррекции по Холму – Бонферрони (набор 4).

Решалась задача построения алгоритма, с максимальной точностью относящего каждый из случаев в упомянутые ранее группы 1 и 2, формируемые согласно оценкам степени тяжести течения по КТ.

Таблица 3
Результаты ROC анализа при диагностике степени тяжести течения пневмонии по 4 наборам показателей (в скобках указано число показателей)

Методы машинного обучения	Набор 1 (16)	Набор 2 (10)	Набор 3 (26)	Набор 4 (12)
Статистически взвешенные синдромы	0,91	0,82	0,9	0,87
Логистическая регрессия	0,89	0,8	0,89	0,86
Линейный дискриминантный анализ	0,86	0,74	0,84	0,81
Метод опорных векторов	0,9	0,75	0,9	0,79
Градиентный бустинг	0,83	0,74	0,85	0,78
Метод ближайших соседей	0,83	0,63	0,88	0,68
Решающий лес	0,89	0,79	0,88	0,84

Таблица 4
Эффективность компьютерной диагностики степени тяжести течения пневмонии по набору показателей в острой стадии заболевания

Метод машинного обучения	Общая точность классификации	Группа 1	Группа 2
Статистически взвешенные синдромы	85%	87%	85%
Логистическая регрессия	88%	77%	91%
Метод опорных векторов	85%	55%	94%
Ансамбль из трех методов	88%	71%	94%

Таблица 5
Эффективность компьютерной диагностики тяжести течения пневмонии по объединенному набору показателей 3 (весь период наблюдений)

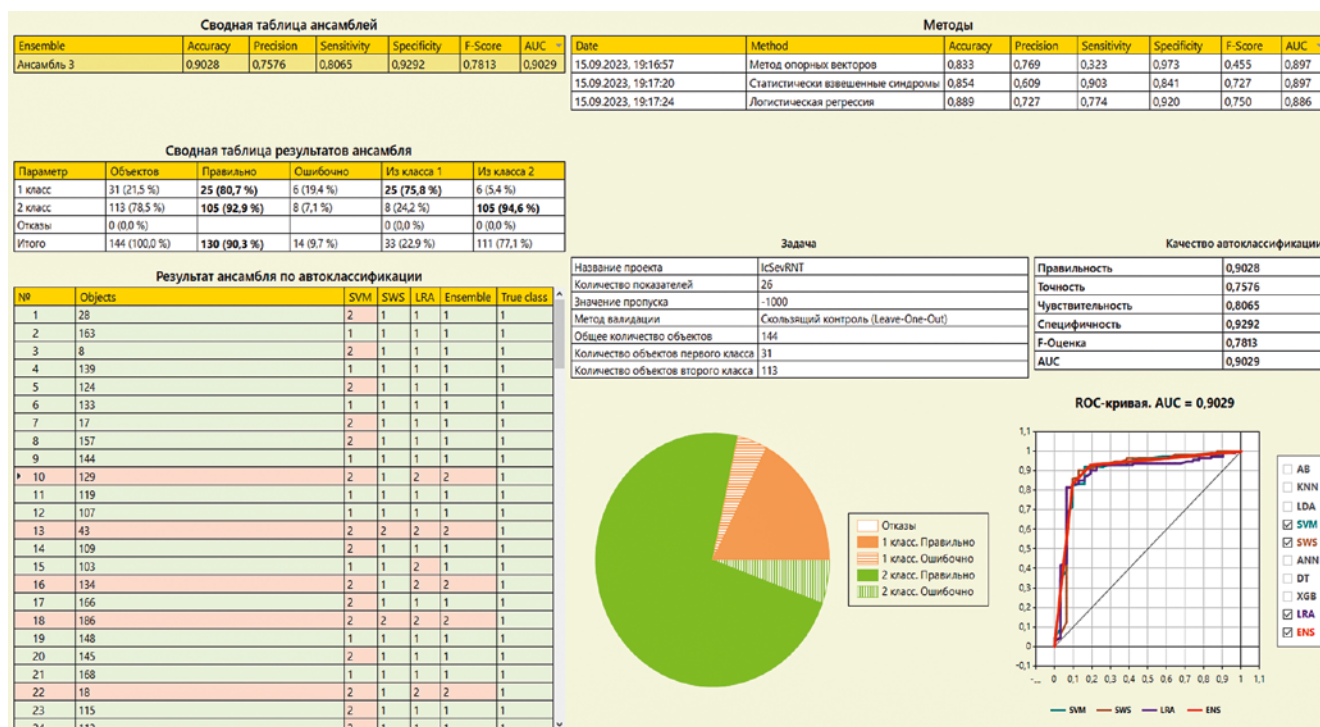
Метод машинного обучения	Общая точность классификации	Группа 1	Группа 2
Статистически взвешенные синдромы	85%	84%	90%
Логистическая регрессия	89%	77%	92%
Метод опорных векторов	85%	32%	97%
Ансамбль из трех методов	90%	76%	95%

В таблице 3 приведены значения ROC AUC для 7 методов машинного обучения, полученные с помощью метода скользящего контроля (LeaveOneOut).

Из таблицы 3 видно, что для всех четырех наборов показателей максимальную эффективность диагностики по значениям показателя ROC AUC демонстрирует метод статистически взвешенных синдромов (СВС). Также высокую эффективность демонстрируют логистическая регрессия и метод опорных векторов.

Также из приведенных в таблице 3 результатов следует, что эффективность диагностики степени тяжести течения пневмонии по показателям в остром периоде (набор 1, 16 показателей) заметно превышает эффективность диагностики по показателям в период реконвалесценции (набор 2, 10 показателей). Более того, показатели в период реконвалесценции не позволяют повысить точность диагностики при использовании их вместе с показателями в остром периоде. Более подробно результаты оценки точности диагностики по набору показателей острой стадии заболевания представлены в таблице 4.

В таблице 4 даны процентные доли правильно классифицированных



случаев из групп 1 и 2 для трех наиболее эффективных методов МО согласно таблице 3. Видно, что наиболее высокая доля правильно классифицированных случаев из группы 2 достигается при использовании метода СВС, поскольку в данном исследовании нас больше интересует группа с тяжелым течением пневмонии. Группа 1 лучше распознается с помощью метода опорных векторов и ансамбля.

На *рисунке 3* представлено «окно», генерируемое используемым нами программным комплексом, в котором приведены результаты автоматической диагностики с использованием набора 3 и ансамбля из трех наиболее эффективных методов: метод опорных векторов, статистически взвешенные синдромы и логистическая регрессия. Ансамбль относит рассматриваемый случай в ту группу, в которую его отнесло большинство алгоритмов. Для ансамбля достигается существенно более высокая чувствительность: 0,93 против значения специфичности – 0,8.

Более подробно результаты оценки точности диагностики по объединенному набору показателей 3 представлены в *таблице 5*.

В *таблице 5* даны процентные доли правильно классифицированных случаев из групп 1 и 2 для трех наиболее эффективных методов согласно таблице 3, а также ансамбля из трех этих методов. Наиболее высокая доля правильно классифицированных случаев из группы 1 достигается для метода СВС (84%). Наиболее высокая точность классификации для случаев из группы 2 достигается для метода опорных векторов и ансамбля из 3-х методов (97 и 95%, соответственно). Однако в группе 1 для ансамбля достигается существенно более высокая чувствительность: 76% для ансамбля по сравнению с 32% для метода опорных векторов.

Отметим также, что для набора показателей, известных в период реконвалесценции наибольшая эффективность диагностики, достигнутая с использованием метода СВС, характеризовалась долей правильно классифицированных случаев из группы 2, равной 80% и долей правильно классифицированных случаев из группы 1, равной 74%, что существенно хуже эффективности по показателям из острого периода.

Для набора из 7 показателей, сохранивших высокую значимость после поправки по Холму – Бонферрони, наибольшая эффективность диагностики с использованием СВС характеризовалась долей правильно распознанных случаев равной 80% и долей правильно классифицированных случаев из группы 1, равной 81%, что ниже результата, достигнутого для набора 1, но несколько лучше результата, достигнутого для показателей из набора 2.

Выводы

1. Проведено сравнение группы из 31 пациента со средней и тяжелой степенью пневмонии (группа 1, соответствует КТ2–4) и группы из 113 пациентов без пневмонии или с легкой степенью пневмонии (группа 2, соответствует КТ0–1) по набору из 105 клинико-лабораторных показателей. Использование стандартных статистических критериев χ^2 и Манна – Уитни, а также поправок на множественное тестирование

по Холму – Бонферрони позволило выделить 7 значимых клинико-лабораторных показателей, наблюдаемых в острый период и 5 значимых клинико-лабораторных показателей, наблюдаемых в период реконвалесценции. Наилучшая точность была достигнута с использованием метода СВС (ROC AUC = 0,87).

2. Метод ОДР позволил выявить значимые различия по 26 показателям, включая 16 показателей, соответствующих острому периоду, и 10 показателей, соответствующих периоду реконвалесценции. Использование этих 26 значимых показателей (*табл. 2*) позволяет повысить качество распознавания: общее значение – 90%, чувствительность – 76%, специфичность – 95%, ROC AUC=0.9.
3. Наиболее высокая эффективность распознавания принадлежности к группам 1 и 2 выявлена для 16 показателей, соответствующих острому периоду. Эффективность была оценена с использованием метода кросс-валидации (LeaveOneOut) на уровне 0,91 по показателю ROC AUC и на уровне 87% и 85% соответственно по чувствительности и специфичности.
4. Эффективность распознавания принадлежности группам 1 и 2 по набору, соответствующему периоду реконвалесценции, оказалась несколько ниже: ROC AUC=0,82.
3. Для более точного прогнозирования степени тяжести течения пневмонии при COVID-19 можно рекомендовать использовать наборы показателей 1, 3 и 4.
5. Все показатели, приведенные в таблицах 1 и 2, ранее отмечались как клинически значимые для оценки степени тяжести течения пневмонии.

Заключение

Исследование статистически значимо подтвердило существование тесной связи между степенью тяжести течения пневмонии при COVID-19 и рядом клинико-лабораторных показателей.

Исследование позволило найти статистически обоснованные значения интервалов клинико-лабораторных показателей, соответствующих средней и тяжелой степени тяжести пневмонии.

Исследование подтвердило возможность прогнозирования результатов оценок степени тяжести течения пневмонии на основе КТ по клинико-лабораторным показателям с использованием методов машинного обучения с эффективностью по ROC AUC около 0.9. Такая точность достигается в том числе только с использованием показателей, соответствующих острому периоду заболевания.

Исследование подтвердило возможность использования технологий машинного обучения для оперативной оценки и прогнозирования степени тяжести течения пневмонии как осложнения при COVID-19, являющейся альтернативой или дополнением к оценкам степени тяжести течения с использованием КТ.

Внедрение в клиническую практику прогнозирования с использованием методов машинного обучения может стать дополнительным вспомогательным инструментом для медицинских работников при рассмотрении

вопроса о выборе тактики лечения больных с пневмонией при COVID-19 и будет способствовать улучшению медицинской помощи госпитализированным пациентам с COVID-19 с пневмонией тяжелой степени тяжести.

Список литературы / References

1. Биличенко Т.Н. Факторы риска, иммунологические механизмы и биологические маркеры тяжелого течения COVID-19 (обзор исследований). *PMJ. Медицинское обозрение*. 2021. Т. 5, № 5. С. 237–244. DOI: <http://doi.org/10.32364/2587-6821-2021-5-5-237-244>
2. Bilichenko T.N. Risk factors, immunological mechanisms and biological markers of severe COVID-19 (review of studies). *RMZh. Medical Review*. 2021. Vol. 5, No. 5. P. 237–244. (In Russ.). DOI: <http://doi.org/10.32364/2587-6821-2021-5-5-237-244>
3. Кудрявцев Ю.С., Берегов М.М., Бердалин А.Б., Лелюк В.Г. Сравнение основных шкал оценки тяжести поражения легких при COVID-19 по данным компьютерной томографии и оценка их прогностической ценности. *Вестник рентгенологии и радиологии*. 2021. Т. 102, № 5. С. 296–303. <https://doi.org/10.20862/0042-4676-2021-102-5-296-303>
4. Kudryavtsev Yu.S., Beregov M.M., Berdalina A.B., Lelyuk V.G. Comparison of the main scales for assessing the severity of lung damage in COVID-19 based on computed tomography data and assessment of their prognostic value. *Bulletin of Roentgenology and Radiology*. 2021. Vol. 102, No. 5. P. 296–303. (In Russ.). <https://doi.org/10.20862/0042-4676-2021-102-5-296-303>
5. Hemraj K.S., Jacob M.J., Kofian V., Sachin D.K., Geetha R.G., & Veliath L.B. Correlation of CT Severity Index with Clinical Severity of COVID-19 Pneumonia and it's Relation to Disease Outcome // *Journal of Advances in Medicine and Medical Research*. 2022. Vol. 34, № 12. P. 16–26. <https://doi.org/10.9734/jamr/2022/v34i1231369>
6. Inoue A., Takahashi H., Ibe T. et al. Comparison of semiquantitative chest CT scoring systems to estimate severity in coronavirus disease 2019 (COVID-19) pneumonia // *Eur Radiol*. 2022. Vol. 32, P. 3513–3524. <https://doi.org/10.1007/s00330-021-08435-2>
7. Almasi N.A., Shahnazari R., Abbasi M.A., Divsalar F., Bayazidi M., Sadatnaseri A. CT severity score in COVID-19 patients, assessment of performance in triage and outcome prediction: a comparative study of different methods // *Egypt. J. Radiol. Nucl. Med*. 2022. Vol. 53, № 1. P. 116. <https://doi.org/10.1186/s43055-022-00781-5> Epub. 2022 May 18. PMID: PMC9116070.
8. Szabó M., Kardos Z., Kostyál L., Tamáska P., Oláh C., Csányi E. and Szekanecz Z. The importance of chest CT severity score and lung CT patterns in risk assessment in COVID-19-associated pneumonia: a comparative study // *Front. Med*. 2023. Vol. 10. P. 1125530. <https://doi.org/10.3389/fmed.2023.1125530>
9. Морозов С.П., Гомболевский В.А., Чернина В.Ю. и др. Прогнозирование летальных исходов при COVID-19 по данным компьютерной томографии органов грудной клетки // *Туберкулез и болезни легких*. 2020. Т. 98, № 6. С. 7–14. <https://doi.org/10.21292/2075-1230-2020-98-6-7-14>
10. Morozov S.P., Gombolevsky V.A., Chernina V. Yu. et al. Prediction of fatal outcomes in COVID-19 based on chest computed tomography data // *Tuberculosis and lung diseases*. 2020. Vol. 98, No. 6. P. 7–14. (In Russ.). <https://doi.org/10.21292/2075-1230-2020-98-6-7-14>
11. Мелехов А.В., Сайфуллин М.А., Петровичев В.С., Агаева А.И., Голубых К.Ю., Никитин И.Г. Сопоставление данных компьютерной томографии с истощами, клиническими и лабораторными характеристиками пациентов с COVID-19 // *Архив внутренней медицины*. 2021. Т. 11, № 6. С. 447–456. <https://doi.org/10.20514/2226-6704-2021-11-6-447-456>

12. Melekhov A.V., Saifullin M.A., Petrovichev V.S., Agaeva A.I., Golubykh K. Yu., Nikitin I.G. Comparison of computed tomography data with outcomes, clinical and laboratory characteristics of patients with COVID-19 // *Archives of Internal Medicine*. 2021. Vol. 11, No. 6. P. 447–456. (In Russ.). <https://doi.org/10.20514/2226-6704-2021-11-6-447-456>
13. Sharma S, Aggarwal A, Sharma RK, Patras E, Singhal A. Correlation of chest CT severity score with clinical parameters in COVID-19 pulmonary disease in a tertiary care hospital in Delhi during the pandemic period // *Egypt J Radiol Nucl Med*. 2022. Vol. 53, № 1. P. 166. DOI: [10.1186/s43055-022-00832-x](https://doi.org/10.1186/s43055-022-00832-x). Epub 2022 Jul 28. PMID: PMC9330926.
14. Wang D, Willis DR, Yih Y. The pneumonia severity index: Assessment and comparison to popular machine learning classifiers // *Int J Med Inform*. 2022. Jul, Vol. 163, P. 104778. DOI: [10.1016/j.ijmedinf.2022.104778](https://doi.org/10.1016/j.ijmedinf.2022.104778). Epub 2022 Apr 26. PMID: 35487075.
15. Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979. Vol. 6, № 2. P. 65–70. JSTOR4615733. MR0538597
16. Senko O.V., Kuznetsova A.V., Matveev I.A., Livinchev I.S. Chapter 8 – Search of regularities in data: optimality, validity, and interpretability, Editor(s): Pandian Vasant, Joshua Thomas, Elias Munapo, Gerhard-Wilhelm Weber. *Advances of Artificial Intelligence in a Green Energy Environment*, Academic Press, 2022, P. 151–171.
17. Сенько О.В., Кузнецова А.В., Воронин Е.М., Крайцова О.А., Борисова Л.Р., Кирилук И.А., Акимкин В.Г. Методы интеллектуального анализа данных в исследованиях эпидемии COVID – 19. *Журн. Белорус. гос. ун-та. Матем. Инф.* 2022. Vol. 1. P. 83–96.
18. Senko O.V., Kuznetsova A.V., Voronin E.M., Kravtsova O.A., Borisova L.R., Kirilyuk I.L., Akimkin V.G. Methods of data mining in COVID-19 epidemic research. *J. of Belarusian State University. Mat. Inf.* 2022. Vol. 1, pp. 83–96. (In Russ.).
19. Cox D.R. The regression analysis of binary sequences (with discussion). *J R Stat Soc B*. 1958; 20(2):215–242. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
20. Cortes C., Vapnik V. Support-Vector Networks. *Machine Learning*, 20, 273–297 (1995) DOI: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018)
21. Hastie, T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. – Springer-Verlag, 2009. ISBN 978-0-387-84857-0.
22. Friedman, J. H. (2002). Stochastic gradient boosting. *Comput. Stat. Data Anal.* 38 (4): 367–378.
23. Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794. New York, NY, USA. ACM.
24. Kuznetsova, A.V., Kostomarov I.V., Senko O.V. Modification of the method of optimal valid partitioning for comparison of patterns related to the occurrence of ischemic stroke in two groups of patients. *Pattern Recognition and Image Analysis* 24 (1), 114–123
25. Гулиев Р.Р., Сенько О.В. и др. Применение оптимальных разбиений для многопараметрического анализа данных в клинических исследованиях. *Математическая биология и биоинформатика* 2016. Т. 11. № 1. С. 46–63. DOI: [10.17537/2016.11.46](https://doi.org/10.17537/2016.11.46)
26. Guliev R.R., Senko O.V., et al. Application of optimal partitions for multiparameter data analysis in clinical trials. *Mathematical biology and bioinformatics* 2016. Vol. 11. No. 1. P. 46–63. (In Russ.). DOI: [10.17537/2016.11.46](https://doi.org/10.17537/2016.11.46)
27. Zweig, Mark H., Campbell, Gregory (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39 (8): 561–577. DOI: [10.1093/clinchem/39.4.561](https://doi.org/10.1093/clinchem/39.4.561). PMID8472349

Статья поступила / Received 22.05.2025
Получена после рецензирования / Revised 30.05.2025
Принята в печать / Accepted 22.07.2025

Сведения об авторах

Сенько Олег Валентинович, д.ф.-м.н., профессор, старший научный сотрудник научной группы математических методов и эпидемиологического прогнозирования¹, ведущий научный сотрудник⁴. E-mail: senkoov@mail.ru, ORCID: 0000-0002-5586-3503

Кузнецова Анна Викторовна, к.б.н., старший научный сотрудник лаборатории математической биофизики³. E-mail: azforus@yandex.ru, ORCID: 0000-0003-2705-1935

Демина Ирина Алексеевна, лаборант-исследователь¹, врач². E-mail: doctor.demira@gmail.com

Воронин Евгений Михайлович, к.м.н., руководитель научной группы математических методов и эпидемиологического прогнозирования¹. E-mail: emvoronin@yandex.ru, ORCID: 0000-0001-5925-7757

Плоскирева Антонина Александровна, профессор РАН, д.м.н., заместитель директора по клинической работе¹. E-mail: antoninna@mail.ru, ORCID: 0000-0002-3612-1889

Посынкина Юлия Романовна, младший научный сотрудник научной группы математических методов и эпидемиологического прогнозирования¹. E-mail: yulya.melnichenko98@gmail.com, ORCID: 0009-0001-3560-7676

Акимкин Василий Геннадьевич, академик РАН, д.м.н., профессор, директор¹. E-mail: crie@pcr.ru, ORCID: 0000-0003-4228-9044

About authors

Senko Oleg V., Dr Physical and Mathematical Sci, professor, senior researcher at Scientific Group of Mathematical Methods and Epidemiological Forecasting¹, leading researcher⁴. E-mail: senkoov@mail.ru, ORCID: 0000-0002-5586-3503

Kuznetsova Anna V., PhD Bio Sci, senior researcher at Laboratory of Mathematical Biophysics³. E-mail: azforus@yandex.ru, ORCID: 0000-0003-2705-1935

Demina Irina A., laboratory assistant-researcher¹, physician². E-mail: doctor.demira@gmail.com

Voronin Evgeny M., PhD Med, head of Scientific Group of Mathematical Methods and Epidemiological Forecasting¹. E-mail: emvoronin@yandex.ru, ORCID: 0000-0001-5925-7757

Plaskireva Antonina A., RAS professor, DM Sci (habil.), deputy director for Clinical Work¹. E-mail: antoninna@mail.ru, ORCID: 0000-0002-3612-1889

Posynkina Iuliia R., junior researcher at Scientific Group of Mathematical Methods and Epidemiological Forecasting¹. E-mail: yulya.melnichenko98@gmail.com, ORCID: 0009-0001-3560-7676

Akimkin Vasily G., RAS academician, DM Sci (habil.), professor, director¹. E-mail: crie@pcr.ru, ORCID: 0000-0003-4228-9044

¹ ФБУН «Центральный научно-исследовательский институт эпидемиологии» Роспотребнадзора, Москва, Россия

² ГБУЗ «Московский многопрофильный научно-клинический центр имени С.П. Боткина Департамента здравоохранения города Москвы, Россия

³ ФБУН «Институт биохимической физики имени Н.М. Эмануэля» Российской академии наук, Москва, Россия

⁴ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия

Автор для переписки: Сенько Олег Валентинович. E-mail: senkoov@mail.ru

Для цитирования: Сенько О.В., Кузнецова А.В., Демина И.А., Воронин Е.М., Плоскирева А.А., Посынкина Ю.Р., Акимкин В.Г. Оценка риска развития средней и тяжелой степени тяжести пневмонии при COVID-19 методами машинного обучения. *Медицинский алфавит*. 2025; (16): 12–20. <https://doi.org/10.33667/2078-5631-2025-16-12-20>

¹ Central Research Institute of Epidemiology, Moscow, Russia

² Moscow Multidisciplinary Scientific and Clinical Center n.a. S.P. Botkin (Botkin hospital), Moscow, Russia

³ Emanuel Institute of Biochemical Physics of Russian Academy of Sciences, Moscow, Russia

⁴ Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

Corresponding author: Senko Oleg V. E-mail: senkoov@mail.ru

For citation: Senko O.V., Kuznetsova A.V., Demina I.A., Voronin E.M., Plaskireva A.A., Posynkina I.R., Akimkin V.G. Assessing the risk of developing moderate and severe pneumonia in COVID-19 patients using machine learning methods. *Medical alphabet*. 2025; (16): 12–20. <https://doi.org/10.33667/2078-5631-2025-16-12-20>

