

А.В. ГУСЕВ,

к.т.н., директор по развитию ООО «К-Скай», г. Петрозаводск, Россия, e-mail: agusev@webiomed.ai, ORCID: 0000-0002-7380-8460

Р.Э. НОВИЦКИЙ,

генеральный директор ООО «К-Скай», г. Петрозаводск, Россия, e-mail: roman@webiomed.ai, ORCID: 0000-0002-2350-977X

ТЕХНОЛОГИИ ПРОГНОЗНОЙ АНАЛИТИКИ В БОРЬБЕ С ПАНДЕМИЕЙ COVID-19

УДК: 004.81

DOI: 10.37690/1811-0193-2020-4-24-33

Гусев А.В., Новицкий Р.Э. Технологии прогнозной аналитики в борьбе с пандемией COVID-19 (ООО «К-Скай», г. Петрозаводск, Россия)

Аннотация. В последнее время новая коронавирусная инфекция или COVID-19, вызванная возбудителем SARS-CoV-2, продолжает быстрое распространение по всему миру. По мнению Всемирной организации здравоохранения (ВОЗ), объявившей эту вспышку пандемией, COVID-19 является серьезной проблемой для общественного здравоохранения, имеющей международное значение. Из-за отсутствия доказанного эффективного лечения и вакцинации против COVID-19 меры предосторожности считаются ВОЗ стратегическими целями и основным способом противодействия пандемии. Руководствам стран рекомендовано принять национальные программы медицинского обслуживания, направленные на оценку и снижение риска распространения инфекции. На этом фоне технологии прогнозной аналитики стали активно использоваться для составления популяционных и персональных прогнозов развития заболеваемости, смертности, оценки тяжести течения болезни и т.д. В данной статье представлен обзор имеющихся разработок и публикаций по теме применения прогнозной аналитики для борьбы с пандемией COVID-19.

Ключевые слова: программное обеспечение, прогнозная аналитика, искусственный интеллект, машинное обучение, dashboard, COVID-19.

UDC: 004.81

Gusev A.V., Novitsky R.E. Predictive analytics technologies in the management of the COVID-19 pandemic (LLC "K-Skai", Petrozavodsk, Russia)

Abstract. Recently, a new coronavirus infection, or COVID-19, caused by the pathogen SARS-CoV-2, has been continuing to spread around the world rapidly. According to the World Health Organization (WHO), which declared this outbreak a pandemic, COVID-19 is a serious public health problem of international concern. Due to the lack of proven effective treatment and vaccination against COVID-19, precautions are considered by WHO to be strategic goals and a primary response to the pandemic. It is recommended that country guidelines adopt national health care programs aimed at assessing and reducing the risk of infection spread. Predictive analytics have begun to be actively used to compile population and personal forecasts of the progression of morbidity, mortality, assess the severity of the course of the disease, etc. This article provides an overview of available developments and publications on the use of predictive analytics in the management of COVID-19 pandemic.

Keywords: software, predictive analytics, artificial intelligence, machine learning, dashboard, COVID-19.

ВВЕДЕНИЕ

Новая коронавирусная инфекция или COVID-19, вызванная возбудителем SARS-CoV-2, продолжает быстрое распространение по всему миру. Всемирная организация здравоохранения (ВОЗ) объявила 11 марта 2020 г. вспышку COVID-19 пандемией. По мнению ВОЗ, она является серьезной проблемой для общественного здравоохранения, имеющей международное значение [1].

Отрицательная динамика роста заболеваемости и смертности привели к острой необходимости улучшения диагностики заболевания, разработке вакцин и новых лекарственных препаратов с целью обнаружения, профилактики и лечения этого смертельно опасного вируса. Существенный рост заболеваемости COVID-19 привел к серьезному давлению на системы здравоохранения практически во всех странах мира, положение которых характеризуется как кризисное [2].

Из-за отсутствия эффективного лечения и вакцинации против COVID-19 меры предосторожности считаются ВОЗ стратегическими целями и основным способом противостояния пандемии. Правительствам рекомендовано принять национальные программы медицинского обслуживания, направленные на **оценку и снижение риска** распространения инфекции, особенно для пациентов с хроническими заболеваниями [1, 2].

На фоне развития пандемии COVID-19 возросла роль прогнозной аналитики, которая является частью более широкого термина «медицинская аналитика». Под медицинской аналитикой мы понимаем использование данных, статистического и количественного анализа и аналитических моделей, включая прогнозное моделирование, которые на основании выявления и понимания ретроспективных данных, помогают понять будущие события и осуществлять поддержку принятия решений. Медицинская аналитика разделяется на несколько видов: описательная аналитика (Descriptive Analytics), прогнозная аналитика (Predictive Analytics) и предписывающая аналитика (Prescriptive Analytics) [3].

Прогнозная аналитика (Predictive Analytics) использует имеющиеся ретроспективные данные и различные прогнозные модели, созданные в том числе с помощью машинного обучения, чтобы помочь ответить на вопрос «Что может случиться?». Прогнозную аналитику можно условно разделить на управленческую и клиническую. **Управленческая прогнозная аналитика** предназначена для поддержки принятия управленческих решений и используется руководителями различного уровня для оценки возможных сценариев развития заболеваний, нагрузки на медицинские организации, потребности в лекарствах и т.д. **Клиническая прогнозная аналитика** предназначена для поддержки принятия врачебных решений, включая анализ медицинских данных пациентов, контроль правильности лекарственной терапии, выбор оптимальной маршрутизации и тактики ведения пациента и т.д.

По данным ряда маркетинговых отчетов последнего времени, пандемия COVID-19 привела к существенному повышению интереса инвесторов к медицинской аналитике. Отчеты 2019 г. давали оценку роста этого рынка к 2027 г. до 18–40 млрд. долл. США [4, 5]. Аналогичные отчеты в апреле 2020 г. повысили его оценку до 80 млрд. долл. в 2027 г. [6], а последний отчет имеет прогноз уже 84 млрд. долл. США в 2027 г., причем на

долю именно прогнозной аналитики приходится 28,6 млрд. долл. США или 34% от всего размера рынка медицинской аналитики [7].

ОБЗОР РАЗРАБОТОК В ОБЛАСТИ ПРОГНОЗНОЙ УПРАВЛЕНСКОЙ АНАЛИТИКИ

В период любой эпидемии, и тем более пандемии, критически важно принимать продуманные и обоснованные стратегические решения в области управления национальным здравоохранением, поскольку как переоценка, так и недооценка ситуации может приводить к плачевным результатам не только в плане зараженных и заболевших, но и в плане экономических потерь, социальной напряженности и многих других глобальных проблем. Принятие решений политиками, руководителями здравоохранения разного уровня должно быть основано на точных и достоверных прогнозах количества инфицированных и заболевших, а также на основе оценки действительной динамики смертности, хотя бы в краткосрочной перспективе.

Отсутствие или игнорирование таких прогнозов в отношении COVID-19, особенно в течение первой волны пандемии, приводило к неопределенности, что, в свою очередь, порождало порой громкие и пугающие заявления, усиливающие глобальную тревогу. Вводимые властями непопулярные меры, основанные в том числе на противоречивых и вселяющих беспокойство прогнозах множества экспертов, вызывали недовольство населения и недоверие к способности системы справиться с пандемией.

В начале февраля 2020 г., еще до официального объявления пандемии, глава ВОЗ заявил, что случаи новой коронавирусной инфекции, зафиксированные в Китае, могут быть лишь «верхушкой айсберга». После этого ведущий эпидемиолог Гонконга в области общественного здравоохранения Габриэль Леунг сообщил, что около 60% населения земли будет инфицировано новым вирусом [8]. В середине февраля 2020 профессор Гарвардской школы общественного здравоохранения Чан Марк Липсич в интервью Wall Street Journal предсказал, что до конца 2020 года будут инфицированы до 70% населения планеты [9]. В середине марта 2020 г. канцлер Германии Ангела Меркель высказала мнение, что до 70% жителей страны могут заразиться коронавирусом [10]. Такие устрашающие прогнозы звучали практически каждый день, мгновенно распространялись по

социальным сетям и СМИ, и в итоге в мире фактически развилось две пандемии: первая – вирусная, а вторая – информационная.

В то же самое время ряд экспертов стал отмечать, что громкие заявления о возможном числе инфицированных и умерших от COVID-19 делаются без проведения достоверных математических прогнозов и скорее являются признаками страха и паники, что может привести к разрушительным экономическим и социальным потрясениям [11]. Немецкий иммунолог и токсиколог, профессор Стефан Хокерц заявил, что COVID-19 не более опасен, чем грипп, просто ему уделяется больше внимания. Опаснее вируса являются страх и паника, создаваемые СМИ, и «авторитарная реакция» многих правительств. Профессор Хокерц также отметил, что большинство так называемых «умерших от коронавируса» на самом деле просто имели положительный результат теста, а умерли от других причин [12].

Представленные противоречивые заявления экспертов только укрепляют наше понимание, что наличие достоверных популяционных прогностических моделей развития ситуации является жизненно-важным аспектом работы руководителей в сфере здравоохранения, политиков и государственных служащих во многих странах. В апреле 2020 г. Президент РФ Владимир Путин, обращаясь к Правительству, заявил: «Для того, чтобы и дальше действовать на упреждение, нам нужен профессиональный прогноз, в том числе на ближайшее время – трёх, семи, десяти дней. Я прошу правительство вести и постоянно обновлять такой прогноз, исходя из реально складывающейся обстановки, и прошу докладывать мне об этом ежедневно» [13]. Некоторые авторы также стали обращать внимание, что средства массовой информации и общественности следует больше опираться на доступные математические прогностические модели, построенные на открытых данных [14].

В истории человечества было много вспышек инфекционных заболеваний, таких как лихорадка денге, малярия, грипп, ВИЧ / СПИД. Создание надлежащей эпидемиологической модели для этих эпидемий – сложная задача. Для прогнозирования развития пандемии COVID-19 было предложено большое количество разнообразных моделей, многие из них были основаны на моделях SIS, SIR и SEIR [15]. Большое количество моделей были опубликованы в открытом доступе.

Модель SIR направлена на прогнозирование 3 показателей численности людей S (susceptible-восприимчивых), I (infected – инфицированных) и R (recovered – выздоровевших). Эта модель была введена в 1927 году, менее чем через 10 лет после пандемии гриппа 1918 года. Ее популярность может быть частично связана с ее простотой, которая позволяет разработчикам предсказать поведение инфекции путем оценки небольшого числа параметров [16].

Модель SEIR имеет четыре прогностических элемента: S (восприимчивые), E (подверженные воздействию), I (инфицированные) и R (выздоровевшие). Основная гипотеза модели SEIR состоит в том, что со временем все участники модели будут играть четыре роли. Модель SEIR имеет некоторые ограничения для реальных ситуаций, но она обеспечивает базовую модель для исследования различных видов эпидемий, в том числе активно применялась для прогнозирования развития пандемии COVID-19 [15].

Yu-Feng Zhao и соав. предложили свой подход на основе 6-ти скользящих серых моделей Verhulst с использованием 7-, 8- и 9-дневных последовательностей данных для прогнозирования суточной тенденции роста числа пациентов, у которых подтверждена инфекция COVID-19 в Китае. Максимальная и минимальная средняя абсолютная ошибка в процентах (MAPE) на этапе тестирования составили 4,72% и 1,65% соответственно. Это означает, что предсказанные результаты демонстрируют высокую надежность [17].

Для создания прогностических моделей в мире появилось много общедоступных материалов, таких как данные Университета Джона Хопкинса, Лаборатории Декарта, Бюро переписи населения США, данные ВОЗ и др.

Многие медицинские центры объявили о том, что будут предоставлять в открытом доступе обезличенные наборы клинических сведений, таких как истории болезни или медицинские изображения, с целью применения машинного обучения и создания цифровых инструментов для борьбы с COVID-19, в том числе и для прогностической аналитики. Например, сеть медицинских клиник HMC Hospitals предоставила доступ к 2157 анонимным историям болезней, содержащим различную информацию о пациентах, получивших лечение от COVID-19, включая подробные описания диагнозов, лечения, нахождения в отделении интенсивной терапии, результаты лабораторной диагностики, данные о выписке или смерти.

Поскольку существующие в открытом доступе материалы о COVID-19 сфокусированы преимущественно на демографических критериях, применение дополнительных обезличенных баз медицинских данных является важным шагом вперед в части создания более точных прогностических инструментов.

В России был запущен специальный портал с открытой информацией по COVID-19, доступный по адресу: <https://стопкоронавирус.рф/information/>. Также открытые данные доступны на портале Единой межведомственной информационно-статистической системы (ЕМИСС, <https://www.fedstat.ru/>), а также «Портале открытых данных Правительства Москвы», <https://data.mos.ru/>.

В открытом доступе стали появляться различные аналитические панели, показывающие динамику пандемии и интегрированные в них прогностические модели. Одна из самых известных и используемых аналитических панелей была предложена университетом Джона Хопкинса, доступ к которой был открыт 23 января 2020. Панель визуализирует

данные, полученные из ВОЗ, центров по контролю и профилактике заболеваний США (Centers for Disease Control and Prevention), Европейского центра профилактики и контроля заболеваний (European Centre for Disease Prevention and Control), национальной комиссии здравоохранения КНР (National Health Commission of the People's Republic of China), трекеров реального времени 1point3acres, DXY и Worldmeters.info, новостного агрегатора BNO, государственных и национальных правительственных департаментов здравоохранения, а также сообщений местных СМИ. Составной частью панели является модель прогнозирования эпидемии, которая выкладывается на GitHub по адресу: <https://github.com/HopkinsIDD/COVIDScenarioPipeline>.

Построенная аналитическая панель стала самым визуализированным набором данных о пандемии (рис. 1). Доступ к панели «Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)»

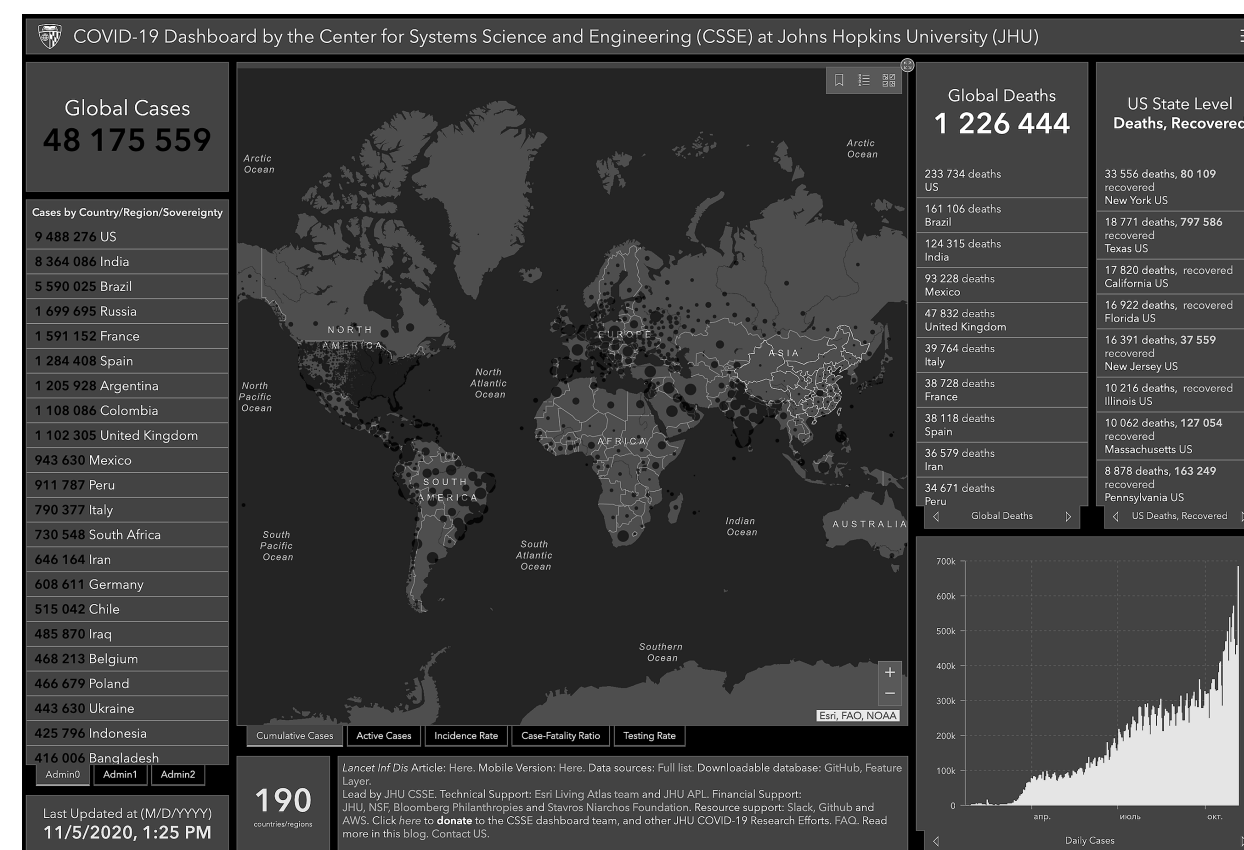


Рис. 1. Аналитическая панель Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU), по данным на 05.11.2020

открыт по адресу: <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>.

ВОЗ также в постоянном режиме анализирует заболеваемость и смертность населения во всех странах мира. В открытом доступе находится аналитическая панель «WHO Coronavirus Disease (COVID-19) Dashboard», <https://covid19.who.int/>, скриншот которой представлен на рис. 2.

Google Cloud в сотрудничестве с Гарвардским глобальным институтом здравоохранения представили «Национальный портал реагирования» (National Response Portal, NRP), доступный по адресу: <https://map.nationalresponseportal.com/portal>. На нем они представили публичные прогнозы случаев COVID-19, смертей и других показателей на ближайшие 14 дней для округов и штатов США. Для этого были разработаны новые модели машинного обучения на основе временных рядов и достоверных открытых эпидемиологических баз данных [18].

Академический медицинский центр «Клиника Кливленда» (Cleveland Clinic), управляющий 11

больницами и 19 медицинскими центрами в различных штатах США, совместно с американской компанией SaS разработали аналитическую панель для принятия решений в связи с пандемией COVID-19 и прогнозированию ее развития в различных разрезах, состоящую из 10 различных прогностических моделей. Некоторые из них опубликованы в открытом доступе на GitHub по адресу: <https://github.com/sassoftware/covid-19-sas>. При этом модели предназначались для различных сценариев, включая полную неэффективность ограничительных мер и т.д. Модели прогнозировали смертность, число пациентов, нуждающихся в госпитализации или аппаратах искусственной вентиляции лёгких и т.д. На основании прогнозов руководство центра принимало решения о разворачивании дополнительных коек, открытии временных больниц и т.д. По мнению руководителей, это позволило центру более эффективно организовывать медицинские процессы, что в свою очередь позволило достичь более низких показателей заболеваемости по сравнению с другими штатами [19].

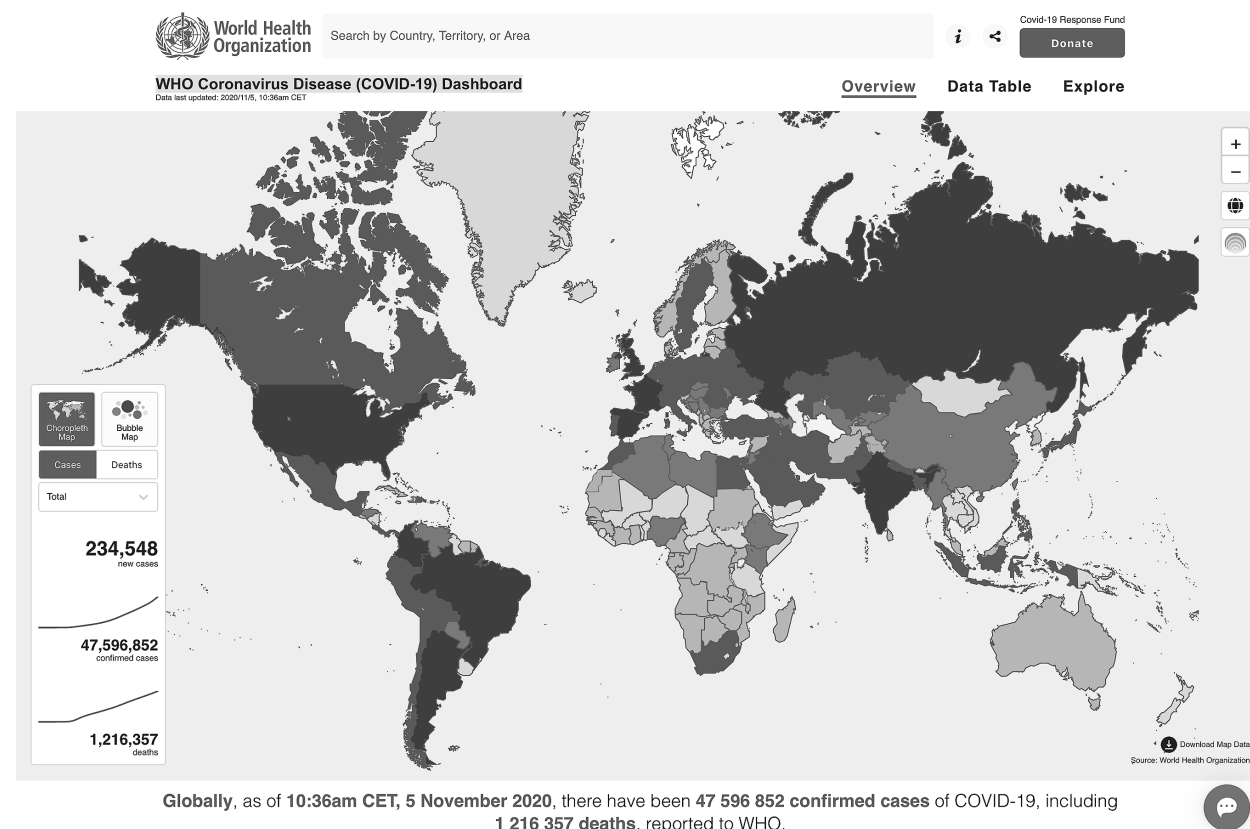


Рис. 2. Аналитическая панель WHO Coronavirus Disease (COVID-19) Dashboard по данным на 05.11.2020

Группа специалистов из Исследовательского центра операций Массачусетского технологического института (MIT) под руководством профессора Димитриса Берцимаса разработала и предоставила в открытом доступе несколько прогностических инструментов, включая популяционные модели и калькуляторы персонального риска для применения медицинскими центрами и политиками США в качестве инструментов борьбы с COVID-19. Все наработки доступны по адресу: <https://www.covidanalytics.io/>.

Модель Калифорнийского Университета в Лос-Анжелесе (UCLA) на основе открытых данных и машинного обучения предназначена для использования политиками и руководителями медицинских клиник с целью предсказания развития эпидемии и принятия решений о необходимости развертывания дополнительных мощностей. Адрес доступа: <https://covid19.uclaml.org/>

Многие штаты, университеты и исследовательские центры также разработали свои аналитические панели, их адреса доступа:

- <http://www.covid19dashboard.us/>;
- <https://covid-19.tacc.utexas.edu/projections/>;
- <https://www.cc.gatech.edu/~badityap/covid.html>;
- <https://covid19.healthdata.org/united-states-of-america>;
- <https://covid-19.bsvgateway.org/>;
- <https://covid19.gleamproject.org/>;
- <https://github.com/shaman-lab/COVID-19-Projection>.

ОБЗОР РАЗРАБОТОК В ОБЛАСТИ ПРОГНОЗНОЙ КЛИНИЧЕСКОЙ АНАЛИТИКИ

Пандемия COVID-19 привела к существенному росту нагрузки на систему здравоохранения. Были мобилизованы все возможные силы, созданы новые инфекционные больницы, пересмотрены правила направления пациентов на обследование и лечение. Основная задача, которую решали организаторы здравоохранения практически во всех странах мира, состояла в перенастройке системы здравоохранения с целью справиться с возросшей нагрузкой из-за обращений пациентов COVID-19 с учетом имеющихся ограниченных ресурсов, таких важных как медицинский персонал, средства индивидуальной защиты, оборудование для интенсивной терапии.

Медицинским организациям, оказывающим помощь больным с COVID-19, приходилось

принимать, обследовать и лечить намного большее количество пациентов, чем было рассчитано в штатной ситуации. Нагрузка на медицинский персонал возросла. В этой связи появилось множество разработок и публикаций, посвященных применению искусственного интеллекта в качестве систем поддержки принятия врачебных решений (СППВР), главным образом в области автоматизированного анализа компьютерной томографии (КТ) грудной клетки с целью сокращения врачебных ошибок и нагрузки на медицинский персонал [21, 22, 23].

Вместе с этим разработки в области прогнозной клинической аналитики также эффективно проявили себя в нескольких аспектах:

1. Как симпомчекеры, которые определяли вероятность, что пациент заразился коронавирусной инфекцией.
2. Как инструменты оценки риска заражения, например, на основании информации о контактах, перемещениях граждан и т.д.
3. Как калькуляторы оценки персонального риска развития осложнений и смерти.
4. Как системы помощи при маршрутизации пациентов и т.д.

При лечении пациентов с COVID-19 важным является быстрая, точная и ранняя клиническая оценка тяжести заболевания. На её основе необходима быстрая и эффективная сортировка пациентов, которая имеет решающее значение как для эффективного лечения пациентов, так и для обоснованного и рационального использования ограниченных ресурсов, таких как аппараты ИВЛ, реанимационные койки, медицинский персонал. Такую персональную индивидуализированную оценку состояния пациента и возможных сценариев развития здоровья предоставляют прогнозные модели, разрабатываемые для применения в реальной клинической практике как СППВР.

В настоящее время в мире разработано множество прогностических инструментов для применения врачами. Bayat и соавт. представили модель, умеющую с точностью 86,4% предсказать диагноз COVID-19 на основании 20 стандартных лабораторных тестов, выполненных в течение 2-дневного периода до даты тестирования на COVID-19 [24].

Yan и соавт. опубликовали в «Nature Machine Intelligence» статью, в которой представили результаты разработки модели на основе машинного обучения, предсказывающей госпитальную летальность пациентов с COVID-19 более чем за 10 дней с точностью выше 90%. На основе

анализа базы данных образцов крови 485 инфицированных пациентов г. Ухань, Китай, авторы выделили 3 входных параметра: высокочувствительный тест на С-реактивный белок, уровень лактатдегидрогеназы крови и относительное содержание лимфоцитов в крови [25].

Liang и соавт. представили модель прогнозирования риска развития критических состояний у пациентов с COVID-19, позволяющую выявлять пациентов, подверженных риску развития серьезных осложнений и проводить им оперативные лечебные мероприятия на ранней стадии. На основе данных о 1590 пациентах из 575 медицинских центров исследователи разработали основанную на алгоритмах глубокого машинного обучения модель выживания, предсказывающую развитие критических состояний в течение 5, 10 и 30 дней, исходя из персональных клинических характеристик при поступлении. Изучив 74 клинические характеристики, авторы выделили 10 ключевых факторов риска, включая патологию, зарегистрированную при рентгеновской визуализации, возраст, одышку и ряд сопутствующих заболеваний [26].

Schalekamp и соавт. с помощью многомерной логистической регрессии разработали калькулятор индивидуальной оценки риска развития критического состояния у пациентов с COVID-19, таких как смерть, госпитализация в отделение интенсивной терапии, необходимость инвазивной вентиляции легких и т.д. Модель анализирует клинические, рентгенологические и лабораторные данные пациента. AUC модели составила 0,77 [27].

Jehi и соавт. представили модель прогнозирования госпитализации пациентов с COVID-19. Используя базу данных лабораторного подтвержденного диагноза COVID-19 у 4536 пациентов, авторы применили метод логистической регрессии оператора наименьшего абсолютного сжатия и выбора «LASSO». AUC модели составила 0,90 [28].

Компании «Digital Health» и «Optum Ventures» разработали инструмент «Buoy COVID-19 Digital Symptom Tracker» для оценки риска заражения COVID-19, используя руководство Центров по контролю и профилактике заболеваний.

В США Управление по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA) предоставило экстренное разрешение для системы оценки медицинских данных на основе искусственного интеллекта «CLEW Medical», которое позволяет выявлять инфицированных COVID-19 с высоким риском дыхательной или

сердечно-сосудистой недостаточности, поступивших в отделение интенсивной терапии. Система «CLEWICU» принимает большой спектр данных о пациентах, включая сведения из электронных медицинских карт (ЭМК) и подключенных медицинских устройств без необходимости ручного ввода информации. В ответ система выдает предупреждение о потенциальном ухудшении состояния пациента для принятия врачебных решений. Кроме выявления пациентов высокого риска, система также определяет пациентов, которые не нуждаются в каких-либо вмешательствах в течение ближайших восьми часов. Принимая решение о разрешении для указанной системы, FDA сообщило, что «CLEWICU» может помочь уменьшить контакт между медицинскими работниками и пациентами с COVID-19. «Благодаря выявлению пациентов с высоким риском и определения приоритетов лечения можно снизить частоту ложных опасений. Это, вероятно, обеспечит лучшую помощь пациентам и снизит нагрузку на персонал отделений интенсивной терапии», — сказано в заявлении FDA, которое выдало разрешение на экстренное использование системы [29].

Система «eCart», используемая в более, чем 20 госпиталях США, применяется для предсказания летальности у пациентов, находящихся в отделениях реанимации. Система интегрирует данные из ЭМК пациентов (например, показатели жизнедеятельности, результаты лабораторных исследований и демографическую информацию) для обеспечения оценки рисков пациентов в реальном времени [30]. Новая версия системы, модернизированная в связи с пандемией COVID-19, использует около 100 различных параметров для работы прогнозных моделей.

В Великобритании в Национальной системе здравоохранения (NHS) внедряется сервис, который автоматически выявляет пациентов высокого риска, например, пожилых людей и пациентов, имеющих хронические заболевания, потенциально опасные для осложненного течения COVID-19. Сервис автоматически персонально информирует их о принадлежности к группе риска с рекомендациями оставаться дома [31].

В Израиле Maccabi Healthcare Services, одна из крупнейших организаций здравоохранения, использует ИИ для определения наиболее подверженных риску тяжелых осложнений COVID-19 из обслуживаемых 2,4 миллионов человек. Для этого разработчики адаптировали уже существовавшую

раннее систему, обученную давать оценку риска по гриппу. Используя миллионы медицинских данных, накопленные в Maccabi за 27 лет, удалось создать систему выявления пациентов высокого риска на основе таких данных, как возраст, индекс массы тела, наличие хронических заболеваний, данные предыдущих госпитализаций и т.д. ИИ-система может автоматически просматривать огромное количество записей и выявлять лиц высокого риска. На основе этой оценки Maccabi автоматически определяет рекомендуемую индивидуальную тактику ведения каждого пациента: следует ли его оставить дома, отправить в карантинный отель или на госпитализацию [32].

Еще одна разработка из Израиля, созданная в интересах Министерства обороны Израиля, тоже использует ИИ и обучение на основании собранных в медицинских клиниках данных для определения пациентов с высоким риском развития осложнений от COVID-19. Прогнозная система на базе этой технологии может предупреждать медицинский персонал о возможном ухудшении состояния пациента, привлекая внимание врачей к пациентам с более высокими рисками прогрессирования болезни и тем самым значительно улучшая эффективность лечения пациента. Модель была разработана исследователями и инженерами подразделения «Systems Missiles and Space Group» компании «IAI». Система может предупредить врачей об ухудшении или улучшении состояния пациента в ближайший день с точностью 80%. При этом в случаях, когда состояние пациента остается прежним, точность модели составляет от 90% до 92% [33].

Компания «Jvion» использует методы машинного обучения для создания алгоритма, который помогает определить вероятность заражения и распространения вируса, а также заражения, требующего госпитализации. Для этого в «Jvion» оценивают индивидуальные факторы риска, минимальные или отсутствующие клинические данные и используют ограниченную информацию, которая не требует специальных медицинских устройств. Дополнительно к медицинским данным, ИИ способен учесть такие факторы как проживание в плотно населённых районах, посещение общественных мероприятий и многое другое. Это позволит применить ИИ для автоматической оценки больших групп населения с целью определения среди них лиц высокого риска без необходимости их личного посещения медицинской организации. Такие инструменты позволяют соединить в одном

как управленческую популяционную аналитику, так и персональные выявление и мониторинг пациентов высокого риска [34].

ОСНОВНЫЕ ПРОБЛЕМЫ И ПУТИ РАЗВИТИЯ ПРОГНОЗНОЙ АНАЛИТИКИ КАК ИНСТРУМЕНТА БОРЬБЫ С ПАНДЕМИЕЙ COVID-19

Основная проблема большинства управленческих прогнозных моделей состоит в том, что они учитывают весьма ограниченный набор данных, таких как статистика заболеваемости и смертности, которые в действительности не являются индикаторами причин развития пандемии, а служат лишь проекциями состоявшихся фактов. В этой связи, многие модели давали и продолжают давать существенные расхождения в реальной точности. Maggie Koerth, Laura Bronner и Jasmine Mithani приводят интересные примеры: Центры по контролю и профилактике заболеваний США (The Centers for Disease Control and Prevention) используют модель, согласно которой в конце марта 2020 в США ожидалось около 200 тыс. смертей. В то же самое время Imperial College London предсказывал, что смертность в США в тот же период времени составит 2,2 млн. человек [17].

Для повышения эффективности прогнозных управленческих моделей целесообразно повышать их предсказательную точность. Для этого, вероятно, они должны обновляться ежедневно с целью коррекции своих параметров на основании оперативных данных, а также, по возможности, учитывать распространенность основных факторов риска COVID-19, действующих ограничительных мер и многих других факторов, которые еще предстоит осознать.

Еще одной существенной проблемой повышения эффективности прогнозной аналитики является отсутствие достаточных исходных данных. Многие открытые наборы ограничены лишь демографическими показателями и общими данными заболеваемости и смертности. Для развития прогнозных инструментов нужны дополнительные входные данные, включая сведения из ЭМК, точные сведения о диагнозе заболевания, лабораторные и иные данные. Такие наборы просто недоступны на уровне регионов или стран в целом, хотя присутствуют во многих локально установленных системах ведения ЭМК, что является одной из главных проблем развития популяционных прогнозных инструментов

[31]. Необходимо развивать программы, направленные на сбор и предоставление открытого доступа к таким материалам.

Важнейшим направлением дальнейших исследований и разработок как в части популяционных прогнозных моделей, так и в моделях для персональной оценки риска, является оценка эффекта различных интервенций в снижение бремени заболеваемости COVID-19.

Руководители и лица, принимающие решения, в настоящее время нуждаются не только и не столько в моделях прогнозирования пандемии, сколько в инструментах, созданных по принципу «что, если?» – другими словами в инструментах предписывающей аналитики. Такие разработки могли бы помочь оценить эффективность различных управленческих решений и сценариев развития заболеваемости, например, таких как ограничительные меры, и содействовать более эффективному и своевременному управлению, в том числе – и отъезду от мер, не имеющих достоверного влияния на клинические исходы.

В условиях пандемии врачи в большей степени нуждаются не в калькуляторах оценки риска, а в надёжных и эффективных системах поддержки принятия врачебных решений, бесшовно

интегрированных в системы ведения ЭМК. В условиях чрезмерной нагрузки и дефицита времени у врачей нет возможности использовать точные, но часто требующие ручного ввода модели оценки рисков. Вместо создания все новых и новых калькуляторов, разработчикам следует сосредоточиться на интеграции указанных инструментов в ЭМК и обеспечении автоматизации их работы. Лишь в случае обнаружения опасной ситуации прогнозные модели должны давать соответствующий сигнал в ЭМК, а система ЭМК должна выводить его на экран врачу или отправлять соответствующее сообщение на мобильный телефон, тем самым обращая внимание врача только в обоснованном случае, который требует врачебного вмешательства.

Финансирование исследования и конфликт интересов. Исследование финансировалось в рамках выполнения гранта компании ООО «К-Скай» с Некоммерческой Организацией «Фонд развития центра разработки и коммерциализации новых технологий» (Фонд «Сколково») № МГО4/20 от 22 мая 2020 г. по теме «Разработка системы поддержки принятия врачебных решений с использованием методов машинного обучения WEBIOMED». Конфликты интересов, связанные с данным исследованием, отсутствуют.

ЛИТЕРАТУРА



1. World Health Organization. Coronavirus disease (COVID-19) pandemic, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
2. Hamed M.A. An overview on COVID-19: reality and expectation. Bull Natl Res Cent. 2020;44(1):86. doi: 10.1186/s42269-020-00341-9. Epub 2020 Jun 1. PMID: 32514228; PMCID: PMC7266424, <https://doi.org/10.1186/s42269-020-00341-9>
3. Khalifa M. Health Analytics Types, Functions and Levels: A Review of Literature. Stud Health Technol Inform. 2018;251:137–140. PMID: 29968621, <https://pubmed.ncbi.nlm.nih.gov/29968621/>
4. Healthcare Analytics Market Size By Product (Descriptive Analytics, Predictive Analytics, Prescriptive Analytics), By Application (Operations Management, Financial Management, Population Health Management, Clinical Management) By End-Use (Hospitals, Clinics, Others), Industry Analysis Report, Regional Outlook, Application Potential, Price Trends, Competitive Market Share & Forecast, 2019–2025, <https://www.gminsights.com/industry-analysis/healthcare-analytics-market>
5. Global Healthcare Analytics Markets, 2019–2024 by Type, Solution, Deployment Model, Application, End-user, and Geography – ResearchAndMarkets.com, <https://www.researchandmarkets.com/reports/4858254/healthcare-analytics-market-forecasts-from-2019>
6. Healthcare Analytics Market Size, Share & Industry Analysis, By Product (Descriptive, Predictive, and Prescriptive), By Application (Financial Analytics, Population Health Analytics, Clinical Analytics, and Operations and Administrative Analytics), By End User (Payers, Providers, and Others) and Regional Forecast, 2019–2026, <https://www.fortunebusinessinsights.com/healthcare-analytics-market-102641>
7. Healthcare Analytics Market by Type (Predictive, Prescriptive), Component (Hardware, Software, and Services), Delivery Mode (Cloud), Application (Clinical, RCM, Claims, Fraud, Risk, PHM), End user (Payer, Provider) and Geography- Global Forecast to 2027, <https://www.meticulousresearch.com/product/healthcare-analytics-market/>
8. Coronavirus 'could infect 60% of global population if unchecked', <https://www.theguardian.com/world/2020/feb/11/coronavirus-expert-warns-infection-could-reach-60-of-worlds-population>
9. Harvard Professor Sounds Alarm on 'Likely' Coronavirus Pandemic: 40% to 70% of World Could Be Infected This Year, <https://www.mediaweek.com/news/harvard-professor-sounds-alarm-on-likely-coronavirus-pandemic-40-to-70-of-world-could-be-infected-this-year/>

10. Coronavirus: Up to 70% of Germany could become infected – Merkel, <https://www.bbc.com/news/world-us-canada-51835856>
11. The Cognitive Bias That Makes Us Panic About Coronavirus, <https://www.bloomberg.com/opinion/articles/2020-02-28/coronavirus-panic-caused-by-probability-neglect>
12. Коронавирус: опасная иллюзия смертности, <https://habr.com/ru/post/494896/>
13. Путин о коронавирусе: «Нам нужен профессиональный прогноз», <https://vostokmedia.com/news/society/13-04-2020/putin-o-koronaviruse-nam-nuzhen-professionalnyy-prognoz>
14. Jewell N.P., Lewnard J.A., Jewell B.L. Predictive Mathematical Models of the COVID-19 Pandemic: Underlying Principles and Value of Projections. JAMA. 2020;323(19):1893–1894. doi:10.1001/jama.2020.6585, <https://jamanetwork.com/journals/jama/fullarticle/2764824>
15. He S., Peng Y. & Sun K. SEIR modeling of the COVID-19 and its dynamics. Nonlinear Dyn 101, 1667–1680 (2020). <https://doi.org/10.1007/s11071-020-05743-y>
16. Tolles J., Luong T. Modeling Epidemics With Compartmental Models. JAMA. 2020;323(24):2515–2516. doi:10.1001/jama.2020.8420, <https://jamanetwork.com/journals/jama/fullarticle/2766672>
17. Zhao Y.-F., Shou M.-H., Wang Z.-X. Prediction of the Number of Patients Infected with COVID-19 Based on Rolling Grey Verhulst Models. Int. J. Environ. Res. Public Health 2020, 17, 4582, <https://doi.org/10.3390/ijerph17124582>
18. Why It's So Freaking Hard To Make A Good COVID-19 Model, <https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/>
19. Google Cloud AI and Harvard Global Health Institute Collaborate on new COVID-19 forecasting model, <https://cloud.google.com/blog/products/ai-machine-learning/google-cloud-is-releasing-the-covid-19-public-forecasts>
20. Cleveland Clinic's COVID-19 strategy driven by data modeling, <https://searchbusinessanalytics.techtarget.com/feature/Cleveland-Clinics-COVID-19-strategy-driven-by-data-modeling>
21. Harmon S.A., Sanford T.H., Xu S. et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun 11, 4080 (2020). <https://doi.org/10.1038/s41467-020-17971-2>
22. Морозов С.П., Гомболевский В.А., Чернина В.Ю., Блохин И.А., Мокиенко О.А., Владзимирский А.В., Белевский А.С., Проценко Д.Н., Лысенко М.А., Зайратьянц О.В., Никонов Е.Л. Прогнозирование летальных исходов при covid-19 по данным компьютерной томографии органов грудной клетки. Туберкулез и болезни легких. – 2020. – Т. 98. – № 6. – С. 7–14.
23. Лучевая диагностика коронавирусной болезни (COVID-19): организация, методология, интерпретация результатов: препринт № ЦДТ – 2020 – I. Версия 2 от 17.04.2020 / сост. С.П. Морозов, Д.Н. Проценко, С.В. Сметанина [и др.] // Серия «Лучшие практики лучевой и инструментальной диагностики». – Вып. 65. – М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2020. – 78 с.
24. Bayat V., Phelps S., Ryono R., Lee C., Parekh H., Newton J., Sedghi F., Etmiani P., Holodny M. A SARS-CoV-2 Prediction Model from Standard Laboratory Tests. Clin Infect Dis. 2020 Aug 12: ciaa1175. doi: 10.1093/cid/ciaa1175. Epub ahead of print. PMID: 32785701; PMCID: PMC7454351, <https://pubmed.ncbi.nlm.nih.gov/32785701/>
25. Yan, L., Zhang, H., Goncalves, J. et al. An interpretable mortality prediction model for COVID-19 patients. Nat Mach Intell 2, 283–288 (2020). <https://doi.org/10.1038/s42256-020-0180-7>
26. Liang W., Yao J., Chen A. et al. Early triage of critically ill COVID-19 patients using deep learning. Nat Commun 11, 3543 (2020). <https://doi.org/10.1038/s41467-020-17280-8>
27. Schalekamp S., Huisman M., van Dijk R.A., Boomsma M.F., Freire Jorge P.J., de Boer W.S., Herder G.J.M., Bonarius M., Groot O.A., Jong E., Schreuder A., Schaefer-Prokop C.M. Model-based Prediction of Critical Illness in Hospitalized Patients with COVID-19. Radiology. 2020 Aug 13:202723. doi: 10.1148/radiol.2020202723. Epub ahead of print. PMID: 32787701; PMCID: PMC7427120, <https://pubmed.ncbi.nlm.nih.gov/32787701/>
28. Jehi L., Ji X., Milinovich A., Erzurum S., Merlino A., Gordon S., Young J.B., Kattan M.W. Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19. PLoS One. 2020 Aug 11;15(8): e0237419. doi: 10.1371/journal.pone.0237419. PMID: 32780765; PMCID: PMC7418996, <https://pubmed.ncbi.nlm.nih.gov/32780765/>
29. FDA greenlights ICU AI for predicting fatal COVID-19 complications, <https://www.fiercebitech.com/medtech/fda-greenlights-icu-ai-for-predicting-fatal-covid-19-complications>
30. Bartkowiak B., Snyder A.M., Benjamin A., Schneider A., Twu N.M., Churpek M.M., Roggin K.K., Edelson D.P. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–1063. doi: 10.1097/SLA.0000000000002665. PMID: 31082902; PMCID: PMC6610875, <https://pubmed.ncbi.nlm.nih.gov/31082902/>
31. Covid-19: NHS datasets 'not sophisticated enough' to flag high risk patients, <https://www.digitalhealth.net/2020/03/covid-19-nhs-datasets-not-sophisticated-enough-to-flag-high-risk-patients/>
32. Israel is using AI to flag high-risk covid-19 patients, <https://www.technologyreview.com/2020/04/24/1000543/israel-ai-prediction-medical-testing-data-high-risk-covid-19-patients/>
33. Israel Aerospace разработала AI-модель для прогнозирования прогрессирования болезни у пациентов с COVID-19, <https://evercare.ru/news/israel-aerospace-razrabotala-ai-model-dlya-prognozirovaniya-progressirovaniya-bolezni-u>
34. How AI can determine which coronavirus patients require hospitalization, <https://thenextweb.com/neural/2020/04/02/ai-can-help-manage-hospital-resources-during-the-coronavirus-crisis-syndication/>