

УДК 616.12:578.834.1"COVID-19"519.688(001.8)

DOI: 10.36604/1998-5029-2023-88-50-58

ВОЗМОЖНОСТИ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ ОЦЕНКИ ИСХОДОВ COVID-19 У ПАЦИЕНТОВ С ЗАБОЛЕВАНИЯМИ СИСТЕМЫ КРОВИ

А.В.Талько¹, В.А.Невзорова², М.З.Ермолицкая³, Ж.В.Бондарева²

¹Государственное бюджетное учреждение здравоохранения «Краевая клиническая больница №2», 690105,
г. Владивосток, ул. Русская, 55

²Федеральное государственное бюджетное образовательное учреждение высшего образования
«Тихоокеанский государственный медицинский университет» Министерства здравоохранения Российской
Федерации, 690002, г. Владивосток, пр-т Острякова, 2

³Федеральное государственное бюджетное учреждение науки «Институт автоматики и процессов
управления» Дальневосточного отделения Российской академии наук, 690041, г. Владивосток, ул. Радио, 5.

РЕЗЮМЕ. Введение. Различные технологии искусственного интеллекта получают широкое применение во многих областях медицины с интеграцией в научно-исследовательскую и практическую работу, включая гематологию. Привлекательность методов машинного обучения обусловлена возможностью исключения субъективного фактора, как оценки состояния пациента, так и результатов обследования. **Цель.** Построение прогностической модели выживаемости пациентов гематологического профиля при заболевании COVID-19. **Материалы и методы.** Ретроспективно проанализированы 144 медицинские карты пациентов со злокачественными и доброкачественными заболеваниями системы крови, получавших лечение в Краевой клинической больнице №2 г. Владивостока. Средний возраст больных составил 64 года. Твердая конечная точка – летальность пациентов от всех причин (46 человек или 32%). В качестве предикторов для построения прогностических моделей использовали такие показатели как тип заболевания (злокачественное, доброкачественное); этап терапии; клинические проявления COVID-19 (есть/нет), симптомы инфекции, статус по шкале ECOG на момент поступления, сопутствующие заболевания, терапия глюкокортикостероидами, использование увлажненного кислорода и осложнения COVID-19. При построении прогностических моделей с бинарным классификатором использовали методы машинного обучения: логистическую регрессию, дерево решения на основе «условного вывода» и «случайный лес». **Результаты.** Были разработаны 3 прогностические модели. Выбор модели зависел от количества включаемых параметров. Согласно F-мере, точность модели «случайный лес» оказалась выше. На основании выбранных методов машинного обучения наличие дыхательной недостаточности, требующей кислородной поддержки, явилось самым значимым предиктором прогнозирования исхода COVID-19. **Заключение.** Проведенное нами исследование позволило выявить значимые предикторы неблагоприятного исхода, на основе которых построены прогностические модели выживаемости пациентов гематологического профиля при заболевании коронавирусной инфекцией.

Ключевые слова: COVID-19, заболевания системы крови, искусственный интеллект, машинное обучение, алгоритм случайного леса.

THE POSSIBILITIES OF DATA MINING METHODS FOR ASSESSING THE OUTCOMES OF COVID-19 IN PATIENTS WITH DISEASES OF THE BLOOD SYSTEM

A.V.Talko¹, V.A.Nevzorova², M.Z.Ermolitskaya³, Zh.V. Bondareva²

¹Regional Clinical Hospital No.2, 55 Russkaya Str., Vladivostok, 690105, Russian Federation

²Pacific State Medical University, 2 Ostryakova Ave., Vladivostok, 690002, Russian Federation

Контактная информация

Ангелина Владимировна Талько, врач-гематолог, Государственное бюджетное учреждение здравоохранения «Краевая клиническая больница №2», 690105, Россия, г. Владивосток, ул. Русская, 55. E-mail: talkang92@mail.ru

Correspondence should be addressed to

Angelina V. Talko, MD, Hematologist, Regional Clinical Hospital No.2, 55 Russkaya Str., Vladivostok, 690105, Russian Federation. E-mail: talkang92@mail.ru

Для цитирования:

Талько А.В., Невзорова В.А., Ермолицкая М.З., Бондарева Ж.В. Возможности методов интеллектуального анализа данных для оценки исходов COVID-19 у пациентов с заболеваниями системы крови // Бюллетень физиологии и патологии дыхания. 2023. Вып.88. С.50–58. DOI: 10.36604/1998-5029-2023-88-50-58

For citation:

Talko A.V., Nevzorova V.A., Ermolitskaya M.Z., Bondareva Zh.V. The possibilities of data mining methods for assessing the outcomes of COVID-19 in patients with diseases of the blood system. *Bulleten' fiziologii i patologii dyhaniâ = Bulletin Physiology and Pathology of Respiration* 2023; (88):50–58 (in Russian). DOI: 10.36604/1998-5029-2023-88-50-58

³*Institute of Automation and Control Processes of the Far Eastern Branch of the Russian Academy of Sciences, 5 Radio Str. Vladivostok, 690041, Russian Federation*

SUMMARY. Introduction. Various artificial intelligence technologies are widely used in many areas of medicine with integration into research and practical work, including hematology. The attractiveness of machine learning methods is due to the possibility of excluding the subjective factor both assessment of the patient's condition and examination results. **Aim.** The construction of a predictive survival model for hematological patients with COVID-19 coronavirus infection. **Materials and methods.** 144 medical records of patients with malignant and benign diseases of the blood system treated at the Regional Clinical Hospital No. 2 in Vladivostok were retrospectively analyzed. The average age of the studied patients was 64 years. The solid endpoint is the mortality of patients from all causes (46 people or 32%). Indicators such as the type of disease (malignant, benign); the stage of therapy; clinical manifestations of COVID-19 (yes/no); symptoms of infection were used as predictors for constructing predictive models; ECOG status at the time of admission; concomitant diseases; glucocorticosteroids therapy; the use of humidified oxygen and complications of COVID-19. When constructing predictive models with a binary classifier, machine learning methods were used: logistic regression, a decision tree based on "conditional inference" and a "random forest". **Results.** 3 predictive models were developed. The choice of the model depended on the number of parameters included. According to the F-measure, the accuracy of the "random forest" model was higher. Based on the selected machine learning methods, the presence of respiratory failure requiring oxygen support was the most significant predictor of forecasting the outcome of COVID-19. **Conclusion.** Our study allowed us to identify significant predictors of an unfavorable outcome, on the basis of which prognostic models of survival of hematological patients with coronavirus infection were built.

Key words: COVID-19, diseases of the blood system, artificial intelligence, machine learning, random forest.

Машинное обучение – раздел информатики и статистики, который генерирует прогностические или описательные модели путем обучения данных [1]. Внедрение методов машинного обучения в медицине началось в XX веке (1970-е гг.) для помощи врачу на этапах диагностики и прогнозирования исходов заболевания [2]. Их использование позволяет анализировать большие объемы информации о пациентах и помогает прогнозировать особенности течения и исходов различных заболеваний [3, 4].

Согласно данным B.Sekeroglu & K.Tuncal [5], изучившим заболеваемость и смертность онкологических пациентов в странах Европы, использовались различные методы машинного обучения. При сравнении моделей для надежного прогнозирования исходов линейная регрессия и регрессия с опорным вектором превосходили другие способы оценки.

Различные технологии искусственного интеллекта получают широкое применение во многих областях медицины с интеграцией в научно-исследовательскую и практическую работу, включая гематологию [6]. Привлекательность методов машинного обучения обусловлена возможностью исключения субъективного фактора, как оценки состояния пациента, так и результатов обследования [7].

Пациенты с болезнями системы крови являются уязвимой группой по заражению коронавирусной инфекцией вследствие более старшего возраста возникновения большинства гематологических заболеваний, а также в силу иммунокомпроментированного статуса как результата патогенеза заболевания, так и использования специфической химиотерапии [8].

Заболеваемость онкогематологических больных COVID-19 значительно усугубляет течение болезни, снижает выживаемость и увеличивает риск смертности [9]. По данным российских исследований [10] корона-

вирусная инфекция у пациентов гематологического профиля осложняет течение основного заболевания в 16,7% случаев, а летальность от COVID-19 у этих пациентов составляет 24,7%. В сложившейся ситуации прогнозирование исходов у пациентов гематологического профиля с совокупной оценкой всех возможных входящих точек, используя технологии, основанные на искусственном интеллекте, могут существенно повлиять на систему принятия решений о тактике их ведения.

Целью данного исследования явилось построение прогностической модели выживаемости пациентов гематологического профиля при заболевании коронавирусной инфекцией COVID-19.

Материалы и методы исследования

В качестве исходных данных были ретроспективно проанализированы 144 медицинские карты проживающих в Приморском крае пациентов со злокачественными и доброкачественными заболеваниями системы крови с подтвержденной методом полимеразной цепной реакции ПНК SARS-CoV-2 коронавирусной инфекцией COVID-19, получавших лечение в Краевой клинической больнице №2 г. Владивостока в период с апреля 2020 по октябрь 2021 года. Возраст больных составил от 18 до 90 лет (средний возраст 64 года). Твердой конечной точкой выбрана летальность пациентов от всех причин (46 человек или 32%).

В качестве предикторов для построения прогностических моделей использовали следующие показатели: тип заболевания (злокачественное, доброкачественное); этап терапии (до начала лечения/в процессе химиотерапии/без лечения); клинические проявления COVID-19 (есть/нет); симптомы инфекции – лихорадка, сухой кашель, утомляемость, боль в горле, одышка, боль в груди, гипертермия; статус по шкале оценки состояния

онкологического больного (Eastern Cooperative Oncology Group или ECOG) на момент поступления; сопутствующие заболевания, такие как сахарный диабет и другие (гипертоническая болезнь, ожирение), терапия глюкокортикостероидами (ГКС), использование увлажненного кислорода и осложнения COVID-19 (двусторонняя пневмония, сепсис, острый респираторный дистресс-синдром (ОРДС), полиорганная недостаточность, дыхательная недостаточность, цитокиновый шторм). Согласно критерию Фишера, выделенные показатели взаимосвязаны с выживаемостью гематологических пациентов при заболевании COVID-19 ($p < 0.05$).

Реализация анализа и обработка данных осуществлялись в программе RStudio (Version 1.0.153). Были использованы следующие статистические критерии: критерий Шапиро-Уилка для проверки на нормальность распределения, критерий Фишера для проверки независимости признаков и критерий Краскела-Уоллиса для проверки равенства медиан нескольких выборок. Результаты считали статистически значимыми при значениях уровня значимости $p < 0.05$. Для оценки связи между определенным исходом и фактором риска рассчитывали отношение шансов (ОШ). При построении прогнозных моделей с бинарным классификатором использовали следующие методы машинного обучения: логистическую регрессию, дерево решения на основе «условного вывода» и «случайный лес». Проверка качества моделей осуществлялась с помощью расчета среднеквадратической ошибки (MSE) и построения матрицы неточностей (confusion matrix), позволяющей определить чувствительность (Sensitivity) – истинный положительный показатель, специфичность (Specificity) – истинный отрицательный

показатель и точность (Accuracy) – доля правильных ответов (количество ответов, по которым классификатор принял правильное решение, деленное на размер выборки). Дополнительно для каждой модели рассчитана F-мера, представляющая собой гармоническое среднее между точностью (Precision) и полнотой (Recall).

У всех пациентов было получено письменное информированное согласие. Исследование было выполнено в соответствии со стандартами надлежащей клинической практики (Good Clinical Practice) и принципами Хельсинкской Декларации.

Исследование было одобрено на заседании локального независимого междисциплинарного комитета по этике (протокол №4 от 19.12.2022 г.) ФГБОУ ВО ТГМУ Минздрава России.

Результаты исследования

Для построения прогнозных моделей исследуемая выборка была предварительно поделена на обучающую и тестовую выборки в соотношении 3:1. Обучающая выборка насчитывала 107 наблюдений (34 случая со злокачественными и 73 с доброкачественными заболеваниями), тестовая выборка – 36 наблюдений (11 и 25 случаев, соответственно).

При построении логистической регрессии в результате прямого отбора предикторов достоверная связь получена с семью показателями. Анализ значений коэффициентов регрессии показал, что наибольшее отрицательное влияние на исход заболевания установлено для следующих показателей: полиорганная недостаточность, боль в горле, сахарный диабет, использование кислородотерапии и ГКС (табл. 1).

Таблица 1

Оценки коэффициентов регрессии логистической модели

Показатели	Коэффициент	p-value
Свободный коэффициент	1,036	<2e-16
Боль в горле	-0,248	0,001309
Гипертермия	0,076	0,017616
ECOG на момент поступления	-0,065	0,032193
Сахарный диабет	-0,182	0,030296
Терапия ГКС	-0,146	0,010997
Полиорганная недостаточность	-0,341	0,000163
Кислородная терапия	-0,178	3,08e-07

Согласно матрице неточностей на обучающей выборке четыре пациента неправильно классифицированы; на тестовой выборке – три пациента (табл. 2).

Деревья решений осуществляют разбиение пространства объектов в соответствии с некоторым набором правил разбиения (splitting rule), в результате чего формируются компактные деревья, не требующие про-

цедуры обрезания. При этом автоматически выполняется отбор информативных предикторов и учитываются возможные взаимодействия между ними. Это метод решений одинаково хорошо применим как к количественным, так и к качественным показателям, что важно в нашем исследовании.

Для построения прогнозных моделей использовали

22 показателя. В результате получили, что на обучающей выборке 8 пациентов неправильно классифицированы, на тестовой выборке все пациенты правильно

классифицированы. Оценки качества модели представлены в таблице 2.

Таблица 2

Проверка качества построенных моделей

Модель	Выборка	MSE	Accuracy	Sensitivity	Specificity	F-мера
Логистическая регрессия	обучающая	0,037	0,963	1	0,948	0,937
	тестовая	0,083	0,917	1	0,880	0,88
Дерево решений на основе «условного вывода»	обучающая	0,075	0,925	0,906	0,933	0,878
	тестовая	0	1	1	1	1
Случайный лес	обучающая	0,009	0,9907	1	0,9865	0,985
	тестовая	0,056	0,944	1	0,92	0,917

Случайный лес (Random Forest) – это набор деревьев решений. Каждое дерево решений предсказывает выходной класс на основе соответствующих переменных-предикторов, используемых в этом дереве. Результат всех деревьев решений в «случайном лесу» записывается, и класс с большинством голосов вычисляется как выходной класс. Нами было построено 500 деревьев.

Согласно матрице ошибок, на обучающей выборке один пациент неправильно классифицирован, на тесто-

вой выборке два пациента неправильно классифицированы (табл. 2).

Модели, полученные с помощью классификационных деревьев, по сравнению с логистической моделью, в данном исследовании являются наиболее качественными. Сравнивая эти две модели следует заметить, что они включают в себя разное количество предикторов (16 и 22, соответственно), но по важности предикторов они согласуются (рис. 1 и 2).

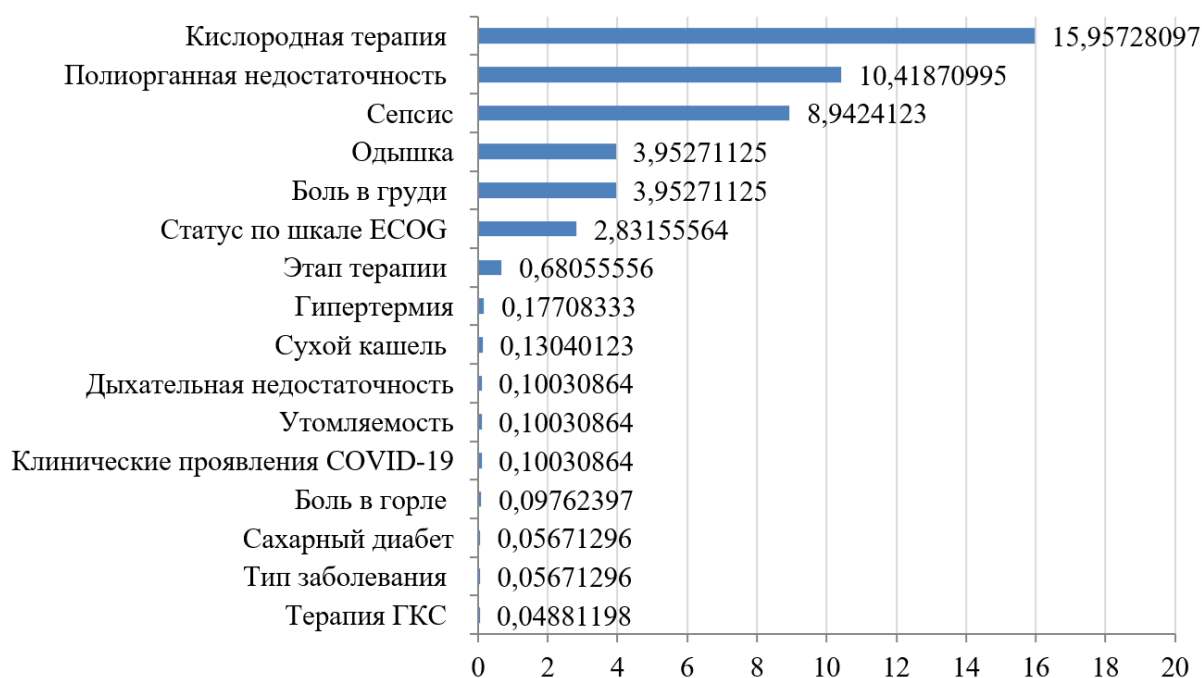


Рис. 1. Значения важности предикторов (Importance), полученных с помощью алгоритма дерева решений на основе «условного вывода».



Рис. 2. Значения важности предикторов по методу «случайный лес».

Наиболее важным показателем при прогнозировании исхода COVID-19 является использование кислородной терапии (включая искусственную вентиляцию легких – ИВЛ) (табл. 3). Подача кислорода была необходима 47% (n=68) пациентам (62 с опухолевыми и 6 с неопухолевыми заболеваниями системы крови), из них 60% (n=41) были на ИВЛ (38 с опухолевыми и 3 с неопухолевыми заболеваниями). Среди 46 умерших 84,8% (n=39) находились на ИВЛ (37 человек с опухолевыми и 2 с неопухолевыми заболеваниями крови). У 8,7% (n=4) больных на ИВЛ причина смерти – прогрессирование/осложнение гематологического заболевания. Из находящихся на ИВЛ 4,9% (n=2) остались в живых (1 женщина с опухолевыми и 1 с неопухолевыми заболеваниями). А 17,1% (n=7) больных умерли без ИВЛ (6 человек с опухолевыми и 1 с неопухолевыми заболеваниями), причина смерти у 4 человек (все с опухолевыми заболеваниями крови) – осложнения COVID-19.

Согласно данным таблицы 3, на втором месте по важности предикторов – наличие полиорганной недостаточности. Всего у 23,6% (n=34) пациентов (32 человека с опухолевыми и 2 с неопухолевыми заболеваниями крови) развилась полиорганная недостаточность, из них 91,2% (n=31) пациентов умерли.

Сепсис установлен у 20% (n=29) пациентов – 12 женщин с опухолевыми заболеваниями системы крови

и 17 мужчин (16 с опухолевыми и 1 с неопухолевыми заболеваниями). Из них 93% (n=27) умерли.

Обнаружена взаимосвязь показателя «сепсис» с увеличением объема поражения легких согласно стадии КТ (p-value=0,0007328).

Обсуждение результатов исследования

Разработка прогностических моделей на основе интеллектуального анализа данных рассматривается в качестве перспективного подхода для создания системы принятия врачебных решений без участия субъективных факторов. В период пандемии COVID-19 беспрецедентной по масштабам охвата, длительности течения, видоизменчивости и вирулентности вируса вопросы прогнозирования течения заболевания приобретают особую актуальность, особенно у пациентов с наличием ко- и полиморбидности.

На основе проведенного анализа и полученных нами результатов можно выделить модель классификации с использованием алгоритма «случайный лес», т.к. на обучающей выборке значения MSE наименьшие, а точность (Ассигасу) и F-меры наибольшие. При проверке на тестовых данных модель «случайный лес» уступает по эффективности модели, полученной с помощью дерева решений на основе «условного вывода», где правильно предсказаны все наблюдения.

Таблица 3

Анализ бинарных предикторов на основе расчета отношения шансов и применения критерия Фишера

Показатели	Отношение шансов (ОШ)	ДИ (95%)	Уровень значимости (p)
Кислород	41,85333	11,92597-146,8812	<2,2e-16
Полиорганная недостаточность	65,44444	17,76051-241,1516	<2,2e-16
Сепсис	65,44444	14,94358-311,3496	9,33e-15
Одышка	14,65079	3,209699-66,87411	7,275e-11
Боль в груди	10,05714	4,180048-24,19736	6,442e-08
Дыхательная недостаточность	7,412308	3,376634-16,27132	1,902e-07
Сахарный диабет	5,411765	1,882444-15,55807	1,899e-3
Терапия ГКС	3,363636	1,605852-7,045511	1,339e-3
Боль в горле	4,326389	1,219885-10,96267	2,613e-3
Другие сопутствующие заболевания	3,791667	1,237931-11,61352	0,01523
Сухой кашель	2,717514	1,181139-6,25234	0,02232
Тип заболевания	3,675214	1,032806-13,07816	0,04896
Циклотинный шторм	9,894737	2,812302-34,81341	0,01875
Двусторонняя пневмония	8,366197	1,895427-36,92743	7,281e-4
Лихорадка	9,246377	2,094839-40,81244	3,72e-4
Утомляемость	3,909091	1,102222-13,86381	0,03019
Осложнения COVID-19	14,59459	1,908296-111,6191	6,431e-4

В результате нашего исследования было разработано 3 прогностические модели. Выбор модели зависел от количества включаемых параметров. В логистическую регрессию на основе коэффициента регрессии включали предикторы методом перебора, в то время как в метод «случайного леса» внесены все 22 признака. Согласно F-мере, точность модели «случайный лес» оказалась выше.

В работе Л.А.Гривкова и соавт. [11] наиболее точный прогноз рецидивов базально-клеточного рака кожи был получен с помощью модели на основе многофакторной линейной регрессии, а прогностическая точность модели стохастического градиентного бустинга была менее заметной.

Метод «случайного леса» показал свою эффективность при оценке прогнозирования вероятности возникновения бронхиальной астмы у детей раннего возраста [12]. Основными методами машинного обучения, используемыми при решении задач в репродуктологии, были метод опорных векторов, алгоритмы «случайный лес» и «дерево решений» [13]. В исследовании P.Nafiz et al. [14] Random Forest превзошел остальные алгоритмы. При анализе методов искус-

ственного интеллекта в диагностике заболеваний мочеполовой системы наиболее распространенными методами как раз являлись «дерево решений» и «случайный лес» [15]. В работе H.L.Wang et al. [16] Random Forest показал самую высокую точность прогнозирования.

В большинстве исследований, направленных на поиск значимых маркеров прогнозирования течения коронавирусной инфекции, чаще использовались методы регрессионного анализа [17–19]. У пациентов с заболеваниями системы крови и полиморбидностью возможности методов интеллектуального анализа данных для оценки исходов COVID-19 были проанализированы впервые. Наибольшая точность в нашем исследовании получилась в методе «случайный лес», что позволяет в дальнейшем применять эти методы машинного обучения в практической работе согласно созданию баз данных пациентов.

На основании выбранных методов машинного обучения наличие дыхательной недостаточности, требующей кислородной поддержки, явилось самым значимым предиктором прогнозирования исхода COVID-19, что коррелирует с данными изученной литературы [20].

Заключение

Проведенное нами исследование позволило выявить значимые предикторы неблагоприятного исхода, на основе которых построены прогностические модели выживаемости пациентов гематологического профиля при заболевании коронавирусной инфекцией. Полученные результаты могут быть использованы в клинической практике для создания системы принятия решений при лечении COVID-19 у пациентов изучаемого профиля.

Конфликт интересов

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи

Conflict of interest

The authors declare no conflict of interest

Источники финансирования

Исследование проводилось без участия спонсоров

Funding Sources

This study was not sponsored

ЛИТЕРАТУРА

1. Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies // Lancet Haematol. 2020. Vol.7, Iss.7. e541–e550. EDN: SXSAHV. [https://doi.org/10.1016/S2352-3026\(20\)30121-6](https://doi.org/10.1016/S2352-3026(20)30121-6)
2. Weiss S.M., Kulikowski C.A., Amarel S., Safir A. A model-based method for computer-aided medical decision making // Artif. Intelligence. 1978. Vol.11, Iss.1-2. P.145–172. [https://doi.org/10.1016/0004-3702\(78\)90015-2](https://doi.org/10.1016/0004-3702(78)90015-2)
3. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction // BMC Med. Res. Methodol. 2019. Vol.19. Article number: 64. EDN: CMNPFK. <https://doi.org/10.1186/s12874-019-0681-4>
4. Гусев А.В., Новицкий Р.Э., Ившин А.А., Алексеев А.А. Машинное обучение на лабораторных данных для прогнозирования заболеваний // Фармакоэкономика. Современная фармакоэкономика и фармакоэпидемиология. 2021. Т.14, №4. С.581–529. EDN: ANFRKW. <https://doi.org/10.17749/2070-4909/farmakoeconomika.2021.115>
5. Sekeroglu B., Tuncal K. Prediction of cancer incidence rates for the European continent using machine learning models // Health Informatics J. 2021; Vol.27, Iss.1. Article number: 1460458220983878. <https://doi.org/10.1177/1460458220983878>
6. Лучинин А.С. Искусственный интеллект в гематологии // Клиническая онкогематология. 2022. Т.15, №1. С.16–27. EDN: ISVFJA. <https://doi.org/10.21320/2500-2139-2022-15-1-16-27>
7. Venkatesh R., Balasubramanian C., Kahappan M. Development of big data predictive analytics model for disease prediction using machine learning technique // J. Med. Syst. 2019. Vol.43, Iss.8. Article number: 272. <https://doi.org/10.1007/s10916-019-1398-y>
8. Piñana J.L., Martino R., García-García I., Parody R., Morales M.D., Benzo G., Gómez-Catalan I., Coll R., De La Fuente I., Luna A., Merchán B., China A., de Miguel D., Serrano A., Pérez C., Díaz C., Lopez J.L., Saez A.J., Bailen R., Zudaire T. et al. Risk factors and outcome of COVID-19 in patients with hematological malignancies // Exp. Hematol. Oncol. 2020; Vol.9. Article number: 21. <https://doi.org/10.1186/s40164-020-00177-z>
9. Liang W., Guan W., Chen R., Wang W., Li J., Xu K., Li C., Ai Q., Lu W., Liang H., Li S., He J. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China // Lancet Oncol. 2020. Vol.21, Iss.3. P.335–377. [https://doi.org/10.1016/S1470-2045\(20\)30096-6](https://doi.org/10.1016/S1470-2045(20)30096-6)
10. Зинченко А.В., Лим В.С., Крюков Е.В., Казаков С.П., Игнатъева Е.В., Рукавицын О.А. COVID-19 в гематологическом стационаре, течение и исходы // Гематология. Трансфузиология. Восточная Европа. 2021. Т.7, №2. С.131–141. EDN: SAURGI. <https://doi.org/10.34883/PI.2021.7.2.001>
11. Гривков Л.А., Шахгельдян К.И., Гельцер Б.И., Котельников В.Н., Апанасевич В.И. Методы машинного обучения в прогнозировании рецидивов базальноклеточного рака кожи после фотодинамической терапии // Тихоокеанский медицинский журнал. 2022. №2(88). P.54–59. EDN: GUIWTI. <https://doi.org/10.34215/1609-1175-2022-2-54-59>
12. Баширов А.Н., Воронов В.И. Прогнозирование вероятности возникновения бронхиальной астмы у детей с применением алгоритма случайного леса // Современные наукоемкие технологии. 2019. №12-2. С.249–255. EDN: ZTDCCD. <https://doi.org/10.17513/snt.37867>
13. Ившин А.А., Багаудин Т.З., Гусев А.В. Искусственный интеллект на страже репродуктивного здоровья: этические и клинические особенности // Акушерство и гинекология. 2021. №5. С.17–24. EDN: MMMTKH. <https://doi.org/10.18565/aig.2021.5.17-24>
14. Hafiz P., Nematollahi M., Boostani R., Namavar Jahromi B. Predicting implantation outcome of in vitro fertilization and intracytoplasmic sperm injection using data mining techniques // Int. J. Fertil. Steril. 2017. Vol.11, Iss.3. P.184–190. <https://doi.org/10.22074/ijfs.2017.4882>
15. Yoruk U., Hargraves B.A., Vasanawala S.S. Automatic Renal Segmentation for MR Urography Using 3D-GrabCut and Random Forests // Magn. Reson. Med. 2017. Vol.79, Iss.3. P.1696–1707. <https://doi.org/10.1002/mrm.26806>
16. Wang H.L., Hsu W.Y., Lee M.H., Weng H.H., Chang S.W., Yang J.T., Tsai Y.H. Automatic machine-learning-based outcome prediction in patients with primary intracerebral hemorrhage // Front. Neurol. 2019. Vol.10: Article number: 910.

<https://doi.org/10.3389/fneur.2019.00910>

17. Мелик-Гусейнов Д.В., Карякин Н.Н., Благоднарова А.С., Климов В.И., Баврина А.П., Другова О.В., Саперкин Н.В., Ковалишова О.В. Регрессионные модели прогнозирования количества летальных исходов при новой коронавирусной инфекции // Современные технологии в медицине. 2020. Т.12, №2. С.6–13. EDN: GUXMYH. <https://doi.org/10.17691/stm2020.12.2.01>

18. Vaid A., Somani S., Russak A.J., De Freitas J.K., Chaudhry F.F., Paranjpe I., Johnson K.W., Lee S.J., Miotto R., Richter F., Zhao S., Beckmann N.D., Naik N., Kia A., Timsina P., Lala A., Paranjpe M., Golden E., Danieleto M., Singh M., Bottinger E.P., Nadkarni G.N., Glicksberg B.S., Percha B., Aberg J.A., Horowitz C.R. et al. Machine Learning to Predict Mortality and Critical Events in a Cohort of Patients With COVID-19 in New York City: Model Development and Validation // J. Med. Internet Res. 2020. Vol.22, Iss.11. Article number: e24018. EDN: AWUCQB. <https://doi.org/10.2196/24018>

19. Yan L., Zhang H., Goncalves J., Xiao Ya., Wang M., Guo Yu., Sun Ch., Tang X., Jing L., Zhang M., Huang X., Xiao Y., Cao H., Chen Ya., Ren T., Wang F., Xiao Ya., Huang S., Tan X., Huang N. et al. An interpretable mortality prediction model for COVID-19 patients // Nat. Mach. Intell. 2020. Vol.2, Iss.5. P.283–288. EDN: UIKCVI. <https://doi.org/10.1038/s42256-020-0180-7>

20. Авдеев С.Н., Царева Н.А., Мерзоева З.М., Трущенко Н.В., Ярошецкий А.И. Практические рекомендации по кислородотерапии и респираторной поддержке пациентов с COVID-19 на дорегистрационном этапе // Пульмонология. 2020. Т.30, №2. С.151–163. EDN: FINWPJ. <https://doi.org/10.18093/0869-0189-2020-30-2-151-163>

REFERENCES

1. Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies. *Lancet Haematol.* 2020; 7(7):e541–e550. [https://doi.org/10.1016/S2352-3026\(20\)30121-6](https://doi.org/10.1016/S2352-3026(20)30121-6)

2. Weiss S.M., Kulikowski C.A., Amarel S., Safir A. A model-based method for computer-aided medical decision making. *Artif. Intelligence* 1978; 11(1-2):145–172. [https://doi.org/10.1016/0004-3702\(78\)90015-2](https://doi.org/10.1016/0004-3702(78)90015-2)

3. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med. Res. Methodol.* 201; 19:64. <https://doi.org/10.1186/s12874-019-0681-4>

4. Gusev A.V., Novitskiy R.E., Ivshin A.A., Alekseev A.A. [Machine learning based on laboratory data for disease prediction]. *FARMAKOEKONOMIKA. Modern Pharmacoeconomics and Pharmacoepidemiology* 2021; 14(4):581–592 (in Russian). <https://doi.org/10.17749/2070-4909/farmakoeconomika.2021.115>

5. Sekeroglu B., Tuncal K. Prediction of cancer incidence rates for the European continent using machine learning models. *Health Informatics J.* 2021; 27(1):1460458220983878. <https://doi.org/10.1177/1460458220983878>

6. Luchinin A.S. [Artificial intelligence in hematology]. *Clinical Oncohematology* 2022; 15(1):16–27 (in Russian). <https://doi.org/10.21320/2500-2139-2022-15-1-16-27>

7. Venkatesh R., Balasubramanian C., Kahappan M. Development of big data predictive analytics model for disease prediction using machine learning technique. *J. Med. Syst.* 2019; 43(8):272. <https://doi.org/10.1007/s10916-019-1398-y>

8. Piñana J.L., Martino R., García-García I., Parody R., Morales M.D., Benzo G., Gómez-Catalan I., Coll R., De La Fuente I., Luna A., Merchán B., China A., de Miguel D., Serrano A., Pérez C., Diaz C., Lopez J.L., Saez A.J., Bailen R., Zudaire T. et al. Risk factors and outcome of COVID-19 in patients with hematological malignancies. *Exp. Hematol. Oncol.* 2020; 9:21. <https://doi.org/10.1186/s40164-020-00177-z>

9. Liang W., Guan W., Chen R., Wang W., Li J., Xu K., Li C., Ai Q., Lu W., Liang H., Li S., He J. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol.* 2020; 21(3):335–377. [https://doi.org/10.1016/S1470-2045\(20\)30096-6](https://doi.org/10.1016/S1470-2045(20)30096-6)

10. Zinchenko A.V., Lim V.S., Kryukov E.V., Kazakov S.P., Ignatieva E.V., Rukavitsyn O.A. [COVID-19 in the hematological hospital, course and outcomes]. *Hematology. Transfusiology. Eastern Europe* 2021; 7(2):131–141 (in Russian). <https://doi.org/10.34883/PI.2021.7.2.001>

11. Grivkov L.A., Shahgeldyan K.I., Geltser B.I., Kotelnikov V.N., Apanasevich V.I. [Machine learning methods in prediction of basal cell skin cancer recurrence after photodynamic therapy]. *Pacific Medical Journal* 2022; (2):54–59 (in Russian). <https://doi.org/10.34215/1609-1175-2022-2-54-59>

12. Bashirov A.N., Voronov V.I. [Prediction of the probability of occurrence of bronchial asthma in children with the use of the random forest algorithm]. *Sovremennye naukoemkie tehnologii = Modern High Technologies* 2019; (12-2):249–255. <https://doi.org/10.17513/snt.37867>

13. Ivshin A.A., Bagaudin T.Z., Gusev A.V. [Artificial intelligence on guard of reproductive health]. *Akusherstvo i Ginekologiya/Obstetrics and Gynecology* 2021; (5):17–24 (in Russian). <https://doi.org/10.18565/aig.2021.5.17-24>

14. Hafiz P., Nematollahi M., Boostani R., Namavar Jahromi B. Predicting implantation outcome of in vitro fertilization and intracytoplasmic sperm injection using data mining techniques. *Int. J. Fertil. Steril.* 2017; 11(3):184–190. <https://doi.org/10.22074/ijfs.2017.4882>

15. Yoruk U., Hargraves B.A., Vasanawala S.S. Automatic Renal Segmentation for MR Urography Using 3D-GrabCut

and Random Forests. *Magn. Reson. Med.* 2017; 79(3):1696–1707. <https://doi.org/10.1002/mrm.26806>

16. Wang H.L., Hsu W.Y., Lee M.H., Weng H.H., Chang S.W., Yang J.T., Tsai Y.H. Automatic machine-learning-based outcome prediction in patients with primary intracerebral hemorrhage. *Front. Neurol.* 2019; 10:910. <https://doi.org/10.3389/fneur.2019.00910>

17. Melik-Huseynov D.V., Karyakin N.N., Blagonravova A.S., Klimko V.I., Bavrina A.P., Drugova O.V., Saperkin N.V., Kovalishena O.V. Regression Models Predicting the number of deaths from the new coronavirus infection. *Sovremennye tehnologii v medicine* 2020; 12(2):6–13. <https://doi.org/10.17691/stm2020.12.2.01>

18. Vaid A., Somani S., Russak A.J., De Freitas J.K., Chaudhry F.F., Paranjpe I., Johnson K.W., Lee S.J., Miotto R., Richter F., Zhao S., Beckmann N.D., Naik N., Kia A., Timsina P., Lala A., Paranjpe M., Golden E., Danieleto M., Singh M., Bottinger E.P., Nadkarni G.N., Glicksberg B.S., Percha B., Aberg J.A., Horowitz C.R. et al. Machine Learning to Predict Mortality and Critical Events in a Cohort of Patients With COVID-19 in New York City: Model Development and Validation. *J. Med. Internet Res.* 2020; 22(11): e24018. <https://doi.org/10.2196/24018>

19. Yan L., Zhang H., Goncalves J., Xiao Ya., Wang M., Guo Yu., Sun Ch., Tang X., Jing L., Zhang M., Huang X., Xiao Y., Cao H., Chen Ya., Ren T., Wang F., Xiao Ya., Huang S., Tan X., Huang N. et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* 2020; 2(5):283–288. <https://doi.org/10.1038/s42256-020-0180-7>

20. Avdeev S.N., Tsareva N.N., Merzhoeva Z.M., Trushenko N.V., Yaroshetskiy A.I. [Practical guidance for oxygen treatment and respiratory support of patients with COVID-19 infection before admission to intensive care unit]. *Pulmologiya* 2020; 30(2):151–163 (in Russian). <https://doi.org/10.18093/0869-0189-2020-30-2-151-163>

Информация об авторах:

Ангелина Владимировна Талько, врач-гематолог, Государственное бюджетное учреждение здравоохранения «Краевая клиническая больница №2»; e-mail: talkang92@mail.ru

Вера Афанасьевна Невзорова, д-р мед. наук, профессор, директор Института терапии и инструментальной диагностики, Федеральное государственное бюджетное образовательное учреждение высшего образования «Тихоокеанский государственный медицинский университет» Министерства здравоохранения Российской Федерации; e-mail: nevzorova@inbox.ru

Марина Захаровна Ермолицкая, канд. биол. наук, доцент, старший научный сотрудник лаборатории информационно-аналитических и управляющих систем и технологий, Федеральное государственное бюджетное учреждение науки «Институт автоматизации и процессов управления» Дальневосточного отделения Российской академии наук; e-mail: ermmmz@mail.ru

Жанна Викторовна Бондарева, канд. мед. наук, доцент Института терапии и инструментальной диагностики, Федеральное государственное бюджетное образовательное учреждение высшего образования «Тихоокеанский государственный медицинский университет» Министерства здравоохранения Российской Федерации; e-mail: bondareva.zhv@tgmu.ru

Author information:

Angelina V. Talko, MD, Hematologist, Primorsky Regional Clinical Hospital No.2; e-mail: talkang92@mail.ru

Vera A. Nevzorova, MD, PhD, DSc (Med.), Professor, Director of the Institute of Therapy and Instrumental Diagnostics, Pacific State Medical University; e-mail: nevzorova@inbox.ru

Marina Z. Ermolitskaya, PhD (Biol.), Associate Professor, Senior Staff Scientist, Institute of Automation and Control Processes FEB RAS; e-mail: ermmmz@mail.ru

Zhanna V. Bondareva, MD, PhD (Med.), Associate Professor, Institute of Therapy and Instrumental Diagnostics, Pacific State Medical University; e-mail: bondareva.zhv@tgmu.ru

Поступила 10.05.2023
Принята к печати 29.05.2023

Received May 10, 2023
Accepted May 29, 2023