

На правах рукописи

МОНАХОВА АНГЕЛИНА АНДРЕЕВНА

**СОВЕРШЕНСТВОВАНИЕ ЭПИДЕМИОЛОГИЧЕСКОГО НАДЗОРА
ЗА НОВОЙ КОРОНАВИРУСНОЙ ИНФЕКЦИЕЙ (COVID-19)
НА ОСНОВЕ ЦИФРОВЫХ ТЕХНОЛОГИЙ УЧЕТА И АНАЛИЗА
ДАННЫХ МОЛЕКУЛЯРНО-БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ**

3.2.2. Эпидемиология

Автореферат

диссертации на соискание ученой степени

кандидата медицинских наук

Москва – 2026

Работа выполнена в Федеральном бюджетном учреждении науки «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека

Научный руководитель

Акимкин Василий Геннадьевич — академик РАН, доктор медицинских наук, профессор

Официальные оппоненты:

Асланов Батырбек Исмелович — доктор медицинских наук, профессор, заведующий кафедрой эпидемиологии, паразитологии и дезинфектологии Федерального государственного бюджетного образовательного учреждения высшего образования «Северо-Западный государственный медицинский университет им. И.И. Мечникова» Министерства здравоохранения Российской Федерации

Зайцева Наталья Николаевна — доктор медицинских наук, доцент, директор Федерального бюджетного учреждения науки «Нижегородский научно-исследовательский институт эпидемиологии и микробиологии им. академика И.Н. Блохиной» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека

Ведущая организация — Федеральное государственное бюджетное военное образовательное учреждение высшего образования «Военно-медицинская академия имени С.М. Кирова» Министерства обороны Российской Федерации

Защита состоится «_____» _____ 2026 г. в _____ на заседании Диссертационного Совета 64.1.010.01 в Федеральном бюджетном учреждении науки «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека по адресу: 111123, Москва, ул. Новогиреевская, д. 3а

С диссертацией можно ознакомиться в библиотеке Федерального бюджетного учреждения науки «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека и на сайте института www.crie.ru

Автореферат разослан «_____» _____ 2026 г.

Ученый секретарь

Диссертационного Совета,
доктор медицинских наук

Николаева Светлана Викторовна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Пандемия новой коронавирусной инфекции (COVID-19) ярко продемонстрировала разрушительные последствия массовых инфекционных болезней, уносящих миллионы человеческих жизней и подрывающих международную экономику. Пандемия не только обнажила проблемы системы здравоохранения большинства стран, но и явилась импульсом научного прогресса во многих областях медицинских и биологических наук [Попова А.Ю., 2021]. Глобальное распространение новой коронавирусной инфекции (COVID-19) способствовало развитию цифровых технологий в эпидемиологии и ускорило создание программных платформ и внедрение молекулярно-биологических исследований в практику общественного здравоохранения [Акимкин В.Г., 2025].

Молекулярно-биологические исследования представляют собой методы научной деятельности, сосредоточенной на изучении структуры и функции молекул, таких как ДНК, РНК и белки, и их роли в жизнедеятельности клеток и организма в целом. Данные исследования позволяют глубже понять механизмы заболеваний на молекулярном уровне, что открывает возможности для разработки новых методов диагностики, лечения и профилактики [Альбертс Б., 2013; Акимкин В.Г., 2025].

Применение полимеразной цепной реакции (ПЦР) усовершенствовало процесс диагностики инфекционных болезней. ПЦР позволяет проводить раннюю и дифференциальную диагностику, а также идентифицировать патогены с высокой специфичностью и чувствительностью. Благодаря возможности быстро масштабировать данную методику, массовая ПЦР диагностика стала эффективным инструментом для выявления возбудителей инфекционных болезней, что особенно актуально в условиях пандемий и эпидемий, когда важна быстрая реакция системы здравоохранения на распространение инфекции и снижение риска контактирования с потенциально инфицированными лицами.

Платформа SOLAR (*System Of Laboratory Aggregation Results* — Платформа агрегирования результатов лабораторных исследований), представляющая собой централизованную базу данных результатов ПЦР исследований для выявления возбудителя новой коронавирусной инфекции (SARS-CoV-2), была разработана и внедрена в ФБУН Центральный НИИ Эпидемиологии Роспотребнадзора в рамках исполнения Постановления Правительства РФ №452 от 27 марта 2021 г. Целью создания Платформы явилось оперативное информирование граждан о результатах лабораторных исследований.

По состоянию на 1 февраля 2026 г. в систему загружено более 200 млн результатов ПЦР-исследований, поступающих от более чем 2 000 лабораторий различной ведомственной принадлежности. Передача данных осуществляется по защищенным каналам связи ФБУН Центральный НИИ Эпидемиологии Роспотребнадзора, откуда они перенаправляются на сервер портала Госуслуг. Таким образом, Платформа SOLAR представляет собой важный инструмент цифровой трансформации эпидемиологического надзора, обеспечивая централизованный сбор и анализ данных о распространении новой коронавирусной инфекции (COVID-19).

Учение об эпидемиологическом надзоре является одним из базовых в теории и практике эпидемиологической науки. Отечественные ученые (Громашевский Л.В., 1941; Беляков В.Д., 1986; Покровский В.И., 1987; Черкасский Б.Л., 1988) в XX веке сформулировали основополагающие теории, определяющие методологию эпидемиологического надзора на современном этапе развития научного познания.

В соответствии с теорией саморегуляции паразитарных систем академика Белякова В. Д., основу развития эпидемического процесса составляют фазовые изменения гетерогенности биологических свойств взаимодействующих популяций возбудителя и человека, базирующиеся на взаимосвязанном процессе саморегуляции. При этом важное регулирующее значение имеют социальные и природные факторы. Согласно данной теории, эпидемиологический надзор

рассматривается как информационно-аналитическая и управленческая система, предполагающая динамическую оценку тенденций развития эпидемического процесса в пространстве и времени, обеспечивающая своевременное вмешательство в его ход с целью снижения заболеваемости населения и ликвидации отдельных инфекций [Беляков В.Д., 1987; Акимкин В.Г., 2024].

Важными элементами системы эпидемиологического надзора за инфекционной заболеваемостью является постановка эпидемиологического диагноза и формирование эпидемиологического прогноза, включающие оценку эпидемиологической ситуации, ее причин и тенденций развития. В качестве начального этапа эпидемиологической диагностики рассматривается оценка проявлений эпидемического процесса по выявлению территорий риска, групп и времени риска, т.е. изучение пространственно-временной структуры и динамики заболеваемости населения [Дубоделов Д.В., 2024].

В современных реалиях информационные технологии шагнули далеко вперед и способны решать сложные задачи эпидемиологического анализа с применением цифровых платформенных решений. В последние годы значительно увеличился объем данных доступных для эпидемиологического анализа, что позволяет успешнее использовать имитационные математические модели для анализа и прогнозирования динамики эпидемического процесса, а также делает актуальным создание эффективных алгоритмов поиска значимых показателей, блоков интерпретации и обработки данных [Михеева И.В., 2026].

В связи с этим необходимы новые подходы, поскольку активное развитие цифровых технологий и цифровизации в целом, а также постоянно усложняющиеся задачи, возникшие перед системой эпидемиологического надзора в период пандемии новой коронавирусной инфекции (COVID-19), создали предпосылки для стремительного развития цифровой трансформации в эпидемиологии, что подчеркивает актуальность рассматриваемой проблемы.

Значительная разнородность по плотности населения, географическим и экономическим условиям субъектов Российской Федерации обусловила необходимость организации работы по осуществлению эпидемиологического анализа с использованием методов кластеризации [Дубоделов Д.В., 2024].

В настоящее время обсуждаются перспективные направления цифровизации, связанные с анализом и структурированием данных с использованием математических методов и подходов, а также с разработкой интеллектуальных алгоритмов. Широкое распространение получили методы машинного обучения (*англ. machine learning*), основанные на машинной обработке подготовленных статистических данных для выявления закономерностей и построения на их основе алгоритмов, позволяющих системе искусственного интеллекта формулировать обоснованные выводы на основе поступающей информации [Гаврилов Д. В., 2021]. Такие подходы повышают эффективность прогнозирования рисков за счет применения масштабных хранилищ данных, независимой идентификации новых предикторов риска и сложных взаимосвязей между ними [Дубоделов Д. В., 2024].

Совершенствование системы эпидемиологического надзора на основе цифровизации является важной научной задачей современной эпидемиологии. Пандемия новой коронавирусной инфекции (COVID-19) подчеркнула значимость цифровой трансформации в системе эпидемиологического надзора как важного дополнения к традиционным методам эпидемиологического анализа данных [Акимкин В.Г., 2023].

Таким образом, современные цифровые технологии открывают новые горизонты для оперативного и точного принятия управленческих решений, направленных на минимизацию рисков для здоровья населения, что делает их незаменимыми в системе эпидемиологического надзора. Кроме того, цифровые технологии могут помочь в поиске инновационных методологических решений по совершенствованию системы эпидемиологического надзора, создают новые возможности для управления эпидемиологической ситуацией и повышают адаптивность системы здравоохранения к меняющимся условиям эпидемиологической обстановки.

Степень разработанности темы исследования

Последствия пандемии новой коронавирусной инфекции (COVID-19) остаются важным предметом интереса эпидемиологической науки. Учитывая это, в разных странах мира, в том числе и Российской Федерации, происходит усовершенствование научных инструментов эпидемиологического надзора за распространением инфекции. Так, в Российской Федерации разработаны научные основы и реализован геномный эпидемиологический надзор за возбудителями инфекционных болезней. Созданы три программные цифровые Платформы, анализирующие значительные объемы данных (англ. Big Data): российская система агрегации данных VGARus, SOLAR и аналитическая платформа для анализа и прогнозирования эпидемиологической обстановки по заболеваемости новой коронавирусной инфекцией (COVID-19) [Акимкин В.Г., 2025].

Также, предложен и разработан метод кластеризации показателей заболеваемости за многолетний период, помогающий повысить информативность и объективность результатов эпидемиологического анализа данных [Дубоделов Д.В., 2024].

Использование результатов ПЦР-исследований в системе эпидемиологического надзора стало ключевым элементом в борьбе с инфекционными болезнями. ПЦР-исследования позволяют обнаруживать патогены на ранних стадиях заболевания, даже до появления клинических симптомов, что дает возможность своевременно начать лечение и предотвратить распространение инфекции. Кроме того, ПЦР-методы дают возможность с высокой точностью различать патогены, что особенно актуально при наличии схожих клинических симптомов (например, гриппа и новой коронавирусной инфекции (COVID-19), что способствует правильному выбору тактики лечения [Акимкин В.Г., 2025].

В рамках эпидемиологического надзора ПЦР-исследования используются для отслеживания мутаций патогенов, что важно для оценки их вирулентности и возможности формирования новых штаммов [Акимкин В.Г., 2021].

Таким образом, актуальность изучения и внедрения в систему эпидемиологического надзора цифровых технологий анализа результатов молекулярно-биологических исследований и математических моделей развития эпидемического процесса на их основе не вызывает сомнения и является важным направлением его совершенствования.

Цель исследования

Совершенствование эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) на основе цифровых технологий учета и анализа данных молекулярно-биологических исследований.

Задачи исследования

1. Провести ретроспективный эпидемиологический анализ заболеваемости новой коронавирусной инфекцией (COVID-19) населения Российской Федерации за 2020–2023 гг. с использованием метода кластеризации субъектов в различные периоды пандемии.
2. Оценить влияние социально-демографических и инфраструктурных характеристик на уровень и динамику заболеваемости новой коронавирусной инфекцией (COVID-19) в субъектах Российской Федерации.
3. Изучить возможности применения системных цифровых технологий учета результатов и анализа данных молекулярно-биологических исследований населения Российской Федерации для повышения информативности анализа эпидемиологической обстановки по новой коронавирусной инфекции (COVID-19).
4. Предложить научно обоснованные пути совершенствования системы эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) на основе внедрения цифровых технологий учета и анализа результатов молекулярно-биологических исследований.

Научная новизна исследования

Углублены научные знания, характеризующие распространение новой коронавирусной инфекции (COVID-19) на территории Российской Федерации в различные периоды эпидемии. Проведен детальный эпидемиологический анализ динамики уровня и структуры заболеваемости населения Российской Федерации новой коронавирусной инфекцией (COVID-19) в период 2020–2023 гг. с использованием метода кластеризации интенсивных показателей заболеваемости.

Показана роль деперсонифицированных сведений ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) для проведения оперативного анализа заболеваемости новой коронавирусной инфекцией (COVID-19) на основе детальной оценки факторов риска и характеристик проявлений эпидемического процесса в различные периоды эпидемии в целях краткосрочного прогнозирования эпидемиологической обстановки по заболеваемости новой коронавирусной инфекцией (COVID-19).

Определена научная обоснованность и эффективность применения методов машинного обучения для проведения ретроспективного эпидемиологического анализа заболеваемости и выявления факторов, влияющих на уровень и динамику эпидемического процесса новой коронавирусной инфекции (COVID-19).

Научно обосновано совершенствование эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19), базирующееся на применении цифровых технологий для обработки и анализа данных молекулярно-биологических исследований.

Теоретическая и практическая значимость исследования

Диссертационное исследование позволяет оценить эффективность практического использования цифровых платформенных решений для оперативного и ретроспективного анализа заболеваемости новой коронавирусной инфекцией (COVID-19) в конкретных регионах и Российской Федерации в целом.

Результаты научного исследования подчеркивают роль молекулярно-биологических исследований в объективизации эпидемиологической информации и оперативности ее получения с целью принятия своевременных управленческих решений.

Показано, что внедрение современных цифровых инструментов позволяет существенно повысить качество и скорость анализа молекулярно-биологических данных, что способствует более точному прогнозированию эпидемиологических тенденций и принятию научно обоснованных управленческих решений на различных уровнях деятельности системы здравоохранения.

Научные данные, полученные в результате диссертационного исследования создают научную основу для использования прогностических моделей в интересах решения будущих эпидемиологических задач в отношении аэрозольных антропонозов с высоким эпидемическим и пандемическим потенциалом распространения возбудителя.

Методология и методы исследования

Цель исследования определила последовательность этапов научной работы, каждый из которых опирался на применение соответствующих научных методов. Работа представляет собой комплексное исследование, включающее не только классические эпидемиологические подходы, но и математические методы, позволяющие глубже рассмотреть механизмы развития эпидемического процесса. На этапе разработки дизайна исследования применялся комплексный подход к оценке возможных источников систематических и случайных ошибок.

В основе эпидемиологического метода применена классическая схема: описательный эпидемиологический анализ распространенности и заболеваемости с определением факторов, влияющих на развитие эпидемического процесса новой коронавирусной инфекции (COVID-19), и последующий аналитический этап для установления причинно-следственных связей. Для этого использованы как дескриптивные статистические методы, так и аналитические методы, включающие методы машинного обучения.

Обработка полученных данных осуществлялась с помощью Microsoft Excel. Проведен корреляционный анализ для оценки связей между разными переменными. Все полученные результаты систематизированы и представлены в главах диссертации.

Положения, выносимые на защиту

1. Применение метода кластеризации субъектов Российской Федерации по уровню заболеваемости населения новой коронавирусной инфекцией (COVID-19) при проведении ретроспективного эпидемиологического анализа, а также анализ интенсивных показателей заболеваемости среди установленных кластеров позволяет выявить особенности проявлений эпидемического процесса новой коронавирусной инфекции (COVID-19).
2. Уровень заболеваемости и динамика показателей, характеризующие проявления эпидемического процесса новой коронавирусной инфекции (COVID-19) в субъектах Российской Федерации, имели выраженную корреляционную зависимость от доли городского населения, плотности населения региона, интенсивности экономического развития (количества активных предприятий), показателей активности миграционных процессов (коэффициента миграционного прироста).
3. Использование цифровой Платформы агрегирования результатов лабораторных исследований SOLAR (*System of laboratory aggregation results*) и современных цифровых технологий позволяет существенно объективизировать и оптимизировать процессы эпидемиологического анализа данных, обеспечить краткосрочное прогнозирование заболеваемости новой коронавирусной инфекцией (COVID-19).
4. Внедрение в практику эпидемиологического надзора платформ цифровизации данных молекулярно-биологических исследований и инструментов машинного обучения позволяют совершенствовать подходы к решению эпидемиологических задач.

Личный вклад автора в получении результатов

Автор принимал непосредственное участие в разработке концепции исследования, формулировке цели, задач и логики построения диссертационной работы. Автором осуществлен анализ эпидемиологической информации, полученной из федеральных информационных систем и статистических отчетов.

Сбор, систематизация и подготовка исходных данных для эпидемиологического анализа выполнены автором в полном объеме. Проведен ретроспективный эпидемиологический анализ заболеваемости новой коронавирусной инфекцией (COVID-19) в субъектах Российской Федерации. Автором осуществлена кластеризация территорий по эпидемиологическим показателям, проведен корреляционный анализ между принадлежностью к кластерам и демографическими характеристиками. Выполнено построение модели на основе метода машинного обучения (*Random Forest*) для оценки значимости факторов и сравнение ее результатов с другими подходами.

Автор лично сформулировал предложения по совершенствованию системы эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19), основанные на применении цифровых платформ, результатов молекулярно-биологических данных и технологий автоматизированного анализа. Все аналитические расчеты, интерпретация результатов, подготовка графического материала, формулировка выводов и написание текста диссертации выполнены автором самостоятельно.

Внедрение в практику

На основе результатов диссертационного исследования получены:

- Свидетельство о государственной регистрации программы для ЭВМ №2026661550 от 22 апреля 2026 г. «EpidSmart — модуль COVID, версия 2.0», предназначенная для оценки эпидемиологической обстановки по новой коронавирусной инфекции (COVID-19), проведения

оперативного и ретроспективного эпидемиологического анализа с выявлением групп, территорий и факторов риска (заявка №2026619201 от 03.04.2026 г.);

– Свидетельство о государственной регистрации базы данных №2026621597 от 09 апреля 2026 г. «EpidSmart — модуль COVID, версия 2.0», предназначенная для обеспечения функционирования программного обеспечения «EpidSmart — модуль COVID, версия 2.0». Содержит информацию о проявлениях эпидемического процесса новой коронавирусной инфекции (COVID-19) в различных возрастных группах населения субъектов Российской Федерации, демографические показатели, результаты молекулярно-генетического мониторинга за возбудителем (SARS-CoV-2).

Материалы диссертационного исследования внедрены в образовательный процесс на кафедре эпидемиологии с курсами молекулярной диагностики и дезинфектологии ФБУН ЦНИИ Эпидемиологии Роспотребнадзора.

Степень достоверности и апробация результатов

Достоверность результатов исследования, ключевых выводов и рекомендаций обеспечена путем аналитического обобщения данных, полученных из открытых и проверенных источников научной литературы. В диссертационной работе применены общепризнанные методы сбора и обработки официальной статистической информации. Представлены масштабные эпидемиологические и молекулярно-биологические исследования, использующие современные методы эпидемиологического, молекулярно-биологического, математического и статистического анализа.

Основные положения и результаты исследования представлены и обсуждены на Конгрессах и научно-практических конференциях различных уровней: Научно-практическая конференция молодых ученых и специалистов ФБУН ЦНИИ Эпидемиологии Роспотребнадзора «Проблемы эпидемиологии, терапии и лабораторной диагностики инфекционных болезней — 2024» (29–30 мая 2024 г., г. Москва); VI Международная научно-практическая конференция по вопросам противодействия инфекционным заболеваниям (11–12 декабря 2025 г., г. Санкт-Петербург); Всероссийский Конгресс с международным «Эпидемиология — 2025» (15–16 октября 2025 г., г. Москва); V Ежегодной конференции по инфекционным болезням «Покровские чтения» (27–28 октября 2025 г., г. Москва); Всероссийский Конгресс с международным «Эпидемиология — 2026» (16–17 апреля 2026 г., г. Москва).

В окончательном виде диссертационная работа апробирована и рекомендована к защите на заседании апробационной комиссии ФБУН ЦНИИ Эпидемиологии Роспотребнадзора (протокол №113 от 14 апреля 2026 г.).

Соответствие диссертации паспорту научной специальности

Научные положения диссертации соответствуют паспорту специальности 3.2.2. «Эпидемиология». Результаты проведенного исследования соответствуют областям исследований: пунктам 2, 5 и 6 паспорта специальности «Эпидемиология».

Публикации

По теме диссертации опубликовано 12 научных работ, в том числе 4 статьи в изданиях, рекомендованных ВАК Российской Федерации для публикации основных научных результатов диссертации по специальности 3.2.2. «Эпидемиология».

Структура и объём диссертации

Диссертационное исследование представлено на 152 листах, состоит из Введения, 6 глав (обзора литературы; главы «Материалы и методы исследования»; 4 глав собственных исследований), заключения, выводов и списка литературы. Диссертация иллюстрирована 8 таблицами и 11 рисунками. Список литературы содержит 120 источников, в том числе 51 — на русском языке и 69 — на английском языке.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Материалы и методы

Диссертационное исследование проведено в 2024–2026 гг. на базе Федерального бюджетного учреждения науки «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека в рамках НИР АААА–А21–121011890131–4 «Научно-методическое обоснование системы эпидемиологического надзора и профилактики заболеваний, вызываемых SARS-подобными коронавирусами» (2021–2025 гг.) и НИР 1025030500172–4–3.3.9 «Совершенствование системы эпидемиологического надзора за инфекционными болезнями в Российской Федерации на основе современных цифровых технологий» (2026–2030 гг.).

В рамках проведенного диссертационного исследования использованы эпидемиологический, молекулярно-биологический, математический и статистический методы исследования.

Выполнен ретроспективный эпидемиологический анализ заболеваемости новой коронавирусной инфекцией (COVID-19) на территории Российской Федерации за период 2020–2023 гг. с использованием отчетной формы № 1035 «Мониторинг о количестве заболевших коронавирусной инфекцией, в том числе внебольничными пневмониями, и летальных исходов» и формы №970 «Информация о случаях инфекционных заболеваний у лиц с подозрением на новую коронавирусную инфекцию».

Для оценки динамики циркулирующих вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2) использованы данные из Российской базы данных VGARus (Российской платформы для агрегирования информации о геномах вирусов (*англ. Virus Genome Aggregator of Russia*), созданного в рамках выполнения Постановлений Правительства Российской Федерации от 23 марта 2021 г. № 448 «Об утверждении Временного порядка предоставления данных расшифровки генома возбудителя новой коронавирусной инфекции (COVID-19)». Для оценки значимости результатов молекулярно-биологических исследований и построения прогностических моделей проведен анализ данных российской информационной системы «SOLAR», созданной в рамках выполнения Постановления Правительства Российской Федерации от 27.03.2021 № 452 «Об обеспечении уведомления физических лиц о результатах исследований на наличие возбудителя новой коронавирусной инфекции (COVID-19) с использованием федеральной государственной информационной системы «Единый портал государственных и муниципальных услуг (функций)» и обмена информацией о результатах таких исследований», а также сведений Федеральной службы государственной статистики (таблица 1).

Эпидемиологический метод

В рамках эпидемиологического метода осуществлялась регистрация новых случаев инфекции, госпитализаций, выздоровлений и летальных исходов, что позволяло отслеживать динамику заболеваемости, выделять периоды ее подъема и снижения, а также сравнивать эпидемиологические показатели между регионами. Значимое внимание уделялось анализу пространственно-временных закономерностей — различий в распространении инфекции по территории, в том числе между городскими и сельскими населенными пунктами, а также учету сезонных колебаний и климатических факторов, влияющих на эпидемический процесс.

В рамках исследования применены описательно-оценочный и аналитический методы эпидемиологического анализа. На основании официальных данных отчетной формы №1035 «Мониторинг о количестве заболевших коронавирусной инфекцией, в том числе внебольничными пневмониями, и летальных исходов» произведен ретроспективный анализ динамики заболеваемости населения субъектов Российской Федерации новой коронавирусной инфекцией (COVID-19) за период 2020–2023 гг. Рассчитаны интенсивные показатели (уровни заболеваемости на 100 тыс. населения) и экстенсивные показатели (распределение случаев по

временным и территориальным характеристикам). Проведено сопоставление уровня заболеваемости новой коронавирусной инфекцией (COVID-19) по полу, возрастным группам и типу населенного пункта (городские и сельские территории) в субъектах Российской Федерации за 2020–2023 гг.

Таблица 1 — Материалы исследования

Источник данных	Содержание данных	Объем сведений для анализа
Форма отчетности Роспотребнадзора №1035 «Мониторинг количества заболевших коронавирусной инфекцией, в том числе внебольничными пневмониями, и летальных исходов»	Мониторинг о количестве заболевших коронавирусной инфекцией, в том числе внебольничными пневмониями, и летальных исходов	1 349 отчетных форм (19 млн показателей)
Форма отчетности Роспотребнадзора №970 «Информация о случаях инфекционных заболеваний у лиц с подозрением на новую коронавирусную инфекцию»	Информация о случаях инфекционных заболеваний у лиц с подозрением на новую коронавирусную инфекцию	796 отчетных форм (17 млн случаев заболевания)
Данные Платформы агрегирования результатов лабораторных исследований SOLAR (System of laboratory aggregation results)	Результаты ПЦР-исследований на наличие возбудителя новой коронавирусной инфекции (SARS-CoV-2)	Более 200 млн исследований
Российская база данных VGARus (Российская платформа для агрегирования информации о геномах вирусов (<i>англ. Virus Genome Aggregator of Russia</i>))	Сведения о вариантах возбудителя новой коронавирусной инфекции (SARS-CoV-2)	298 тыс. результатов секвенирования (194 тыс. полных геномов и 104 тыс. фрагментов генома возбудителя)
Данные Федеральной службы государственной статистики	Инфраструктурные и социально-демографические сведения	Более 10 тыс. показателей

Целью данного анализа явилось выявление различий в эпидемиологической ситуации между указанными группами, а также потенциальных факторов, влияющих на интенсивность распространения инфекции. Методология проведенного эпидемиологического анализа опиралась на классические эпидемиологические подходы, описанные в трудах российских ученых конца XX – начала XXI века: Беляков В.Д., Черкасский Б.Л. и Покровский В.И., а также на современные методы статистического анализа больших данных.

Молекулярно-биологический метод

Для подтверждения случаев новой коронавирусной инфекции (COVID-19) в 2022–2023 гг. использован метод полимеразной цепной реакции (ПЦР), обеспечивающий выявление РНК возбудителя даже в незначительных концентрациях (использованы данные о более 200 млн исследований) (таблица 1). ПЦР-диагностика оставалась основным методом лабораторного подтверждения инфицирования, широко применялась в амбулаторной и стационарной практике и регламентировалась нормативными актами. Также учтены результаты изотермической амплификации нуклеиновых кислот, в частности LAMP (*англ. loop-mediated isothermal amplification*), позволяющие сократить время анализа за счет проведения реакции при постоянной температуре без применения сложного оборудования, и методы секвенирования нового поколения (NGS), направленные на определение полной нуклеотидной последовательности генома возбудителя новой коронавирусной инфекции (SARS-CoV-2). Полученные геномные данные (более 298 тыс. результатов секвенирования) (таблица 1) использовались для отслеживания мутаций, появления новых вариантов вируса (например, Alpha, Delta, Omicron SARS-CoV-2), и для оценки возможного влияния данных вариантов на

эпидемиологические показатели, включая тяжесть течения инфекции и эффективность вакцинации. В настоящем исследовании проанализированы данные ПЦР диагностики новой коронавирусной инфекции (COVID-19) за 2022–2023 гг. на основе данных федеральной платформы SOLAR, обеспечивающей централизованный сбор результатов ПЦР-обследования по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2). К системе подключено более 2 000 организаций из 85 субъектов Российской Федерации. В едином реестре аккумулировано свыше 200 млн результатов, с ежедневным добавлением около 30 тыс. новых записей (в период 2020–2023 гг.) (таблица 1).

Математический метод

Метод кластеризации (k-means). Использован для кластеризации временных рядов ежедневной заболеваемости. Алгоритм k-means можно описать в несколько шагов: 1) выбор числа кластеров (k). На первом этапе необходимо определить количество кластеров, на которые будут разделены данные. Этот параметр задают вручную, и его правильный выбор напрямую влияет на качество кластеризации; 2) инициализация центроидов. Алгоритм случайным образом выбирает k начальных точек, называемых центроидами. Эти точки служат временными центрами кластеров; 3) назначение объектов кластерам. Каждый объект в наборе данных назначается к кластеру, центроид которого находится ближе всего; 4) обновление центроидов. После назначения объектов кластерам вычисляют новые центроиды. Каждый центроид перемещается в среднюю точку всех объектов, принадлежащих его кластеру; 5) повторение. Шаги 3 и 4 повторяют до тех пор, пока центроиды не перестанут значительно изменяться, что свидетельствует о достижении сходимости. В некоторых реализациях алгоритм также может быть завершён при достижении заданного числа итераций. Метод реализован с использованием библиотеки Matplotlib (версия 3.8.0).

Методы анализа линейных зависимостей. В рамках исследования использованы методы анализа линейных зависимостей для оценки связи между отдельными социально-демографическими и инфраструктурными характеристиками субъектов Российской Федерации и уровнем заболеваемости новой коронавирусной инфекцией (COVID-19). Рассчитан коэффициент корреляции Пирсона, позволяющий количественно оценить степень линейной зависимости между двумя количественными признаками. Значения коэффициента варьируют от -1 до $+1$ и характеризуют направление и силу взаимосвязи. В расчетах проанализированы показатели плотности населения, уровня урбанизации и обеспеченности транспортной инфраструктурой в сопоставлении с интенсивными показателями заболеваемости новой коронавирусной инфекцией (COVID-19) в период 2020–2023 гг. Оценку силы корреляционной связи проводили по шкале Чеддока (0-0,1 – очень слабая, 0,1-0,3 – слабая, 0,3-0,5 – средняя, 0,5-0,7 – умеренно сильная, 0,7-0,9 – высокая, 0,9-0,99 – очень высокая).

Метод машинного обучения «Random Forest». В рамках исследования применен метод «Random Forest» — ансамблевый алгоритм машинного обучения на основе «решающих деревьев». Данный метод позволяет выявлять нелинейные зависимости и учитывать сложные взаимодействия между множеством факторов. Алгоритм формирует совокупность «решающих деревьев», каждое из которых обучается на случайной выборке данных и случайном подмножестве признаков, после чего прогнозы усредняются. Метод использован для анализа влияния социально-демографических и инфраструктурных показателей (уровня урбанизации, плотности населения, доли пожилого населения, количества ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) и др.) на уровень заболеваемости новой коронавирусной инфекцией (COVID-19) по субъектам Российской Федерации в 2020–2023 гг.

Статистический метод

Для обработки полученных данных применен комплекс методов статистического анализа: описательная статистика, графический анализ, а также методы параметрической и

непараметрической статистики. Также использованы приемы математического моделирования и анализа временных рядов с задействованием библиотек SciPy, Pandas и Statsmodels. При изучении динамики заболеваемости новой коронавирусной инфекцией (COVID-19) населения Российской Федерации в 2020–2023 гг. использованы как дневные, так и недельные интенсивные показатели числа зарегистрированных случаев. Сезонная декомпозиция временных рядов произведена на основе ежедневных данных с применением метода локально взвешенной линейной регрессии (LOESS), реализованного в модуле statsmodels.tsa.seasonal. Для дополнительного анализа сезонных колебаний регистрируемой заболеваемости построены графики автокорреляционной функции с использованием библиотеки Matplotlib (версия 3.8.0). Доверительный интервал (95% ДИ) рассчитывался с использованием точного метода Клоппера-Пирсона (онлайн калькулятор EPITOOLS). Расчет стандартного отклонения осуществлен при помощи программы Microsoft Office Excel.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ОБСУЖДЕНИЕ

Ретроспективный эпидемиологический анализ заболеваемости населения новой коронавирусной инфекцией (COVID-19) в Российской Федерации в 2020–2023 гг. с использованием метода кластеризации интенсивных показателей

В период с 2020 по 2023 гг. в Российской Федерации прослеживались восемь эпидемических периодов распространения новой коронавирусной инфекции (COVID-19), различавшихся по интенсивности заболеваемости, длительности и особенностям циркуляции генетических вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2). Разделение эпидемического процесса новой коронавирусной инфекции (COVID-19) на условные периоды позволило учитывать специфику каждого из них при проведении последующего анализа (таблица 2).

Первый эпидемический подъем новой коронавирусной инфекции (COVID-19), пришедшийся на март – август 2020 г., характеризовался началом активной циркуляции вируса среди населения, не имеющего иммунитета. Максимальное недельное значение заболеваемости зафиксировано 4 мая 2020 г. (на 19-й неделе), достигнув 51,41 случая на 100 тыс. населения. Реализация мер, действовавших в период с конца марта по май 2020 г., привела к замедлению распространения инфекции, снижению нагрузки на систему здравоохранения и постепенной стабилизации эпидемической ситуации. После частичной отмены ограничений заболеваемость продолжила снижаться, и к августу 2020 г. показатель составил минимум 23,83 случая на 100 тыс. населения.

С началом учебного года, на 37-й неделе 2020 г. (7 сентября), отмечен **второй эпидемический подъем**. Максимальное недельное значение заболеваемости во втором периоде пришлось на 52-ю неделю 2020 г. (21 декабря) и составило 136,69 случая на 100 тыс. населения. В дальнейшем наблюдалась тенденция к снижению интенсивности эпидемического процесса новой коронавирусной инфекции (COVID-19), и к марту 2021 г. недельный показатель стабилизировался на уровне минимум 41,15 случая на 100 тыс. населения.

С мая 2021 г. регистрировался **третий эпидемический подъем** заболеваемости новой коронавирусной инфекцией (COVID-19), напрямую связанный с распространением варианта Delta SARS-CoV-2. На 29-й неделе 2021 г. (12 июля) показатель достиг максимума 118,8 случая на 100 тыс. населения. К середине августа – началу сентября 2021 г. недельная заболеваемость новой коронавирусной инфекцией (COVID-19) стабилизировалась на минимальном уровне 91,32 случая на 100 тыс. населения.

В сентябре 2021 г. начался **четвертый эпидемический подъем**. Максимальный недельный показатель заболеваемости зарегистрирован на 45-й неделе 2021 г. (1 ноября) и составил 191,87 случая на 100 тыс. населения. К концу декабря 2021 г. эпидемический процесс новой коронавирусной инфекции (COVID-19) стабилизировался на уровне минимум 85,33 случая на 100 тыс. населения.

В январе 2022 г. отмечен **пятый эпидемический подъем**, обусловленный активным распространением варианта Omicron SARS-CoV-2. Данный подъем сопровождался резким увеличением числа новых случаев, и на 7-й неделе 2022 г. (14 февраля) зарегистрирован максимальный недельный показатель заболеваемости новой коронавирусной инфекцией (COVID-19) за весь период наблюдений — 905 случаев на 100 тыс. населения.

На этой фазе эпидемического процесса доминировала генетическая линия BA.1 Omicron SARS-CoV-2, которая в конце февраля – начале марта 2022 г. была вытеснена линией BA.2 Omicron SARS-CoV-2. Смена доминирующего варианта происходила на фоне устойчивого снижения заболеваемости, и уже в начале мая 2022 г. средний недельный показатель составил $21,75 \pm 1,36$ случая на 100 тыс. населения.

Шестой подъем заболеваемости зарегистрирован в августе 2022 г. и связан с распространением линии BA.5 Omicron SARS-CoV-2. Максимальные показатели зарегистрированы на 38-й неделе 2022 г. (12 сентября), что совпало с влиянием сезонных факторов. С ноября 2022 г. отмечалась стабилизация эпидемического процесса, а средний недельный показатель составил $26,36 \pm 1,94$ случая на 100 тыс. населения.

С середины января 2023 г. формировался **седьмой эпидемический подъем** заболеваемости новой коронавирусной инфекцией (COVID-19), обусловленный доминированием генетической линии XBB SARS-CoV-2. Наибольший недельный показатель заболеваемости при этом зарегистрирован на 8-й неделе 2023 г. (13 февраля) и составил 64,64 случая на 100 тыс. населения. С конца мая 2023 г. эпидемический процесс новой коронавирусной инфекции (COVID-19) стабилизировался, а средний недельный показатель снизился до $2,90 \pm 0,53$ случая на 100 тыс. населения.

Восьмой подъем заболеваемости новой коронавирусной инфекцией (COVID-19) зарегистрирован в сентябре 2023 г. Пик заболеваемости пришелся на 50-ю неделю 2023 г. (11 декабря 2023 г.) и составил 71,95 случая на 100 тыс. населения. Учитывая существенные различия периодов заболеваемости, а также представленные различия в структуре циркулирующих генетических вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2), целесообразно проведение отдельного анализа уровня и динамики заболеваемости и факторов, оказывающих влияние на эпидемический процесс новой коронавирусной инфекции (COVID-19) для каждого из восьми периодов отдельно.

Для всестороннего изучения динамики заболеваемости населения новой коронавирусной инфекцией (COVID-19) и влияния различных факторов на эпидемический процесс в различных субъектах Российской Федерации в разные периоды пандемии применен метод кластерного анализа. Выявлены регионы, в которых доля пропущенных значений превышала 1% от общего числа дней наблюдения. Данные субъекты исключены из основного кластерного анализа и объединены в отдельную группу — кластер 4.

В кластер 4 вошли 23 субъекта Российской Федерации, доля пропусков в которых варьировала от 1 до 17%. Данные регионы выделены в отдельную группу, поскольку неполная характеристика данных не позволяет корректно сравнивать их динамику с остальными субъектами и исключены из дальнейшего анализа. Кластеризация в данном исследовании охватывает период с 01 января 2020 г. по 01 января 2024 г. В результате нами выделено четыре кластера (0–3) (таблица 2), характеризующиеся различной интенсивностью эпидемического процесса новой коронавирусной инфекции (COVID-19).

Кластер 0 включает субъекты с устойчиво низким уровнем заболеваемости (средний недельный показатель заболеваемости колебался от $16,72 \pm 0,97$ в 7 периоде до $125,89 \pm 17,07$ на 100 тыс. населения в 5 периоде), в которых дневные показатели на протяжении всего периода наблюдений оставались значительно ниже среднероссийских значений.

В регионах, отнесенных к **кластеру 1**, заболеваемость имела преимущественно низкий уровень (средний недельный показатель заболеваемости колебался от $29,23 \pm 2,41$ в 7 периоде до $219,26 \pm 24,52$ на 100 тыс. населения в 5 периоде), однако для них были характерны отдельные

периоды умеренного роста интенсивности эпидемического процесса новой коронавирусной инфекции (COVID-19).

В **кластер 2** объединены субъекты, демонстрировавшие преимущественно высокий уровень заболеваемости населения новой коронавирусной инфекцией (COVID-19) (средний недельный показатель заболеваемости колебался от $41,36 \pm 10,54$ в 7 периоде до $260,31 \pm 38,94$ на 100 тыс. населения в 5 периоде). Для данного кластера характерна выраженная динамика подъемов на фоне смены доминирующих генетических вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2).

В **кластер 3** отнесены субъекты Российской Федерации, характеризующиеся стабильно высоким уровнем заболеваемости населения новой коронавирусной инфекцией (COVID-19) на протяжении всего анализируемого периода с 2020 по 2023 гг. (средний недельный показатель заболеваемости колебался от $51,03 \pm 6,95$ в 7 периоде до $538,08 \pm 69,98$ на 100 тыс. населения в 4 периоде). Для данных регионов отмечены выраженные подъемы интенсивности эпидемического процесса, а также высокая чувствительность к смене генетических вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2), что приводило к более заметным колебаниям дневных и недельных показателей заболеваемости.

Таблица 2 — Характеристика периодов эпидемического процесса новой коронавирусной инфекции (COVID-19) на территории Российской Федерации в 2020–2023 гг. с учетом распределения субъектов по кластерам

Номер периода	Начало периода	Конец периода	Длительность (неделя)	Недельные показатели заболеваемости по Российской Федерации (мин., макс., среднее) на 100 тыс.	Доминирующий вариант SARS-CoV-2	Кластер 3 (высокий уровень заболеваемости)		Кластер 2 (преимущественно высокий уровень заболеваемости)		Кластер 1 (преимущественно низкий уровень заболеваемости)		Кластер 0 (низкий уровень заболеваемости)	
						Кол-во субъектов	Средний недельный показатель заболеваемости (на 100 тыс)	Кол-во субъектов	Средний недельный показатель заболеваемости (на 100 тыс)	Кол-во субъектов	Средний недельный показатель заболеваемости (на 100 тыс)	Кол-во субъектов	Средний недельный показатель заболеваемости (на 100 тыс)
1	-	31.08.2020	-	макс - 51,41, среднее - $33,28 \pm 4,14$	Уханьский	2	87,88 $\pm 13,85$	7	47,34 $\pm 4,13$	19	37,77 $\pm 2,92$	34	19,53 $\pm 0,87$
2	01.09.2020	09.05.2021	35	мин - 25,41, макс - 136,69, среднее - $74,28 \pm 11,74$	Уханьский	1	217,14 $\pm 51,72$	4	181,23 $\pm 22,56$	23	80,96 $\pm 3,33$	34	46,70 $\pm 1,50$
3	10.05.2021	12.09.2021	18	мин - 40,90, макс - 118,80, среднее - $85,63 \pm 13,22$	Delta	1	283,99 $\pm 77,07$	2	189,97 $\pm 29,75$	21	109,24 $\pm 6,16$	38	58,36 $\pm 2,51$
4	13.09.2021	09.01.2022	17	мин - 75,96, макс - 191,87, среднее - $140,04 \pm 17,32$	Delta	1	538,08 $\pm 69,98$	9	242,60 $\pm 14,50$	18	167,15 $\pm 6,52$	34	106,25 $\pm 3,28$
5	10.01.2022	25.07.2022	29	мин - 14,00, макс - 907,46, среднее - $186,19 \pm 97,55$	BA.1, BA.2 (Omicron)	1	502,47 $\pm 249,72$	15	260,31 $\pm 38,94$	28	219,26 $\pm 24,52$	18	125,89 $\pm 17,07$
6	26.07.2022	09.01.2023	24	мин - 15,66, макс - 256,47, среднее - $42,83 \pm 33,41$	BA.6 (Omicron)	1	162,66 $\pm 18,82$	12	93,11 $\pm 9,75$	18	93,00 $\pm 9,65$	21	60,87 $\pm 7,73$
7	10.01.2023	28.08.2023	33	мин - 1,98, макс - 64,64, среднее - $23,47 \pm 7,57$	BA.2, XBB (Omicron)	7	51,03 $\pm 6,95$	1	41,36 $\pm 10,54$	16	29,23 $\pm 2,41$	38	16,72 $\pm 0,97$
8	29.08.2023	-	-	мин - 5,43, макс - 71,95, среднее - $31,00 \pm 10,87$	XBB (Omicron)	2	64,76 $\pm 19,01$	8	58,47 $\pm 8,48$	19	35,38 $\pm 3,82$	33	17,58 $\pm 1,38$

Таким образом, динамика средних суточных значений интенсивного показателя заболеваемости новой коронавирусной инфекцией (COVID-19) в кластерах не только отражает общую эпидемиологическую ситуацию в Российской Федерации в 2020-2023 гг., но и позволяет более детально проследить территориальные различия и специфику развития эпидемического процесса на уровне субъектов (рисунок 1, таблица 2).

Различия между среднесуточными показателями заболеваемости новой коронавирусной инфекцией (COVID-19) в субъектах, относящихся к кластерам 0 и 1, оказалась наиболее выраженной во втором эпидемическом периоде, когда значения в кластере 1 ($12,09 \pm 0,17$ на 100 тыс. населения) превышали значения кластера 0 ($3,28 \pm 0,08$ на 100 тыс. населения) в 3,7 раза ($p < 0,01$). В шестом периоде различия были менее значительными ($8,72 \pm 0,33$ и $13,3 \pm 0,53$ на 100 тыс. населения в 0 и 1 кластере, соответственно) и различались в 1,5 раза ($p < 0,01$), что свидетельствует об аналогичных тенденциях в динамике эпидемического процесса в регионах с минимальными уровнями заболеваемости новой коронавирусной инфекцией (COVID-19) (рисунок 1, таблица 2).

Сравнение кластеров 1 и 2 показало, что в периодах 1, 6, 7 и 8 различия между средними суточными показателями заболеваемости новой коронавирусной инфекцией (COVID-19) были минимальными. Наиболее близкие значения зарегистрированы в шестом периоде — $13,30 \pm 0,53$ и $13,33 \pm 0,54$ на 100 тыс. населения, соответственно (различия — около 0,2%), тогда как в восьмом периоде различие достигало 63% — $5,19 \pm 0,23$ и $8,49 \pm 0,48$ на 100 тыс. населения, соответственно (рисунок 1, таблица 2).

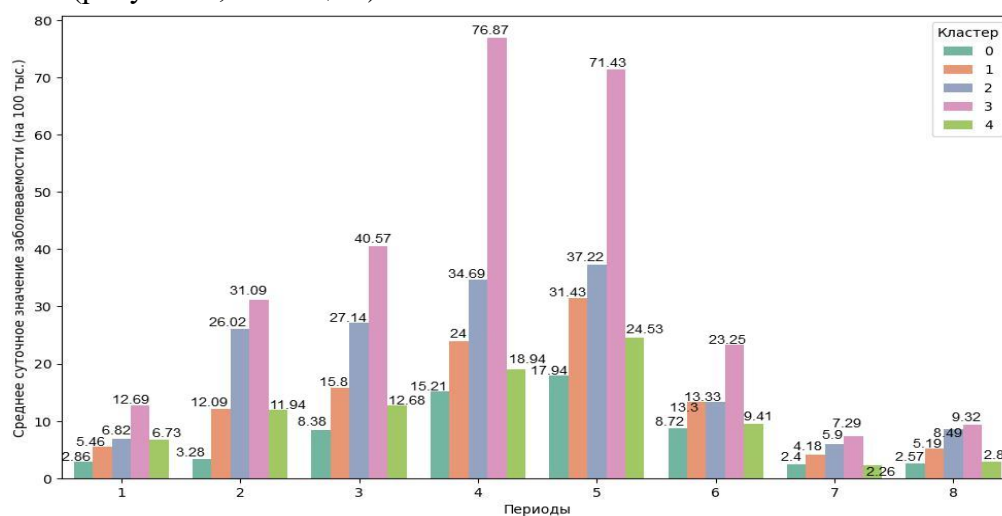


Рисунок 1 — Динамика суточных интенсивных показателей заболеваемости новой коронавирусной инфекцией (COVID-19) в кластерах в каждом из периодов пандемии 2020–2023 гг.

В периодах 2, 3, 4 и 5 различия между кластерами 1 и 2 возрастали, что отражает неоднородность распространения доминирующих генетических вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2) среди населения различных регионов Российской Федерации. В частности, во втором периоде показатели кластера 2 ($26,02 \pm 1,25$ на 100 тыс. населения) превышали значения кластера 1 ($12,09 \pm 0,17$ на 100 тыс. населения) почти в 2 раза ($p < 0,01$), что совпадает с активным ростом заболеваемости новой коронавирусной инфекцией (COVID-19) осенью 2020 г. (рисунок 1, таблица 2).

Различия между кластерами 2 и 3, где объединены регионы со средним и высоким уровнями заболеваемости, также были существенными. В восьмом периоде зафиксирован минимальный разрыв (9,77%), тогда как в четвертом периоде значения кластера 3 ($76,87 \pm 3,73$ на 100 тыс. населения) превышали показатели кластера 2 ($34,69 \pm 0,80$ на 100 тыс. населения) почти в 2 раза ($p < 0,01$), что подтверждает устойчивое эпидемиологическое неблагополучие в субъектах, относящихся к кластеру 3, где на протяжении нескольких периодов наблюдались выраженные эпидемические подъемы заболеваемости новой коронавирусной инфекцией (COVID-19) (рисунок 1, таблица 2). Можно предположить, что наиболее значимые факторы, определившие принадлежность этих территорий к кластеру с наиболее интенсивным развитием эпидемического процесса новой коронавирусной инфекции (COVID-19), связаны с особенностями расселения,

плотностью населения, высокой степенью урбанизации и выраженной транспортной нагрузкой, и связанностью между собой большинства из них.

В состав кластера 2 входили субъекты Российской Федерации, в которых уровень заболеваемости новой коронавирусной инфекцией (COVID-19) на протяжении значительной части пандемии был выше среднего ($6,82 \pm 0,25$, $26,02 \pm 1,25$, $27,14 \pm 1,67$, $34,69 \pm 0,80$, $37,22 \pm 2,15$, $13,33 \pm 0,54$, $5,90 \pm 0,58$, $8,49 \pm 0,48$ на 100 тыс. населения с 1 по 8 периоды, соответственно), но не столь стабильно и экстремально высок, как в кластере 3 (рисунок 1). Данные территории характеризовались выраженными подъемами заболеваемости в отдельные периоды, формируя своего рода «вторую линию» распространения инфекции после регионов с максимальной эпидемической нагрузкой. Динамика заболеваемости в кластерах 2 и 3 нередко была схожей: в отдельные периоды показатели заболеваемости практически совпадали.

В отличие от регионов преимущественно относившихся к кластеру 2, субъекты, входившие в кластер 1, демонстрировали преимущественно низкие уровни заболеваемости на протяжении большей части наблюдаемого периода ($5,46 \pm 0,16$, $12,09 \pm 0,17$, $15,80 \pm 0,33$, $24,00 \pm 0,36$, $31,43 \pm 1,34$, $13,30 \pm 0,53$, $4,18 \pm 0,14$, $5,19 \pm 0,23$ на 100 тыс. населения с 1 по 8 периоды, соответственно). Относительно невысокие значения заболеваемости в этих регионах могут быть связаны с меньшей плотностью населения, особенностями расселения, более низкой интенсивностью миграционных потоков или своевременным введением ограничительных мер.

Наиболее специфичным оказался кластер 0, включавший разнородную группу субъектов Российской Федерации, где на протяжении всех периодов наблюдений не отмечено резких и выраженных подъемов заболеваемости — $2,86 \pm 0,05$, $3,28 \pm 0,08$, $8,38 \pm 0,14$, $15,21 \pm 0,18$, $17,94 \pm 0,94$, $8,72 \pm 0,33$, $2,4 \pm 0,06$, $2,57 \pm 0,08$ на 100 тыс. населения с 1 по 8 периоды, соответственно. Для данных территорий характерна либо географическая изолированность, ограничивающая интенсивность внешних контактов, либо быстрое и эффективное реагирование региональных органов власти на угрозу распространения возбудителя новой коронавирусной инфекции (SARS-CoV-2), что позволило избежать значительных резких эпидемических подъемов заболеваемости новой коронавирусной инфекцией (COVID-19) в разные периоды пандемии.

Оценка влияния социально-демографических и инфраструктурных характеристик, влияющих на уровень и структуру заболеваемости новой коронавирусной инфекцией (COVID-19) в субъектах Российской Федерации

С целью проверки гипотезы о влиянии социально-демографических и инфраструктурных характеристик на уровень и динамику заболеваемости новой коронавирусной инфекцией (COVID-19), а также на распределение субъектов Российской Федерации по кластерам, проведен корреляционный анализ между кластерной принадлежностью субъектов и рядом показателей, отражающих плотность населения, уровень урбанизации и транспортную доступность.

Собраны данные по 62 субъектам Российской Федерации, включающие социально-экономические, демографические и транспортные показатели. На основе результатов кластерного анализа введена метрика «Сумма по кластерам» — интегральный кластерный индекс, отражающий принадлежность региона к определенным кластерам. Данный показатель представляет собой сумму по кластерам за 8 исследуемых периодов по каждому субъекту Российской Федерации. Высокое значение данного индекса указывает, что регион попадал в кластеры с более высоким уровнем заболеваемости новой коронавирусной инфекцией (COVID-19).

На первом этапе проведен корреляционный анализ с расчетом коэффициента корреляции Пирсона между «Суммой по кластерам» и исследуемыми показателями. Так как все рассматриваемые переменные являются количественными (большой частью близкими к непрерывным), расчет коэффициентов Пирсона в данном случае вполне допустим. Проведен корреляционный анализ с расчетом коэффициента корреляции Пирсона между интегральным кластерным индексом «Сумма по кластерам» и исследуемыми показателями. Коэффициент

корреляции Пирсона отражает исключительно линейную зависимость. Если реальная связь имеет нелинейный или пороговый характер, значение r может оказаться близким к нулю, что может создать ошибочное впечатление об отсутствии зависимости.

Применение модели «Random Forest» позволило выделить четыре главных фактора, влияющих на величину интегрального кластерного индекса: 1) **Доля городского населения** оказалась самым значимым признаком (около 52% совокупной важности), и полностью согласуется с результатами проведенного корреляционного анализа, который также выявил значимую положительную связь с интегральным кластерным индексом. Высокая доля городского населения существенно повышает риск распространения новой коронавирусной инфекции (COVID-19), вероятно, вследствие высокой интенсивности социальных контактов, нагрузки на систему здравоохранения и других факторов. 2) **Плотность населения** — второй по значимости признак (28%), показатель которого высоко коррелировал с долей городского населения и числом предприятий (как показывал корреляционный анализ: например, $r = 0,89$ между плотностью населения и числом предприятий, что является высоким показателем). Тем не менее, модель «Random Forest» распределила важность между ними, подтверждая, что плотность населения напрямую влияет на уровень заболеваемости: в плотнонаселенных регионах распространение инфекций происходит существенно быстрее и интенсивнее. 3) **Количество активных предприятий** — третий по вкладу признак (19%). Экономическая активность региона, связанная с численностью предприятий, очевидно сопряжена с количеством населения и урбанизацией, поэтому ее влияние могло частично отражать те же процессы (например, крупные экономические центры – это обычно мегаполисы с высокой плотностью населения и интенсивным транспортным трафиком). Тем не менее, модель «Random Forest» сочла этот признак менее информативным по сравнению с предыдущими. Вероятно, количество предприятий служит косвенным индикатором уровня экономического развития и деловой активности населения, что влияет на распространение инфекционного заболевания. 4) **Коэффициент миграционного прироста** — четвертый значимый фактор (1%).

Корреляция Пирсона измеряет только линейную степень связи между двумя переменными. «Решающие деревья» не предполагают заранее определенной формы зависимости между признаками и целевыми переменными. Они рекурсивно разделяют признаки на области, адаптируясь к сложным и нелинейным взаимосвязям между ними. В нашем случае модель «дерево решений» выявила влияние коэффициента миграционного прироста на интегральный кластерный индекс (не проявившийся в корреляционном анализе). Для выявления сложных зависимостей применение модели «дерево решений» более эффективно, тогда как традиционный расчет корреляций может упустить важные эффекты. В отличие от этого, значимость признаков в модели «дерево решений» представлена набором чисел (по одному на каждый признак), показывающих вклад в модель. Данные числа характеризуют относительную важность, но не имеют направления влияния. Например, высокая важность доли городского населения не указывает напрямую, повышает урбанизация заболеваемость или снижает. Кроме того, модель «решающее дерево» учитывает все признаки одновременно. Важность признака отражает его вклад с учетом остальных факторов, что дает более целостное представление о его роли в многомерной системе.

Возможность применения цифровых технологий учета результатов и анализа данных молекулярно-биологических исследований населения Российской Федерации для повышения информативности анализа эпидемиологической обстановки по новой коронавирусной инфекции (COVID-19)

Цифровая трансформация процессов анализа данных стала неотъемлемой частью эпидемиологического надзора, особенно в период пандемии и явилась предпосылкой для разработки цифровой платформы SOLAR, представляющей важный инструмент для обработки и анализа данных молекулярно-биологических исследований. Платформа создана в рамках

Постановления Правительства РФ №452 от 27 марта 2021 г. в ФБУН Центральный НИИ Эпидемиологии Роспотребнадзора с целью оперативного информирования граждан о результатах ПЦР-исследований, направленных по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2). К ноябрю 2021 г. завершена интеграция всех лабораторий различного ведомственного подчинения, функционирующих на территории Российской Федерации, в систему SOLAR.

Платформа SOLAR (схема представлена на рисунке 2) интегрирует данные, поступающие от более чем 2 000 лабораторий по всей Российской Федерации, и позволяет в реальном времени обрабатывать миллионы результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2). При помощи данной Платформы происходит быстрая передача данных в личные кабинеты пользователей через портал Госуслуг, при этом от момента загрузки результатов диагностической лабораторией до получения информации пользователем проходит не более 30 минут.

Информирование региональных органов Роспотребнадзора о деперсонифицированных результатах диагностического лабораторного обследования дает дополнительную возможность контролировать степень охвата населения обследованием. Так, в соответствии с постановлением Главного государственного врача Российской Федерации от 16 октября 2020 г. №31 «О дополнительных мерах по снижению рисков распространения COVID-19 в период сезонного подъема заболеваемости острыми респираторными вирусными инфекциями и гриппом» был установлен норматив обследования населения (не менее 150 исследований на 100 тыс. населения, среднесуточно за 7 дней), Платформа SOLAR стала важным инструментом контроля достижения этого целевого показателя. С целью определения перспектив использования в рамках оперативного и ретроспективного эпидемиологического анализа заболеваемости деперсонифицированных данных, поступающих из программной платформы SOLAR, для оценки динамики и прогнозирования эпидемического процесса новой коронавирусной инфекции (COVID-19) среди населения Российской Федерации, решены две исследовательские задачи: определение наличия и величины корреляции, а также временного интервала между динамикой официально зарегистрированных случаев заболеваемости новой коронавирусной инфекцией (COVID-19) совокупного населения Российской Федерации и динамикой количества ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) в 5–8 периоды пандемии (10.01.2022–17.12.2023 гг.). Выбор периодов пандемии для изучения связан с началом полноценного функционирования платформы SOLAR к началу 5-го периода пандемии (с декабря 2021 г.) и ограничивается концом 2023 г., когда объем проводимых ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) значительно снизился.

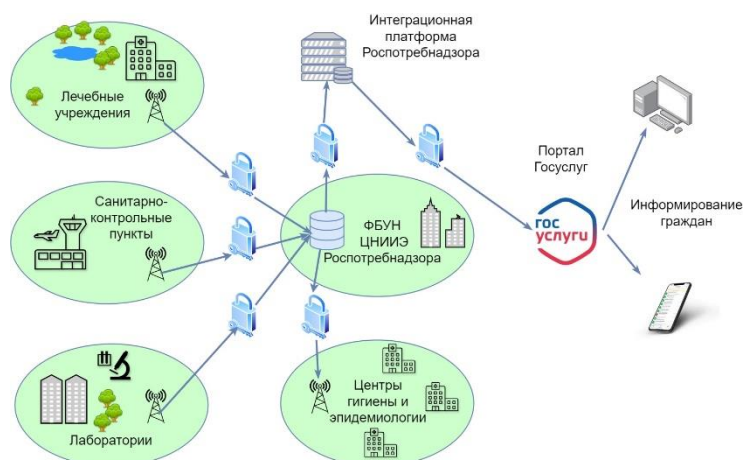


Рисунок 2 — Принцип функционирования программной платформы SOLAR в различных системах федеральных органов исполнительной власти

Для определения корреляции и временного интервала между динамикой заболеваемости новой коронавирусной инфекцией (COVID-19) совокупного населения Российской Федерации и динамикой количества ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) использованы сведения о количестве ежедневно регистрируемых новых случаев заболевания по данным отраслевых форм статистического наблюдения и программной платформы SOLAR за 5–8 периоды пандемии (10.01.2022–17.12.2023 гг.).

В 5-м периоде пандемии новой коронавирусной инфекции (COVID-19) максимальный коэффициент корреляции между количеством заболевших и количеством ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) получен при величине интервала, равном 9 дням, при этом он составил 0,93 (очень высокая связь по шкале Чеддока). Это означает, что динамика количества проведенных ПЦР-исследований опережала регистрацию новых случаев заболевания новой коронавирусной инфекцией (COVID-19) в среднем на 9 дней в 5-м периоде пандемии (10.01–25.07.2022 г.). При этом зависимость между анализируемыми показателями с высокой степенью точности (коэффициент детерминации равен 0,86, что соответствует высокой силе связи по шкале Чеддока) в этом периоде может быть описана линейной функцией. Это дает возможность утверждать о наличии прямой линейной зависимости между показателем количества ПЦР-исследований и количеством заболевших в 5-м периоде пандемии (10.01–25.07.2022 гг.). Полученные результаты свидетельствуют о возможности использования показателя количества ПЦР-исследований в 5-м периоде пандемии для краткосрочного прогнозирования динамики регистрации новых случаев заболевания.

В 6–8 периодах (26.07.2022–17.12.2023 гг.) пандемии наблюдалась выраженная недельная цикличность в обоих анализируемых временных рядах, что нашло свое отражение на графиках экстремумов взаимной корреляции. Здесь и далее под термином экстремум мы используем общепринятое определение — значение непрерывной функции, являющееся ее максимумом или минимумом. При этом максимальный коэффициент корреляции между динамикой количества заболевших и количеством ПЦР-исследований в 6–8 периодах получен при величине интервала, равном 1 дню, и составил 0,67 (умеренно сильная связь по шкале Чеддока), 0,72 (высокая связь по шкале Чеддока) и 0,50 (умеренно сильная связь по шкале Чеддока), соответственно.

В 5-м периоде (10.01–25.07.2022 гг.) пандемии новой коронавирусной инфекции (COVID-19) максимальный коэффициент корреляции между динамикой количества заболевших и долей положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) получен при величине интервала, равном 6 дням, при этом он составил 0,97 (очень высокая связь по шкале Чеддока). Полученные значения свидетельствуют о том, что динамика доли положительных результатов опережала регистрацию новых случаев заболевания новой коронавирусной инфекцией (COVID-19) в среднем на 6 дней в 5-м периоде пандемии (10.01–25.07.2022 гг.).

В 6-м периоде (26.07.2022–09.01.2023 гг.) пандемии максимальное значение коэффициента корреляции (0,98 – очень высокая связь по шкале Чеддока) между динамикой регистрации новых случаев заболевания новой коронавирусной инфекцией (COVID-19) и динамикой доли положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) получено при значении интервала, равном 8 дням (динамика доли положительных результатов опережает динамику регистрации новых случаев заболевания в среднем на 8 дней). Построенная линейная модель взаимосвязи временных рядов показала высокую точность, коэффициент детерминации составил 0,96 (очень высокая связь по шкале Чеддока).

Полученные результаты свидетельствуют о возможности использования показателя доли положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) для краткосрочного прогнозирования динамики регистрации новых случаев заболевания в 6-м периоде пандемии (26.07.2022–09.01.2023 гг.).

В 7-м периоде пандемии (10.01–28.08.2023 гг.) максимальное значение коэффициента корреляции, равное 0,99 (очень высокая связь по шкале Чеддока), получено при величине временного интервала, равном одному дню, что свидетельствует о том, что динамика доли положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) опережала динамику регистрации новых случаев заболевания в среднем на 1 день. Такой показатель подтверждает высокую информативность доли положительных ПЦР-результатов, который можно использовать для оперативного прогнозирования заболеваемости, особенно в условиях высокой динамичности распространения вируса. Данный показатель имеет важное значение для более точного планирования противоэпидемических мероприятий, таких как временные ограничения или мобилизация медицинских ресурсов.

В 8-м периоде пандемии (29.08–17.12.2023 гг.) максимальный коэффициент корреляции между динамикой количества заболевших новой коронавирусной инфекцией (COVID-19) и долей положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) получен при величине интервала, равном 8 дням, при этом он составил 0,98 (очень высокая связь по шкале Чеддока). Полученные значения свидетельствуют о том, что динамика доли положительных результатов опережала регистрацию новых случаев заболевания в среднем на 8 дней в восьмом периоде пандемии. Этот результат подтверждает, что доля положительных результатов ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) является чувствительным индикатором, который может прогнозировать рост заболеваемости с высокой точностью за несколько дней до того, как новые случаи заболевания будут зарегистрированы. При этом зависимость между анализируемыми показателями с высокой степенью точности (коэффициент детерминации равен 0,95 – очень высокая связь по шкале Чеддока) в этом периоде может быть описана линейной функцией, что также подтверждает высокую прогностическую способность данного показателя в эпидемиологических моделях.

Важной особенностью платформы SOLAR является включение в систему данных о результатах лабораторных исследований у лиц, не имеющих клинических проявлений заболевания и не обращавшихся за медицинской помощью. Интеграция результатов ПЦР-исследований лиц с бессимптомными формами инфекции в единый контур наблюдения обеспечила более полное представление о масштабах циркуляции возбудителя новой коронавирусной инфекции (SARS-CoV-2) и позволила выявлять потенциальных носителей вируса на более ранних этапах. Полученные результаты подчеркивают важность массового ПЦР-обследования как инструмента выявления скрытой циркуляции вируса и демонстрируют, что включение данных о бессимптомных носителях значительно повышает эффективность оперативного эпидемиологического анализа, позволяя корректнее и эффективнее оценивать реальные уровни заболеваемости и выстраивать более обоснованные стратегии профилактики и контроля распространения инфекции.

Совершенствование эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) на основе цифровых технологий учета и анализа данных молекулярно-биологических исследований

Использование цифровой платформы SOLAR в системе эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) позволило существенно трансформировать информационную, диагностическую и управленческую подсистемы, обеспечив их интеграцию на основе единых цифровых решений и больших массивов данных молекулярно-биологических исследований. Совершенствование информационной подсистемы эпидемиологического надзора связано, прежде всего, с централизацией и унификацией потоков данных о результатах ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2). В отличие от фрагментарного сбора информации, характерного для начального

этапа пандемии, внедрение SOLAR обеспечило централизованный сбор данных из всех субъектов Российской Федерации, минимизацию дублирования и снижение вероятности ошибок при передаче сведений. Диагностическая подсистема эпидемиологического надзора была существенно усовершенствована благодаря внедрению платформы SOLAR. Систематизация данных о результатах ПЦР-исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) способствовала более точному определению начала и завершения эпидемических подъемов, объективизации эпидемиологического диагноза и повышению достоверности расчета интенсивных показателей заболеваемости. Кроме того, интеграция данных SOLAR с аналитическими инструментами (в том числе методами кластеризации и машинного обучения) позволила выявлять скрытые закономерности развития эпидемического процесса, оценивать вклад различных факторов и формировать научно обоснованные прогнозы.

Совершенствование управленческой подсистемы выразилось в повышении оперативности и обоснованности принимаемых решений. Наличие актуальной информации в режиме, приближенном к реальному времени, обеспечило возможность быстрого реагирования на изменение эпидемиологической ситуации. Разработка и внедрение платформы SOLAR позволили получить инструмент для оценки эффективности проводимых противоэпидемических мероприятий, распределения ресурсов, а также планирования дополнительных мер в субъектах с неблагоприятной динамикой показателей. Платформа SOLAR (рисунок 3) может интегрироваться с другими информационными системами, что обеспечивает более широкий охват и глубину анализа эпидемиологической ситуации.

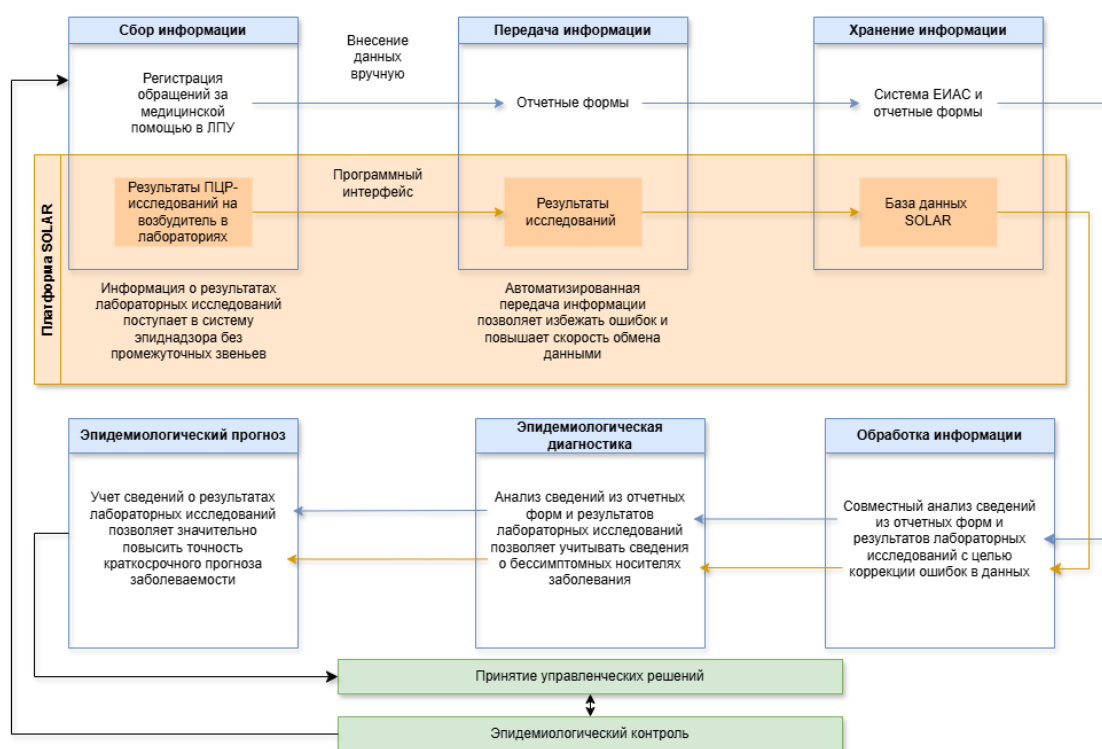


Рисунок 3 — Использование данных Платформы SOLAR для совершенствования системы эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19)

Использование данных Платформы SOLAR, а именно показателя доли положительных результатов ПЦР-исследований, продемонстрировал высокую прогностическую способность для краткосрочного прогнозирования динамики регистрации новых случаев заболевания. Максимальные коэффициенты корреляции в разные периоды пандемии достигали значений от 0,97 до 0,99 (очень высокая связь по шкале Чеддока) при различных величинах временного интервала, что подтверждает информативность этого показателя для оперативного планирования противоэпидемических мероприятий. Совершенствование системы эпидемиологического

надзора за новой коронавирусной инфекцией (COVID-19) на основе цифровых технологий учета и анализа данных молекулярно-биологических исследований позволяет объективизировать эпидемиологический анализ, повысить точность расчета показателей и обеспечить научно обоснованное прогнозирование развития эпидемического процесса.

Использование больших массивов стандартизированных лабораторных данных в сочетании с методами математического моделирования и машинного обучения формирует предпосылки для перехода к проактивной модели эпидемиологического надзора, ориентированной на предупреждение осложнения ситуации и своевременное принятие управленческих решений.

ВЫВОДЫ

1. В динамике заболеваемости новой коронавирусной инфекцией (COVID-19) населения Российской Федерации в период 2020–2023 гг. отмечалось восемь эпидемических циклов заболеваемости, характеризовавшихся особенностями проявлений эпидемического процесса и обусловленных спецификой доминирующих вариантов возбудителя новой коронавирусной инфекции (SARS-CoV-2). Динамика заболеваемости новой коронавирусной инфекцией (COVID-19) в период с апреля 2020 г. по январь 2022 г. имела тенденцию к росту. Наиболее высокий уровень заболеваемости новой коронавирусной инфекцией (COVID-19) зарегистрирован в пятом периоде пандемии в январе — феврале 2022 г. — 905 случаев на 100 тыс. населения (вариант Омикрон SARS-CoV-2), после чего наметилась тенденция к снижению уровня заболеваемости.

2. Кластерный анализ временных рядов заболеваемости новой коронавирусной инфекцией (COVID-19), проведенный в 62 субъектах Российской Федерации, позволил выделить группы территорий достоверно различающиеся по интенсивности проявлений эпидемического процесса. Наиболее высокая интенсивность эпидемического процесса регистрировалась в кластере, включающем в различные периоды г. Москву, г. Санкт-Петербург, Хабаровский край, где средние показатели заболеваемости новой коронавирусной инфекцией (COVID-19) превышали общероссийские в среднем 2,8–4,5 раза ($p < 0,01$).

3. На основе применения корреляционного анализа и использования метода машинного обучения установлено, что уровень заболеваемости и характер распространения новой коронавирусной инфекции (COVID-19) в субъектах Российской Федерации имел выраженную зависимость от доли городского населения, плотности населения региона, интенсивности экономического развития (количества активных предприятий), показателей активности миграционных процессов (коэффициента миграционного прироста).

4. Цифровая Платформа агрегирования результатов лабораторных исследований SOLAR (*System of laboratory aggregation results*) является эффективным инструментом для агрегации, обработки и анализа данных ПЦР исследований по выявлению РНК возбудителя новой коронавирусной инфекции (SARS-CoV-2) среди населения Российской Федерации. Интеграция данных о результатах лабораторных исследований среди бессимптомных носителей вируса посредством Платформы SOLAR позволила обеспечить контроль уровня охвата населения ПЦР-обследованием и оперативную оценку эпидемиологической обстановки на основе показателей заболеваемости, структуры и динамики данных лабораторных исследований.

5. Совершенствование системы эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) на основе цифровизации технологий учета и анализа результатов молекулярно-биологических исследований позволяет существенно дополнять сведения о заболеваемости населения Российской Федерации, достигать высокой информативности о характере анализируемой эпидемиологической ситуации, обеспечивать достоверное краткосрочное прогнозирование эпидемиологической обстановки в интересах оперативности и адресности принимаемых управленческих решений.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

По результатам, полученным в данном диссертационном исследовании, при проведении ретроспективного эпидемиологического анализа нами рекомендуется применение метода кластеризации территорий по интенсивным показателям при изучении заболеваемости населения Российской Федерации новой коронавирусной инфекцией (COVID-19) для получения объективной сравнительной оценки развития эпидемического процесса в различных субъектах Российской Федерации. Данные молекулярно-биологических методов исследования и кластеризация субъектов Российской Федерации по интенсивным показателям заболеваемости населения могут применяться в качестве индикатора эффективности проведения противоэпидемических мероприятий с учетом особенностей каждого субъекта Российской Федерации. Для автоматизации рутинных процессов, проведения комплексного анализа эпидемиологических данных с использованием методов машинного обучения, а также надежного хранения и поиска информации о заболеваемости новой коронавирусной инфекцией (COVID-19) на территории Российской Федерации целесообразно использовать современные цифровые платформы, в том числе с агрегированными результатами молекулярно-биологических исследований (SOLAR).

ПЕРСПЕКТИВЫ ДАЛЬНЕЙШЕЙ РАЗРАБОТКИ ТЕМЫ

Результаты, полученные в рамках диссертационного исследования, касающиеся применения кластерного анализа для изучения ретроспективных данных по заболеваемости новой коронавирусной инфекцией (COVID-19), открывают новые перспективы для углубленного анализа региональных особенностей эпидемического процесса и других инфекций с аэрозольным механизмом передачи, с высоким эпидемическим и пандемическим потенциалом распространения возбудителя. Внедрение математических методов моделирования и алгоритмов машинного обучения в данный контекст позволит значительно повысить точность прогнозирования эпидемической обстановки и эффективность управления эпидемиологическими рисками, что в свою очередь, будет способствовать разработке более целенаправленных и научно обоснованных стратегий эпидемиологического надзора. Дополнение существующей структуры Платформы SOLAR данными молекулярно-биологических исследований, полученными при изучении заболеваемости другими инфекционными агентами, позволит выстроить высокотехнологичную систему эпидемиологического надзора, основанную на современных цифровых решениях.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. **Монахова, А.А.** Роль использования результатов ПЦР-исследований, как важного индикатора текущей эпидемиологической ситуации / А.А. Монахова, Д.В. Дубоделов // Сборник тезисов V Ежегодной конференции по инфекционным болезням "Покровские чтения", Москва, 27–28 октября 2025 года. – Москва: Общество с ограниченной ответственностью "Медицинское Маркетинговое Агентство", 2025. – С. 56.
2. **Монахова, А.А.** Цифровые технологии в эпидемиологическом надзоре за новой коронавирусной инфекцией (COVID-19) / А.А. Монахова, Д.В. Дубоделов // Сборник тезисов V Ежегодной конференции по инфекционным болезням "Покровские чтения", Москва, 27–28 октября 2025 года. – Москва: Общество с ограниченной ответственностью "Медицинское Маркетинговое Агентство", 2025. – С. 57.
3. **Монахова, А.А.** Эпидемиологический анализ и особенности коронавирусной инфекции (COVID-19) за период 2020-2022 гг. / А.А. Монахова, Д.В. Дубоделов // Сборник тезисов V Ежегодной конференции по инфекционным болезням "Покровские чтения", Москва, 27–28 октября 2025 года. – Москва: Общество с ограниченной ответственностью "Медицинское Маркетинговое Агентство", 2025. – С. 57.
4. **Монахова, А.А.** Эпидемиологический анализ и особенности коронавирусной инфекции (COVID-19) за период 2022-2023 гг./ А.А. Монахова, Д.В. Дубоделов // Сборник тезисов V

Ежегодной конференции по инфекционным болезням "Покровские чтения", Москва, 27–28 октября 2025 года. – Москва: Общество с ограниченной ответственностью "Медицинское Маркетинговое Агентство", 2025. – С. 58.

5. **Монахова, А.А.** Эпидемиологический надзор за новой коронавирусной инфекцией (COVID-19) / А.А. Монахова, Д.В. Дубоделов // Сборник тезисов V Ежегодной конференции по инфекционным болезням "Покровские чтения", Москва, 27–28 октября 2025 года. – Москва: Общество с ограниченной ответственностью "Медицинское Маркетинговое Агентство", 2025. – С. 58.

6. Опыт использования машинного обучения Random Forest для оценки влияния социально-экономических показателей регионов на заболеваемость новой коронавирусной инфекцией (COVID-19) / Д.В. Дубоделов, **А.А. Монахова**, Н. В. Сычева [и др.] // Сборник материалов VI Международной научно-практической конференции по вопросам противодействия инфекционным заболеваниям, Санкт-Петербург, 11–12 декабря 2025 года. – Саратов: Амирит, 2025. – С. 85-86.

7. **Монахова, А.А.** Возможности использования платформы solar в системе эпидемиологического надзора / А.А. Монахова, Д.В. Дубоделов, В.Г. Акимкин // Эпидемиология - 2025: Сборник тезисов Конгресса с международным участием, Москва, 15–16 октября 2025 года. – Москва: Центральный научно-исследовательский институт эпидемиологии Роспотребнадзора, 2025. – С. 90-91.

8. **Монахова, А.А.** Результаты эпидемиологического анализа динамики заболеваемости COVID-19 в Российской Федерации в 2020-2023 годах / А.А. Монахова, Д.В. Дубоделов, В.Г. Акимкин // Эпидемиология - 2025: Сборник тезисов Конгресса с международным участием, Москва, 15–16 октября 2025 года. – Москва: Центральный научно-исследовательский институт эпидемиологии Роспотребнадзора, 2025. – С. 91-92.

9. Эпидемиологическая и экономическая эффективность массового ПЦР-обследования населения Российской Федерации в период эпидемии COVID-19 / В.Г. Акимкин, Д.В. Дубоделов, А.С. Есьман [и др.] // Эпидемиология и вакцинопрофилактика. – 2025. – Т. 24, № 6. – С. 28-35. – DOI 10.31631/2073-3046-2025-24-6-28-35.

10. Использование интегративных показателей на основе сведений о результатах лабораторных исследований для оценки эффективности действующей системы эпидемиологического надзора / Д.В. Дубоделов, В.Г. Акимкин, Т.А. Семенов [и др.] // Эпидемиология и инфекционные болезни. Актуальные вопросы. – 2025. – Т. 15, № 1. – С. 35-42. – DOI 10.18565/epidem.2025.15.1.35-42.

11. Цифровые решения (VGARus, SOLAR, «Epidsmart - модуль COVID») в системе эпидемиологического надзора за новой коронавирусной инфекцией (COVID-19) / В.Г. Акимкин, Д.В. Дубоделов, А.С. Есьман [и др.] // Эпидемиология и инфекционные болезни. Актуальные вопросы. – 2025. – Т. 15, № 1. – С. 43-50. – DOI 10.18565/epidem.2025.15.1.43-50.

12. Значение бессимптомных и легких форм заболевания в развитии эпидемического процесса новой коронавирусной инфекции (COVID-19) / В.Г. Акимкин, А.С. Есьман, Д.В. Дубоделов [и др.] // Эпидемиология и инфекционные болезни. Актуальные вопросы. – 2026. – Т. 16, № 2. – С. 6-12. – DOI 10.18565/epidem.2026.16.2.6-12.