

НАО «Карагандинский медицинский университет»

УДК 616.9:578.834.1: 616-092.6-084-07

На правах рукописи

**СУЛТАНБЕКОВА АЙДАНА АСКАРОВНА**

**«Детекция SARS-CoV-2 в назальных мазках на основе время-пролетной  
масс-спектрометрии (MALDI-MS)»**

8D10100 – Медицина

Диссертация на соискание степени  
доктора философии (PhD)

Отечественные научные консультанты  
доктор PhD,  
ассоциированный профессор  
И.А. Кадырова

Зарубежный научный консультант:  
доктор PhD  
С.В. Егоров  
(Онтарио: Институт  
Майкла Г. ДеГрота)

Республика Казахстан  
Караганда, 2024

## СОДЕРЖАНИЕ

|   |    |
|---|----|
| <b>НОРМАТИВНЫЕ ССЫЛКИ</b> .....   | 3  |
| <b>ОБОЗНАЧЕНИЕ И СОКРАЩЕНИЯ</b> .....   | 4  |
| <b>ВВЕДЕНИЕ</b> .....   | 5  |
| <b>1 ОБЗОР ЛИТЕРАТУРЫ</b> .....   | 11 |
| 1.1 Молекулярно-генетические методы диагностики КВИ.....  | 12 |
| 1.2 Серологические методы диагностики КВИ .....   | 15 |
| 1.3 Опыт применения MALDI-TOF MS для диагностики вирусов .....  | 17 |
| 1.4 Опыт применения MALDI-TOF MS для диагностики КВИ .....  | 20 |
| <b>2 МАТЕРИАЛЫ ИССЛЕДОВАНИЯ</b> .....   | 25 |
| 2.1 Молекулярно-генетическая детекция патогенов.....  | 26 |
| 2.2 Проведение MALDI-TOFMS.....   | 28 |
| 2.3 Препроцессинг спектров.....   | 28 |
| 2.4 Машинное обучение и статистический анализ.....  | 29 |
| <b>3 РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ</b> .....  | 33 |
| 3.1 Характеристика основной группы по данным анкетирования.....   | 33 |
| 3.2 Характеристика контрольной группы.....  | 40 |
| 3.3 Сравнительный анализ контрольной и основной групп.....  | 44 |
| 3.4 Характеристика групп в соответствии с ПЦР тестированием.....  | 46 |
| 3.5 Результаты комбинированной работы методов MALDI-TOF MS и ML.....  | 47 |
| 3.5.1 Применение существующего пайплайн MALDI-TOF MS для выделения образцов SARS-CoV-2+, собранных в Казахстане (Анализ I)  | 47 |
| 3.5.2 Самостоятельное обучение алгоритмов, используя собственные данные MALDI-TOF MS для идентификации SARS-CoV-2 среди прочих ОРИ и асимптоматических мазков (Анализ II).....                        | 56 |
| 3.6 Перспективы применения MALDI-TOF MS/ML для диагностики SARS-CoV-2 в назальных мазках.....   | 64 |
| <b>ЗАКЛЮЧЕНИЕ</b> .....   | 66 |
| <b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b> .....   | 70 |
| <b>ПРИЛОЖЕНИЕ А</b> – Акт внедрения .....   | 79 |
| <b>ПРИЛОЖЕНИЕ Б</b> - Свидетельство об авторском праве.....   | 81 |
| <b>ПРИЛОЖЕНИЕ В</b> – Стандартная операционная процедура «Правила анкетирования пациентов».....   | 82 |
| <b>ПРИЛОЖЕНИЕ Г</b> – Стандартная операционная процедура «Забор назального мазка».....  | 86 |
| <b>ПРИЛОЖЕНИЕ Д</b> – Масс-спектрометрия, выполненная с использованием стратегии Nachtigall и соавт. на образцах из Казахстана (Анализ I).....  | 90 |
| <b>ПРИЛОЖЕНИЕ Е</b> – Пики MALDI-TOF MS, обнаруженные в объединенном наборе данных, содержащем образцы как из Казахстана, так и из Южной Америки и присутствующие хотя бы в 80% каждой подгруппы..... | 94 |

## **НОРМАТИВНЫЕ ССЫЛКИ**

В настоящей диссертации использованы ссылки на следующие стандарты:

Закон Республики Казахстан. О науке: принят 18 февраля 2011 года, №407-IVЗРК.

Кодекс Республики Казахстан. О здоровье народа и системе здравоохранения: принят 7 июля 2020 года, №360-VI ЗРК.

Закон Республики Казахстан. О биологической безопасности Республики Казахстан: принят 21 мая 2022 года, №122-VII ЗРК.

ГОСТ 7.32-2017. Отчет о научно-исследовательской работе. Структура и правила оформления.

ГОСТ 7.1-2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления.

## ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

|              |  |
|--------------|--|
| AUC          | – Area Under the Curve (площадь под кривой)  |
| COVID-19     | – CoronaVirusDisease 2019 (коронавирусная болезнь 2019)  |
| DT           | – decision tree  |
| FAMD         | – factor analysis of mixed data (факторный анализ смешанных данных)  |
| KNN          | – k-nearest neighbors  |
| LDA          | – linear discriminant analysis   |
| LG           | – logistic regression  |
| MALDI-TOF MS | – Matrix Assisted Laser Desorption/Ionization time of Flight Mass Spectrometry (матрично-активированная лазерная десорбция/ионизация время-пролетная масс-спектрометрия) |
| ML           | – machine learning (машинное обучение)   |
| NB           | – naïve Bayes  |
| NPV          | – negative predictive value (прогностическая ценность ложного результата)  |
| PPV          | – positive predictive value (прогностическая ценность положительного результата)   |
| PCA          | – principal component analysis (Метод главных компонент)   |
| RF           | – random forest  |
| ROC          | – Receiver operating characteristic (Рабочая характеристика приемника)   |
| SARS-COV-2   | – severe acute respiratory syndrome-related coronavirus 2 (тяжелый острый респираторный синдромом связанный с Коронавирус 2)   |
| SARS-COV-2+  | – основная группа пациентов с коронавирусной инфекцией   |
| SVM-L        | – support vector machine with linear kernel  |
| SVM-R        | – support vector machine with radial basis function kernel   |
| XGBoost      | – gradientboostingmethod   |
| КВИ          | – коронавирусная инфекция  |
| ОРИ          | – острые респираторные инфекции  |
| ОТ-ПЦР       | – полимеразная цепная реакция с обратной транскрипцией   |
| Пайплайн     | – алгоритм проведения биоинформатического анализа, типовых задач   |
| СОП          | – стандартная операционная процедура   |

## ВВЕДЕНИЕ

### **Актуальность исследования**

Диссертационная работа выполнена в рамках научного проекта грантового финансирования Министерства образования и науки РК (AP09259123) «Обнаружение SARS-CoV-2 в назальных мазках с помощью MALDI-MS и методов машинного обучения» и соответствует приоритетному направлению развития науки «Науки о жизни и здоровье».

Возникновение нового коронавируса в 2019 году создало значительные трудности для мирового сообщества и систем здравоохранения. COVID-19 продолжает быть проблемой общественного здравоохранения [1] и особенно опасен для регионов с ограниченной инфраструктурой здравоохранения, которые страдают от недостаточного эпидемиологического контроля [2]. В связи с тем, что пандемия выявила множество проблем в организации медицинских и диагностических процедур, которые были направлены на лечение пациентов, в странеслучился коллапс в стационарах и лабораториях в период пика эпидемии [3].

Лабораториями в 2023 году проведен большой набор исследований, которые направлены на выявление компонентов вируса SARS-CoV-2 или же на обнаружение косвенных признаков его наличия в организме [4-6].

На данный момент метод ПЦР с обратной транскрипцией (ОТ-ПЦР) является наиболее эффективным способом обнаружения коронавиральной РНК в биоматериале. В настоящее время использование ОТ-ПЦР является экономически выгодным способом диагностики. Однако, его применение требует значительных временных и финансовых затрат [7-9].

Проведение серологических тестов стало особенно актуальным в данный момент. Для диагностики поствакцинального иммунитета, а также стадийности течения иммунной реакции после контакта с инфекционным агентом, используется определение антител IgG, IgA, IgM к различным структурным элементам вирусной частицы SARS-CoV-2 (нуклеокапсид, RBD, Spike- protein) [10]. Ограничениями данного подхода к диагностике являются невозможность обнаружения инфекции в первые дни болезни, отсутствие контакта с носителем патогена и перекрестная реактивность, приводящая к ошибочным выводам.

Метод нейтрализации имеет ряд особенностей, которые делают его высокочувствительным и специфичным. Эта методика позволяет определить антитела, которые участвуют в нейтрализации вируса и предотвращении клеточной инфекции [11]. Метод нейтрализации является трудоемким, требует наличия определенных технических условий для проведения анализа, а также имеет ряд особенностей, которые необходимо учитывать при его применении [12].

Известно, что использование метода выделения вирусной культуры для постановки диагноза нецелесообразно, так как для того, чтобы проявились очевидные цитопатические эффекты SARS-CoV-2 в клеточных линиях, необходимо не менее 3 дней [13]. С другой стороны, уровень оснащённости лабораторий должен соответствовать третьему уровню биобезопасности, что

нехарактерно для большинства медицинских учреждений, так как они не имеют в наличии соответствующих помещений.

На начальных стадиях пандемии нехватка доступных, простых и быстрых методов диагностики, которые могли бы быть модифицированы для адаптации к новому вирусу, побуждала к поиску альтернативных диагностических инструментов. Согласно исследованиям, проведенным в области молекулярной и иммунохроматографической диагностики, метод масс-спектрометрии с лазерной десорбцией/ионизацией с использованием матрицы (MALDI MS) является перспективным способом поиска патогенов ОРВИ, который может быть использован в качестве альтернативы молекулярным и иммунохроматографическим методам [14]. Несмотря на то, что в литературе было описано множество способов, основанных на MALDI MS [14, p. 154922-1-15422-8], в данной работе рассматривался подход в использовании комбинации MALDI-TOF MS и алгоритмов машинного обучения для обнаружения SARS-CoV-2 непосредственно в назальных мазках [15-26]. Такой подход дает преимущество по времени выполнения анализа, т.к. этап пробоподготовки отсутствует. Это в свою очередь может привести к искажению полученных данных, т.к. объект в таком исследовании имеет изменчивые характеристики в популяции (состав носовой слизи, географическое расположение, иммунный ответ хозяина, достаточная вирусная нагрузка).

Таким образом, проблема комбинированного использования MALDI-TOF MS/ML является актуальной, однако данные о воспроизводимости и эффективном использовании описанных методик отсутствуют [15, p. 1168-1172; 16, p. 661358-1-661358-11]. Исходя из актуальности исследования была определена гипотеза и цель.

#### **Гипотезы:**

1. Производительность оригинальных моделей машинного обучения обученных на данных MALDI-TOF MS может отличаться в различных исследовательских группах из-за изменчивых характеристик в популяции.
2. Увеличение количества и биологического разнообразия коллекции назальных мазков, с последующим дообучением позволит оптимизировать производительность комбинированного метода MALDI ToF MS/ML.

#### **Цель исследования**

Оценить и оптимизировать экспресс-метод определения вируса SARS-CoV-2 в материале, полученном из назального мазка при помощи времяпролетной масс-спектрометрии и методов машинного обучения.

#### **Задачи исследования**

Поставленная цель определила следующие задачи исследования:

1. Сформировать коллекцию назальных мазков и базу клинических данных пациентов с COVID-19, ОРВИ и условно-здоровых пациентов.
2. Провести анализ назальных мазков молекулярно-генетическими методами исследования (ОТ-ПЦР).
3. Провести MALDI-TOF спектрометрический анализ назальных мазков и препроцессинг полученных пиков.

4. Оценить характеристики производительности оригинальной модели на казахстанских данных.

5. Обучить собственные модели машинного обучения с целью оптимизации экспресс-методики детекции SARS-CoV-2 в назальных мазках.

6. Сравнить производительность ML первичных разработчиков и собственных обученных моделей.

**Объект исследования.** Объектом исследования явились назальные мазки больных с COVID-19, ОРИ и участников контрольной группы.

**Предмет исследования.** Предметом исследования является анализ эффективности, точности, возможностей и ограничений метода детекции SARS-CoV-2 в назальных мазках с использованием время-пролетной масс-спектрометрии (MALDI-TOF MS).

**Хронологический период исследования.** Для ответа на поставленные задачи было проведено наблюдательное аналитическое исследование в течение 2020-2022 годов в г. Караганде.

#### **Научная новизна:**

Научные положения и новизна исследования являются полностью новыми. Научная новизна диссертации заключается в следующем:

1. Впервые идентифицировано 120 пиков обеспечивающих вклад в детекцию SARS-CoV-2, и сгенерирована матрица интенсивности масс-спектрометрических пиков назальных мазков пациентов Казахстана и Южной Америки.

2. Впервые оценена производительность оригинальных алгоритмов машинного обучения первичных разработчиков на данных Южной Америки и Казахстана (ROC AUC: DT –0.62, KNN - 0.56, NB - 0.65, RF - 0.67, SVM-L - 0.74, SVM-R –0,60, XGBoost - 0.62).

3. Впервые обучены собственные модели машинного обучения на матрице интенсивности объединенных данных (ROC AUC: DT –0.972, KNN - 0.92, NB - 0.801, RF - 0.93, SVM-L - 0.94, SVM-R - 0.983, XGBoost - 0.958).

#### **Основные положения, выносимые на защиту:**

1. Алгоритмы машинного обучения первичных разработчиков не могут использоваться без предварительной коррекции, ROC AUC оригинальных моделей на данных Казахстана составила: DT –0.62, KNN - 0.56, NB - 0.65, RF - 0.67, SVM-L - 0.74, SVM-R –0,60, XGBoost - 0.62.

2. После внесения дополнительных данных и переобучения ML достигнута лучшая производительность, ROC AUC собственных моделей машинного обучения составила: DT –0.972, KNN - 0.92, NB - 0.801, RF - 0.93, SVM-L - 0.94, SVM-R - 0.983, XGBoost - 0.958.

3. Оптимизация комбинированного метода MALDI-TOF MS/ML улучшает производительность модели: ROC AUC анализ I: RF - 0.67, SVM-R – 0,60; ROC AUC анализ II: RF- 0.93, SVM-R –0,98.

#### **Практическая значимость:**

1. Диссертация имеет практическую ценность, которая заключается в оптимизации нового эффективного и экономически выгодного метода детекции SARS-CoV-2 в назальных мазках.

2. Результаты проведенного исследования открывают перспективы развития таргетного лечения заболеваний, вызванных вирусами ввиду их быстрой и экономически выгодной диагностики.

3. Экспресс-методика может быть использована для дифференциальной диагностики пациентов, поступивших в стационар по экстренным показаниям с симптомами простуды к основному заболеванию. О чем свидетельствует акт внедрения в клиническую практику клиники НАО «МУК».

4. Комбинированное использование MALDI-TOF MS и методов машинного обучения может быть применено для скрининга населения с целью выявления SARS-CoV-2 и ОРВИ.

#### **Внедрение в практику.**

Реализация исследований подтверждена актами внедрения в эксплуатацию в перечень методик Научно-исследовательской лаборатории, в клиническую практику клиники НАО Медицинского университета Караганды (Приложение А).

#### **Личный вклад автора**

Автором собрана коллекция назальных мазков пациентов с COVID-19, ОРВИ и условно-здоровых пациентов, показана эпидемиологическая характеристика собранной выборки, впервые оценена производительность оригинальных алгоритмов машинного обучения первичных разработчиков на данных Южной Америки и Казахстана, разработаны собственные модели машинного обучения на матрице интенсивности объединенных данных, получены сведения об эффективности, точности, возможности и ограничении метода детекции SARS-CoV-2 в назальных мазках с использованием время-пролетной масс-спектрометрии (MALDI-TOF MS) и машинного обучения. Экспериментально доказана возможность детекции SARS-CoV-2 в назальных мазках при помощи комбинированного подхода MALDI-TOF MS и методов машинного обучения.

Достоверность результатов исследований была достигнута современными методами исследования и использованием современного лабораторного оборудования, программного обеспечения, валидированных протоколов исследования (ПЦР). В диссертации даны цель и задачи эксперимента, порядок и методика проведения эксперимента, анализ результатов. Теоретические исследования (машинное обучение) основаны на разработке и анализе обучающих и прогнозирующих алгоритмов, которые могут принимать решения на основе данных.

#### **Апробация работы**

Основные положения диссертации доложены и обсуждены на:

1. SARS-COV-2 Detection in MALDI-TOF Mass Spectra by Machine Learning. 11th Int'l Conference on Mathematical Modeling in the Physical Sciences (5-8 сентября, 2022).

2. Detection of SARS-COV-2 in nasopharyngeal swabs with MALDI-TOF MS and machine learning approach. 10th International Congress of the Asia Pacific Society of Infection Control 2022 (APSIC 2022) (28-31 июля, 2022).

3. Application of MALDI-MS to Differentiate SARS-CoV-2 and Non-SARS-CoV-2 Symptomatic Infections in the Early and Late Phases of the Pandemic. *irc 2023 XVII international research conference proceedings* (24-25 июля, 2023).

4. Обнаружение SARS-COV-2 методами машинного обучения и MALDI-TOF MS. VI-я Центрально-Азиатская конференция по медицинскому образованию (20-22 апреля, 2023).

5. Применение MALDI-TOF MS и машинного обучения для SARS-COV-2 и других респираторных инфекций. Международная конференция «Астана Биотех 2024» (сентябрь 12-13, 2024)

### **Публикации**

По материалам диссертации опубликовано 4 научные работы, из них:

– 3 в научных изданиях, рекомендованных Комитетом по контролю в сфере образования и науки МОН РК- в данных статьях автор диссертационной работы проводил систематический обзор литературы, концептуализацию проводимых методов исследования, анализ данных, лабораторные исследование, а также участвовал в написании и редактировании текста; в статье «Опыт применения время пролетной масс-спектрометрии (MALDI-TOF) для детекции вирусных патогенов, в частности коронавируса. Систематический обзор» является единственным автором.

– 1 публикация в международном научном издании, входящем в информационную базу WOS (Q2):

Application of MALDI-TOF MS and machine learning for the detection of SARS-CoV-2 and non-SARS-CoV-2 respiratory infections// *ASM Journals. Microbiology Spectrum*. – 2024. Vol. 12, No. 5, в статье были продемонстрированы результаты лабораторной и аналитической работы автора диссертации, автором была написана значительная часть статьи.

Получено свидетельство о государственной регистрации прав на объект авторского права №24868 от 7 апреля 2022. «Разработка экспресс-метода определения вируса в материале, полученном из назального мазка при помощи времяпролетной масс-спектрометрии и методов машинного обучения SARS-COV-2»(Приложение Б).

**Структура диссертации.** Диссертация состоит из введения, 3 разделов, заключения, списка использованных источников, содержащего 122 наименований, приложения, проиллюстрирована 13 таблицами и 29 рисунками.

**Во введении** обосновывается выбор темы исследования, ее актуальность, сформулирована цель исследования, определяющая ее задачи, объект и предмет исследования, раскрыта научная новизна, теоретическая и практическая значимость работы, выделены этапы исследования.

**В первом разделе** проанализирована ситуация лабораторной диагностики коронавирусной инфекции, показана роль технологии MALDI-TOF-MS в определении вирусов в том числе SARS-COV-2. Определено, что существует необходимость быстрой и достоверной лабораторной диагностики вирусов, в том числе SARS-CoV-2.

**Во втором разделе** показан дизайн исследования, описаны материалы и методы проведения исследования.

**В третьем разделе** приведены данные, полученные при проведении лабораторных и аппаратных методов исследования. Приведены результаты исследования методами ОТ-ПЦР и MALDI-TOF MS в комбинации с машинным обучением.

**В заключении** охарактеризованы результаты исследования, включающие основные выводы и практические рекомендации по итогам диссертационного исследования

## 1 ОБЗОР ЛИТЕРАТУРЫ

В декабре 2019 года в городе Ухань, Китай произошло событие, которое стало причиной вспышки «вирусной пневмонии». Причиной этого явления стал ранее неизвестный штамм коронавируса, который после изучения получил название SARS-CoV-2 [27]. Пандемия коронавируса была объявлена ВОЗ 11 марта 2020 года. Данный вирус относится к роду Betacoronavirus. В его составе присутствует коронаобразная оболочка и одноцепочная РНК. Данная структура вируса состоит из нуклеокапсидного белка, который покрывает геномную РНК и образует так называемый спиральный нуклеокапсид (N). Структура ядра формируется из мембранных белков (M), которые находятся в внутриклеточной мембране и взаимодействуют с внутренними нуклеопротеинами, образуя ядро. Белок, который является оболочкой вируса, состоит из оболочечного белка (E), мембранного белка и спайк-белка (S), который выступает в роли оболочки вируса [28, 29]. Важно отметить то, что общество уже не в первый раз сталкивается с пандемией коронавируса. Ранее отмечались вспышки вируса SARS-CoV-1 в 2003 году, MERS-Co V - в 2012 и 2020 годах, а также пандемия SARS COV-2 в 2019 году стала третьей по счету. Важным этапом в борьбе с вирусом является его обнаружение, в связи с этим лабораторные работники несут большую ответственность [30].

Уровень организации медицинских и диагностических процедур, которые существовали во время пандемии, показал множество недочетов. Это привело к коллапсу стационарной и лабораторной службы по всему миру в период наибольшей заболеваемости [3, p. 100096].

В настоящее время метод ПЦР с обратной транскрипцией (ОТ-ПЦР), является наиболее эффективным способом для обнаружения коронавирусной РНК в биоматериале. Исходя из того факта, что метод ОТ-ПЦР является высокоточным, он требует больших затрат на проведение диагностики [7, p. 2312; 8, p. 2281-2285; 9, p. 131S-137S].

MALDI-TOF MS – метод, который используется в микробиологии, позволяет быстро и точно определить культивируемые микроорганизмы. Данная методика была впервые предложена для идентификации бактерий в 1975 году. В 2004 году в результате исследований была создана база данных, которая позволяет определить бактерии [31]. Метод привлекает своей точностью, высокоэффективностью и низкой затратностью [32].

Согласно известным данным, MALDI-TOF масс-спектрометр состоит из трех основных частей: лазера, анализатора масс и детектора. При воздействии лазерного излучения молекулы подвергаются десорбции и ионизации, что происходит за счет переноса заряда при поглощении энергии кратковременного лазерного импульса. Десорбированные и ионизированные молекулы в процессе своего движения по анализатору масс сначала ускоряются в электрическом поле, затем выбрасываются через металлическую трубку и сталкиваются с детектором, расположенным на конце трубки. Благодаря этому биомолекулы могут быть разделены в соответствии с их массой и зарядом. В результате этого образуются так называемые «отпечатки пальцев» – характерные молекулярные

профили микроорганизма. Затем полученные спектральные пики детектируются в соответствующей базе данных для идентификации микроорганизма [33, 34], например, Bruker BioTyper, bioMerieux VITEK MS и Andromas [35-37]. Для проведения опыта образец смешивают с матрицей. Она состоит из небольших молекул кислоты, которые обладают сильным оптическим поглощением на длине волны используемого лазера. В настоящее время этот метод широко применяется в целях идентификации бактерий, грибов и насекомых [32, с. 97-104].

Благодаря технологии MALDI-TOF MS появилась возможность идентификации бактерий на уровне вида и рода, а также на уровне подвидов. Данная методика позволяет проводить анализ как целых бактериальных клеток, так и их отдельных компонентов [38]. Данный метод может помочь выявить устойчивость к антибактериальным средствам, а также отслеживать изменения в метаболизме бактерий. MALDI-TOF MS, в свою очередь, успешно используется в анализе MassARRAY, который является заменой флуоресцентных зондов при проведении генетических исследований. Также для мониторинга мутаций можно использовать MassARRAY [35, p. 116250].

Несмотря на это, не стоит полностью игнорировать недостатки методики. Стоит иметь в виду, что спектры некоторых близкородственных видов бактерий могут быть идентифицированы неправильно [39]. Вместе с тем, чтобы получить спектр хорошего качества, необходимо иметь хорошо изолированные и чистые колонии, а также высокую концентрацию клеток в образце (10<sup>5</sup>-10<sup>9</sup> клеток на одно пятно). Это увеличивает длительность всего процесса, так как требуется включение этапа культивирования. Кроме того, культуральная среда и матрица, выбранная для анализа влияют на итоговый результат [40]. Наиболее существенным недостатком является база данных, постоянно пополняемая новыми спектрами, которые получены для возбудителей, точно идентифицированными другими методами, в частности редких видов бактерий [41]. Решающими факторами по вопросу определения вируса, помимо скудных спектральных библиотек, является то, что они трудно размножаются даже в идеальных условиях [35, p. 116250]. А также относительно низкое содержание белка и более высокая молекулярная масса вирусных белков создают трудности для работы с MALDI-TOF MS [42]. Тем не менее, в литературе сообщается о положительном опыте использования MALDI-TOF MS для исследования вирусов, следовательно, этот вопрос требует дальнейшего изучения.

### **1.1 Молекулярно-генетические методы диагностики КВИ**

В начале 80-х годов прошлого века был разработан метод полимеразной цепной реакции. Он основан на методе репликации ДНК in-situ, который позволяет быстро и эффективно производить амплификацию целевой ДНК с помощью полимерных олигонуклеотидных праймеров, а также термостабильной ДНК-полимеразы [35, p. 116250; 43]. Новая версия метода ПЦР, которая называется ПЦР в реальном времени, позволяет отслеживать результаты реакции амплификации в реальном времени после каждого цикла

[35, p. 116250; 37, p. 1607-1622]. Существуют различные факторы, которые могут влиять на чувствительность и эффективность методов ПЦР. Это может быть не только присутствие обогащающих компонентов в культуральной среде или наличие раствора для извлечения ДНК, но и температура, и циклические условия, а также концентрация праймеров или матриц [35, p. 116250; 43, p. 97-106]. С помощью данного метода можно провести диагностику бактерий и выявить вирусы. При работе с различными вирусами чаще всего используют количественный-ПЦР анализ. В случае проведения количественного-ПЦР анализа для обнаружения амплифицированных вирусных нуклеиновых кислот используют специальные флуоресцентные зонды, которые содержат в себе маленькую нуклеотидную последовательность, прикрепленную к индикатору с флуоресцентным спектром. Данный метод имеет очень высокую чувствительность. Однако, следует помнить о том, что исследуемый материал должен быть абсолютно чистым. В результате наличия в пробе различных примесей можно получить ложный результат [35, p. 116250; 44]. В особенности данный аспект следует учитывать для РНК-содержащих вирусов, которые могут быть идентифицированы с помощью трансформации РНК в кДНК при помощи обратной транскрипции. Показано, что в случае высокой концентрации вирусной РНК в исследуемом образце возможно ингибирование ПЦР с обратной транскрипцией [35, p. 116250; 45].

В настоящий момент, метод ПЦР с обратной транскрипцией (ОТ-ПЦР) является наиболее эффективным и распространенным способом для обнаружения SARS-CoV-2. ОТ-ПЦР – это наиболее надежный и быстрый метод, который позволяет получить результаты уже через несколько часов с высокой вероятностью [46, 47]. Принцип ОТ-ПЦР основан на преобразовании РНК в комплементарную ДНК (кДНК) при помощи фермента обратной транскрипции, который способствует превращению РНК в кДНК. Далее следует амплификация образца кДНК с помощью полимеразной цепной реакции, которая включает в себя использование ген-специфичных праймеров и флуоресцентно-меченных зондов гидролиза. Во время первого этапа разрабатываются «шаблоны» ДНК, которые будут использованы на втором этапе для создания образцов. При повторении термических циклов второй этап увеличивает количество копий ДНК. Для того, чтобы амплифицировать только выбранную область генома, генно-специфические праймеры используют вторую реакцию. В то же время зонды производят флуоресцентные сигналы при каждой успешной амплификации генов, что обеспечивает количественную систему реакции [48, 49].

Готовые диагностические наборы, включающие в себя ферменты обратной транскрипции и амплификации, наборы праймеров и зондов для создания определенных областей вирусного генома, а также разрешенные реагенты с отрицательным, положительным и внутренним контролем, хорошо подходят для проведения диагностики SARS-CoV-2. Предпринимаемый контроль обеспечивает качество диагностического процесса и проводится таким же образом, как и клинические пробы [50, 51].

Несмотря на то, что метод является одним из самых эффективных, он

имеет свои недостатки. Из-за того, что SARS-CoV-2 содержит РНК-содержащий вирус, он легко поддается мутации, которая может привести к изменению структуры матрикса или даже к изменениям в консервативных участках генома [35, p. 116250]. В результате изменений в последовательности матричной цепи, происходит несовпадение с последовательностью зонда, что приводит к ложным результатам количественных ПЦР, которые требуют постоянной коррекции зонда [35, p. 116250; 48, p. 247-248; 52]. Изменения в консервативных регионах могут снизить аффинность связывания матрицы с праймером, что способствует низкой эффективности реакции [35, p. 116250; 48, p. 247-248]. В связи с высокой чувствительностью ПЦР-анализа, он становится уязвимым для перекрестного загрязнения, что может привести к ложноположительным результатам [35, p. 116250; 49, p. 104215]. На каждом этапе исследования возможна контаминация образцов. Например, при сборе и транспортировке проб, в процессе выделения генетического материала, а также на этапах очистки генома или после проведения ПЦР-дезинфекции с помощью аэрозоля очищенного генома или продуктов его обработки [35, p. 116250]. Самое масштабное исследование по тестированию на коронавирусы оценило частоту ложноотрицательных результатов в 41% при диагностических тестах ОТ-ПЦР, используемых в Китае [53].

В целях обнаружения коронавируса был оценен еще один метод, основанный на амплификации нуклеиновых кислот. Он получил название петлевая изотермическая амплификация нуклеотидов (LAMP). Это высокоточный и быстрый метод исследования ДНК, который может обнаружить только шесть копий целевой последовательности в пределах одной выборки, при условии, что температура на момент проведения анализа находится в пределах стабильной [54]. Применяя данный метод, можно получить от двух до трех разных пар праймеров, которые используются для последовательной амплификации мишени. Это приводит к накоплению 109 копий ДНК-мишени менее чем за час. Установлено, что в качестве конечных продуктов выступают стержне-петлевые ДНК, состоящих из нескольких инвертированных последовательностей мишени, которые образуют структурный ампликон с консистенцией, подобной цветной капусте [54, p. e63-1-e63-6]. Обратная транскриптаза LAMP (ОТ-LAMP) – это диагностический метод, который может быть использован для проведения ПЦР в реальном времени, и его можно использовать в мобильных устройствах, что значительно увеличивает скорость и упрощает проведение диагностики SARS-CoV-2 [55, 56]. Отличие ОТ-ПЦР от ОТ-LAMP заключается в том, что последний не требует изменения температурных параметров и более устойчив к ингибиторам, связанным с образцом, которые могут влиять на эффективность метода [55, p. 22727-22734; 57, 58]. В ОТ-LAMP есть возможность проведения диагностики на месте, в том числе и тестирование образцов слюны, всего за 1-2 часа, что значительно сокращает время исследования генетического материала [59].

Вместе с тем, в связи с непредсказуемостью результатов, которые могут

быть получены даже при использовании специфических и правильных (*in silico*) праймеров LAMP, перед их использованием для диагностики пациентов необходимо провести ряд изменений в дизайне самих праймеров [60-62]. Одного только выравнивания *in silico* недостаточно, и необходима оптимизация процесса в условиях лаборатории [60, р. 2826-1-2826-9; 62, р. 729-734]. По результатам исследований, проведенных на сегодняшний день, можно сделать вывод о том, что необходимо выполнить несколько попыток создания различных вариантов дизайна праймеров, прежде чем будет найден оптимальный набор для проведения ОТ-LAMP диагностики SARS-CoV-2 [63].

Высокопроизводительное секвенирование/метагеномное секвенирование следующего поколения (mNGS) – метод, благодаря которому китайские ученые за 5 дней завершили идентификацию и анализ генома SARS-CoV-2. Идентификация вируса в тот период стала ключевым звеном в борьбе с коронавирусом [64]. Протокол, используемый в mNGS, состоит из пяти последовательных этапов: извлечение нуклеиновой кислоты, создание библиотеки, секвенирование генов, анализ биологической информации и интерпретация. При условии, что последовательность нуклеиновой кислоты имеет высокую степень гомологии с последовательностью SARS-CoV-2, можно говорить о том, что данный случай может быть подтвержден. Существует несколько ограничений на секвенирование генов, которые включают в себя длительное время выполнения, сложность процедуры, недостаточную стандартизацию и высокий уровень квалификации сотрудников, занимающихся поиском и анализом данных [65]. В настоящее время секвенирование ДНК проводится только в некоторых медицинских лабораториях [66].

Таким образом, каждая методика имеет свои достоинства и не лишена недостатков. Подводя итог, можно отметить, что метод ОТ-ПЦР имеет ряд положительных сторон. Он имеет возможность проведения в течение 3-5 часов (в зависимости от региона), дает результаты уже на первой неделе заболевания, но имеет некоторые недостатки: для проведения процедуры необходимо специальное оборудование, а также персонал, который прошел обучение. Также есть случаи получения ложных результатов. ОТ-LAMP, основанный на амплификации нуклеиновых кислот тест, применяется для обнаружения вируса у пациента в период болезни (Point-of-Care Testing), имеет относительно небольшое время обнаружения (1-2 часа), является информативным при первой неделе заболевания, но имеет сведения о ложноположительных результатах. При секвенировании гена возможно обнаружить одновременно несколько возбудителей, определить мутации штаммов вируса, идентифицировать новые виды патогенов и выявить их генетическую структуру. Необходимо иметь в наличии специализированную лабораторию, персонал с опытом работы и квалификацией, а сам анализ должен занимать длительное время (3-5д) [66, р. 172-181].

## **1.2 Серологические методы диагностики КВИ**

Чаще всего в серологической диагностике COVID-19 используется непрямой метод иммуноферментного анализа (ИФА). Самым

распространенным и часто используемым серологическим тестом является иммуноферментный анализ, который имеет средний срок обнаружения от 2 до 8 часов [67]. Его изобретение относится к периоду с 1960-х годов и широко применяется в период с 1970-х годов до настоящего времени. Основной процесс этого метода заключается в нанесении вирусного белка (N, S), субъединицы белка S (S1) или домена белка (RBD) на твердофазный носитель, который связывается с сывороточными антителами, связанными с ферментом, для получения хромогенной реакции. Исходя из этого принципа, исследователи создали устройства, которые могут быстро и эффективно проводить анализ в условиях пандемии. Это устройство ИФА представляет собой автоматизированную систему для определения уровня IgA, которые характерны для SARS-CoV-2 в слюне человека, и его можно использовать менее чем за 15 минут для стандартного иммуноферментного анализа [68]. Isho и соавторы разработали тест иммуноферментного анализа для обнаружения специфичных к S-белку и RBD-специфических IgG/IgA/IgM в сыворотке пациентов [69]. Тест, который разработали Mahallawi и соавторов основан на обнаружении антител против S-белка SARS-CoV-2 и субъединицы S1/S2 в сыворотке. Они обнаружили, что полноразмерный белок S показал самую сильную реактивность к IgG, в то время как S1 показал самую высокую специфичность [70]. По литературным данным в настоящее время на рынке существует множество коммерческих наборов для обнаружения SARS-CoV-2, нацеленные на различные типы антигенов и антител [71, 72].

Как и иммуноферментный анализ, хемилюминесцентный иммуноанализ (CLIA) представляет собой высокопроизводительный анализ с более высокой точностью и низким отношением сигнал/шум [73, 74]. Хемилюминесцентного иммуноанализа в основном дает результаты через 0,5-2 ч [75]. Принцип хемилюминесцентного иммуноанализа аналогичен принципу иммуноферментного анализа, где используется высокая аффинность связывания между вирусными антигенами и антителами хозяина. Однако, разница заключается в том, что хемилюминесцентный иммуноанализ использует химическую реакцию для создания светящегося химического зонда для обнаружения положительного сигнала [76].

Применительно к работе с SARS-CoV-2 использовался магнитно-хемилюминесцентный иммуноферментный анализ (MCLIA) для обнаружения антител против N- и S-антигенов SARS-CoV-2 [77]. Ма и соавт. использовали метод хемилюминесцентного иммуноанализа для выявления RBD-специфических IgA, IgM и IgG в крови пациентов с инфекцией SARS-CoV-2 с чувствительностью 98,6, 96,8 и 96,8% соответственно и специфичностью 98,1, 92,3, и 99,8% соответственно. При комбинированном обнаружении трех антител чувствительность и специфичность увеличились до 99,5 и 100% соответственно [78]. В настоящее время точность результатов теста хемилюминесцентного иммуноанализа выше, чем у других методов.

Иммуноанализ с латеральным потоком (LFIA) – недорогой, простой, быстрый и портативный тест. Этому методу мгновенной диагностики требуется всего 3-30 минут для завершения всех процессов, а также небольшое

количество образцов [79, 80]. Иммуноанализ с латеральным потоком представляет бумажную платформу, где происходит обнаружение и анализ. Взяв IgG, специфичный к S-белку, каплю цельной крови добавляют в прокладку для образца, а затем пропускают через прокладку для высвобождения конъюгата, содержащую IgG, специфичный для S-белка, и контрольный IgG (конъюгированный с окрашенной или флуоресцентной частицей). В сопровождении потока раствора окрашенные комплексы частиц с конъюгатами связываются со специфическими антителами, иммобилизованными в области обнаружения [78, р. 773-774; 81, 82]. Метод широко используется в биомедицинских, сельскохозяйственных, пищевых и экологических науках. В большинстве тестов иммуноанализа с латеральным потоком для диагностики COVID-19 используется коллоидное золото, меченное антигеном SARS-CoV-2. Другая группа исследователей сообщает об устройстве, использующем флуоресцентные наночастицы в качестве сигнала для антител против SARS-CoV-2, демонстрирующем 100% специфичность и чувствительность за счет качественного обнаружения антител IgM, IgA и IgG против SARS-CoV-2 [83]. Ни и соавторы обнаружили, что EDTA-K2 может хелатировать коллоидное золото, это может привести адсорбированию коллоидного золота на пластине для высвобождения конъюгата, что значительно повышает чувствительность и специфичность диагностики [84]. Это означает то, что данный метод может применяться в целях проведения крупномасштабной серологической и быстрой диагностики COVID-19. На фоне удобства и простоты в использовании, тест имеет склонность давать ложные результаты при сравнении с другими методами тестирования [85, 86].

Дополнительно, существует проблема серологических тестов – это перекрестная реактивность, что приводит к ложноположительным результатам. Кроме того, выявление зараженных на ранних стадиях заболевания проблематично из-за необходимости времени для образования антител, однако, метод годен для пост-вакцинированного анализа COVID-19.

### **1.3 Опыт применения MALDI-TOF MS для диагностики вирусов**

Общие лабораторные методы обнаружения вирусов разделяют на прямые – электронная микроскопия, культивирование вирусов с дальнейшей детекцией, молекулярно-генетические методы исследования - ПЦР, секвенирование и косвенные – реакция торможения гемагглютинации (РТГА), реакцию связывания комплемента (РСК), реакция пассивной гемагглютинации (РПГА) и реакция обратной пассивной гемагглютинации (РОПГА), различные варианты ИФА [31, р. 84-89; 32, с. 97-104].

В настоящий период MALDI-TOF MS применяется для обнаружения вирусов только в исследовательских целях, а вопрос о создании базы данных вирусов пока находится в стадии обсуждения. Ключевым фактором в ранней диагностике и лечении вирусных инфекций является их быстрое распознавание, что позволит проводить прицельную терапию и профилактику, а также борьбу со вспышками и эпидемиями [31, р. 84-89; 32, с. 97-104].

К сожалению, вирусы ОРВИ и грипп продолжают представлять собой

серьезную угрозу для здоровья человека. На необходимость в средствах для быстрой и точной диагностики, позволяющей быстро и эффективно выявить различные виды вирусов, указывает прошедшая около 10 лет назад межвидовая передача вируса птичьего гриппа человеку и вспышка варианта гриппа H1N1.

Schwahn и соавторы провели протеотипирование вируса гриппа с обнаружением пептидных отпечатков триптических гидролизатов целого вируса. Им удалось дифференцировать уникальный для вакцины штамм H1N1 от сезонных штаммов типа А (H1N1), что свидетельствует о высокой чувствительности метода при типировании и субтипировании вируса [35, p. 116250; 87].

Проект Majchrzykiewicz-Koehorst и соавторов, был направлен на разработку нового способа идентификации вируса гриппа А с помощью MALDI-TOF MS. Вирус выделялся и культивировался в течение нескольких дней после сбора назофарингиальных мазков и трахеальных смывов, затем он был заморожен и хранился в замороженном виде. После того, как пептиды были извлечены из организма, их подвергли MALDI-TOF MS и анализировали с помощью программного обеспечения MASCOT. После проведения экспериментов все штаммы вируса гриппа А (восемь H1N1 и два H3N2) были идентифицированы, благодаря идентификации пептидов, которые были получены из нуклеопротеиновой ткани, не только по их типу, но и по подтипу. Процедура была выполнена в течение 3 ч. Все действия, включая анализ данных и получение результатов, были выполнены в течение этого времени. Таким образом, авторы предложили новый метод подготовки образцов для идентификации вируса на основе масс-спектрометрии, как в монокультуре, так и в смешанных образцах [32, с. 97-104; 88].

Musajі и соавторы продемонстрировали, что вирус гриппа можно обнаружить, подвергая интактный вирус воздействию MALDI-TOF MS. Спектры соответствующих нейраминидазы и гемагглютинина в белках, а также во всем вирусе гриппа были использованы авторами для сравнения [35, p. 116250; 89].

Calderaro и соавторы исследовали полиовирус серотипы 1, 2, 3 на платформе MALDI-TOF масс-спектрометрии. Целью их работы была идентификация вирусов и обнаружение специфических вирусных биомаркеров [90]. В исследовании авторы использовали эталонные штаммы полиовируса Сэбина 1, 2 и 3 типов. Были использованы монослой конфлюэнтных клеток LLCMK2 (линия эпителиальных клеток почки макаки-резус LLC-MK2; ATCC CCL-7) для культивирования и титрования методом предельных разведений. Анализ спектров очищенного вириона проводился в низкомолекулярном диапазоне (2-20 кДа) и в диапазоне высоких молекулярных масс (17-90 кДа). В низкомолекулярном диапазоне было выделено два значимых пика около 3800 и 7600 m/z. Это соответствует массе белков VPg (геном-ассоциированный вирусный белок) и белков полиовируса VP4. Во втором варианте было определено наличие трех пиков: около 26 000 m/z (капсидный белок VP3), 30 000 m/z (капсидный белок VP2) и 33 000 m/z (капсидный белок VP1). Спектры были импортированы в программное обеспечение ClinPro Tools. Далее для

выявления отличительных пиков, способных дифференцировать 3 серотипа, были применены 3 контролируемые статистические алгоритмы: Supervised NeuralNetwork, GeneticAlgorithm и QuickClassifier. Затем, чтобы подтвердить валидность метода, авторы создали собственную базу данных вирусов с выше указанными серотипами и провели исследование замороженных фрагментов полиовирусов типов 1, 2 и 3 из различных клинических образцов. Проведенный анализ дал положительные результаты в идентификации полиовируса. А также дал основание полагать, что структурный белок VP4 может быть вирусным биомаркером для идентификации штаммов полиовируса на уровне серотипа [32, с. 97-104; 90, р. 6803-1-6803-9].

Еще одной глобальной проблемой здравоохранения остается вирус гепатита В (ВГВ). Болезнь имеет высокие показатели заболеваемости и смертности. Гепатит В может проходить в виде бессимптомного носительства, переходить в хроническую форму, приводить к развитию цирроза печени и гепатоцеллюлярной карциномы [91-93].

Разные группы исследователей описывали возможности масс-спектрометрии для определения ВГВ. Tian и соавторы исследовали пептидомное профилирование в процессе инфицирования ВГВ. В качестве исследуемого материала была использована сыворотка крови. Анализ данных проводился с помощью программ flexAnalysis и ClinProTools 2.1 (Bruker Daltonics, GmbH, Бремен, Германия). Разработанная модель состояла из двух пептидов 2882,89 m/z и 4476,12 m/z. Разработка позволила отличить инфицированных ВГВ от здоровых людей с чувствительностью 95,5% и специфичностью 95,4% при проведении анализа перекрестной валидации. Анализ пептидомной модели, а также перекрестная валидация проводилась с помощью статистических алгоритмов GeneticAlgorithm, Supervised NeuralNetwork и QuickClassifier [32, с. 97-104; 91, р. 56-61].

Сао и соавторы в своей работе, используя MALDI-TOF MS, обнаружили новые специфические сывороточные биомаркеры для диагностики цирроза печени. В результате авторы предложили классификаторы для прогнозирования цирроза печени у пациентов с ВГВ. В исследовании участвовали больные хроническим гепатитом В, с циррозом печени и здоровые добровольцы. После проведения масс-спектрометрии на MALDI-TOF MS Autoflex (Bruker Daltonics, Германия) обработка и анализ данных проводился на flexAnalysis (Bruker Daltonics, Германия), ClinProtTools 2.2 (Bruker Daltonik, Германия), и Waikato Environment for Knowledge Analysis (WEKA). Для разработки классификаторов использовалось машинное обучение (NB, MLP, SVM, DT, CART). Более высокую точность при определении больных с циррозом печени показали NB и MLP. Исследователи получили следующие результаты: чувствительность - 90,9%, специфичность - 94,9% и предположили, что данный подход может стать методом скрининга больных с циррозом печени, без необходимости проведения биопсии печени [32, с. 97-104; 93, р. 814876-1-814876-6].

Группа во главе с Ren исследовали полипептиды 4210 Да и 1866 Да в качестве биомаркеров сывороточного метаболита гепатопатии при вирусном гепатите В. Проводилась масс-спектрометрия сыворотки крови с помощью

MALDI-TOF MS (Clin-ToFII, BioyongTechnologyCompany). Измерялись уровни полипептидов 4210 Да и 1866 Да по пяти группам диагнозов (хронический гепатит В, цирроз печени, гепатоцеллюлярная карцинома, здоровый контроль, группа с естественным клиренсом ВГВ). Между пятью группами наблюдались значительные различия в уровнях 4210 Да и 1866 Да. Наименьшие показатели были в группе здорового контроля. Таким образом, авторы рекомендуют использовать полипептиды 4210 Да и 1866 Да для контроля лечения и прогнозирования клинических результатов возникновения и прогрессирования хронического гепатита, цирроза печени и гепатоцеллюлярной карциномы, вызванных ВГВ [32, с. 97-104; 92, р. 16982-1-16982-11].

Saí и соавторы разработали новый тест генотипирования 18 типов ВПЧ высокого риска на основе платформы MALDI-TOF с использованием рекомбинационных плазмид. В работе авторы продемонстрировали положительные результаты своей методики двухлучного анализа для обнаружения и типирования всех 18 типов ВПЧ высокого риска, сопоставили свои данные с результатами, полученными с помощью анализа ВПЧ 4800 Roche Cobas (Cobas) (соответствие между двумя анализами составило 82,6%). В своей работе авторы пришли к выводу, что MALDI-TOF MS является высокопроизводительным методом диагностики всех генотипов ВПЧ [32, с. 97-104; 94]. Несколько иные цели преследовали Hernández-Rosas и соавторы при изучении вируса папилломы человека (ВПЧ) на основе MALDI-TOF MS. Ими было проведено генотипирование папилломавируса [32, с. 97-104; 94, р. 825-1-825-8; 95]. Была определена распространенность 24 генотипов ВПЧ у мексиканских женщин с помощью времяпролетной масс-спектрометрии с лазерной десорбцией/ионизацией. Авторы определили факторы риска ВПЧ, а также пришли к выводу, что распространенность генотипов ВПЧ высока даже при нормальных цитологических данных. Основываясь на полученных данных, исследователи предложили проведение скрининга ДНК ВПЧ для поддержания эпидемиологического надзора за генотипом ВПЧ и профилактики заболевания [32, с. 97-104; 94, р. 825-1-825-8; 95, р. 48-55].

Можно резюмировать, что технология MALDI-TOF масс-спектрометрии демонстрирует высокие результаты в диагностике разных вирусов, как самостоятельно, так и совместно с методами машинного обучения.

#### **1.4 Опыт применения MALDI-TOF MS для диагностики КВИ**

В разгар пандемии COVID-19 ученые из разных стран занимались поиском и разработкой быстрых и эффективных методов диагностики SARS-CoV-2, в частности, оценивали возможности масс-спектрометрии для обнаружения вирусов разных биоматериалах.

В качестве биоматериала для работы с масс-спектрометрией использовались назофарингиальные мазки, кровь, слюна, а также мокрота и бронхолегочные смывы.

Использование крови в качестве биоматериала для исследования может продемонстрировать более достоверную диагностику, несмотря на то, что забор крови является инвазивным методом [96]. В свою очередь в опубликованных

исследованиях материалом крови рассматривались плазма [97] или сыворотка [96, p. 4782-4786; 98, 99].

Анализируя плазму крови, исследовательская группа во главе с Lazari продемонстрировала интересные результаты. Пики MALDI-TOFMS при 11 633, 11 476 и 11393 m/z соответствовали фрагментам сывороточного амилоида A2. Дополнительно, в этой работе исследователи оценивали наиболее подходящую матрицу среди  $\alpha$ -циано-4-гидроксикоричной кислоты, синапиновой кислоты и дигидробензойной. Авторы пришли к выводу, что  $\alpha$ -циано-4-гидроксикоричная кислота позволяет выявлять большее количество характерных пиков [97, p. e202000946-1-e202000946-11].

При разработке метода прогнозирования течения и исхода заболевания, Gomila и соавторы использовали сыворотку крови пациентов с различными клиническими проявлениями болезни. Выбор характерных пиков был основан на определении медианной интенсивности каждого пика для каждой группы. Это позволило увидеть значительные различия в группах между критическим и легким течением заболевания. В группе с критическим течением было выявлено 16 характерных пиков и их интенсивность была в 4 раза больше группы с легким течением. При сравнении групп с тяжелым и критическим течением таких пиков всего 5 (11 532, 4875, 2979, 3067 и 6294 m/z). Для построения модели прогнозирования течения заболевания авторы использовали машинное обучение и обнаружили зависимость профиля пептидома в сыворотке от тяжести течения COVID-19 со 100% точностью. Так, была предложена возможность прогнозирования развития заболевания у нестабильных пациентов. Согласно этому исследованию, пик 11 532 m/z соотносится с сывороточным белком амилоид A2 (SAA2), который является одним из главных маркеров тяжелого течения короновиральной инфекции. Данный факт был предложен для сортировки и оказания соответствующей помощи пациентам с короновиральной инфекцией [98, p. ofab222-1-fab22-7].

В одной из работ проводился анализ метаболитов образцов для выявления бессимптомных больных. Предварительно была проведена экстракция сыворотки. Для получения результатов использовали комбинацию методов машинного обучения, так называемый «voting algorithm». Для выявления бессимптомных больных исследователями была создана панель биомаркеров, содержащая десять метаболитов и липидов (фосфолипиды и аминокислоты, пурины и нуклеозиды, глутаминовая кислота, тирозин, таурин, ксантин и уридин). Результаты были получены благодаря выбору «voting algorithm» (SVM, XGB и MLP), и были достигнуты более высокая точность, специфичность и чувствительность (0,934, 0,945, 0,913 соответственно). Высокая производительность позволила проанализировать 384 образца в одной мишени MALDI-TOF в течение 10 минут. Авторы рекомендуют определение метаболического профиля сыворотки на основе MALDI-TOF MS в сочетании с машинным обучением как метод диагностирования пациентов с бессимптомным течением заболевания [99, p. 746134-1-746134-8].

Yan и соавторы утверждают, что образцы сыворотки могут быть использованы для быстрой диагностики, скрининга и рутинного обнаружения

вируса SARS-CoV-2 методом MALDI-TOF MS совместно с моделью машинного обучения логистической регрессии (LR). Разработка показала следующие результаты: специфичность = 100%, чувствительность = 98%, точность = 99%. Для работы было отобрано 25 характерных пиков из 386 полученных после обработки масс-спектрометрией. Валидация метода проводилась на других группах людей, сформированных специально для контрольного тестирования (больные туберкулезом, пациенты со схожими клиническими симптомами с COVID-19 и здоровые люди) [96, p. 4782-4786]. Данные, полученные в ходе исследования, показали, что метод имеет большой потенциал для проведения скрининга пациентов с COVID-19, а также проведения эпиднадзора и диагностики достаточно большого количества людей [32, с. 97-104].

Слюна как объект исследования вызывает интерес неинвазивностью забора метода и простотой получения материала [100-102]. После применения методов машинного обучения на данных характерных для пиков слюны, полученных в MALDI-TOF масс-спектрометрии, возможно отнести образец к группе с наличием вируса в слюне или к группе здоровых [102, p. 708-1-708-14]. Подобный способ подойдет для быстрой идентификации зараженных пациентов. Вместе с тем, применяя масс-спектрометрию возможно идентифицировать белки, интенсивность которых возрастает у зараженной когорты, а также оценить иммунный ответ хозяина [100, p. 1-14; 101, p. 31-40]. Chivte и соавторы приняли решение при работе со слюной выбрать диапазон массы на заряд от 2000 до 200 000. Пики были получены от пациентов с положительным ПЦР тестом на короновирус и группой здорового контроля. Показатели, которые были уникальными для людей с положительным результатом на COVID-19, присутствовали в диапазоне от 33 000 m/z до 51 000 m/z и от 65 000 m/z до 120 000 m/z [101, p. 31-40]. Для проведения измерений вирус был культивирован, затем проведенная масс спектрометрия показала следующие данные: интенсивность фрагментов S-белка S1 (79 000 m/z), S2b (62 000–72 000 m/z) и S2a (13 500 m/z). Далее при сравнении слюны группы с положительным SARS-CoV-2 и здоровых добровольцев было отмечено значительное повышение пика характерного для S1. Наличие этого пика авторы предлагают считать индикатором на COVID-19 [32, с. 97-104; 100, p. e202000946-1-e202000946-11].

Биоматериалом в поиске альтернативного либо дополнительного метода ПЦР тесту чаще всего являлись носоглоточные образцы. Использование носоглоточного мазка обусловлено тем, что метод ПЦР проводится на биоматериале с аналогичной области, поэтому более точное сравнение возможно при использовании образцов из носоглотки. Имеются работы [15, p. 1168-1172; 16, p. 661358-1-661358-11; 17, p. 8219-1-8219-9; 32, с. 97-104], где определение вируса проводилось с назальных мазков инфицированных людей. Nachtigall с соавторами провели масс-спектрометрический анализ с помощью времяпролетного MALDI-TOF (Autoflex, Bruker) с импульсным азотным лазером. Диапазон масс составлял от 3 до 20 кДа. В качестве матрицы использовалась  $\alpha$ -циано-гидроксикоричная кислота. Файлы MALDI-TOF MS с

помощью MSconvert из набора ProteoWizard были конвертированы в mzML, а затем обработаны в R MALDIquant и MALDIquantForeign. Для дальнейшей классификации вируса и контрольных образцов были выбраны алгоритмы машинного обучения (NB, SVM-L, SVM-R, KNN, DT и RF). Алгоритм SVM-R с MALDI-TOFMS продемонстрировали высокую точность (93,9%) и возможность надежного применения для обнаружения SARS-CoV-2 [32, с. 97-104]. Аналогичную работу проделали Tran и соавторы на MALDI-TOFMS (Shimadzu 8020 Shimadzu Scientific Instruments, Columbia, MD). В работе также использовалась  $\alpha$ -циано-гидроксикоричная кислота, но в самом рабочем процессе имелись отличия. Диапазон масс составлял от 2 до 20 кДа. Выделение пиков масс-спектров проводилось на программном обеспечении MALDIquant (Shimadzu Scientific Instruments, Columbia, MD). Этап классификации проходил с помощью машинного обучения на базе автоматизированной платформы Machine Intelligence Learning Optimizer (MILO) (MILO ML, LLC, Сакраменто, Калифорния). MILO определил две очень многообещающие модели машинного обучения (DNN и GBM), демонстрирующие точность 98,3%. В своем исследовании авторы продемонстрировали высокую производительность MALDI-TOFMS с улучшенным машинным обучением для тестирования на COVID-19 [17, р. 8219-1-8219-9; 32, с. 97-104]. Росса и соавторы для работы использовали назофарингиальные мазки и собрали спектральные данные между 2 до 20 кДа на MicroFlex LT версии 3.4 (Bruker Daltonics, Бремен, Германия). Анализ спектров проходил по разным подходам и затем результаты сравнивались. Была создана «внутренняя» база данных, состоящая из 9 проб COVID-19 положительных, 8 – отрицательных и 3 других респираторных вирусов. Альтернативный анализ проходил в программном обеспечении Flex Analysis v3.4 и ClinPro tools, где использовался алгоритм машинного обучения KNN [32, с. 97-104; 102, р. 708-1-708-14]. Результаты, полученные с помощью «внутренней» базы данных, соответствовали: точность = 38%, специфичность = 26,5%, чувствительность = 62,5%. В то время как результаты, полученные с помощью метода машинного обучения 37,6, 71,7, 61,7%, соответственно. Группа во главе с Deulofeu также занималась идентификацией вируса на масс-спектрометре (аппарат AutoflexmaXwithTime-Of-Flight (TOF) analyzer from Bruker Daltonics (Bruker Daltonics, Bremen, Germany)), но в качестве матрицы была использована синапиновая кислота. Для анализа были использованы алгоритмы машинного обучения XBOOST и SVM. В работе авторы особое внимание уделили влиянию транспортной среды на конечные результаты детекции вируса. После забора назофарингиальных мазков зонды помещались в транспортные среды DeltaSwab-ViCUM или DeltaSwab-Virus. В результате авторы пришли к мнению, что транспортная среда имеет влияние на достоверное определение вируса. Образцы собранные в транспортную среду DeltaSwab-ViCUM были определены по представленной методике с точностью=87%, специфичностью=96% и чувствительностью=65% [16, р. 661358-1-661358-11; 32, с. 97-104].

Абсолютно другая методология обнаружения коронавируса в носоглоточном материале представлена при оценке масс-спектрометра MALDI-

TOF MassARRAY. Было проведено секвенирование гена коронавируса для оценки возможностей MALDI-TOF MassARRAY в сравнении с методом ПЦР [30, с. 36-43; 99, р. 746134-1-746134-8; 103]. Обе работы пришли к единому мнению: время, необходимое для определения двух таргетов (ORF1ab, N) методом ПЦР, меньше, но масс-спектрометр может определить до 5 мишеней (ORF1ab, ORF1, N1, N2, N3). В процессе определения вируса возможно определить и его мутацию. Это является важным преимуществом масс-спектрометра. Группа ученых также с целью определения мутация представила разработку, названную технология минисеквенирования мПЦР-МС, объединив ПЦР и MALDI-TOF. В качестве биоматериала использовались мокрота, мазок из носоглотки и бронхолегочные смывы [29, р. 198-206; 104].

Rashidi и соавторы в своей работе провели оценку платформ разных платформ машинного обучения (Machine Intelligence Learning Optimizer [MILO], MILO ML, LLC, Sacramento, CA и Microsoft ML.NET, Microsoft, Redmond, WA) для скрининга КВИ. В своей работе авторы отмечают необходимость качественных входных данных при работе с машинным обучением, но особенно важна достоверность самих алгоритмов машинного обучения [105].

Согласно изученной литературе в будущем возможно использование масс-спектрометрии в качестве альтернативного метода ОТ-ПЦР. Обязательным условием для широкого использования метода является указание единого протокола работы [103, р. 113991]. Среди недостатков метода можно отметить отсутствие единой базы данных для идентификации вирусов, некоторые научные группы создают свои для обработки материала [106, 107]. Процесс подготовки композиций культуры также нуждается в стандартизации: выбор матрицы и транспортной среды влияет на эффективность обнаружения [16, р. 661358-1661358-11; 35, р. 116250]. Тем не менее ученые предсказывают широкое использование MALDI-TOF масс-спектрометрии для выявления SARS-CoV-2, после устранения выше изложенных препятствий и стандартизации рабочего процесса [35, р. 116250].

Известно, что в условиях пандемии, когда важно проводить диагностики в максимально короткие сроки выбор метода играет ключевую роль. Таким образом, после проведения анализа данных, можно заключить, что возможно применение время-пролетной масс-спектрометрии для крупномасштабного тестирования/скрининга населения в развивающихся странах, где ограничены возможности лабораторной службы. Технология уже зарекомендовала себя в области диагностики бактерий, ее потенциал для быстрой и точной диагностики вирусов описывается разными исследователями. После устранения недостатков, стандартизации протокола работы возможно повсеместное внедрение технологии в диагностическую практику в клинической вирусологии.

## 2 МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Для решения поставленных задач были обследованы 203 пациента с симптомами ОРВИ, включая COVID-19, а также 101 условно-здоровый участник. Забор назальных мазков проводился в период с 25 мая 2020 года по 20 декабря 2023 года. Письменное согласие было получено от всех участников в присутствии координатора исследования. Диагноз ОРВИ был поставлен на основе наличия хотя бы одного из следующих симптомов: жар, заложенность носа, кашель, боль в горле и/или лимфаденопатия. Участники были рекрутированы среди пациентов Карагандинской областной клинической больницы и Городского Центра первичной медико-санитарной помощи г. Караганды. Лабораторная часть проводилась на базе Научно-исследовательской лаборатории НАО МУК. Работа проводилась в соответствии с этическими протоколами, утвержденными комитетом по биоэтике НАО «Карагандинский медицинский университет» №45 от 06.04.2020 г., №20 от 19.05.2021 г.

Группой 1 явились пациенты с симптомами острой респираторной инфекцией, обратившиеся на фильтр поликлиник, а также пациенты, с COVID-19 инфекционных госпиталей г. Караганда (SARS-CoV-2+). Забор материала осуществлялся при поступлении в стационар (на фильтр). Клинический осмотр, постановка диагноза, первичное лабораторное исследование в рамках клинического протокола, лечение осуществлялось врачами инфекционного госпиталя, и/или палат интенсивной терапии инфекционного госпиталя, амбулаторной службой для бессимптомных пациентов в соответствии с клиническим протоколом (таблица 1).

Таблица 1 – Критерии включения и исключения

| Критерии включения в основную и группу сравнения                                 | Критерии исключения для всех групп   |
|--|--|
| подтвержденная результатами ПЦР коронавирусная инфекция для основной группы      | пациенты с ВИЧ   |
| симптомы острой респираторной инфекции, неподтвержденная коронавирусная инфекция | пациенты с туберкулезом  |
| возраст 18-65 лет  | пациенты, получавшие терапию иммуноглобулинами за последние 3 месяца   |
|  | пациенты с онкологическим заболеванием   |
|  | пациенты, имеющие иммунодефицитные состояние (первичные иммунодефициты, вторичные иммунодефициты на фоне тяжелых инфекций, осложненных воспалительных/гнойных заболеваний, распространенных ожогов, тяжелых травм и оперативных вмешательств в период менее 6 месяцев) |
|  | отказ пациента или его законного представителя от участия в исследовании   |

Группа 2 представлена участниками исследования без признаков острого респираторного заболевания.

Информированное согласие на проведение исследования подписывалось пациентом или его законным представителем.

Всеми участниками исследования заполнялись анкеты, включающие в себя вопросы о хронических заболеваниях, посещаемых странах и т.д. Для проведения анкетирования участников исследования была разработана анкета и Стандартная операционная процедура (СОП) «Правила анкетирования пациентов» (Приложение В). На основе данных, полученных из анкет, была сформирована база данных, оценено количество пропущенных ответов. Полученная информация использовалась для описания групп.

Материалом исследования являлись назальные мазки пациентов с COVID-19, ОРВИ и участников контрольной группы без симптомов ОРВИ или коронавирусной инфекции. Забор назального мазка производился, согласно разработанному СОП "Забор назального мазка" (Приложение Г). Использовались тампоны из АБС-пластика, палочки с флокированным концом из нейлона или же флок-тампоны из дакрона или полиэстера. Мазок проводился путем введения тампона в носовые ходы параллельно нёбу, равноудаленно от ноздри и внешнего отверстия уха и оставался на несколько секунд для впитывания выделений. Далее тампон сразу помещали в стерильную пробирку, содержащую 1 мл транспортной среды (RNA/DNAShield, PBS). Далее образцы хранились при  $-40^{\circ}\text{C}$  в лаборатории НАО МУК.

## **2.1 Молекулярно-генетическая детекция патогенов**

Для проведения молекулярно-генетического анализа (ОТ-ПЦР) с целью детекции SARS-CoV-2 использовался метод экстракции нуклеиновых кислот на основе магнитных частиц (RealBest Sorbitus, Vector-Best, Новосибирск, Россия). ОТ-ПЦР проводилась в режиме реального времени с помощью Real-BestRNASARS- Набор CoV-2 (Новосибирск: Vector-Best), таргетными локусами согласно протоколу производителя выступали участки RdRp и N вируса SARS-CoV-2. Флюоресценция по каналу FAM отображала амплификацию внутреннего контроля пробы, в то время как по каналу ROX флюоресценция характеризовала амплификацию SARS-CoV-2.

Идентификация возбудителей ОРВИ включала выделение ДНК и РНК при помощи MagMAXViral/PathogenKit (ThermoFisherScientific, USA) и дальнейшую постановку 1-ступенчатой ОТ-ПЦР TaqPath™ 1-StepRT-qPCR, CG (ThermoFisherScientific, USA) используя следующие TaqManassay: Bordetellapertussis, Chlamydia pneumoniae, Mycoplasma pneumoniae, Influenza H3/N2, Influenza A/H1, Influenza B virus, Human enterovirus, Human respiratory syncytial virus A, Human coronavirus NL 64 (ThermoFisherScientific, USA). Также использовалась мультиплексная панель AmpliSensORVI-scrin-FL для выявления следующих патогенов: Human Adenovirus B, C, E; Human bocavirus, Human coronavirus NL 63, 229 E, HKU 1/OC43, Human metapneumovirus, Human rhinovirus; Human parainfluenzavirus 1-4,

Humanrespiratorysyncytialvirus(AmpliSens, Москва, Россия) (таблица 2, рисунок 1). Для проведения молекулярно-генетического исследования использовались образцы, которые были собраны в транспортную среду RNA/DNA Shield.

Таблица 2 – Респираторная ПЦР панель. RVSA- респираторно-синцитиальный вирус А, RVS респираторно-синцитиальный вирус

| Патоген         | Таргет                    | ID набора            |
|-----------------|---------------------------|----------------------|
| Бактерии        | Bordetella pertussis      | Ba06439623_s1        |
|                 | Chlamydomphila pneumoniae | Ba06439616_s1        |
|                 | Mycoplasma pneumoniae     | Ba06439620_s1        |
| Вирусы          | Adenovirus B,C,E          | Amlisens ORVI-screen |
|                 | Bocavirus                 | Amlisens ORVI-screen |
|                 | Coronavirus HKU1          | Amlisens ORVI-screen |
|                 | Coronavirus 229E          | Amlisens ORVI-screen |
|                 | Metapneumovirus           | Amlisens ORVI-screen |
|                 | Parainfluenza 1           | Amlisens ORVI-screen |
|                 | Parainfluenza 2           | Amlisens ORVI-screen |
|                 | Parainfluenza 3           | Amlisens ORVI-screen |
|                 | Parainfluenza 4           | Amlisens ORVI-screen |
|                 | Coronavirus OC43          | Amlisens ORVI-screen |
|                 | RSV                       | Amlisens ORVI-screen |
|                 | Covid 19                  | Vector Best          |
|                 | Influenza A/H1            | Vi99990009_po        |
|                 | Coronavirus NL63          | Vi06439673_s1        |
|                 | Influenza B virus         | Vi99990012_po        |
|                 | Human Enterovirus         | Vi06439631_s1        |
| RSVA            | Vi99990014                |                      |
| Influenza H3/N2 | Vi99990010_po             |                      |

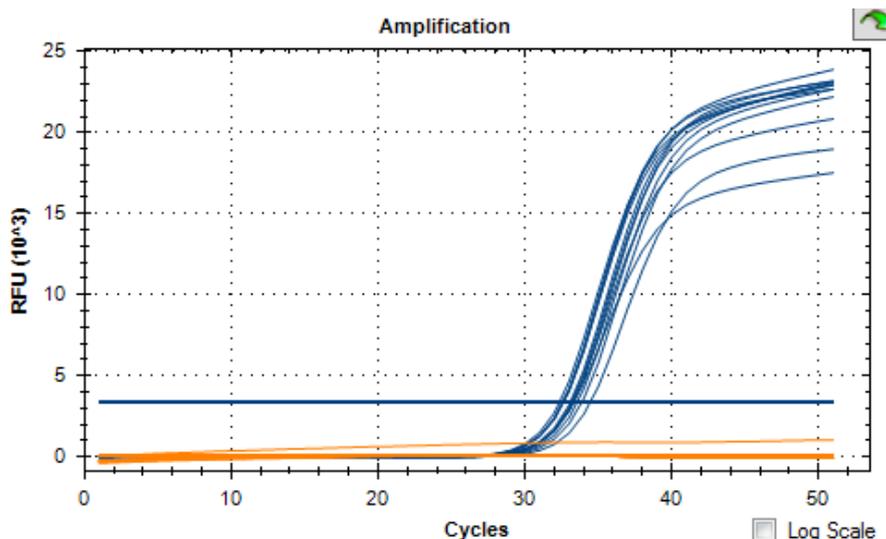


Рисунок 1 – Пример кривых амплификаций ОТ-ПЦР при тестировании назальных мазков участников контрольной группы

Примечани – Синяя линия – ькривая амплификации внутреннего контроля, оранжевая линия – кривая амплификации SARS-CoV-2

## 2.2 Проведение MALDI-TOF MS

Проведение масс-спектрометрии проводилось на приборе MicroFlex LT v. 3.4 (Bruker Daltonics, Бремен, Германия), оснащенном импульсным УФ-лазером (лазер N2 с длиной волны 337 нм, энергией импульса 150 микроджоулей, длительностью импульса 3 нс и частотой повторения 20 Гц). Образец в транспортной среде PBS оттаивался при комнатной температуре. Потом смесь наносилась на стальную мишень объемом 0,5 мкл в пятикратном повторении, высушивалась на воздухе. Затем наносился первый слой матрицы НССА (0,5 мкл). После высушивания покрывался второй слой матрицы. Раствор матрицы НССА содержит  $\alpha$ -циано-4-гидроксикоричную кислоту, разбавленную в ацетонитриле, 2,5% трифтормасляной кислотой и безнуклеазной водой. Далее после полного высыхания на воздухе мишень загружалась в прибор. Спектры генерировались путем суммирования 500 отдельных спектров (10\*50 «выстрелов») в диапазоне от 3 до 20 кДа. Спектры, полученные от каждого образца, сохранялись для дальнейшего анализа.

## 2.3 Препроцессинг спектров

Препроцессинг исходных MALDI-TOFMS (Bruker) файлов проводился в среде R (версия 4.3.0) с использованием библиотек MALDIquantForeign и MALDIquant [108]. Спектры претерпевали процедуру тримминга в диапазоне 3-15,5 кДа, затем данные были преобразованы по методу «Squareroottransform» и сглажены с помощью метода Savitzky–Golay. Коррекция базовой линии была выполнена с использованием алгоритма TopHat, нормализация интенсивности проводилась через калибровку общего ионного тока. Обнаружение пиков проводилось с отношением сигнал/шум равным 2 и с полуразмером окна, равным 10; пики были сгруппированы с допуском 0,003. Группировка пиков проводилась в два этапа для избегания дополнительных различий в калибровке. Сначала спектры каждой группы группировались отдельно, и проводилась фильтрация пиков, сохраняя только те, которые присутствовали в 80% спектров каждой группы. Затем все пики группировались вместе. Полученная матрица интенсивности пиков использовалась для последующих анализов. Пример рабочего процесса препроцессинга данных представлен на рисунке 2.

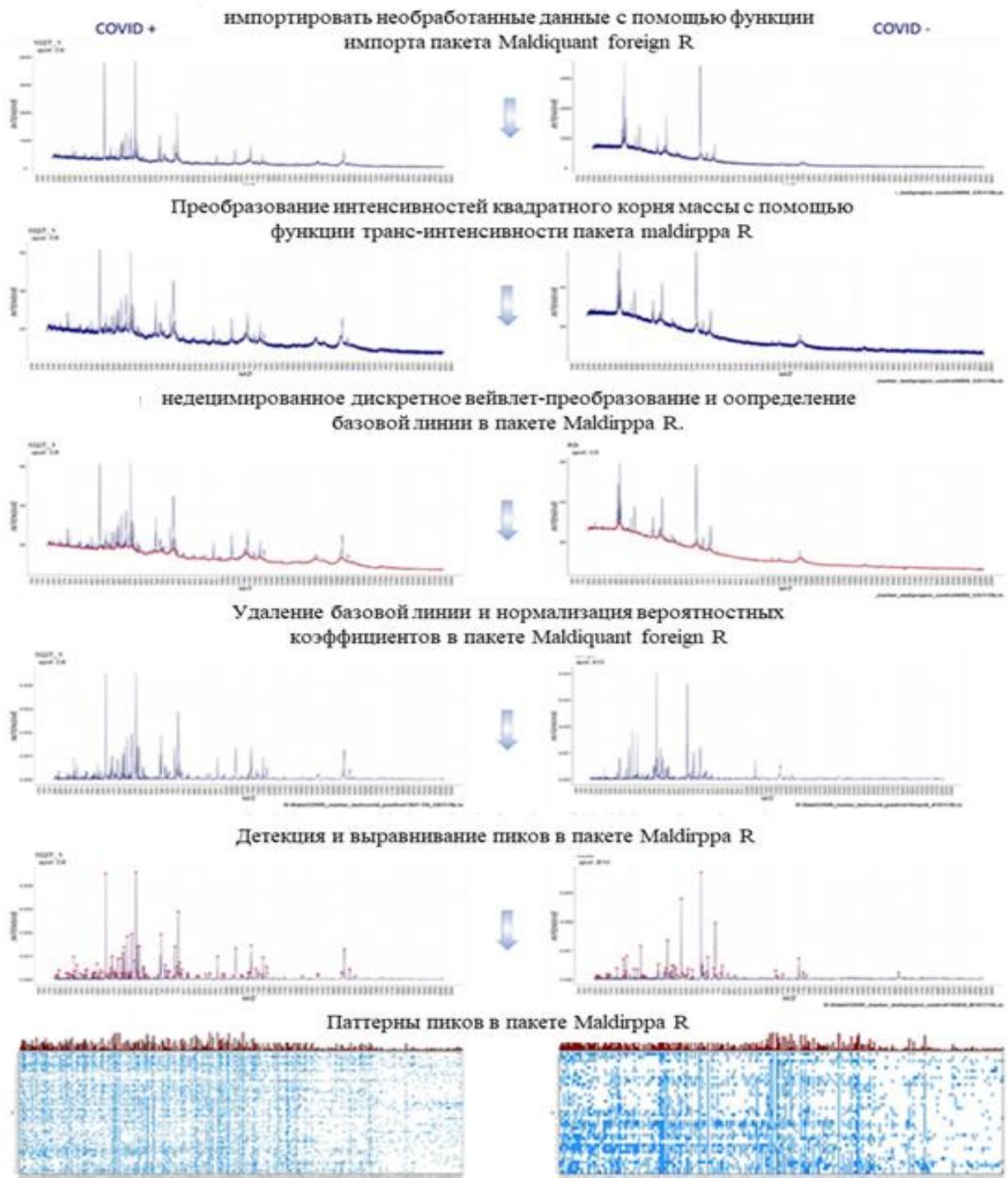


Рисунок 2 – Рабочий процесс препроцессинга спектров MALDI-TOF MS

## 2.4 Машинное обучение и статистический анализ

На начальном этапе были рассмотрены существующие методы ML и определены оптимальные алгоритмы для решения поставленной задачи.

Random Forests представляет собой мощный и гибкий метод машинного обучения, который сочетает в себе принципы бутстрепа и агрегирования для построения множества деревьев решений и их последующего ансамблирования. Этот метод является одним из наиболее популярных и эффективных методов для задач классификации и регрессии. RF работает путем создания множества деревьев решений на подмножествах обучающего набора данных, выбранных с

заменой, и использования случайных подмножеств признаков для разделения узлов. Эти индивидуальные деревья затем "голосуют" или усредняют свои прогнозы для получения окончательного результата [18; 109].

Decision Trees (DT) являются одним из наиболее популярных и широко применяемых методов в области машинного обучения. Они предоставляют графическое и структурированное представление решений, основываясь на определенных признаках, что делает их легко интерпретируемыми. Дерево решений строится рекурсивным разбиением обучающего набора данных на подмножества на основе значений признаков. Каждое разбиение представляет собой узел дерева, а исходящие из него ветви соответствуют различным значениям признака. Листовые узлы (или конечные узлы) дерева представляют собой прогнозы модели. В контексте классификации это могут быть метки классов, а в контексте регрессии – численные значения [18; 110].

Метод  $k$ -nearestneighbors(KNN) базируется на фундаментальной концепции: объекты, принадлежащие к одной категории, часто локализируются в непосредственной близости друг от друга в многомерном пространстве признаков. При наличии возможности квантификации этой пространственной близости, можно провести классификацию новых, не ранее классифицированных объектов, руководствуясь информацией о их пространственных "соседях" [18].

Несмотря на интуитивную простоту метода KNN, ему присущи ряд важных аспектов, которые требуют тщательного рассмотрения. Одним из ключевых параметров метода является число соседей, обозначаемое как  $k$ . В то время как слишком низкие значения  $k$  делают метод уязвимым для аномалий и шума, чрезмерно высокие значения могут привести к потере локальной структуры данных. Однако, благодаря его простоте, интуитивности и адаптивности, метод KNN остается привлекательным инструментом для классификационных задач, особенно на начальных этапах исследования данных [18; 111-115].

Naïve Bayes(NB) классификатор – это вероятностный метод машинного обучения, основанный на применении теоремы Байеса с "наивным" предположением о независимости признаков. Этот метод имеет долгую историю и широко используется во многих областях, особенно в задачах текстовой классификации [18].

NB классификатор предполагает, что наличие определенного признака в классе не связано с наличием какого-либо другого признака. Несмотря на то что это предположение редко выполняется в реальных данных, классификатор демонстрирует высокую эффективность во многих задачах [18].

При обучении модели вероятности классов и условные вероятности признаков при заданных классах оцениваются на основе обучающего набора данных. Для классификации нового объекта вычисляются апостериорные вероятности принадлежности объекта к каждому из классов, и объект присваивается к классу с максимальной апостериорной вероятностью [18; 116].

XGBoost обеспечивает высокую производительность и скорость, этот метод градиентного бустинга деревьев решений стал золотым стандартом в

многих задачах классификации и регрессии. XGBoost строит ансамбль деревьев решений последовательным добавлением деревьев, где каждое следующее дерево пытается исправить ошибки предыдущих. Отличительной особенностью XGBoost является его способность к градиентному бустингу, оптимизируя обобщенную линейную модель на этапе построения каждого дерева [18; 117].

SVM – это метод машинного обучения, который строит гиперплоскость или набор гиперплоскостей в многомерном пространстве для разделения классов данных. Изначально разработанный для задач бинарной классификации, SVM с течением времени был адаптирован и для многоклассовой классификации и регрессии. Основная идея SVM заключается в поиске такой гиперплоскости, которая максимизирует зазор между двумя классами. Эта гиперплоскость называется "оптимальной разделяющей гиперплоскостью", так как она стремится максимизировать расстояние до ближайших точек каждого класса, известных как опорные векторы. Чтобы расширить возможности SVM для работы с нелинейно разделимыми данными, используются различные функции. Путем преобразования исходных признаков в новое пространство, зачастую более высокой размерности, SVM способен находить сложные границы разделения [18; 118].

Вышеперечисленные алгоритмы ML были использованы для создания предсказывающих моделей.

Было реализовано семь алгоритмов машинного обучения, шесть из которых использовались в предыдущем исследовании [15, p. 1168-1172]: NB (Naive Bayes), DT (Decision Tree - Quinlan's C5.0 algorithm), RF (Random Forest), KNN (k-Nearest Neighbors), SVM-L (Support Vector Machine with linear kernel), SVM-R (Support Vector Machine with radial basis function kernel) и дополнительные методы XGBoost (eXtreme Gradient Boosting). У алгоритма XGBoost есть преимущества перед другими моделями (гибкость, точность, устойчивость к переобучению), поэтому с целью получения лучших результатов алгоритм XGBoost был включен в исследование. Для оценки воспроизводимости анализ 1 был выполнен *in silico* на данных SARS-CoV-2+ и ОРИ из Южной Америки (данные были в свободном доступе) в соответствии с ранее использованным протоколом [15, p. 1168-1172; 18]. Второй этап исследования включал собственное обучение моделей ML на доступных данных из Южной Америки и Казахстана (полученные нашей исследовательской группой), таким образом, данные состояли из трех подгрупп, а именно образцы контрольной группы в дополнение к SARS-CoV-2+ и ОРИ [18].

Первоначально весь набор данных был разделен на две подгруппы: обучающую выборку, состоящую из 80% образцов, и тестовую группу, которая составляла оставшиеся 20%. Процесс обучения был проведен в соответствии с работой Nachtigall et al. [15, p. 1168-1172], используя пятикратную (внешнюю) вложенную перекрестную (пять раз) десятикратную (внутреннюю) кросс-валидацию со случайным стратифицированным разбиением. Для оптимизации производительности каждого алгоритма было протестировано 20 комбинаций гиперпараметров во внутреннем цикле подхода кросс-валидации, используя

метод случайного поиска. Этот процесс был повторен 20 раз для обеспечения надежности и устойчивости модели. Наилучшие модели были выбраны на основе площади под кривой (ROC AUC), которая является общим показателем для оценки производительности моделей бинарной классификации, с использованием пакета *Caret* в среде R [18]. ROC-анализ обеспечивает наиболее полное доступное описание диагностической точности, поскольку он оценивает и представляет все комбинации чувствительности и специфичности, которые может обеспечить диагностический тест. ROC-кривая наиболее полезна на ранних стадиях оценки нового диагностического теста. Полученная суммарная мера точности, такая как площадь под кривой (AUC), определяет присущую тесту способность различать больных и здоровых людей, является общей мерой способности теста различать наличие или отсутствие определённого состояния [119]. Кроме того, оценка производительности модели проводилась с использованием других метрик классификации, включая F-меру, полноту, точность, специфичность, чувствительность, а также положительную и отрицательную прогностическую ценность в пакете *yardstick* в среде R [18], различия между подгруппами оценивались в среде R с помощью критерия Краскела-Уоллиса, непараметрического статистического теста, используемого для сравнения средних значений трех или более независимых выборок.

Статистический анализ для оценки характеристики групп проводился в среде R (*DataExplorer: AutomateDataExplorationandTreatment. R package version 0.8, Tools for Data Diagnosis, Exploration, Transformation. R package version 0.5.0*). Оценка распределения данных проводилась на основе критерия Shapiro-Wilks, Q-Q plot [18].

С целью обнаружения вероятных группирующих факторов пациентов был проведен анализ FAMD. Факторный анализ смешанных данных – это метод главных компонент, предназначенный для анализа набора данных, содержащего как количественные, так и качественные переменные. Это дает возможность анализировать сходство между наблюдениями, принимая во внимание смешанные типы переменных. Кроме того, можно исследовать связь между всеми переменными, как количественными, так и качественными. Алгоритм FAMD можно условно рассматривать как смесь анализа главных компонент и анализа множественных соответствий [18].

С целью увеличения объема выборки, а также оценки способности оригинальных моделей, описанных Nachtigall [15, p. 1168-1172] предсказывать наличие SARS-CoV-2 в назальных мазках, собранных в Казахстане были использованы скрипты на языке R, полученные от Южноамериканских коллег, а также данные, находящиеся в свободном доступе. Данные и код хранятся в репозитории GitHub по ссылке [https://github.com/dimbage/ML\\_MALDI-TOF\\_SARS-CoV-2](https://github.com/dimbage/ML_MALDI-TOF_SARS-CoV-2) [18].

### 3 РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

#### 3.1 Характеристика группы 1 по данным анкетирования

В исследуемой группе проанкетировано 203 пациента. Возрастные критерии включения ограничивались диапазоном 18-65 лет.

Описательная характеристика показателя “Возраст” для этой группы представлена в таблице 3.

Таблица 3 – Описательная характеристика показателя “Возраст” в группе 1

| Признак              | Количество | Среднее | Стандартное отклонение | Квартиль 1 | Квартиль 2 | Квартиль 3 | Квартиль 4 |
|----------------------|------------|---------|------------------------|------------|------------|------------|------------|
| Возраст (полных лет) | 182        | 48,8    | 13,1                   | 39,0       | 50,5       | 60,0       | 65         |

В группе 1 среднее значение для показателя “Возраст” составило 48,8 лет со стандартным отклонением 13,1. При анкетировании наблюдалось 10,34% пропущенных значений. Графическое распределение показателя “Возраст” в группе 1 показано на рисунке 3 [18].

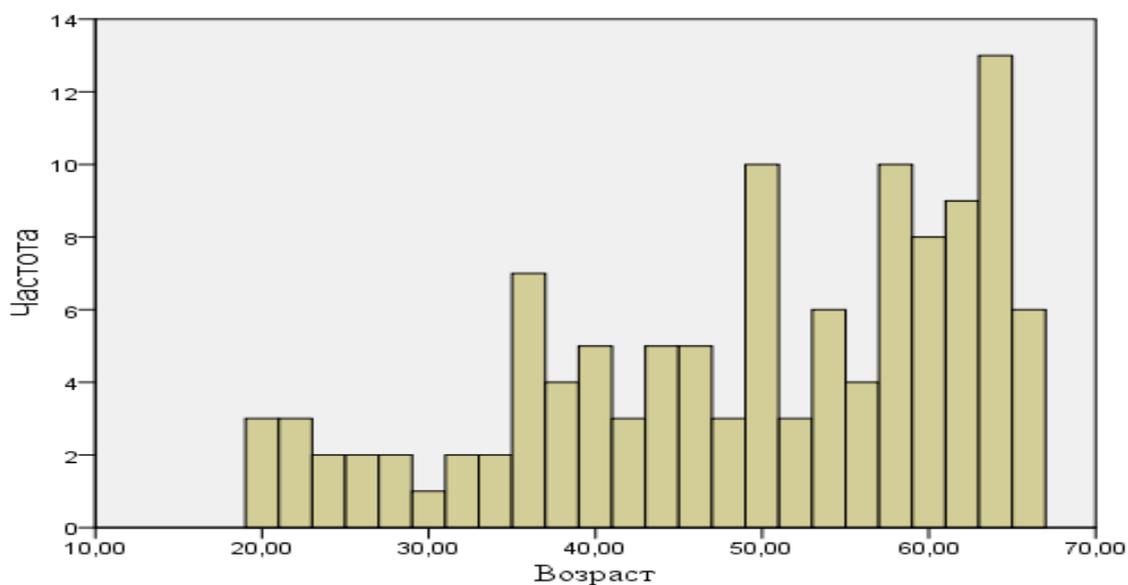


Рисунок 3 – Распределение показателя «Возраст» в группе 1

Результаты анализа качественных данных для группы 1 представлены в таблице 4.

Таблица 4 – Результаты анализа качественных данных в группе 1

| Признаки  | Значение          | Абсолютное количество (человек) | Процент (%) |
|---|-------------------|---------------------------------|-------------|
| 1   | 2                 | 3                               | 4           |
| Пол   | Женский           | 114                             | 56,16       |
|   | Мужской           | 88                              | 43,35       |
|   | Не желаю сообщать | 1                               | 0,49        |
| Курение   | Нет               | 188                             | 92,61       |
|   | Да                | 15                              | 7,39        |
| Беременность  | Нет               | 200                             | 98,52       |
|   | Да                | 3                               | 1,48        |
| Иммунодефицит, включая ВИЧ  | Нет               | 200                             | 98,52       |
|   | Не знаю           | 2                               | 0,99        |
|   | Да                | 1                               | 0,49        |
| Сердечно-сосудистые заболевания, включая гипертонию   | Нет               | 102                             | 50,25       |
|   | Да                | 99                              | 48,77       |
|   | Не знаю           | 2                               | 0,99        |
| Сахарный диабет   | Нет               | 164                             | 80,79       |
|   | Да                | 39                              | 19,21       |
| Болезнь печени  | Нет               | 181                             | 89,16       |
|   | Да                | 22                              | 10,84       |
| Болезнь почек   | Нет               | 152                             | 74,88       |
|   | Да                | 50                              | 24,63       |
|   | Не знаю           | 1                               | 0,49        |
| Хр. неврологическое или нервно-мышечное заболевание   | Нет               | 163                             | 80,30       |
|   | Да                | 40                              | 19,70       |
| Онкопатология   | Нет               | 187                             | 92,11       |
|   | Да                | 15                              | 7,39        |
|   | Не знаю           | 1                               | 0,49        |
| Хроническая болезнь легких  | Нет               | 182                             | 89,66       |
|   | Да                | 20                              | 9,85        |
|   | Не знаю           | 1                               | 0,49        |
| Наличие контакта (ов) с инфицированным  | Нет               | 109                             | 53,69       |
|   | Да                | 73                              | 35,96       |
|   | Не знаю           | 21                              | 10,34       |
| Посещение стран, регионов с неблагоприятной эпид. обстановкой за последние 30 дней до симптомов | Нет               | 201                             | 99,01       |
|   | Да                | 2                               | 0,99        |
| Температура   | Да                | 145                             | 71,43       |
|   | Нет               | 58                              | 28,57       |
| Головная боль   | Нет               | 144                             | 70,94       |
|   | Да                | 56                              | 27,59       |
|   | Не знаю           | 3                               | 1,48        |
| Боль в горле  | Нет               | 141                             | 69,46       |
|   | Да                | 58                              | 28,57       |
|   | Не знаю           | 4                               | 1,97        |



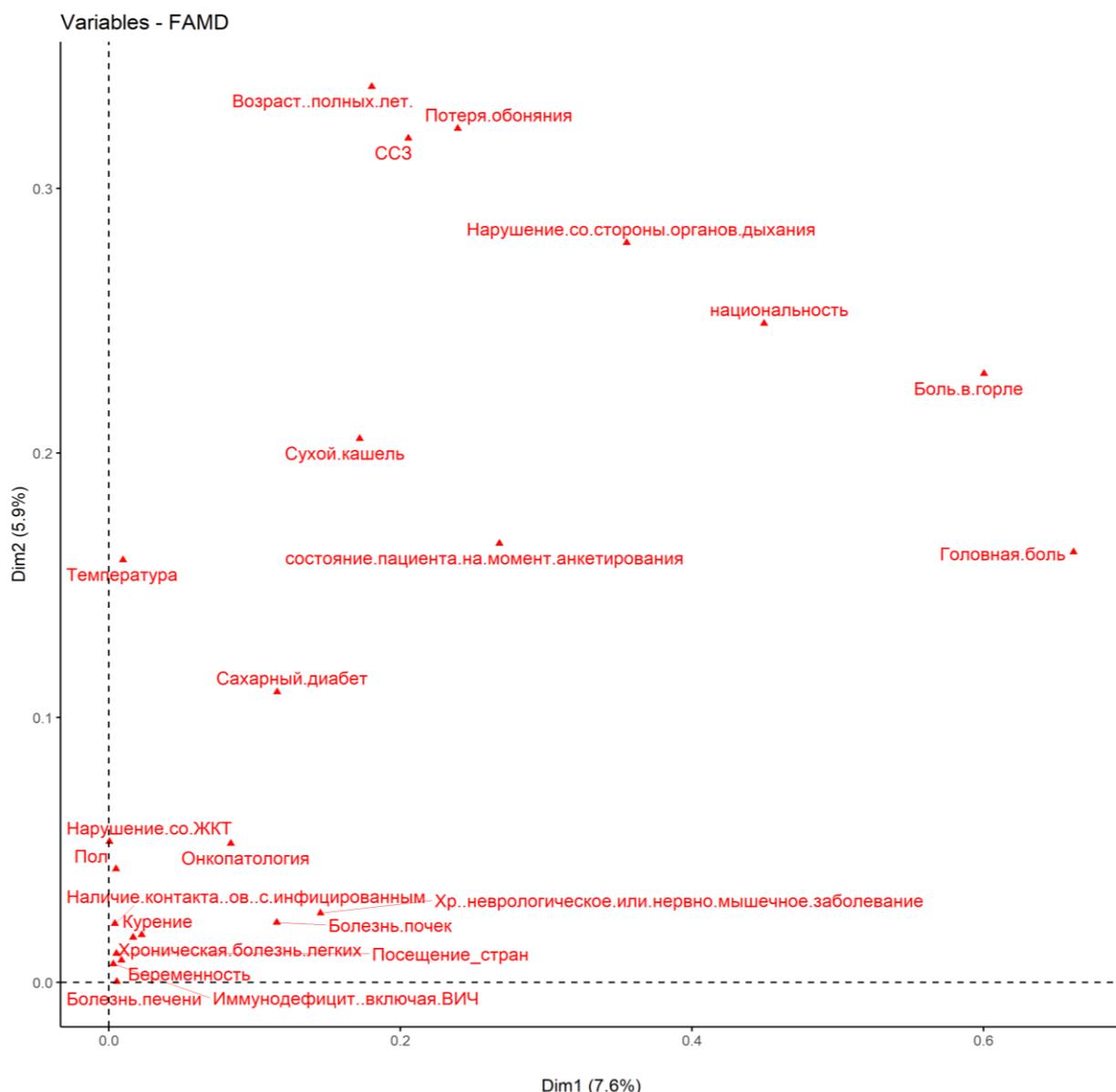


Рисунок 5 – Факторный анализ смешанных данных в группе 1

В соответствии с рисунками 4, 5, с целью обнаружения вероятных группирующих факторов пациентов в группе 1 был проведен анализ FAMD. Были выявлены топ 3 фактора с максимальным вкладом в первую компоненту: "головная боль" (рисунок 6), "боль в горле" (рисунок 7) и "этническая принадлежность (национальность)" (рисунок 8). Три фактора с максимальным вкладом во 2 компоненту – «возраст» (рисунок 9), «потеря обоняния» (рисунок 10) и «сердечно-сосудистые заболевания» (рисунок 11). Согласно этим компонентам, было проведено распределение наблюдений (людей) и их группирование [18].

FAMD анализ показал, что преобладающее большинство пациентов находятся в пределах одной группы/кластера за исключением единичных пациентов [18].

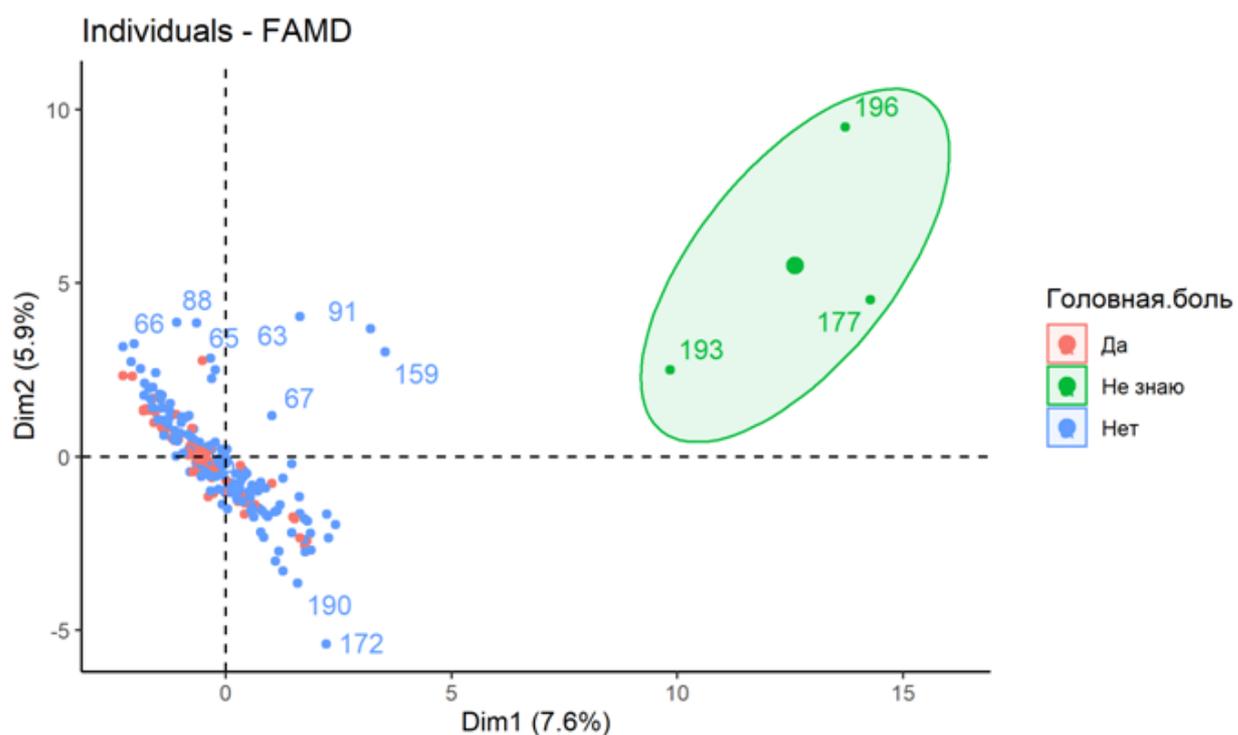


Рисунок 6 – Распределение пациентов группы 1 по показателю «головная боль»

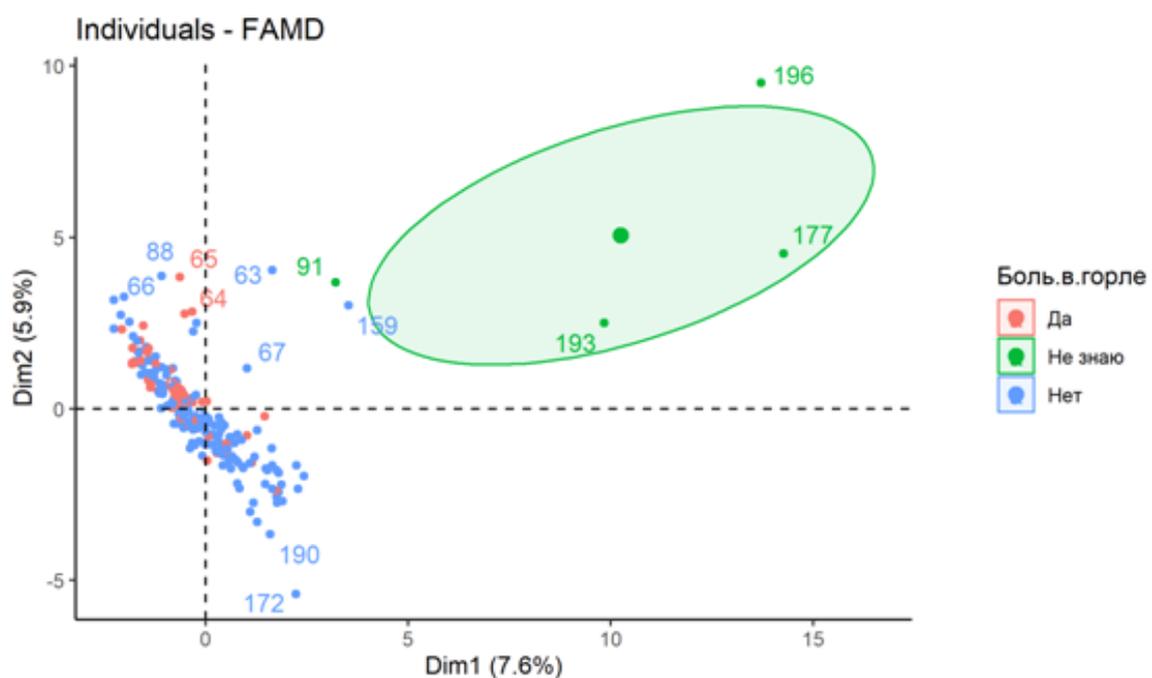


Рисунок 7 – Распределение пациентов группы 1 по показателю «боль в горле»

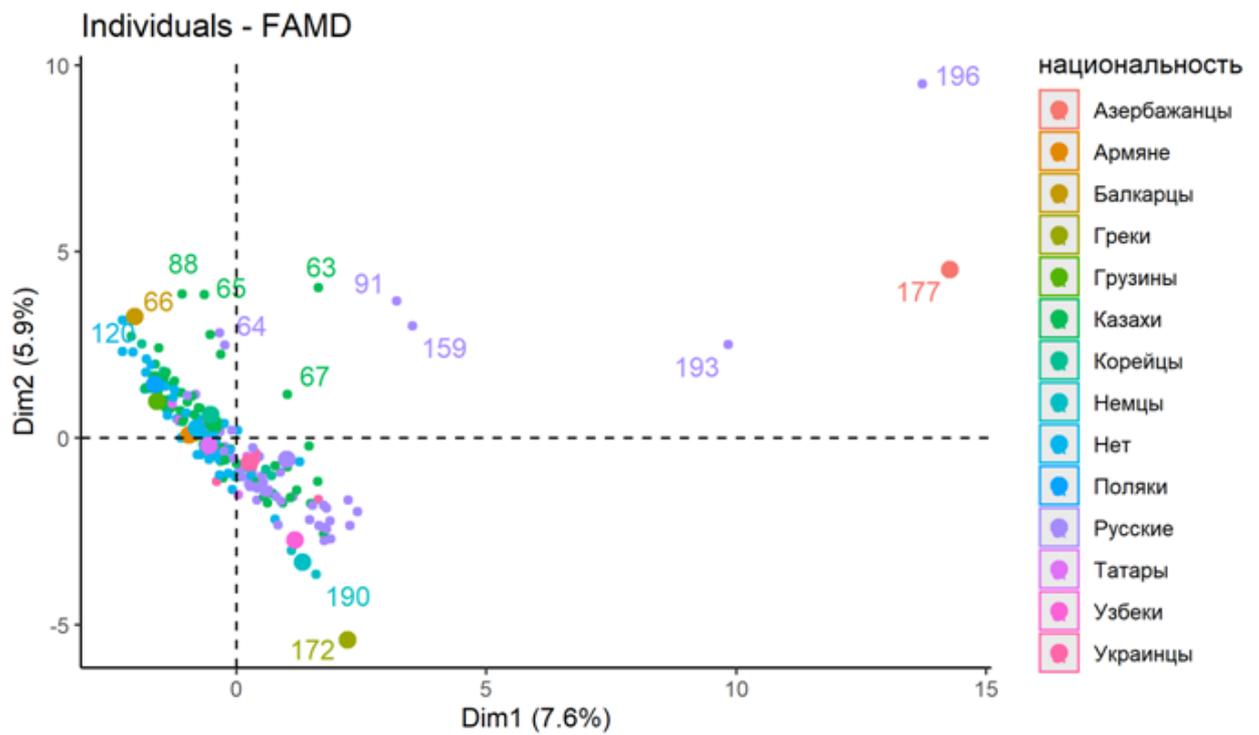


Рисунок 8– Распределение пациентов группы 1 по показателю «национальность»

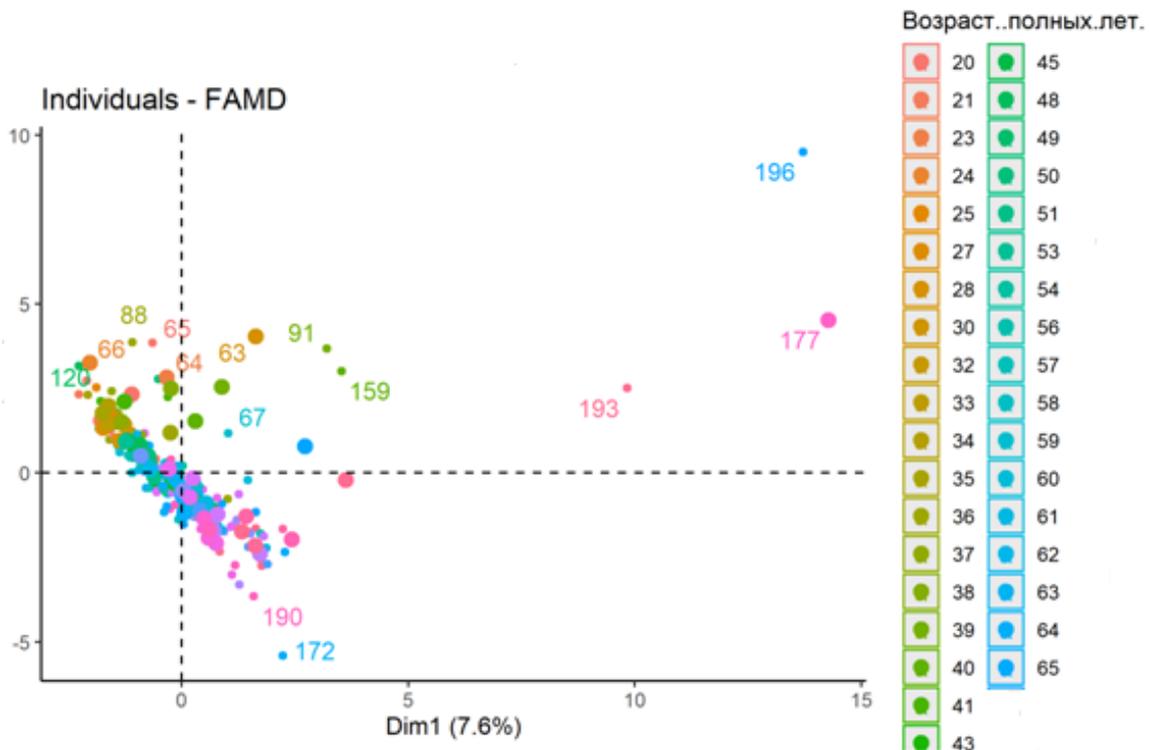


Рисунок 9 – Распределение пациентов группы 1 по показателю «возраст»

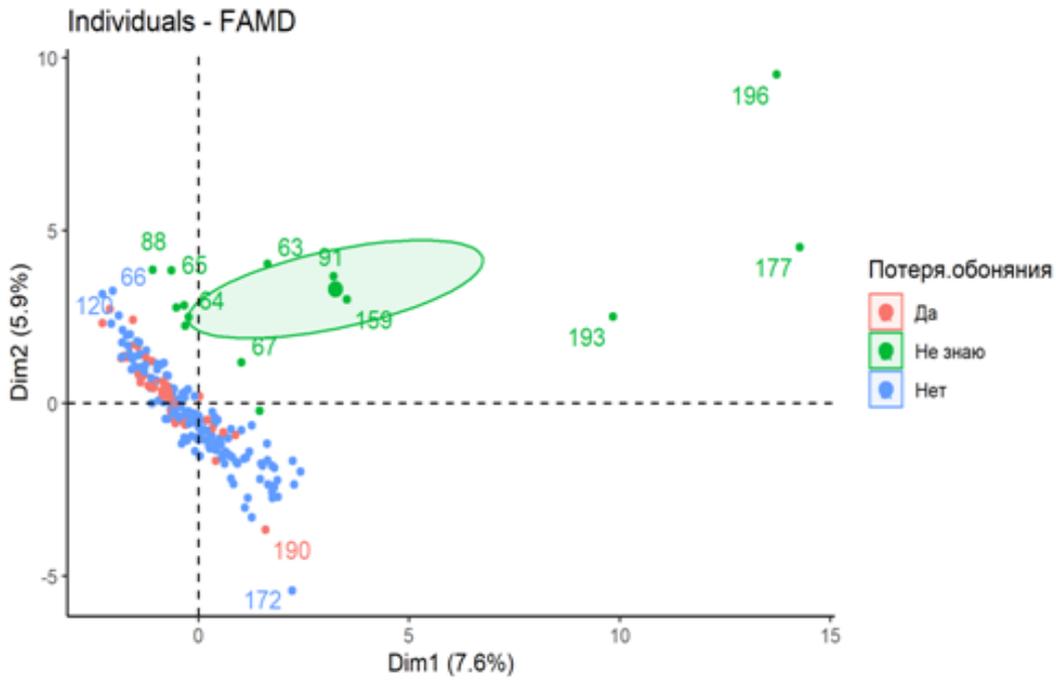


Рисунок 10 – Распределение пациентов группы 1 по показателю «потеря обоняния»

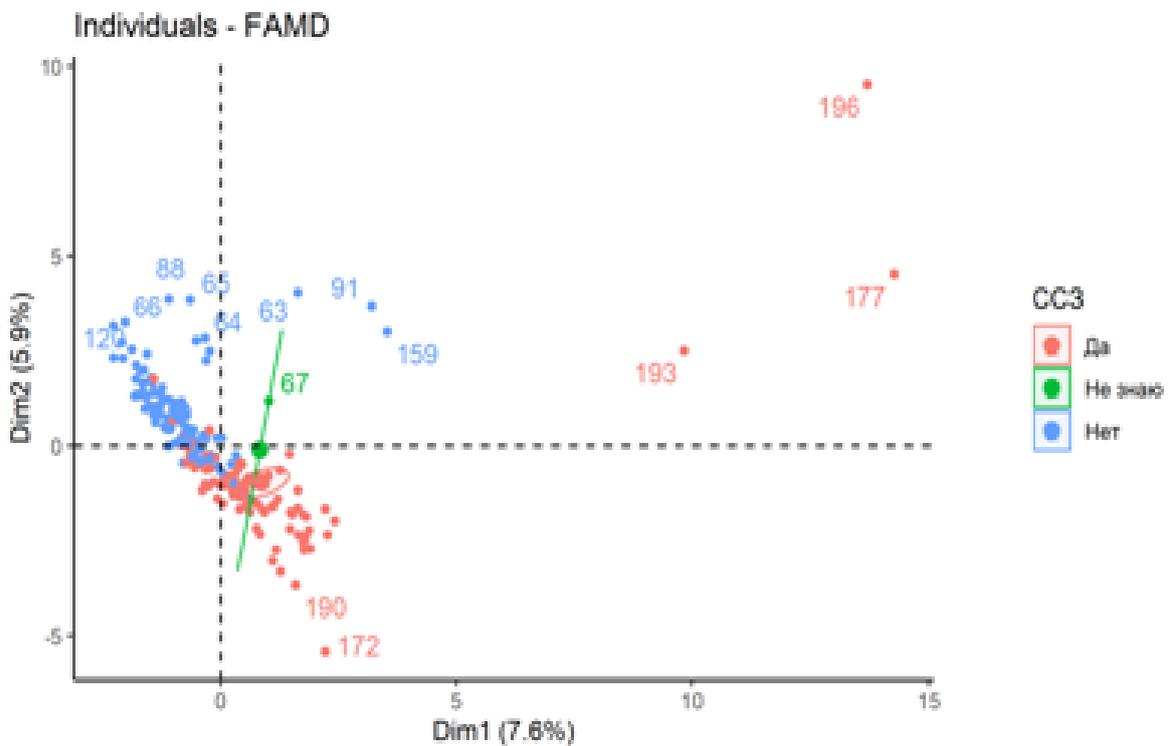


Рисунок 11 – Распределение пациентов группы 1 по показателю «сердечно-сосудистые заболевания»

### 3.2 Характеристика Группы 2

В исследуемой группе проанкетирован 101 пациент. Возрастные критерии включения ограничивались диапазоном 18-65 лет. Описательная характеристика показателя “Возраст” для группы 2 представлена в таблице 5 [18].

Таблица 5 – Описательная характеристика показателя “Возраст” в группе 2

| Признак              | Количество | Среднее | Стандартное отклонение | Квартиль 1 | Квартиль 2 | Квартиль 3 | Квартиль 4 |
|----------------------|------------|---------|------------------------|------------|------------|------------|------------|
| Возраст (полных лет) | 101        | 42,3    | 13,3                   | 32,2       | 41,5       | 57,5       | 65         |

В группе 2 среднее значение для показателя “Возраст” составило 42,3 лет со стандартным отклонением 13,3. При анкетировании пропущенных значений не наблюдалось. Графическое распределение показателя “Возраст” в группе 2 показано на рисунке 12 [18].

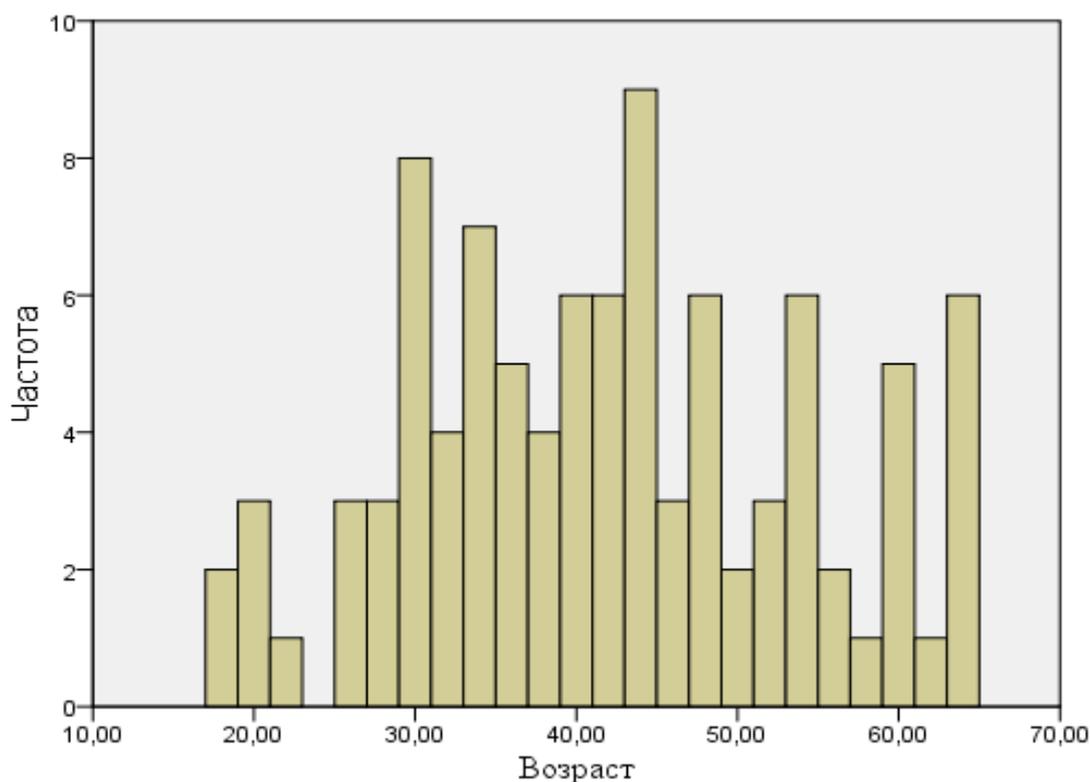


Рисунок 12 – Распределение показателя «Возраст» в группе 2

Результаты анализа качественных данных группы 2 приведены в таблице 6. С целью обнаружения вероятных группирующих факторов в группе 2 был проведен анализ FAMD (рисунок 13). Группирующие факторы показаны на рисунке 14 [18].

Таблица 6 – Результаты анализа качественных данных в группе 2

| Признак  | Значение  | Абсолютное количество | %            |
|--|---|-----------------------|--------------|
| Пол  | жен   | 61                    | 60,40        |
|  | муж   | 40                    | 39,60        |
| Курение  | не указано  | 78                    | 77,23        |
|  | нет   | 20                    | 19,80        |
|  | да  | 3                     | 2,97         |
| Этническая принадлежность (национальность)                                 | казах   | 59                    | 58,42        |
|  | русский   | 30                    | 29,70        |
|  | татарин   | 5                     | 4,95         |
|  | немец   | 2                     | 1,98         |
|  | украинец  | 2                     | 1,98         |
|  | азербайджанец                                     | 1                     | 0,99         |
|  | грек<br>ингуш                                     | 1<br>1                | 0,99<br>0,99 |
| Беременность   | нет   | 101                   | 100,00       |
| Иммунодефицит, включая ВИЧ   | не указано  | 77                    | 76,24        |
|  | нет   | 24                    | 23,76        |
| Сердечно-сосудистые заболевания, включая гипертонию                        | нет   | 86                    | 85,15        |
|  | не указано  | 8                     | 7,92         |
|  | да  | 7                     | 6,93         |
| Сахарный диабет  | нет   | 93                    | 92,08        |
|  | не указано  | 8                     | 7,92         |
| Болезни ЖКТ  | нет   | 73                    | 72,28        |
|  | не указано  | 20                    | 19,80        |
|  | да  | 8                     | 7,92         |
| Болезнь почек  | нет   | 91                    | 90,10        |
|  | не указано  | 8                     | 7,92         |
|  | да  | 2                     | 1,99         |
| Заболевания опорно-двигательного аппарата                                  | нет   | 88                    | 87,13        |
|  | не указано  | 7                     | 6,93         |
|  | да  | 6                     | 5,94         |
| Онкопатология  | нет   | 90                    | 89,11        |
|  | не указано  | 8                     | 7,92         |
|  | да  | 3                     | 2,97         |
| Хроническая болезнь легких   | нет   | 92                    | 91,09        |
|  | не указано  | 8                     | 7,92         |
|  | да  | 1                     | 0,10         |
| Посещение стран, регионов с неблагоприятной эпидемиологической обстановкой | нет   | 89                    | 88,11        |
|  | не указано  | 8                     | 7,92         |
|  | да, за 2 недели до получения 1 компонента вакцины | 3                     | 2,97         |
|  | да, между получением 1 и 2 компонента вакцины     | 1                     | 0,10         |
| Перенесенная КВИ   | нет   | 66                    | 65,35        |
|  | да  | 31                    | 30,69        |
|  | не указано  | 4                     | 3,96         |



Рисунок 13 – Факторный анализ смешанных данных в группе 2

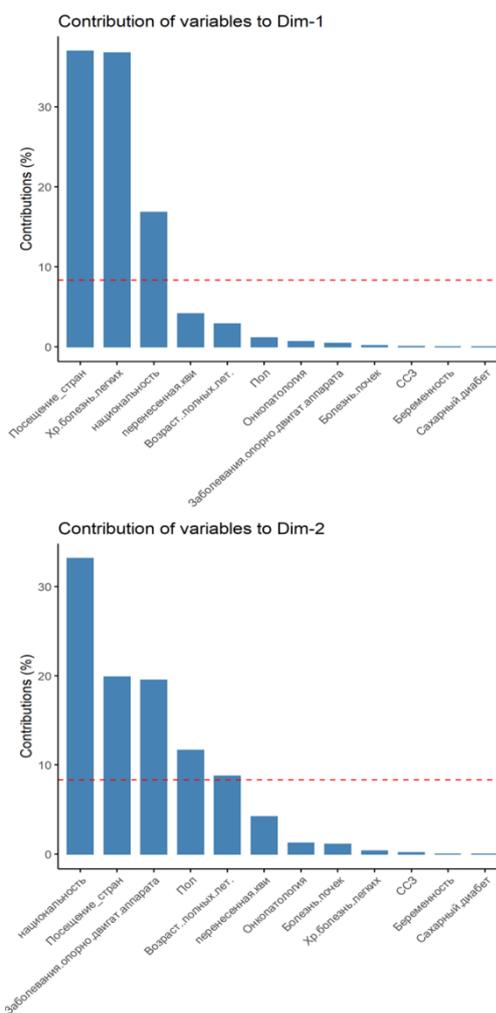


Рисунок 14 – Факторы, группирующие пациентов группы 2

Были выявлены топ 3 фактора с максимальным вкладом в первое измерение: "посещение стран" (рисунок 15), "хроническая болезнь легких" (рисунок 16) и "этническая принадлежность (национальность)" (рисунок 17). Три фактора с максимальным вкладом во второе измерение: "болезни почек", "онкопатология" и "этническая принадлежность (национальность)". Согласно этим компонентам, было проведено распределение наблюдений (людей) и их группирование [18].

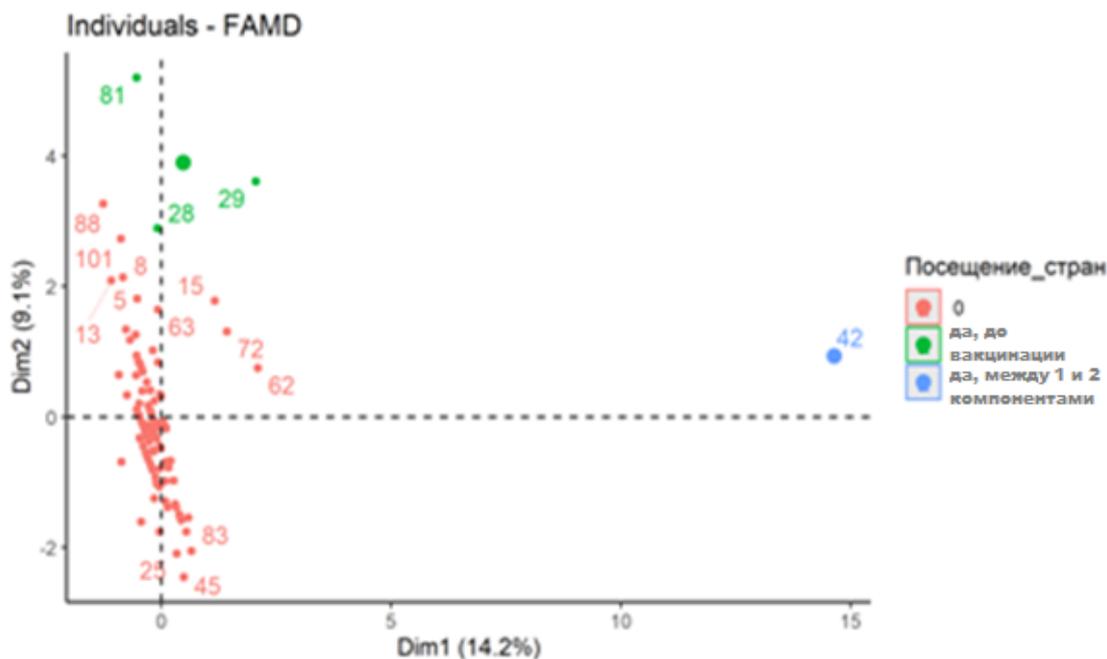


Рисунок 15 – Распределение пациентов группы 2 по показателю «посещение стран с неблагоприятной эпидемиологической обстановкой»

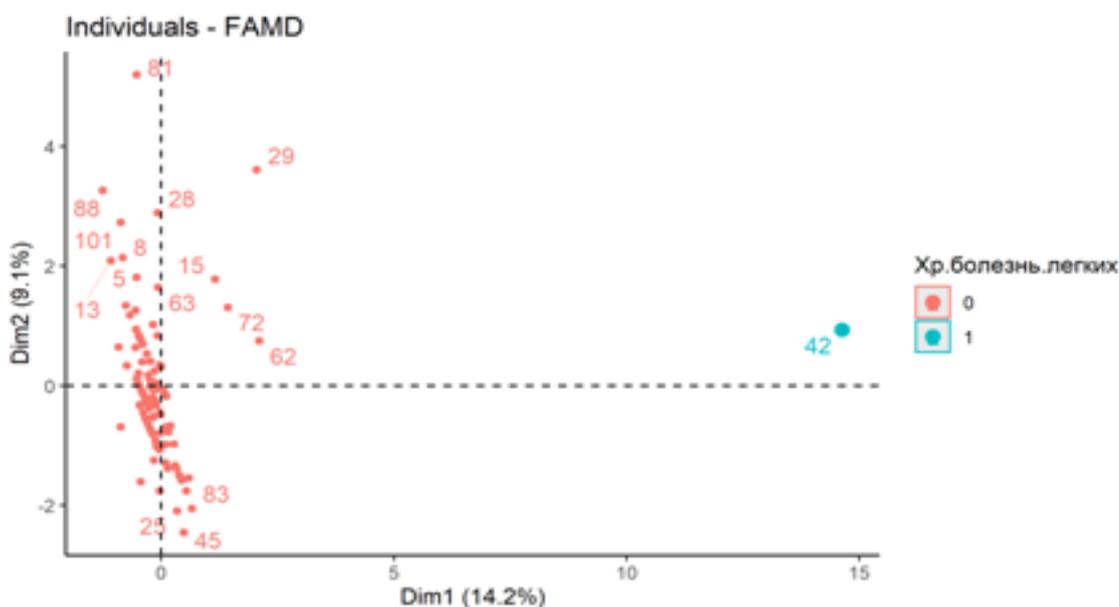


Рисунок 16 – Распределение пациентов группы 2 по показателю «хроническая болезнь легких»

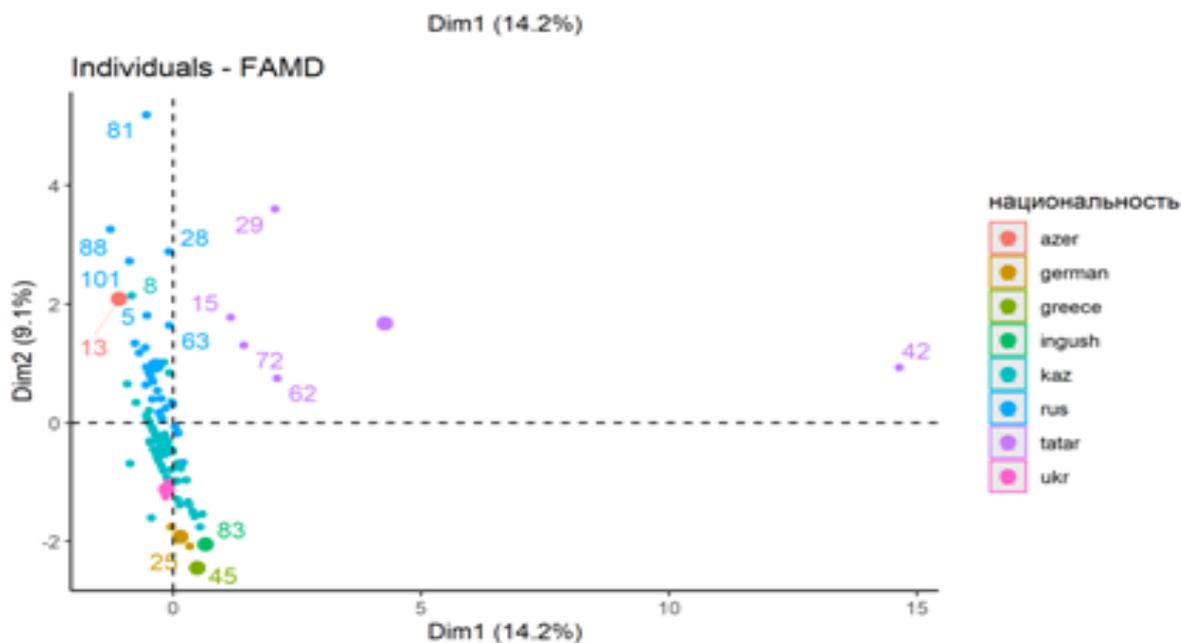


Рисунок 17 – Распределение пациентов группы 2 по показателю «этническая принадлежность (национальность)»

Согласно результатам FAMD, внутри группы 2 дополнительных кластеров выявлено не было.

### 3.3 Сравнительный анализ Группы 1 и Группы 2

Данные групп были подвергнуты сравнительному анализу для оценки сопоставимости. Полученные результаты статистического анализа представлены в таблице 7 [18]. Для сравнения групп использовался тест Манна-Уитни для количественных данных и определение хи-квадрата для качественных.

Таблица 7 – Сравнительный анализ показателей группы 1 и группы 2

| Признак              | Контрольная группа (n=101) | Основная группа (n=203) | P          | N   |
|----------------------|----------------------------|-------------------------|------------|-----|
| 1                    | 2                          | 3                       | 4          | 5   |
| Пол:                 | -                          | -                       | 0,692      | 304 |
| – женский            | 61 (60,4%)                 | 114 (56,2%)             | -          | -   |
| – мужской            | 40 (39,6%)                 | 88 (43,3%)              | -          | -   |
| – не желаю сообщать  | 0 (0,00%)                  | 1 (0,49%)               | -          | -   |
| Возраст (полных лет) | 42,0 [33,0;52,0]           | 48,8 [39,0;60,0]        | 0,001<br>* | 283 |
| Курение:             | -                          | -                       | 0,405      | 226 |
| – да                 | 3 (13,0%)                  | 15 (7,39%)              | -          | -   |
| – нет                | 20 (87,0%)                 | 188 (92,6%)             | -          | -   |
| Беременность:        | -                          | -                       | 0,553      | 304 |
| – да                 | 0 (0,00%)                  | 3 (1,48%)               | -          | -   |
| – нет                | 101 (100%)                 | 200 (98,5%)             | -          | -   |

Продолжение таблицы 7

| 1  | 2          | 3           | 4       | 5   |
|--|------------|-------------|---------|-----|
| Иммунодефицит, включая ВИЧ:  | -          | -           | 1,000   | 227 |
| – да   | 0 (0,00%)  | 1 (0,49%)   | -       | -   |
| – не знаю  | 0 (0,00%)  | 2 (0,99%)   | -       | -   |
| – нет  | 24 (100%)  | 200 (98,5%) | -       | -   |
| Сердечно-сосудистые заболевания, включая гипертонию:                               | -          | -           | <0,001* | 296 |
| – да   | 7 (7,53%)  | 99 (48,8%)  | -       | -   |
| – не знаю  | 0 (0,00%)  | 2 (0,99%)   | -       | -   |
| – нет  | 86 (92,5%) | 102 (50,2%) | -       | -   |
| Сахарный диабет:   | -          | -           | <0,001* | 296 |
| – да   | 0 (0,00%)  | 39 (19,2%)  | -       | -   |
| – нет  | 93 (100%)  | 164 (80,8%) | -       | -   |
| Болезни ЖКТ:   | -          | -           | 0,981   | 284 |
| – да   | 8 (9,88%)  | 22 (10,8%)  | -       | -   |
| – нет  | 73 (90,1%) | 181 (89,2%) | -       | -   |
| Болезнь почек:   | -          | -           | <0,001* | 296 |
| – да   | 2 (2,15%)  | 50 (24,6%)  | -       | -   |
| – не знаю  | 0 (0,00%)  | 1 (0,49%)   | -       | -   |
| – нет  | 91 (97,8%) | 152 (74,9%) | -       | -   |
| Заболевания опорно-двигательного аппарата:   | -          | -           | 0,005*  | 297 |
| – да   | 6 (6,38%)  | 40 (19,7%)  | -       | -   |
| – нет  | 88 (93,6%) | 163 (80,3%) | -       | -   |
| Онкопатология:   |            |             | 0,324   | 296 |
| – да   | 3 (3,23%)  | 15 (7,39%)  | -       | -   |
| – не знаю  | 0 (0,00%)  | 1 (0,49%)   | -       | -   |
| – нет  | 90 (96,8%) | 187 (92,1%) | -       | -   |
| Хроническая болезнь легких:  | -          | -           | 0,006*  | 296 |
| – да   | 1 (1,08%)  | 20 (9,85%)  | -       | -   |
| – не знаю  | 0 (0,00%)  | 1 (0,49%)   | -       | -   |
| – нет  | 92 (98,9%) | 182 (89,7%) | -       | -   |
| Посещение стран, регионов с неблагоприятной эпид. обстановкой за последние 30 дней | -          | -           | 1,000   | 292 |
| – да   | 0 (0,00%)  | 2 (0,99%)   | -       | -   |
| – нет  | 89 (100%)  | 201 (99,0%) | -       | -   |
| * – Статистически значимые различия  |            |             |         |     |

Группы сопоставимы по следующим категориальным признакам: пол ( $p=0,692$ ), курение ( $p=0,405$ ), беременность ( $p=0,553$ ), иммунодефицит, включая ВИЧ ( $p=1,000$ ), болезни ЖКТ ( $p=0,981$ ), онкопатология ( $p=0,324$ ), посещение стран и регионов с неблагоприятной эпидемиологической обстановкой за последние 30 дней до симптомов ( $p=1,000$ ) [18].

Несмотря на наличие статистически значимого различия в медиане возрастов для группы 1 (48,8 лет) и группы 2 (42 года), следует отметить, что медианы возраста находятся в одной возрастной группе от 40 до 50 лет с разницей в 6,8 лет. Данный момент также будет учтен при дальнейшем анализе [18].

В группе 1 наблюдалось превалирование хронических заболеваний: сердечно-сосудистые заболевания, включая гипертонию ( $p < 0,001$ ), сахарный диабет ( $p < 0,001$ ), болезнь почек ( $p < 0,001$ ), заболевания опорно-двигательного аппарата ( $p = 0,005$ ), хроническая болезнь легких ( $p = 0,005$ ), что связано с этиопатогенезом и тяжестью течения COVID-19 [18].

### 3.4 Характеристика групп в соответствии с ПЦР тестированием

В результате ПЦР-тестирования образцы назальных мазков были распределены по трем группам: условно-здоровый контроль, ОРИ и SARS-CoV-2. Асимптоматическое носительство патогенов, идентификация возбудителей других ОРИ среди контрольной группы и пациентов с COVID-19, а также не соответствие пиков требованиям контроля качества при проведении MALDI-TOF MS (Ascore) привело к перераспределению пациентов в группах или удалению образцов из анализа (таблица 8) [18].

Таблица 8 – Описательная характеристика пациентов в группах

| Характеристика  | Всего,<br>N = 252 | SARS-<br>CoV-2,<br>N = 115 | ОРИ,<br>N=98 | Контроль,<br>N=39 | p-<br>value * |
|---|-------------------|----------------------------|--------------|-------------------|---------------|
| Мужской пол, n (%)  | 114 (45.2%)       | 55 (47.8%)                 | 38 (38,8%)   | 21(53.8%)         | <0.001        |
| Принадлежность к казахскому этносу, n (%)                                   | 116 (46%)         | 31 (26.9%)                 | 61 (62.2%)   | 24(61.5%)         | <0.001        |
| Сопутствующие заболевания   | 104 (41.2%)       | 73 (63.4%)                 | 15 (15.3%)   | 14(35.9%)         | <0.001        |
| * – различия между группами оценивались с использованием хи-квадрат Пирсона |                   |                            |              |                   |               |

Большинство (72,4%) участников группы ОРИ дали положительный результат на распространенные респираторные патогены, отличный от SARS-CoV-2, согласно ПЦР-тесту. Образцы были взяты примерно на 3-й день (медиана, 1 и 3 квартили [2-4]) и на 5-й день (медиана, 1 и 3 квартили [3-7]) после начала симптомов для участников с SARS-CoV-2+ и ОРИ соответственно (таблица 9) [18].

Таблица 9 – Характеристика групп ОРИ и SARS-CoV-2+

| Группа      | Процент ПЦР-подтвержденных случаев в группе | День заболевания, Медиана (интерквартильный размах) | Степень тяжести |            |          |
|-------------|---|---|-----------------|------------|----------|
|             |   |   | легкая          | средняя    | тяжелая  |
| ОРИ         | 72.4%                                       | 3.0 (2.0;4.0)                                       | 27 (27,6%)      | 70 (71,4%) | 1(0,94%) |
| SARS-CoV-2+ | 100%  | 5.0 (3.0; 7.0)                                      | 10 (8.7%)       | 98 (85.2%) | 7 (6.1%) |

### 3.5 Результаты комбинированной работы методов MALDI-TOF MS и ML

Основная цель исследования заключалась в оценке способности комбинации методов MALDI-TOF MS и ML обнаруживать инфекцию SARS-CoV-2 в гетерогенной смеси образцов SARS-CoV-2+, ОРИ и условно-здорового контроля. Схематически алгоритм исследования изображен на рисунке 18 [18].

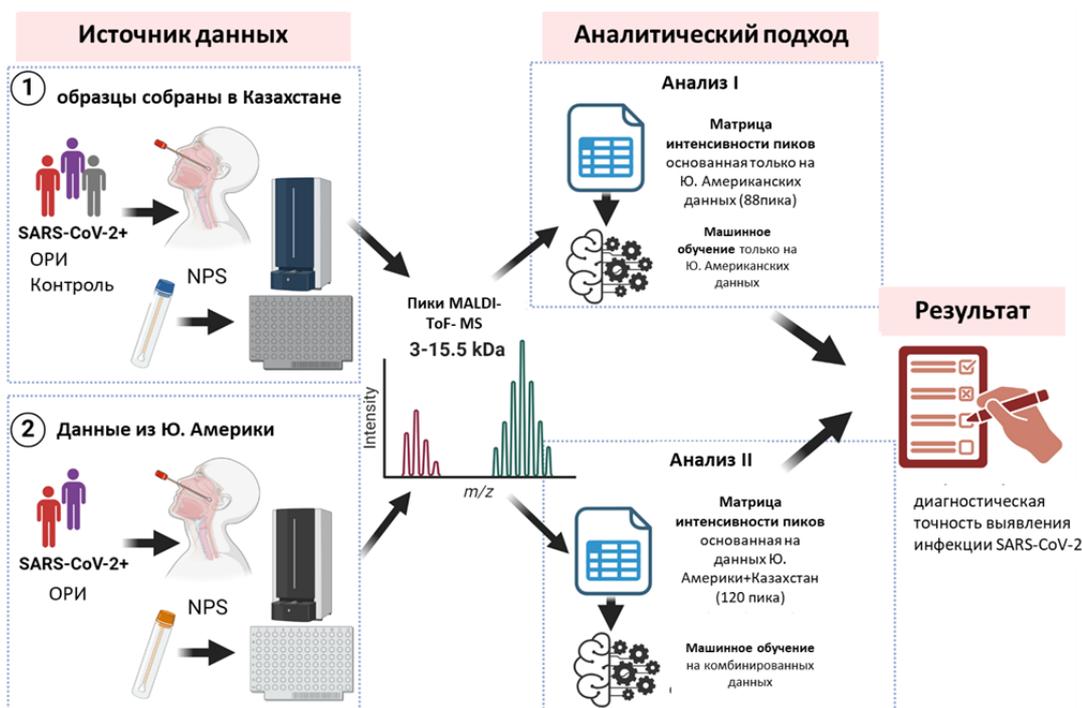


Рисунок 18 – Дизайн проведенного исследования

Применение методов машинного обучения включал два независимых поданализа (рисунок 18). В первом поданализе была оценена эффективность пайплайн машинного обучения Nachtigall и др. на комбинированной выборке образцов, взятых как из исследования Nachtigall (данные собраны в трех странах Южной Америки в 2020 году), так и на собственно полученных данных (данные собраны в 2020 и 2022 годах); в этом анализе модели обучались только на оригинальных данных из Южной Америки. Во втором поданализе алгоритмы ML были обучены по новой, добавив спектры, предоставленные образцами из Казахстана, и был применен этот переобученный алгоритм машинного обучения к комбинированной выборке образцов [18].

#### 3.5.1 Применение существующего пайплайн MALDI-TOF MS для выделения образцов SARS-CoV-2+, собранных в Казахстане (Анализ I)

Для того, чтобы оценить, насколько хорошо обученные модели Nachtigall [15, p. 1168-1172] могут различать образцы SARS-CoV-2+ в наборе данных из Казахстана, были повторены шаги для данных из Казахстана: выбора пиков MALDI-TOF MS (рисунок 19), применения алгоритмов ML и оценки ML.

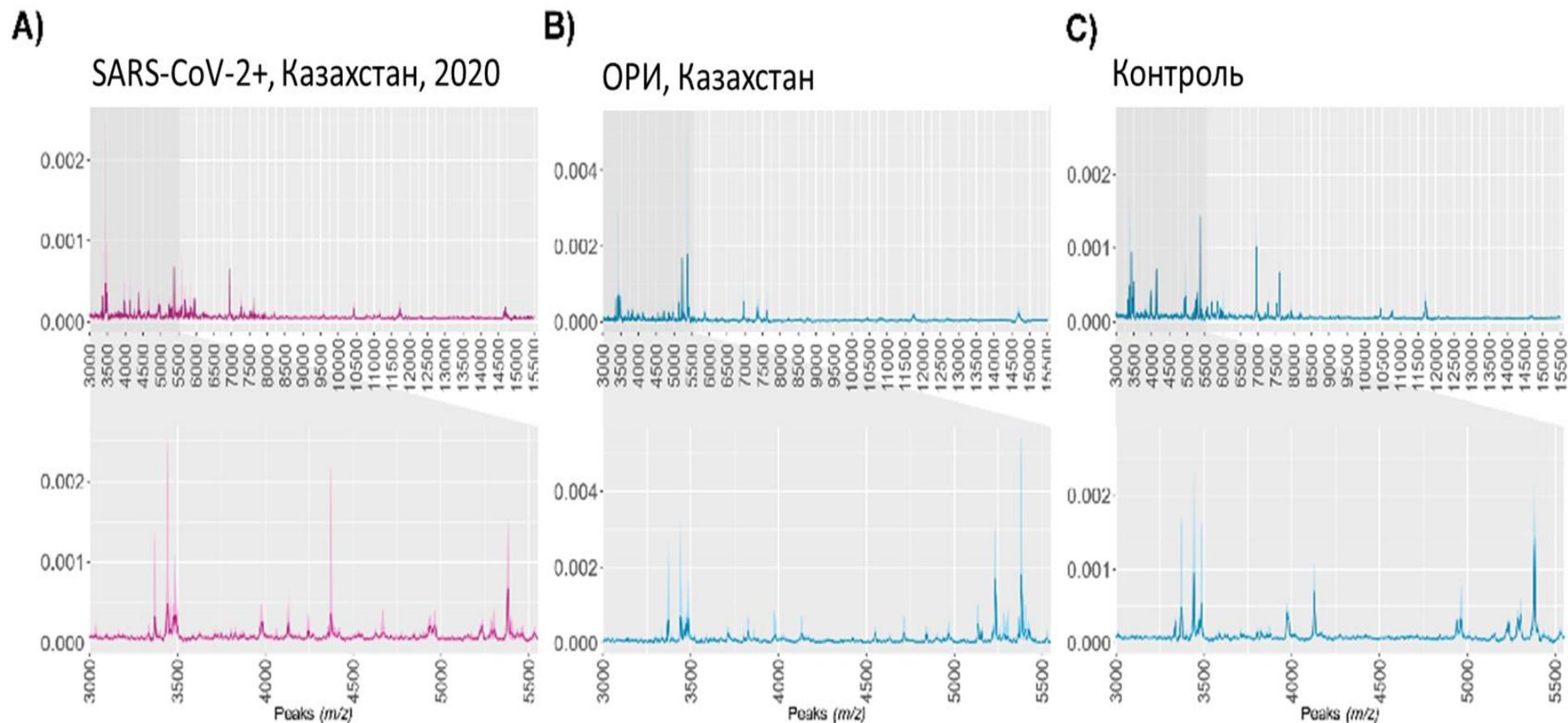


Рисунок 19 – Ось абсцисс –  $m/z$  , ось ординат – интенсивность

Примечания:

1. Сравнение образцов с использованием существующего пайплайна Nachtigall.
2. А-С. спектры MALDI-TOFMS от пациентов с SARS-CoV-2+ (А), ОРИ(В) и здорового контрольного образца из Казахстана (С).
3. Центральная линия указывает медианное значение спектров, в то время как затененная область с обеих сторон представляет интерквартильный интервал.
4. Вставки показывают диапазон от 3000 до 5500  $m/z$ , охватывающий 80% (62/88) определенных пиков

В частности, внимание было сфокусировано на тех же пиках MALDI-TOF MS, которые использовали Nachtigall и др. [15, p. 1168-1172] в своем анализе. Всего было обнаружено 88 пиков, которые затем были использованы в Анализе I. Эти пики были получены с использованием препроцессинга спектров, включая преобразование и сглаживание спектров, удаление базовой линии, калибровку спектров, обнаружение пиков и обработку пиков, что позволило использовать «наши» данные в качестве входных данных для моделей ML предоставленных южноамериканскими коллегами [18].

Для подготовки комбинированного набора данных, включающего образцы из Южной Америки и Казахстана (рисунок 19), (Приложение Д) была построена матрица интенсивности для 88 пиков, идентичных тем, что использовались в оригинальном исследовании Nachtigall [15, p. 1168-1172]. Р-значения были рассчитаны с использованием критерия Краскела-Уоллиса. Определение различий в группах по пикам необходимо для выявления пиков с наибольшим вкладом в детекцию SARS-CoV-2. Для всех образцов были проведены PCA и построение дендрограммы с целью уменьшения многомерности данных. Очевидной кластеризации по подгруппам обнаружено не было, что подчеркивает необходимость более чувствительного подхода для выявления многомерных различий в наборе данных [18]. На рисунках 20, 21, 22 продемонстрировано отсутствие явной кластеризации по подгруппам.

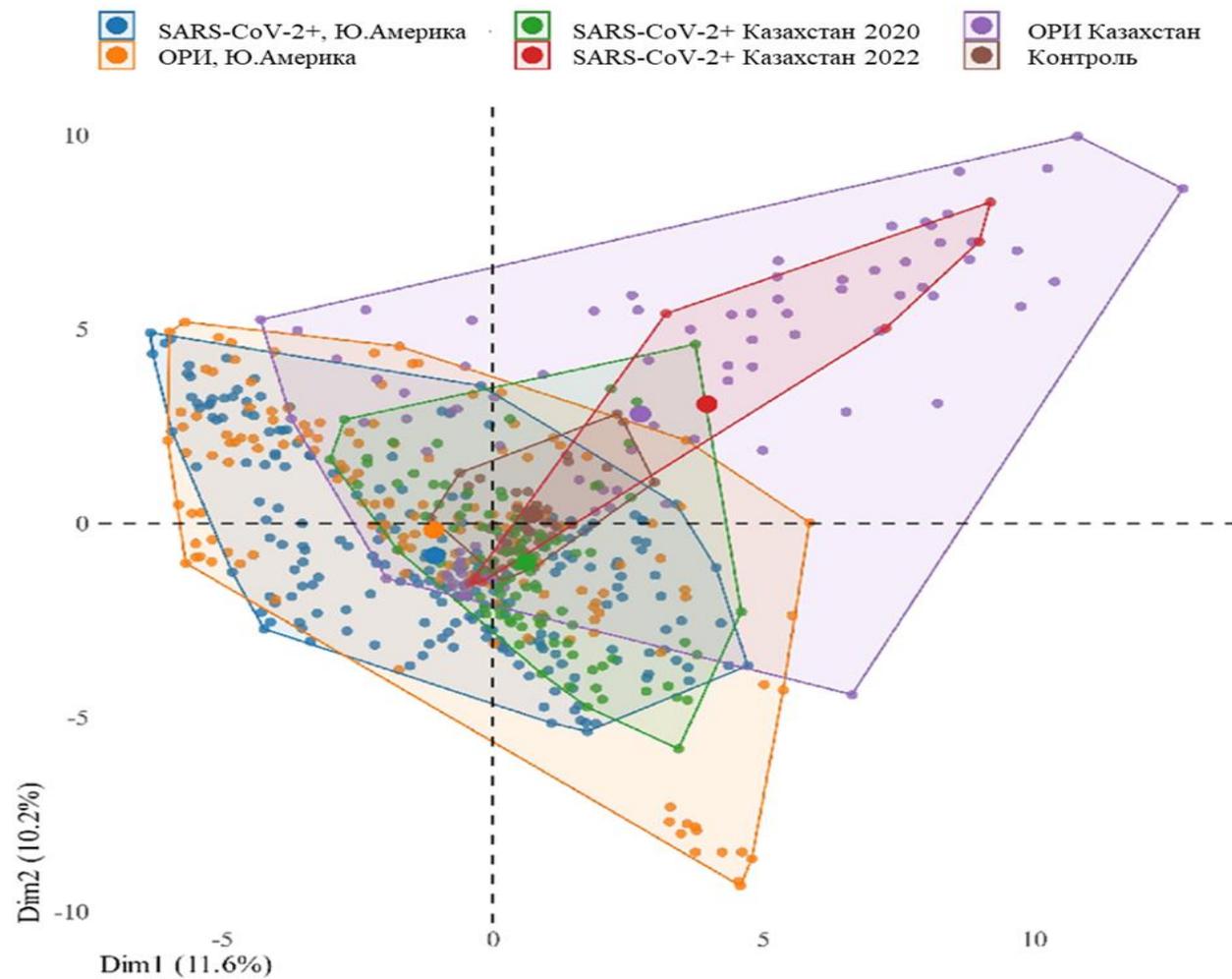


Рисунок 20 – PCA отображающий образцы, разделенные по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа I

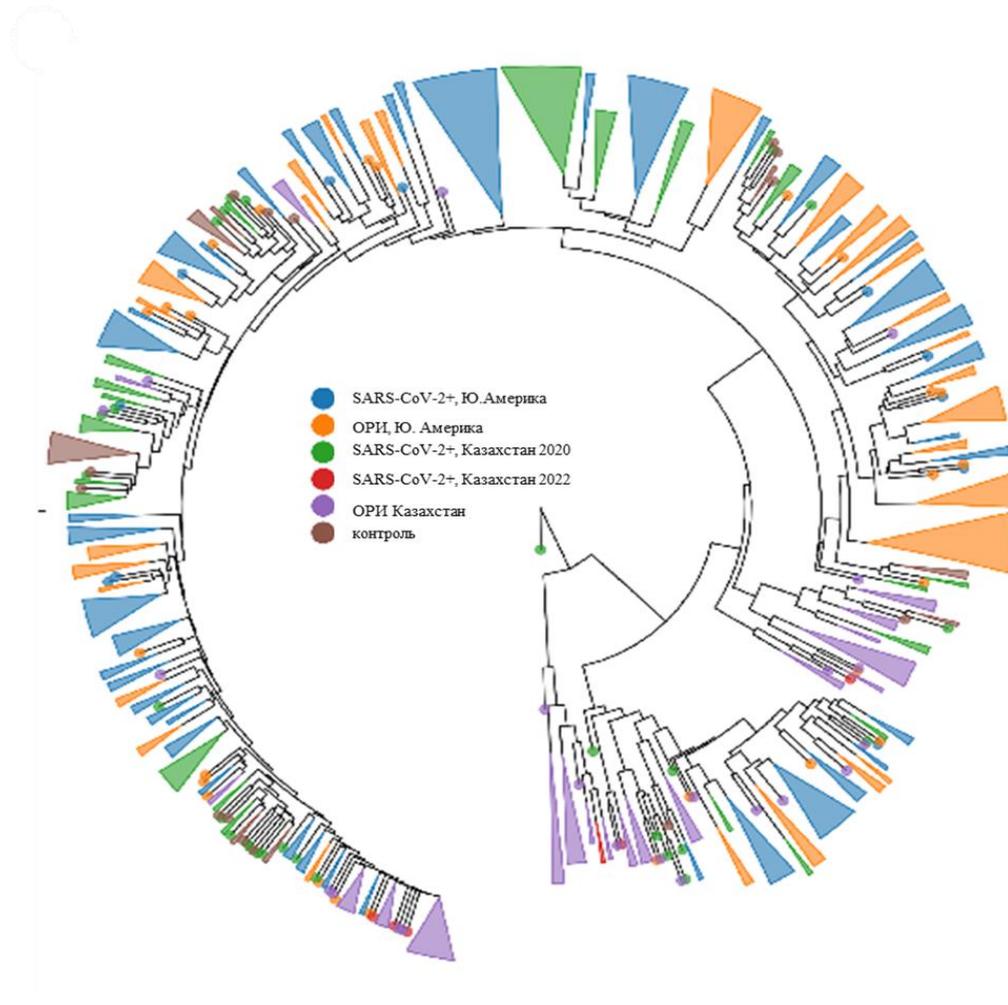


Рисунок 21 – Дендрограмма образцов, разделенных по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа I

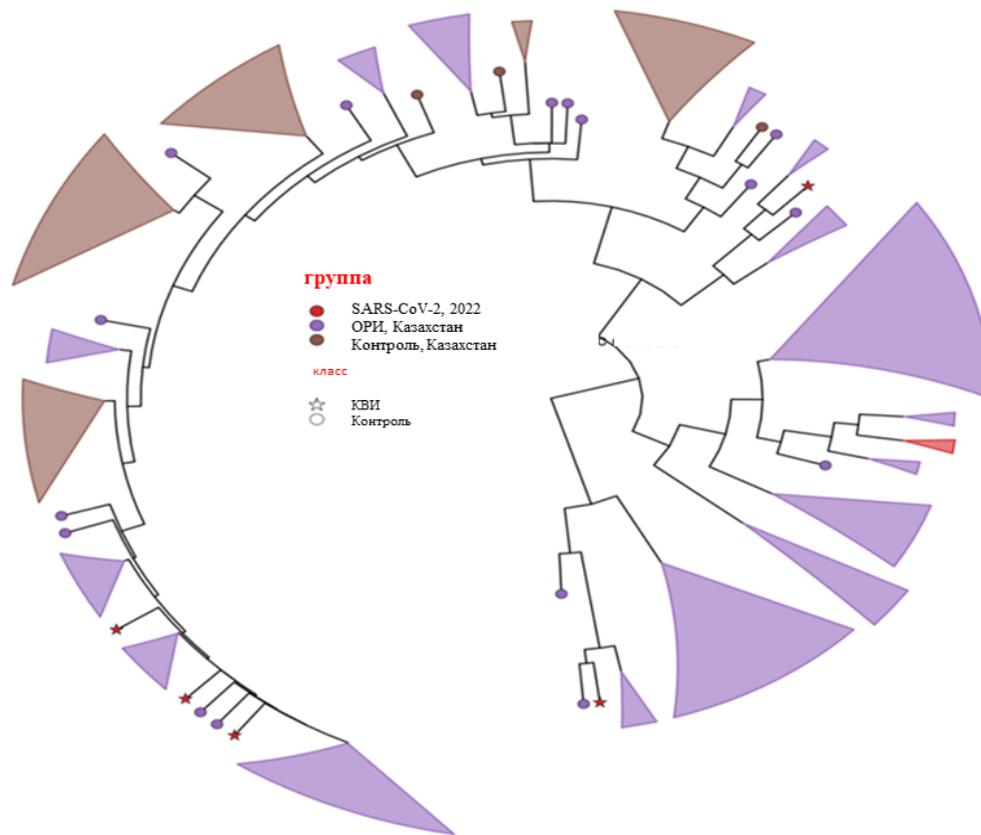


Рисунок 22 – Дендрограмма образцов, построенная на основе полученных масс-спектров, разделенных по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа I

Далее были непосредственно применены алгоритмы машинного обучения Nachtigall и др. (обученный на оригинальном наборе данных из Южной Америки) к комбинированному набору данных Южная Америка-Казахстан.

В соответствии с ранее полученными результатами [15, p. 1168-1172], при тестировании только южноамериканских образцов SVM-R показал наивысшую площадь под кривой ROC AUC, хотя и другие модели имели схожие характеристики высокой производительности (таблица 10, рисунок 23) для классификации случаев SARS-CoV-2 и не-SARS-CoV-2. Этот анализ повторял эксперимент по тестированию, проведенный Nachtigall и соавт. Каждый показатель представлен как медиана и межквартильный размах. Каждая модель тестировалась 5 раз на 20% набора данных Южной Америки [18].

Таблица 10 – Производительность семи ML-моделей при тестировании только на наборе данных Южной Америки (Анализ I)

| Классификационные метрики | DT                         | KNN                        | NB                         | RF                         | SVM-L                      | SVM-R                      | XGBoost                    |
|---------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| F мера                    | 0.95<br>[0.94;<br>0.95]    | 0.90<br>[0.88;<br>0.94]    | 0.88<br>[0.86;<br>0.88]    | 0.95<br>[0.93;<br>0.96]    | 0.93<br>[0.91;<br>0.94]    | 0.96<br>[0.95;<br>0.98]    | 0.93<br>[0.93;<br>0.93]    |
| Полнота                   | 0.976<br>[0.953;<br>0.976] | 0.952<br>[0.951;<br>0.953] | 0.905<br>[0.857;<br>0.907] | 0.976<br>[0.976;<br>0.977] | 0.953<br>[0.951;<br>0.976] | 0.953<br>[0.952;<br>0.953] | 0.952<br>[0.907;<br>0.976] |
| Точность                  | 0.94<br>[0.93;<br>0.95]    | 0.88<br>[0.85;<br>0.93]    | 0.85<br>[0.85;<br>0.86]    | 0.94<br>[0.92;<br>0.95]    | 0.92<br>[0.89;<br>0.93]    | 0.96<br>[0.94;<br>0.97]    | 0.92<br>[0.92;<br>0.92]    |
| Специфичность             | 0.90<br>[0.87;<br>0.90]    | 0.83<br>[0.68;<br>0.90]    | 0.81<br>[0.80;<br>0.87]    | 0.87<br>[0.83;<br>0.90]    | 0.83<br>[0.81;<br>0.87]    | 0.97<br>[0.97;<br>1.00]    | 0.84<br>[0.83;<br>0.93]    |
| Чувствительность          | 0.976<br>[0.953;<br>0.976] | 0.952<br>[0.951;<br>0.953] | 0.905<br>[0.857;<br>0.907] | 0.976<br>[0.976;<br>0.977] | 0.953<br>[0.951;<br>0.976] | 0.953<br>[0.952;<br>0.953] | 0.952<br>[0.907;<br>0.976] |
| PPV                       | 0.93<br>[0.91;<br>0.93]    | 0.89<br>[0.80;<br>0.93]    | 0.86<br>[0.86;<br>0.89]    | 0.91<br>[0.89;<br>0.93]    | 0.89<br>[0.87;<br>0.91]    | 0.98<br>[0.97;<br>1.00]    | 0.89<br>[0.89;<br>0.95]    |
| NPV                       | 0.96<br>[0.94;<br>0.96]    | 0.93<br>[0.89;<br>0.94]    | 0.85<br>[0.81;<br>0.86]    | 0.96<br>[0.96;<br>0.97]    | 0.94<br>[0.92;<br>0.96]    | 0.94<br>[0.93;<br>0.94]    | 0.93<br>[0.88;<br>0.96]    |
| ROC AUC                   | 0.99<br>[0.99;<br>0.99]    | 0.87<br>[0.83;<br>0.93]    | 0.92<br>[0.91;<br>0.93]    | 0.99<br>[0.97;<br>1.00]    | 0.97<br>[0.97;<br>0.97]    | 1.00<br>[0.99;<br>1.00]    | 0.98<br>[0.98;<br>0.99]    |

Впоследствии были оценены производительность тех же алгоритмов машинного обучения на образцах из Казахстана. Каждая модель была протестирована 5 раз на 20% выборке образцов из Казахстана. Использовалась пятикратная (внешняя) вложенная перекрестная (пять раз) десятикратная (внутренняя) кросс-валидация со случайным стратифицированным разбиением. В соответствии с рисунками 23, 24, здесь наблюдался широкий разброс в способности моделей ML различать образцы SARS-CoV-2.

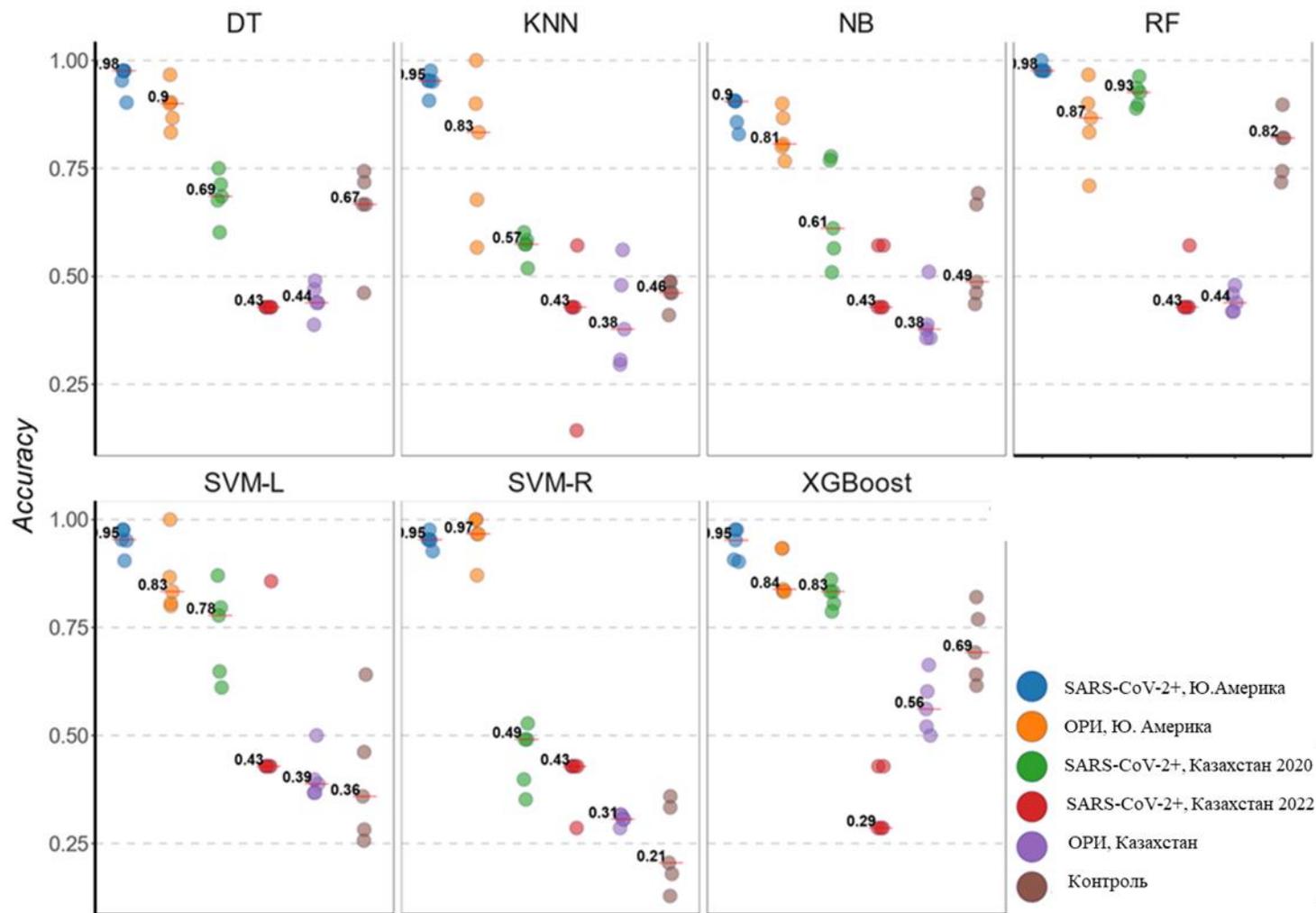


Рисунок 23 – Метрики точности для каждой из семи моделей ML, обученных на данных MALDI-TOF MS из Южной Америки (Анализ I) для различения подгрупп исследования

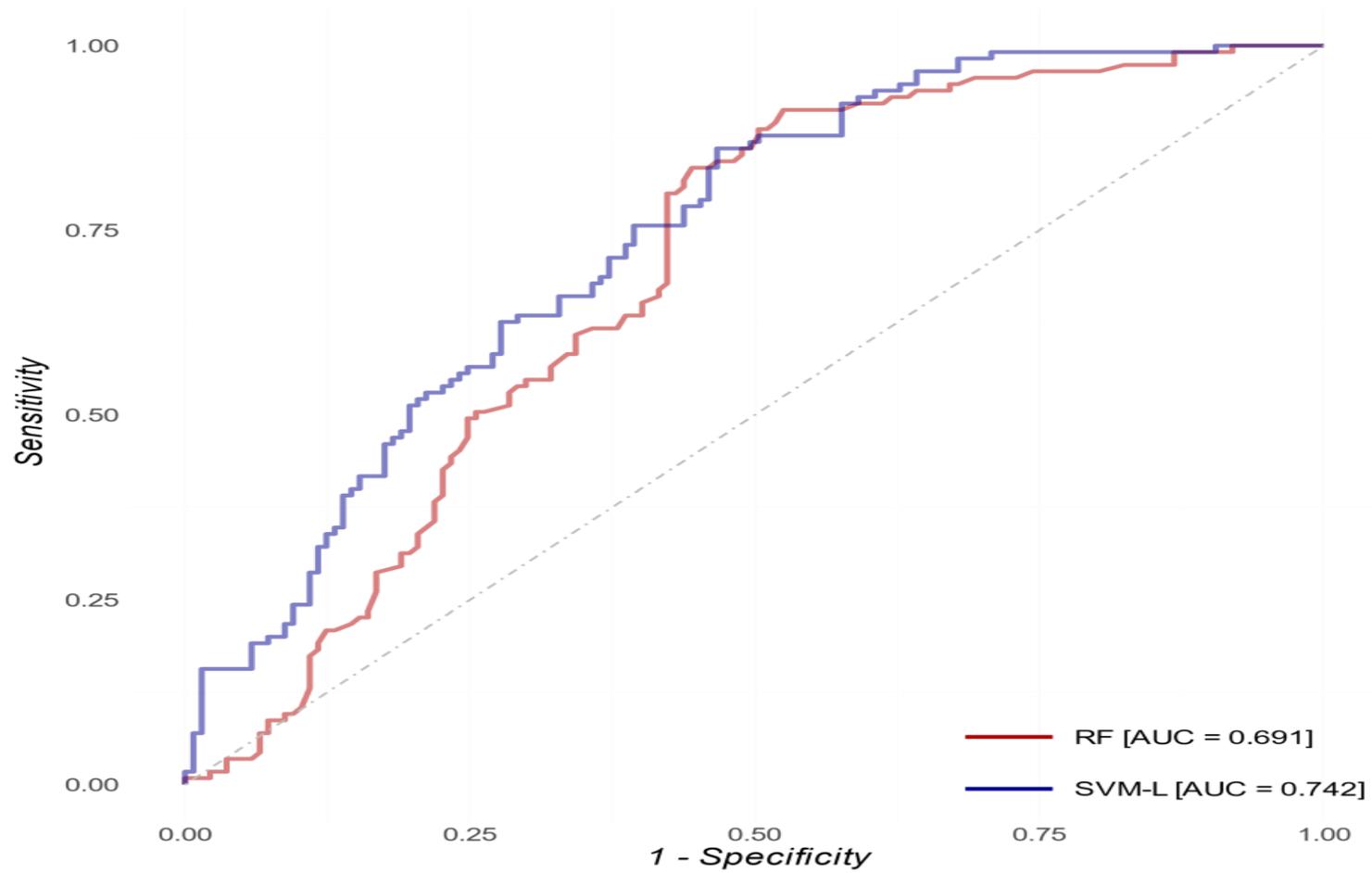


Рисунок 24 – Кривые ROC для наилучших по производительности алгоритмов RF и SVM-L (Анализ I)

RF показал наивысший процент правильно идентифицированных образцов 2020-SARS-CoV-2+ (91%) (таблица 11), однако точность детекции SARS-CoV-2 2022 года была <60% для всех моделей, что схоже с точностью идентификации ОРИ. Раздельное тестирование групп SARS-CoV-2 разных годов имело принципиальную важность. В период пандемии вирус мутировал несколько раз, соответственно было необходимо оценить способность моделей, обученных в 2020 году определять образцы с вирусом SARS-CoV-2 2022 года без предварительного дообучения. RF определил образцы контрольной группы с точностью 82%, что является наивысшим показателем среди всех моделей для этой подгруппы [18].

Таблица 11 – Значения ROC AUC для сравнения всех образцов SARS-CoV-2+ Казахстана (KZ), собранных в 2020 и 2022 годах, по сравнению с контрольными группами (либо ОРИ, либо объединенная группа ОРИ+контроль)

| Попарное сравнение |                        |                                 |                        |                                 |                                 |   |
|--------------------|------------------------|---------------------------------|------------------------|---------------------------------|---------------------------------|---|
| ML                 | SARS-CoV-2/2021 vs ОРИ | SARS-CoV-2/2022 vs ОРИ+контроль | SARS-CoV-2/2022 vs ОРИ | SARS-CoV-2/2022 vs ОРИ+контроль | SARS-CoV-2 [2020 + 2022] vs ОРИ | SARS-CoV-2 [2020+ 2022] vs ОРИ+контроль |
| DT                 | 0.65 [0.65; 0.66]      | 0.64 [0.63; 0.64]               | 0.46 [0.42; 0.48]      | 0.43 [0.38; 0.44]               | 0.64 [0.64; 0.64]               | 0.62 [0.62; 0.63]                       |
| KNN                | 0.57 [0.55; 0.64]      | 0.57 [0.55; 0.61]               | 0.51 [0.47; 0.53]      | 0.52 [0.48; 0.52]               | 0.57 [0.54; 0.63]               | 0.56 [0.55; 0.61]                       |
| NB                 | 0.68 [0.65; 0.70]      | 0.66 [0.66; 0.68]               | 0.52 [0.51; 0.52]      | 0.51 [0.50; 0.51]               | 0.67 [0.64; 0.69]               | 0.65 [0.65; 0.67]                       |
| RF                 | 0.71 [0.70; 0.73]      | 0.69 [0.68; 0.70]               | 0.48 [0.44; 0.48]      | 0.43 [0.41; 0.44]               | 0.70 [0.69; 0.71]               | 0.67 [0.67; 0.68]                       |
| SVM-L              | 0.74 [0.72; 0.75]      | 0.76 [0.72; 0.76]               | 0.57 [0.56; 0.58]      | 0.54 [0.52; 0.57]               | 0.73 [0.71; 0.74]               | 0.74 [0.71; 0.75]                       |
| SVM-R              | 0.60 [0.57; 0.62]      | 0.60 [0.59; 0.63]               | 0.58 [0.58; 0.59]      | 0.61 [0.59; 0.64]               | 0.60 [0.57; 0.62]               | 0.60 [0.59; 0.63]                       |
| XG-Boost           | 0.65 [0.63; 0.67]      | 0.64 [0.62; 0.67]               | 0.30 [0.29; 0.35]      | 0.29 [0.26; 0.32]               | 0.62 [0.61; 0.65]               | 0.62 [0.60; 0.66]                       |

### 3.5.2 Самостоятельное обучение алгоритмов, используя собственные данные MALDI-TOF MS для идентификации SARS-CoV-2 среди прочих ОРИ и асимптоматических мазков (Анализ II)

Была создана комбинированная база данных пиков, полученных из Южной Америки, а также масс-спектрограмм собственных образцов (n=615). Всего было идентифицировано 120 пиков обеспечивающих вклад в детекцию SARS-CoV-2 и была построена матрица интенсивности (Приложение E). Р-значения были рассчитаны с использованием критерия Краскела-Уоллиса. Для всех образцов были проведены PCA и иерархическая кластеризация с целью уменьшения размерности. Очевидной кластеризации по подгруппам обнаружено не было, что подчеркивает необходимость более чувствительного подхода для выявления многомерных различий в наборе данных.

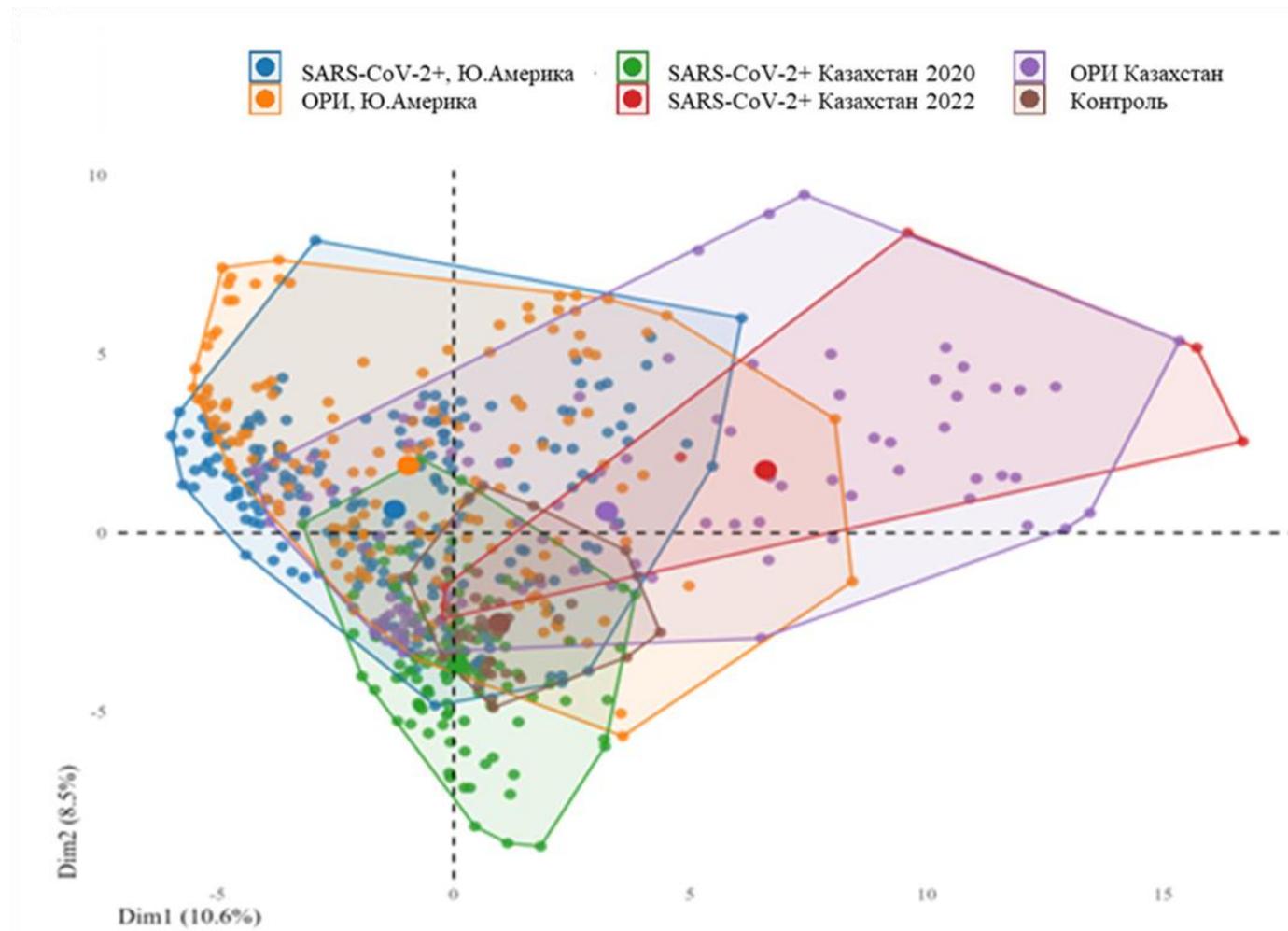


Рисунок 25 – PCA отображающий образцы, разделенные по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа II

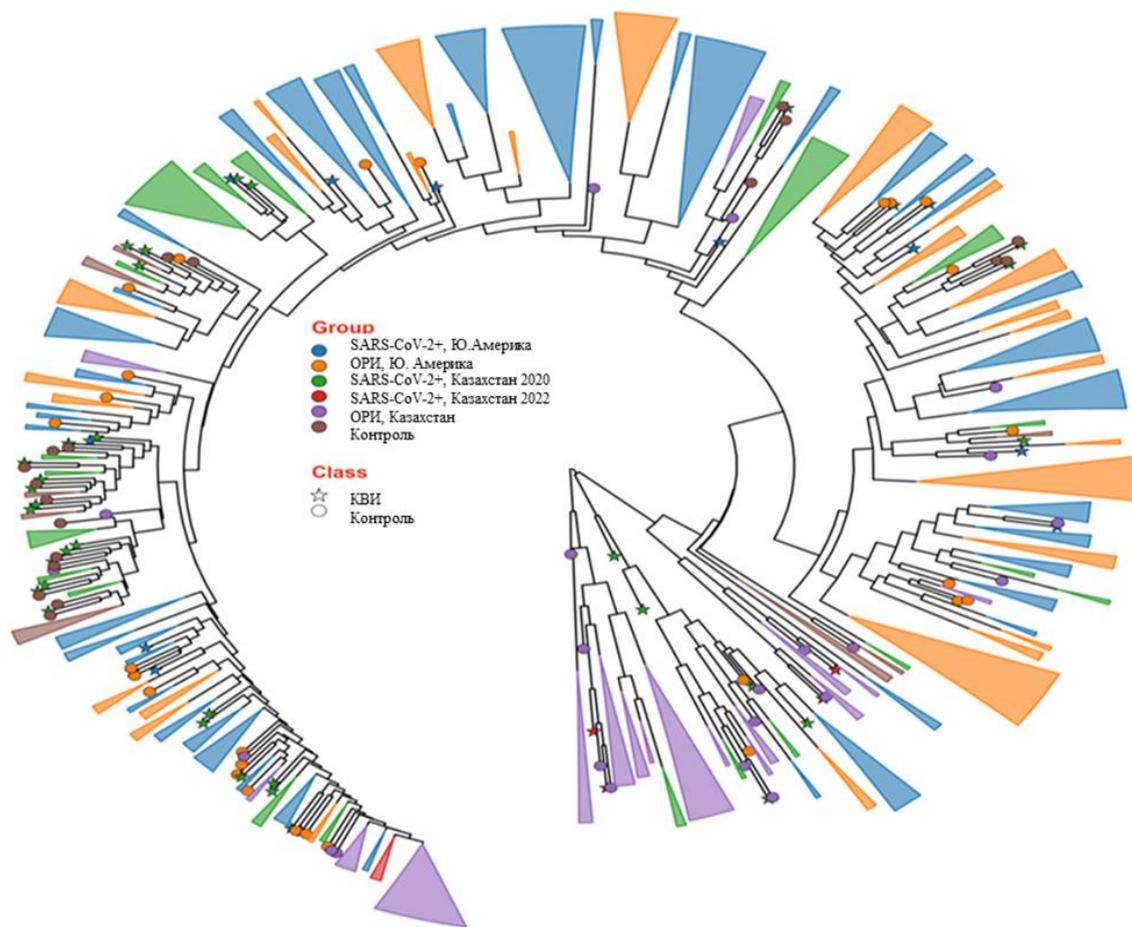


Рисунок 26 – Дендрограмма образцов, разделенных по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа II

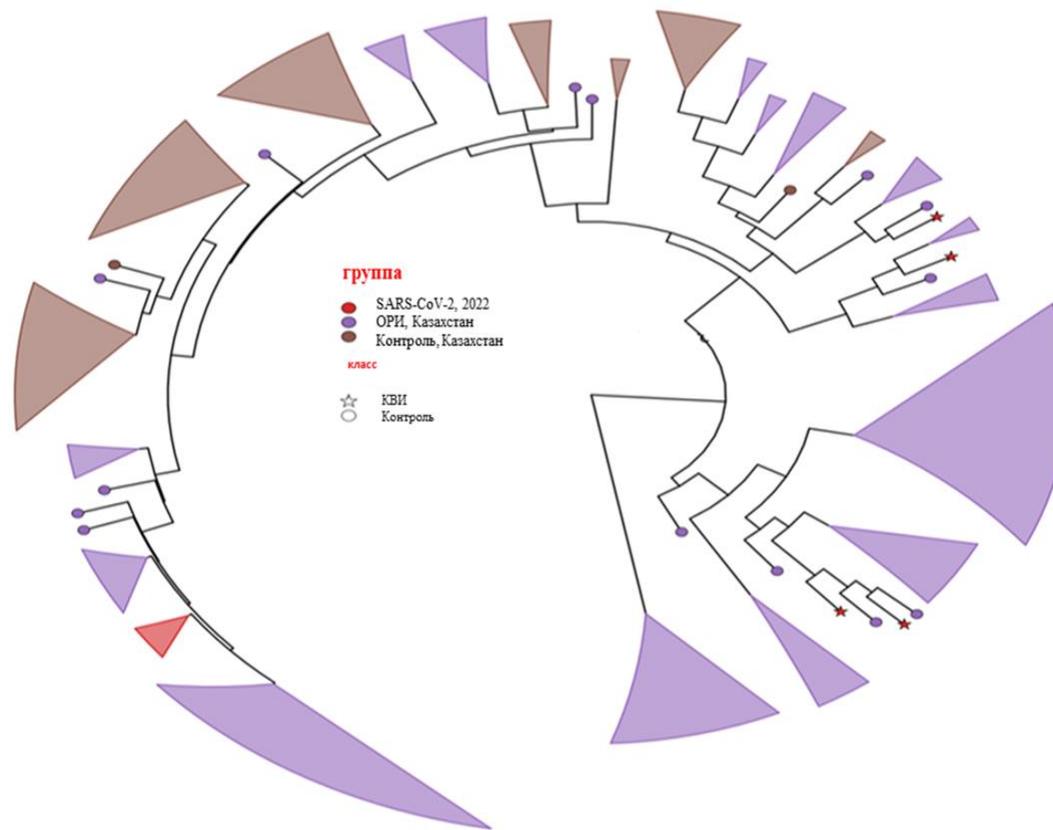


Рисунок 27 – Дендрограмма образцов, разделенных по подгруппам, из объединенного набора данных на основе матрицы интенсивности пиков для Анализа II

На рисунках 25, 26, 27 продемонстрировано отсутствие явной кластеризации по подгруппам. Проведенный анализ показал, что базовые методы анализа не выявляют различий и указывает на необходимость использования методов машинного обучения.

Затем было проведено обучение моделей машинного обучения на комбинированной выборке, состоящей из 120 пиков, из которых 53 перекрывались с 88 пиками из исследования Nachtigall [15, p. 1168-1172]. Для проведения данного анализа группы 2020-SARS-CoV-2+ и 2022-SARS-CoV-2+ были объединены, т.к. выборка группы 2022-SARS-CoV-2+ составляла 7 образцов. Провести машинное обучение и дальнейшее тестирование невозможно на таком малом количестве участников.

Была оценена производительность обученного алгоритма машинного обучения на наборе комбинированных данных Южная Америка-Казахстан. Все модели продемонстрировали схожие характеристики высокой производительности при различении образцов SARS-CoV-2+. SVM-R и DT незначительно превзошли другие пять моделей в определении инфекции SARS-CoV-2 от ОРИ и контроль с показателями ROC AUC 0,983 [0,958, 0,987] и 0,972 [0,966, 0,979] соответственно (рисунки 28, 29, таблица 12). SVM-R, в частности, разделил образцы SARS-CoV-2+, ОРИ и контроль из Казахстана с точностью 88,0, 95,0 и 78,0% соответственно (рисунок 28). Оба алгоритма, SVM-R и DT, также были очень точными в различении подгрупп ОРИ и контроль (таблица 12). Каждая модель была протестирована 5 раз на 20% выборке образцов из Казахстана (остальные 80% использовались для обучения ML-моделей) [18].

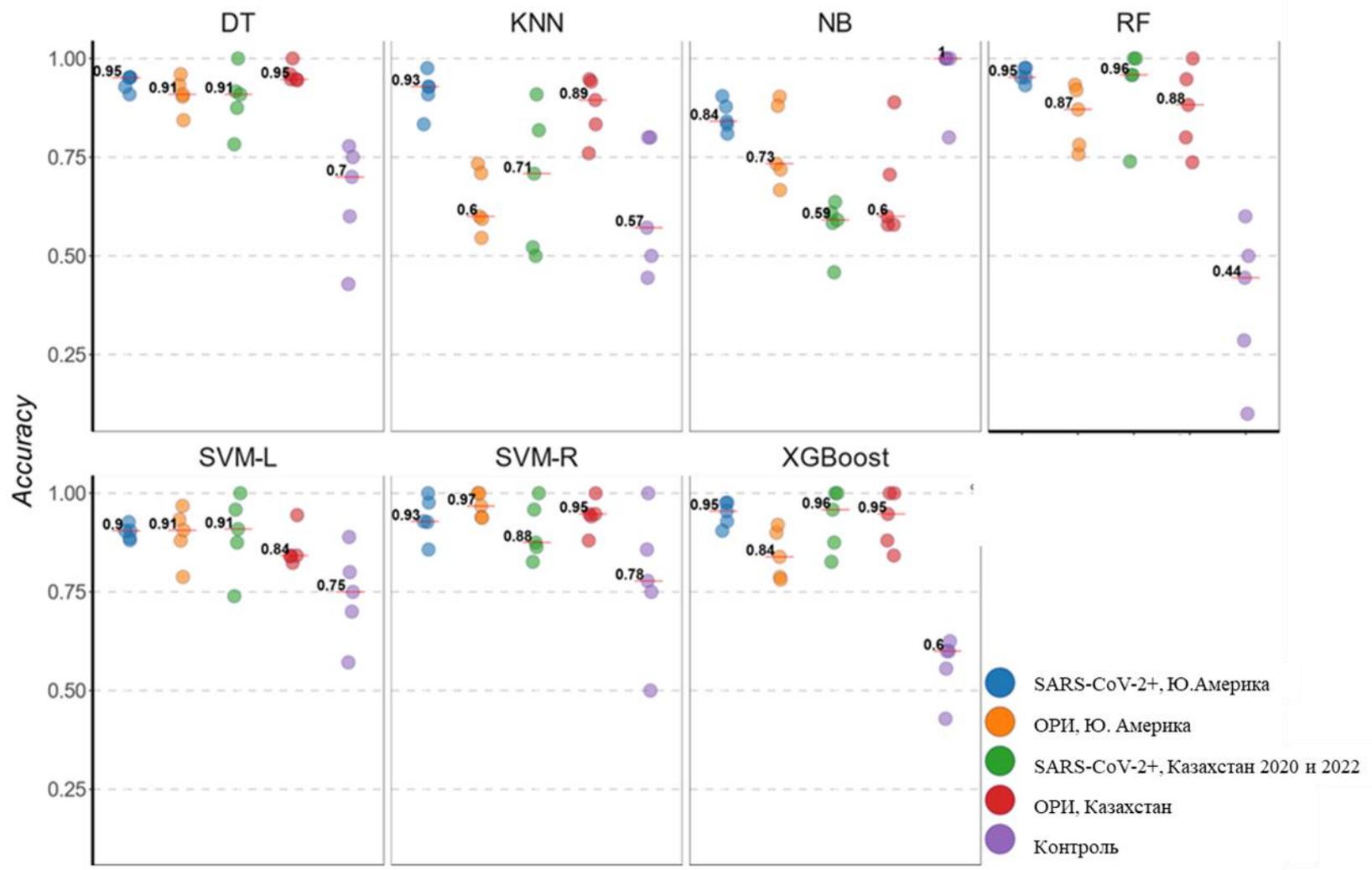


Рисунок 28 – Показатели точности для каждой из семи моделей ML, обученных на объединенном наборе данных Южной Америки и Казахстана (Анализ II в текущем исследовании) для дифференциации исследуемых подгрупп

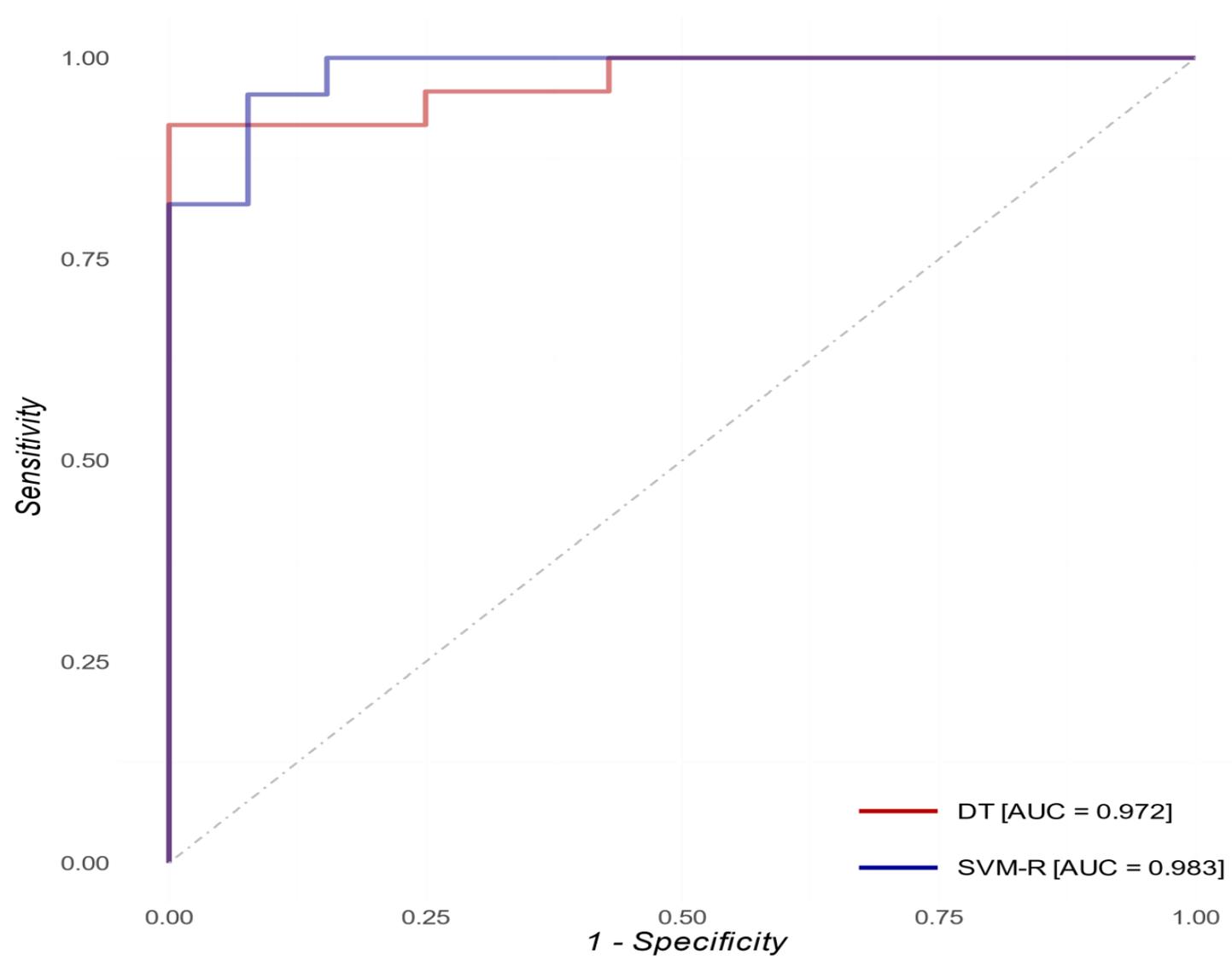


Рисунок 29 – Кривые ROC для наиболее эффективных алгоритмов SVM-R и DT (Анализ II)

Таблица 12 – Значения ROC AUC семи ML-моделей, обученных в Анализе II для дифференциации каждой подгруппы в наборе данных Казахстана (KZ)

| Попарное сравнение | DT             | KNN            | NB             | RF             | SVM-L          | SVM-R          | XGBoost        |
|--------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| SARS-CoV-2         | 0.972          | 0.92           | 0.801          | 0.93           | 0.94           | 0.983          | 0.958          |
| против остальных   | [0.966; 0.979] | [0.88; 0.92]   | [0.769; 0.808] | [0.90; 0.96]   | [0.93; 0.95]   | [0.958; 0.987] | [0.943; 0.970] |
| OPI                | 0.996          | 0.95           | 0.90           | 0.984          | 0.964          | 0.995          | 0.986          |
| против остальных   | [0.979; 1.000] | [0.93; 0.97]   | [0.83; 0.92]   | [0.983; 1.000] | [0.954; 0.984] | [0.964; 0.996] | [0.981; 0.996] |
| контроль           | 0.979          | 0.898          | 0.82           | 0.972          | 0.986          | 0.995          | 0.981          |
| против остальных   | [0.979; 0.984] | [0.886; 0.955] | [0.79; 0.85]   | [0.966; 0.984] | [0.976; 0.992] | [0.990; 0.996] | [0.979; 0.988] |
| Микро-усреднение   | 0.977          | 0.93           | 0.785          | 0.951          | 0.965          | 0.991          | 0.960          |
|                    | [0.966; 0.981] | [0.90; 0.94]   | [0.775; 0.834] | [0.925; 0.968] | [0.958; 0.966] | [0.965; 0.991] | [0.956; 0.977] |
| Макро-усреднение   | 0.971          | 0.91           | 0.806          | 0.962          | 0.961          | 0.983          | 0.972          |
|                    | [0.965; 0.981] | [0.90; 0.92]   | [0.794; 0.857] | [0.941; 0.972] | [0.955; 0.961] | [0.955; 0.986] | [0.951; 0.977] |

В исследовании оценивалась возможность применения MALDI-TOF MS и ML в клинической лаборатории для различения SARS-CoV-2 от других OPI, особенно в контексте минимальной специализированной подготовки образцов. Первоначальное применение оригинального пайплайн MALDI-TOF MS/ML, обученного на образцах из Южной Америки [15], показало сниженную эффективность при идентификации образцов из Казахстана. Дообучение моделей ML с учетом информации о пиках MALDI-TOF из разнообразного набора казахстанских образцов, включая SARS-CoV-2+, объекты OPI и условно-доровые контрольные образцы, привело к значительному улучшению точности обнаружения (таблица 13). Рассматривая это как доказательство концепции, результаты данной работы подтверждают высокий потенциал применения MALDI-TOF MS/ML, особенно на ранних стадиях респираторных эпидемий/пандемий, когда доступна ограниченная информация о идентичности инфекционного патогена, и в условиях с ограниченными ресурсами, где альтернативные методы могут быть еще недоступны [18].

Таблица 13 – Характеристики ROC AUC моделей RF и SVM-R, обученных в Анализе I и Анализе II

| Анализ I   |                  | Анализ II        |                  |
|--|------------------|------------------|------------------|
| RF   | SVM-R            | RF               | SVM-R            |
| 0,67 [0,67;0,68]   | 0,60 [0,59;0,63] | 0,93 [0,90;0,96] | 0,98 [0,96;0,99] |
| * – Результаты представлены в виде медианы и межквартильного интервала |                  |                  |                  |

### **3.6 Перспективы применения MALDI-TOF MS/ML для диагностики SARS-CoV-2 в назальных мазках**

Полученные в данном исследовании результаты демонстрируют потенциал использования комбинированного подхода MALDI-TOF MS/ML для быстрой идентификации SARS-CoV-2 среди других острых респираторных инфекций [18].

Значение данного подхода заключается в возможности быстрой и недорогой диагностики COVID-19, особенно в условиях ограниченных ресурсов, где стандартные молекулярно-генетические методы могут быть недоступны или требуют значительных затрат времени и средств. MALDI-TOF MS позволяет получить характеристики белкового состава биологического образца за считанные минуты, а применение машинного обучения дает возможность выявлять специфические паттерны, отличающие SARS-CoV-2 от других инфекций. Такая комбинация решает вопросы необходимости получения чистых культур при одиночной работе с MALDI-TOF MS, где концентрация микроорганизма в смеси определяет точность диагностики [18].

В то же время, данный метод имеет определенные ограничения. Возможное влияние разных факторов, таких как сопутствующие инфекции, стадия и тяжесть заболевания, на состав белков в образце, что может вносить дополнительный "шум" в данные [100, p. 1-14; 120-122]. Поэтому необходима дальнейшая оптимизация пробоподготовки и условий анализа [18]. В частности, требуется тщательная оптимизация метода и валидация моделей машинного обучения на локальных выборках, чтобы повысить их точность. Согласно результатам проведенной работы, модели, обученные на данных из других стран, могут демонстрировать сниженную эффективность [18]. В целях оптимизации рекомендуется увеличение размера выборки и большее географическое разнообразие объединенной обучающей выборки, а также большее количество общих выборок повысило надежность и гибкость модели. Кроме того, добавление данных из разных популяций обеспечило более широкое представление о возможных генетических и экологических вариациях респираторного микробиома. А также для оптимизации производительности методов машинного обучения необходимо применять комбинаций гиперпараметров во внутреннем цикле подхода кросс-валидации, используя метод случайного поиска. Для алгоритма XGBoost подбор осуществляется при выборе тюнинг-параметров  $\lambda$  и  $\alpha$ .

Так, оптимизация комбинированного метода MALDI-TOF MS/ML включала следующие шаги: 1 – увеличение размера выборки и большее географическое разнообразие объединенной обучающей выборки; 2 – введение группы здоровых пациентов; 3 – обнаружение дополнительных пиков; 4 – дообучение моделей машинного обучения.

Из-за относительно малого размера выборки группы ОРИ дальнейшая стратификация этой группы по идентифицированным с помощью мультиплексной ПЦР патогенам не была проведена. Важность использования MALDI-TOF MS/ML для идентификации различных ОРИ следует рассмотреть

в контексте меняющегося постпандемического ландшафта ОРИ [18; 100, р. 1-14].

Дополнительное изучение ОРИ потенциально дает возможность быстрой диагностики пациентов, рутинно обращающихся в поликлиники, возможность ставить диагноз по этиологии и применять таргетную терапию для вирусных заболеваний, так же, как и для бактериальных инфекций. Соответственно данный подход может способствовать разработке новых лекарственных средств для направленного лечения вирусной патологии. Возможность точной идентификации возбудителя в будущем даст повышенную лояльность к проведению вакцинации.

На сегодняшний день, в условиях быстрой мутации вирусов, готовности ВОЗ к новой пандемии, описанный инструмент предоставляет перспективу быстрого реагирования и независимости от сторонних ресурсов.

Таким образом, результаты данной работы демонстрируют принципиальную возможность использования описанного подхода для диагностики COVID-19. Полученные данные могут послужить основой для разработки экспресс-тестов на базе MALDI-TOF MS/ML, что будет способствовать более эффективному контролю распространения коронавирусной инфекции [18].

## ЗАКЛЮЧЕНИЕ

Диссертация содержит новые научно-обоснованные результаты, использование которых обеспечивает решение важной прикладной задачи - возможность быстрой и экономически выгодной детекции SARS-CoV-2 в назальных мазках при помощи MALDI-TOF MS/ML.

Несомненно, пандемия короновиральной инфекции 2019 года возникла неожиданно и внесла большие изменения во все сферы жизни. Быстроераспространение вируса и невозможность совладать со вспышкой, затем и эпидемией указала на проблемы, существующие в системе лабораторной службы во многих странах мира. Это явление стимулировало научные сообщества на поиски дополнительных методов диагностики, способных отвечать запросам в период пандемий. Проведение этой работы также было мотивировано сложившимися обстоятельствами в период коллапса системы здравоохранения в Казахстане.

В «острый» период пандемии для выявления SARS-CoV-2 применялись и серологические, и молекулярно-генетические методы лабораторной диагностики. Ввиду очевидных преимуществ в диагностике коронавируса ОТ-ПЦР является золотым стандартом. Однако, необходимость импорта специфических реагентов, поступающих из других стран, заставляет искать альтернативные способы диагностики.

Целью диссертационной работы являлась оценка и оптимизация экспресс-метода определения вируса SARS-CoV-2 в материале, полученном из назального мазка при помощи MALDI-TOF MS и методов машинного обучения.

В ходе выполнения работы был проведен забор назальных мазков, анкетирование участков, проведены ОТ-ПЦР, MALDI-TOF MS, препроцессинг данных и машинное обучение.

На первом этапе исследования была собрана коллекция биоматериала для Группы 1 (203 человека) и Группы 2 (101 человек). Анализ анкетных данных позволил определить наиболее значимые факторы и сравнить группы между собой. Так, группы сопоставимы по следующим показателям: пол ( $p=0,692$ ), беременность ( $p=0,553$ ), курение ( $p=0,405$ ), иммунодефицит, включая ВИЧ ( $p=1,000$ ), онкопатология ( $p=0,324$ ), болезни ЖКТ ( $p=0,981$ ), посещение стран и регионов с неблагоприятной эпидемиологической обстановкой за последние 30 дней до симптомов ( $p=1,000$ ). Хронические заболевания преобладали в группе 1: сердечно-сосудистые заболевания, включая гипертонию ( $p<0,001$ ), болезни почек ( $p<0,001$ ), сахарный диабет ( $p<0,001$ ), хроническая болезнь легких ( $p=0,005$ ), заболевания опорно-двигательного аппарата ( $p=0,005$ ). Возрастная характеристика групп имела значимые различия ( $p<0,0001$ ) при медиане 42 года для контрольной группы и 48,8 лет для COVID-позитивной группы. Несмотря на наличие статистически значимого различия в медиане возрастов для группы 2 (42 года) и группы 1 (48,8 лет), следует отметить, что медианы возраста находятся в одной возрастной группе от 40 до 50 лет с

разницей в 6,8 лет. Таким образом, ключевые показатели группы 1 и группы 2 сопоставимы [18].

Далее был проведен ОТ-ПЦР всех назальных мазков. Согласно ПЦР-тесту, большинство (72,4%) участников ОРИ дали положительный результат на распространенные респираторные патогены, отличные от SARS-CoV-2.

Затем был проведен MALDI-TOF MS анализ образцов, получены характерные спектры для каждой группы. После препроцессинга данных была построена матрица интенсивности пиков. Для всех образцов были проведены PCA и построение дендрограммы с целью уменьшения многомерности данных. Согласно проведенному анализу очевидной кластеризации по подгруппам обнаружено не было. Проведенный анализ подчеркивает необходимость более чувствительного подхода для выявления многомерных различий в наборе данных. Таким образом, для реализации поставленных задач использовался метод машинного обучения.

Этап машинного обучения включал 2 вида анализа. Сначала была оценена производительность обученного алгоритма машинного обучения *in silico*, затем на объединенных данных Южной Америки и Казахстана. При использовании оригинальных моделей на данных из Казахстана все модели продемонстрировали показатели низкой производительности. Далее после переобучения моделей на объединенной выборке все модели продемонстрировали характеристики высокой производительности при верификации образцов SARS-CoV-2+. SVM-R и DT незначительно превзошли другие пять моделей в определении инфекции SARS-CoV-2 от ОРИ и контроль с показателями ROC AUC 0,983 [0,958, 0,987] и 0,972 [0,966, 0,979] соответственно. SVM-R, в частности, разделил образцы SARS-CoV-2+, ОРИ и контроль из Казахстана с точностью 88,0, 95,0 и 78,0% соответственно. Оба алгоритма, SVM-R и DT, также были очень точными в различении подгрупп ОРИ и контроля [18].

Исходя из этого, полученные данные указывают на то, что MALDI-TOF MS/ML является высокопроизводительным и инновационным диагностическим инструментом для определения различных ОРИ. При этом, следует помнить о важности проведения дополнительной работы в области оптимизации и валидации перед тем, как начать полноценное внедрение MALDI-TOF MS/ML в повседневную практику врачей. Особенно это актуально в связи с быстрым изменением эпидемиологической ситуации. Для повышения точности диагностики, проводимой с помощью MALDI-TOF MS, необходимо исследовать и выявлять ключевые компоненты спектров, которые могут быть использованы в качестве биомаркеров для различных ОРИ.

Необходимо принимать во внимание, что технологии, такие как MALDI-TOF MS, должны адаптироваться к местным потребностям и особенностям. Для успешной адаптации данной технологии необходимо знать характер распространения инфекции и вариабельность патогенов в разных регионах. В результате, для того, чтобы эффективно использовать технологии машинного обучения в диагностике ОРИ, необходимо тесное взаимодействие

между исследовательскими, клиническими и государственными учреждениями. Данная система, основанная на интеграции и согласованной работе, позволит не только повысить качество диагностики, но и оперативно реагировать на появление новых инфекционных агентов.

Таким образом, были сформированы следующие **выводы**:

1. Собранные данные от участников группы 1 и группы 2, полученные при анализе FAMD, демонстрируют сопоставимость по ключевым критериям: пол ( $p=0,692$ ), беременность ( $p=0,553$ ), курение ( $p=0,405$ ), иммунодефицит, включая ВИЧ ( $p=1,000$ ), онкопатология ( $p=0,324$ ), болезни ЖКТ ( $p=0,981$ ), посещение стран и регионов с неблагоприятной эпидемиологической обстановкой за последние 30 дней до симптомов ( $p=1,000$ ), возраст ( $p<0,0001$ ), но при этом медианы возраста находятся в одной возрастной группе. Эти данные дают основание использовать биоматериал назальных мазков от участников групп для проведения масс-спектрометрии.

2. Согласно ПЦР-тесту в группе с ОРИ 72,4% дали положительный результат на распространенные респираторные патогены: Bordetellapertussis, Chlamydia pneumoniae, Mycoplasma pneumoniae, Influenza H3/N2, Influenza A/H1, Influenza B virus, Human enterovirus, Human respiratory syncytial virus A, Human coronavirus NL63, Human Adenovirus B, C, E; Human bocavirus, Human coronavirus NL 63, 229 E, HKU 1/OC43, Human metapneumovirus, Human rhinovirus; Human parainfluenza virus 1-4, Human respiratory syncytial virus. Процент ПЦР- подтвержденных случаев SARS-CoV-2 в основной группе равен 100.

3. По результатам проведения масс-спектрометрии и препроцессинга идентифицировано 120 пиков, обеспечивающих вклад в детекцию SARS-CoV-2 и построена матрица интенсивности из объединенных данных (Южная Америка, Казахстан).

4. При оценке характеристик производительности оригинальной модели на Казахстанских данных были получены низкие показатели, ROC AUC составила DT -0.62, KNN - 0.56, NB - 0.65, RF - 0.67, SVM-L - 0.74, SVM-R - 0,60, XGBoost - 0.62.

5. Дообученные модели машинного обучения показали высокую производительность собственных моделей машинного обучения, ROC AUC составила: DT -0.972, KNN - 0.92, NB - 0.801, RF - 0.93, SVM-L - 0.94, SVM-R - 0.983, XGBoost - 0.958.

6. Оптимизация комбинированного метода MS\ML позволила достичь высоких результатов производительности модели. ROC AUC оригинальной модели на данных Казахстана: ROC AUC анализ I: RF - 0.67, SVM-R - 0,60; ROC AUC анализ II: RF- 0.93, SVM-R - 0,98.

*Практические рекомендации:*

1. Диссертация имеет практическую ценность, которая заключается в оптимизации нового эффективного и экономически выгодного метода детекции SARS-CoV-2 в назальных мазках.

2. Результаты проведенного исследования открывают перспективы

развития таргетного лечения заболеваний, вызванных вирусами ввиду их быстрой и экономически выгодной диагностики.

3. Результаты проведенного исследования могут быть использованы для дифференциальной диагностики пациентов, поступивших в стационар по экстренным показаниям с симптомами простуды к основному заболеванию, о чем свидетельствует акт внедрения в клиническую практику Клиники НАО «МУК».

4. Комбинированное использование MALDI-TOF MS и методов машинного обучения может быть применено для скрининга населения с целью выявления SARS-CoV-2 и ОРВИ.

Таким образом, актуальность диссертации обоснована значительным по объему обзором и предложенной на его анализе гипотезы исследования: «Возможность детекции SARS-CoV-2 в назальных мазках при помощи MALDI-TOF MS и методов машинного обучения». Обоснование проблемы: использование MALDI-TOFMS и методов машинного обучения для быстрой и экономически выгодной диагностики ОРВИ и COVID-19. Содержание диссертации отражает тему исследования. Все разделы диссертации выполнены в методической последовательности и логически взаимосвязаны, с соблюдением правила внутреннего единства. Практическая значимость и научная новизна соответствует поставленной цели, задачам и названию диссертации.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 COVID-19 epidemiological update // <https://www.who.int/>. 1.11.2023.
- 2 Kadyrova I., Yegorov S., Negmetzhanov B. et al. High SARS-CoV-2 seroprevalence in Karaganda, Kazakhstan before the launch of COVID-19 vaccination // *PLoS One*. – 2022. – Vol. 17, Issue 7. – P. e0272008-1e0272008-11.
- 3 Yegorov S., Goremykina M., Ivanova R. et al. Epidemiology, clinical characteristics, and virologic features of COVID-19 patients in Kazakhstan: A nation-wide retrospective cohort study // *The Lancet Regional Health – Europe*. – 2021. – Vol. 4. – 100096.
- 4 Yüce M., Filiztekin E., Özkaya K.G. COVID-19 diagnosis - A review of current methods // *BiosensBioelectron*. – 2021. – Vol. 172. – P. 112752.
- 5 Taleghani N., Taghipour F. Diagnosis of COVID-19 for controlling the pandemic: A review of the state-of-the-art // *BiosensBioelectron*. – 2021. – Vol. 174. – P. 112830.
- 6 Böger B., Fachi M., Vilhena R. et al. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19 // *American Journal of Infection Control*. – 2021. – Vol. 49(1). – P. 21-29.
- 7 Dramé M., TabueTeguo M., Proye E. et al. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? // *J Med Virol*. – 2020. – Vol. 92(11). – P. 2312-2313.
- 8 Garg A., Ghoshal U., Patel S.S. et al. Evaluation of seven commercial RT-PCR kits for COVID-19 testing in pooled clinical specimens // *J Med Virol*. – 2021. – Vol. 93(4). – P. 2281-2286.
- 9 Torretta S., Zuccotti G., Cristofaro V. et al. Diagnosis of SARS-CoV-2 by RT-PCR Using Different Sample Sources: Review of the Literature // *Ear Nose Throat J*. – 2021. – Vol. 100, Suppl. 2. – P. 131S-138S.
- 10 Rostami A., Sepidarkish M., LeeFlang M.M.G. et al. SARS-CoV-2 seroprevalence worldwide: a systematic review and meta-analysis // *Clin Microbiol Infect*. – 2021. – Vol. 27(3). P. 331-340.
- 11 Muruato A.E., Fontes-Garfias C.R., Ren P. et al. A high-throughput neutralizing antibody assay for COVID-19 diagnosis and vaccine evaluation // *Nat Commun*. – 2020. – Vol. 11(1). – P. 4059-1-4059-6.
- 12 Tan C.W., Chia W.N., Qin X. et al. A SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage of ACE2-spike protein-protein interaction // *Nat Biotechnol*. – 2020. – Vol. 38(9). – P. 1073-1078.
- 13 Zhu N., Zhang D., Wang W. et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019 // *N Engl J Med*. – 2020. – Vol 382(8). – P. 727-733.
- 14 Spick M., Lewis H.M., Wilde M.J. et al. Systematic review with meta-analysis of diagnostic test accuracy for COVID-19 by mass spectrometry // *Metabolism Clinical and Experimental*. – 2022. – Vol. 126. – P. 154922-1-15422-9.
- 15 Nachtigall F.M., Pereira A., Trofymchuk O.S. et al. Detection of SARS-CoV-2 in nasal swabs using MALDI MS // *Nat Biotechnol*. – 2020. – Vol. 38, Issue 10. – P. 1168-1173.

16 Deulofeu M., García-Cuesta E., Peña-Méndez E.M. et al. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining MALDI-TOF MS and Artificial Intelligence // *Front Med (Lausanne)*. – 2021. – Vol. 8. – P. 661358-1-661358-12.

17 Tran N.K. et al. Novel application of automated machine learning with MALDI-TOF-MS for rapid high-throughput screening of COVID-19: a proof of concept // *Sci Rep*. – 2021. – Vol. 11, Issue 1. – P. 8219-1-8219-10.

18 Кадырова И.А. Обнаружение SARS-COV-2 в назальных мазках с помощью MALDI MS и методов машинного обучения: отчет о НИР (заключительный) // <https://muk.qmu.kz/ru/nauka/nauchnye-proekty>. 10.05.2024.

19 Hall M.A. Correlation-based feature selection for discrete and numeric class machine learning // *Proceed. of 17th internat. conf. on Machine Learning (ICML)*. – Уолтем (Массачусетс), 2000. – P. 359-366.

20 Haixiang G., Yijing L., Shang J. et al. Learning from class-imbalanced data: review of methods and applications // *Expert Systems with Applications*. – 2017. – Vol. 73. – P. 220-239.

21 Kassambara A., Mundt F. factoextra: Extract and visualize the results of multivariate data analyses // <https://rpkggs.datanovia.com/factoextra>. 20.10.2023.

22 Kruse R., Borgelt C. Data Mining with Graphical Models. in *Discovery Science // Lecture Notes in Computer Science*. – 2002. – Vol. 2534. – P. 2-11.

23 Vapnik V.N. *The Nature of Statistical Learning Theory*. – Cham: Springer, 2005. – 311 p.

24 Fix E. et al. Discriminatory analysis. Nonparametric discrimination: consistency properties // *Int. Stat. Rev.* – 1989. – Vol. 57. – P. 238-247.

25 Fakiola M., Mishra A., Rai M. et al. Classification and regression tree and spatial analyses reveal geographic heterogeneity in genome wide linkage study of Indian visceral leishmaniasis // *PLoS ONE*. – 2010. – Vol. 5(12). – P. e15807.

26 Kajdanowicz T., Kazienko P. Boosting-based multi-label classification // *J. Univers. Comput. Sci.* – 2013. – Vol. 19(4). – P. 502-520.

27 WHO Coronavirus disease (COVID-19) pandemic // <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. 20.04.2022.

28 Ji T., Liu Z., Wang G. et al. Detection of COVID-19: A review of the current literature and future perspectives // *Biosensors and Bioelectronics*. – 2020. – Vol. 166. – P. 1-18.

29 Masters P.S. Coronavirus genomic RNA packaging // *Virology*. – 2019. – Vol. 537. – P. 198-207.

30 Кадырова И.А., Султанбекова А.А., Барханская В.И. и др. Применение MALDI-MS масс-спектрометрии для детекции SARS-CoV-2 в назофарингиальных мазках. Систематический обзор // *Наука и Здравоохранение*. – 2022. – Т. 24, №5. – С. 36-44.

31 Cobo F. Application of MALDI-TOF Mass Spectrometry in Clinical Virology: A Review // *The Open Virology Journal*. – 2013. – №7. – P. 84-90.

32 Султанбекова А.А. Опыт применения время пролетной масс-спектрометрии (MALDI-TOF) для детекции вирусных патогенов, в частности

коронавируса. Систематический обзор // Фармация Казахстана. – 2023. – №4(249). – С. 97-105.

33 van Belkum A, Welker M, Erhard M, Chatellier S. Biomedical mass spectrometry in today's and tomorrow's clinical microbiology laboratories // J Clin Microbiol. – 2012. – Vol. 50. – P. 1513-1517.

34 Баранцевич Е.П., Баранцевич Н.Е. Применение MALDI-TOF масс-спектрометрии в клинической микробиологии // Трансляционная медицина. – 2014. – №3. – С. 23-28.

35 B. Buszewski E. et al. A new approach to identifying pathogens, with particular regard to viruses, based on capillary electrophoresis and other analytical techniques // Trends in Analytical Chemistry. – 2021. – Vol. 139. – P. 116250.

36 Dingle T.C., Butler-Wu S.M. MALDI-TOF mass spectrometry for microorganism identification // Clin. Lab. Med. – 2013. – Vol. 33. – P. 589-609.

37 Buszewski B., Rogowska A., Pomastowski P. et al. Railean-Plugaru, Identification of microorganisms by modern analytical techniques // J. AOAC Int. – 2017. – Vol. 100. – P. 1607-1623.

38 Bourassa L., Butler-Wu S.M. Chapter 2 - MALDI-TOF Mass Spectrometry for Microorganism Identification // Methods in Microbiology. – 2015. – Vol. 42. – P. 37-85.

39 Croxatto A., Prod'hom G., Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology // FEMS Microbiol. Rev. – 2012. – Vol. 36. – P. 380-407.

40 Pomastowski P., Kupczyk W., Jackowski M. et al. Evaluation of intact cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for capillary electrophoresis detection of controlled bacterial clumping // J. Anal. Bioanal. Tech. – 2015. – Issue 13. – P. 1-7.

41 Hou T.Y., Chiang-Ni C., Teng S.H. Current status of MALDI-TOF mass spectrometry in clinical microbiology // J. Food Drug Anal. – 2019. – Vol. 27. – P. 404-414.

42 Gould O., Ratcliffe N., Krol E. et al. Breath analysis for detection of viral infection, the current position of the field // J. Breath Res. – 2020. – Vol. 14. – P. 041001.

43 Adzitey F., Huda N., Ali G.R.R. Molecular techniques for detecting and typing of bacteria, advantages and application to foodborne pathogens isolated from ducks // Biotech. – 2013. – Vol. 3. – P. 97-107.

44 Bergallo M., Costa C., Gribaudo G. et al. Evaluation of six methods for extraction and purification of viral DNA from urine and serum samples // New Microbiol. – 2006. – Vol. 29. – P. 111-119.

45 Drost C., Panning M., Guenther S. et al. False-negative results of PCR assay with plasma of patients with severe viral hemorrhagic fever // J. Clin. Microbiol. – 2002. – Vol. 40. – P. 4394-4395.

- 46 Green K., Winter A., Dickinson R. et al. What tests could potentially be used for the screening , diagnosis and monitoring of COVID-19 and what are their advantages and disadvantages? // <https://www.cebm.net/covid-19/what>. 04.04.2023.
- 47 Yüce M., Filiztekin E. et al. COVID-19 diagnosis – A review of current methods // *Biosens Bioelectron.* – 2021. – Vol. 172. – P. 112752-1-112752-16.
- 48 Landry M.L., Eid T., Bannykh S. et al. False negative PCR despite high levels of JC virus DNA in spinal fluid: implications for diagnostic testing // *J. Clin. Virol.* – 2008. – Vol. 43. – P. 247-249.
- 49 Artika I.M., Wiyatno A., Ma'roef C.N. Pathogenic viruses: molecular detection and characterization // *Infect. Genet. Evol.* – 2020. – Vol. 81. – P. 104215.
- 50 Nolan T., Hands R.E., Bustin S.A. Quantification of mRNA using real-time RTPCR // *Nat. Protoc.* – 2006. – Vol. 1. – P. 1559-1582.
- 51 Du´s-Ilnicka I., Szymczak A., Małodobra-Mazur M. Tokarski, M. Role of Laboratory Medicine in SARS-CoV-2 Diagnostics. Lessons Learned from a Pandemic // *Healthcare.* – 2021. – Vol. 9. – P. 915-1-915-14.
- 52 Cattoli G., De Battisti C., Marciano S. et al. False-negative results of a validated real-time PCR protocol for diagnosis of newcastle disease due to genetic variability of the matrix gene // *J. Clin. Microbiol.* – 2009. – Vol. 47. – P. 3791-3792.
- 53 Younes N. et al. Challenges in Laboratory Diagnosis of the Novel Coronavirus SARS-CoV-2 // *Viruses.* – 2020. – Vol. 12. – P. 582-1-582-27.
- 54 Notomi T., Okayama H., Masubuchi H. et al. Loop-mediated isothermal amplification of DNA // *Nucleic Acids Res.* – 2000. – Vol. 28. – P. e63-1-e63-7.
- 55 Ganguli A., Mostafa A., Berger J. et al. Rapid isothermal amplification and portable detection system for SARS-CoV-2 // *Proc. Natl. Acad. Sci.* – 2020. – Vol. 117. – P. 22727-22735.
- 56 Janíková M., Hodosy J., Boor P. et al. Loop-mediated isothermal amplification for the detection of SARS-CoV-2 in saliva. *Microb // Biotechnol.* – 2021. – Vol. 14. – P. 307-316.
- 57 Rödel J., Egerer R., Suleyman A. et al. Use of the variplex™ SARSCoV-2 RT-LAMP as a rapid molecular assay to complement RT-PCR for COVID-19 diagnosis // *J. Clin. Virol.* – 2020. – Vol. 132. – P. 104616.
- 58 Nagura-Ikeda M. et al. Clinical Evaluation of Self-Collected Saliva by Quantitative Reverse Transcription-PCR (RT-qPCR), Direct RT-qPCR, Reverse Transcription–Loop-Mediated Isothermal Amplification, and a Rapid Antigen Test to Diagnose COVID-19 // *J. Clin. Microbiol.* – 2020. – Vol. 58. – P. 1-9.
- 59 Quantitative Synthetic SARS CoV 2 RNA: nsp9, nsp12 (RdRp) // <https://www.atcc.org/products/vr-3279sd>. 15.07.2021.
- 60 Lu R., Wu X., Wan Z. et al. A novel reverse transcription loop-mediated isothermal amplification method for rapid detection of sars-cov-2 // *Int. J. Mol. Sci.* – 2020. – Vol. 21. – P. 2826-1-2826-10.
- 61 Annamalai P., Kanta M., Ramu P. et al. A simple colorimetric molecular detection of novel coronavirus (COVID-19), an essential diagnostic tool for pandemic screening // *medRxiv.* – 2020. – Vol. 91. – P. 1-20.

- 62 Park G.-S. et al. Development of Reverse Transcription Loop-Mediated Isothermal Amplification Assays Targeting Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) // *J. Mol. Diagn.* – 2020. – Vol. 22. – P. 729-735.
- 63 Huang W.E., Lim B., Hsu C.C. et al. RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2 // *Microb. Biotechnol.* – 2020. – Vol. 13. – P. 950-961.
- 64 Ren L.L., Wang Y.M., Wu Z.Q. et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study // *Chin Med J (Engl.)*. – 2020. – Vol. 133. – P. 1015-1024.
- 65 Tao Y., Fu Q., Mo Q. Advantages and challenges of metagenomics nextgeneration sequencing (mNGS) in the detection of 2019 novel coronavirus // *Chin J Lab Med.* – 2020. – Vol. 43. – P. 217-220.
- 66 Chen Z. et al. Clinical laboratory evaluation of COVID-19 // *Clinica Chimica Acta.* – 2021. – Vol. 519. – P. 172-182.
- 67 Mekonnen D., Mengist H.M., Derby A. et al. Diagnostic Accuracy of Serological Tests and Kinetics of Severe Acute Respiratory Syndrome Coronavirus 2 Antibody: A Systematic Review and Meta-Analysis // *Rev. Med. Virol.* – 2020. – Vol. 31. – P. e2181.
- 68 Varadhachary A., Chatterjee D., Garza J. et al. Salivary Anti-SARS-CoV-2 IgA as an Accessible Biomarker of Mucosal Immunity against COVID-19 // <https://www.medrxiv.org/content/10.1101/2020.08.07.20170258v1>. 10.05.2023.
- 69 Isho B., Abe K.T., Zuo M. et al. Persistence of Serum and Saliva Antibody Responses to SARS-CoV-2 Spike Antigens in COVID-19 Patients // *Sci. Immunol.* – 2020. – Vol. 5. – P. eabe5511.
- 70 Mahallawi W.H. A Serological Assay to Detect Human SARS-CoV-2 Antibodies // *J. Taibah Univ. Med. Sci.* – 2021. – Vol. 16. – P. 57-62.
- 71 Lassaunière R., Frische A., Harboe Z.B. et al. Evaluation of Nine Commercial SARS-CoV-2 Immunoassays // [doi:10.1101/2020.04.09.20056325](https://doi.org/10.1101/2020.04.09.20056325)
- 72 Whitman J.D., Hiatt J., Mowery C.T. et al. Test Performance Evaluation of SARS-CoV-2 Serological Assays // <https://www.medrxiv.org/content/10.1101/2020.04.22.20056325>.
- 73 Vashist S.K. Vitro Diagnostic Assays for COVID-19: Recent Advances and Emerging Trends // *Diag. (Basel)*. – 2020. – Vol. 10, Issue 4. – P. 202-1-202-7.
- 74 Lin S.B., Zheng Z.X., Zhang R. Application and Evaluation of Chemiluminescence Immunoassay in Blood Screening // *Zhongguo Shi Yan Xue Ye Xue Za Zhi.* – 2019. – Vol. 27. – P. 569-572.
- 75 Tan Y., Pang W., Jia X. et al. Comparison of the Performance of a Chemiluminescence Assay and an ELISA for Detection of Anti-GBM Antibodies // *Ren. Fail.* – 2020. – Vol. 42. – P. 48-53.
- 76 Padoan A., Cosma C., Sciacovelli L. et al. Analytical Performances of a Chemiluminescence Immunoassay for SARSCoV-2 IgM/IgG and Antibody Kinetics // *Clin. Chem. Lab. Med.* – 2020. – Vol. 58. – P. 1081-1088.
- 77 Cai X., Chen J., Hu J. et al. A Peptide-Based Magnetic Chemiluminescence Enzyme Immunoassay for Serological Diagnosis of (COVID-19) // <https://www.medrxiv.org/content/10.1101/2020.02.22.20026617v1>. 10.05.2023.

78 Ma H., Zeng W., He H. et al. Serum IgA, IgM, and IgG Responses in COVID-19 // *Cell Mol Immunol.* – 2020. – Vol. 17. – P. 773-775.

79 Nicol T., Lefeuvre C., Serri O. et al. Assessment of SARS-CoV-2 Serological Tests for the Diagnosis of COVID-19 through the Evaluation of Three Immunoassays: Two Automated Immunoassays (Euroimmun and Abbott) and One Rapid Lateral Flow Immunoassay (NG Biotech) // *J. Clin. Virol.* – 2020. – Vol. 129. – P. 104511-1-104511-8.

80 Carrio A., Sampedro C., Sanchez-Lopez J.L. et al. Automated Low-Cost Smartphone-Based Lateral Flow Saliva Test Reader for Drugs-Of-Abuse Detection // *Sensors (Basel).* – 2015. – Vol. 15. – P. 29569-29593.

81 Deeks J.J., Dinnes J., Takwoingi Y. et al. Antibody Tests for Identification of Current and Past Infection with SARS-CoV-2 // *Cochrane Database Syst. Rev.* – 2020. – Vol. 6. – P. CD013652.

82 Ghaffari A., Meurant R., Ardakani A. COVID-19 Serological Tests: How Well Do They Actually Perform? // *Diagnostics (Basel).* – 2020. – Vol. 10. – P. 453-1-453-14.

83 Montesinos I., Gruson D., Kabamba B. et al. Evaluation of Two Automated and Three Rapid Lateral Flow Immunoassays for the Detection of Anti-SARS-CoV-2 Antibodies // *J. Clin. Virol.* – 2020. – Vol. 128. – P. 104413-1-104413-7.

84 Hu X. et al. EDTA-K2 Improves the Detection Sensitivity of SARS-CoV-2 IgM and IgG Antibodies by Chelating Colloidal Gold in the Immunochromatographic Assay // *Int. J. Nanomedicine.* – 2021. – Vol. 16. – P. 715-724.

85 Goudouris E.S. Laboratory Diagnosis of COVID-19 // *J. Pediatr.* – 2021. – Vol. 97. – P. 7-12.

86 La Rosa Fabian C., Urquizo Briceno L. Anti-SARS-Cov-2 IgA in Current Scenario of IgM and IgG Rapid Test: a New Alternative for the Diagnostic of COVID-19 // *SN Compr. Clin. Med.* – 2020. – Vol. 2. – P. 2167-2169.

87 Schwahn A.B., Wong J.W.H., Downard K.M. Rapid differentiation of seasonal and pandemic H1N1 influenza through proteotyping of viral neuraminidase with mass spectrometry // *Anal. Chem.* – 2010. – Vol. 82. – P. 4584-4590.

88 Majchrzykiewicz-Koehorst J.A., Heikens E., Trip H. et al. Rapid and generic identification of influenza A and other respiratory viruses with mass spectrometry // *J Virol Methods.* – 2015. – Vol. 213. – P. 75-83.

89 Musaji A., Fahlman R., Charlton C. Mass spectrometry of influenza virus using clinically available MALDI-TOF platform // *J. Clin. Virol.* – 2016. – Vol. 82. – P. S45.

90 Calderaro A., Arcangeletti M., Rodighiero I. et al. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry applied to virus identification // *Scientific Reports.* – 2014. – Vol. 4. – P. 6803-1-6803-10.

91 Tian L., Wang Y., Xu D. et al. The differential diagnostic model for serous peptidomics in HBV carrier established by MALDI-TOF-MS analysis // *Clinical Biochemistry.* – 2014. – Vol. 47. – P. 56-62.

92 Ren Y., Yang L., Li M. et al. 4210 Da and 1866 Da polypeptides as potential biomarkers of liver disease progression in hepatitis B virus patients // *Scientific Reports*. – 2021. – Vol. 11. – P. 16982-1-16982-12.

93 Cao Y., He K., Cheng M. et al. Two Classifiers Based on Serum Peptide Pattern for Prediction of HBV-Induced Liver Cirrhosis Using MALDI-TOF MS // *BioMed Research International*. – 2013. – Vol. 2013. – P. 814876-1-814876-7.

94 Cai X., Guan Q., Huan Y. et al. Development of high-throughput genotyping method of all 18 HR HPV based on the MALDI-TOF MS platform and compared with the Roche Cobas 4800 HPV assay using clinical specimens // *BMC Cancer*. – 2019. – Vol. 19. – P. 825-1-825-9.

95 Hernández-Rosas F., Orozco-Hernández E., Maza-Sánchez L. et al. Prevalence and correlation of human papillomavirus genotypes with clinical factors in cervical samples from Mexican women // *Exp Biol Med*. – 2021. – Vol. 246, Issue 1. – P. 48-56.

96 Yan L., Yi J., Huang C. et al. Rapid Detection of COVID-19 Using MALDI-TOF-Based Serum Peptidome Profiling // *Analytical Chemistry*. – 2021. – Vol. 93. – P. 4782-4787.

97 Lazari L.C., Ghilardi F.R., Fernandes L.R. et al. Prognostic accuracy of MALDI-TOF mass spectrometric analysis of plasma in COVID-19 // *Life science alliance*. – 2022. – Vol. 4, Issue 8. – P. e202000946-1-e202000946-12.

98 Gomila R.M., Martorell G., Fraile-Ribot P.A. et al. Use of Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry Analysis of Serum Peptidome to Classify and Predict Coronavirus Disease // *Open Forum Infectious Diseases*. – 2021. – Vol. 8, Issue 6. – P. ofab222-1-fab22-8.

99 Wan Q., Chen M., Zhang Z. et al. Machine Learning of Serum Metabolic Patterns Encodes Asymptomatic SARS-CoV-2 Infection // *Frontiers in Chemistry*. – 2021. – Vol. 9. – P. 746134-1-746134-9.

100 Iles R.K., Zmuidinaite R., Iles J.K. et al. Development of a Clinical MALDI-ToF Mass Spectrometry Assay for SARS-CoV-2: Rational Design and Multi-Disciplinary Team Work // *Diagnostics*. – 2020. – Vol. 10. – P. 1-15.

101 Chivte P., LaCasse Z., Seethi V.D.R. et al. MALDI-ToF protein profiling as a potential rapid diagnostic platform for COVID-19 // *Journal of Mass Spectrometry and Advances in the Clinical Lab*. – 2021. – Vol. 21. – P. 31-41.

102 Costa J., Ferreira E.C., Santos C. COVID-19, Chikungunya, Dengue and Zika Diseases: An Analytical Platform Based on MALDI-TOF MS, IR Spectroscopy and RT-qPCR for Accurate Diagnosis and Accelerate Epidemics Control // *Microorganisms*. – 2021. – Vol. 9, Issue 4. – P. 708-1-708-15.

103 Rocca M.F., Zintgraff J.C., Dattero M.E. et al. A combined approach of MALDI-TOF mass spectrometry and multivariate analysis as a potential tool for the detection of SARS-CoV-2 virus in nasopharyngeal swabs // *Journal of Virological Methods*. – 2020. – Vol. 286. – P. 113991.

104 Zhao F., Lu J., Lu B. et al. A Novel Strategy for the Detection of SARS-CoV-2 Variants Based on Multiplex PCR-Mass Spectrometry Minisequencing Technology // *Microbiology Spectrum*. – 2021. – Vol. 9, Issue 3. – P. 1-10.

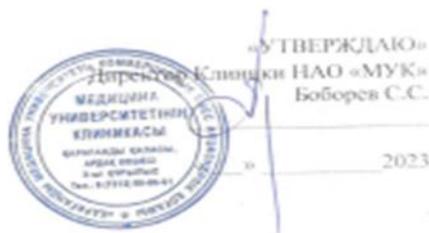
- 105 Rashidi H.H., Pepper J., Howard T. et al. Comparative performance of two automated machine learning platforms for COVID-19 detection by MALDI-TOF-MS // *PLoS ONE*. – 2022. – Vol. 17, Issue 7. – P. e0263954.
- 106 Wandernoth P., Kriegsmann K., Groh-Mohanu C. et al. Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by Mass Spectrometry // *Viruses*. – 2020. – Vol. 12. – P. 1-8.
- 107 Rybicka M., Miłosz E., Bielawski K.P. Superiority of MALDI-TOF Mass Spectrometry over Real-Time PCR for SARS-CoV-2 RNA Detection // *Viruses*. – 2021. – Vol. 13. – P. 1-11.
- 108 Gibb S., Strimmer K. MALDIquant: a versatile R package for the analysis of mass spectrometry data // *Bioinformatics*. – 2012. – Vol. 28, Issue 17. – P. 2270-2271.
- 109 Rigatti S.J. Random Forest // *J Insur Med*. – 2017. – Vol. 47, Issue 1. – P. 31-39.
- 110 Karalis G. Decision Trees and Applications // *Adv Exp Med Biol*. – 2020. – Vol. 1194. – P. 239-242.
- 111 Lê S., Josse J., Husson F. FactoMineR: An R Package for Multivariate Analysis // *Journal of Statistical Software*. – 2008. – Vol. 25, Issue 1. – P. 1-18.
- 112 Yu G., Smith D.K., Zhu H. et al. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data // *Methods in Ecology and Evolution*. – 2017. – Vol. 8, Issue 1. – P. 28-36.
- 113 Yu G. Using ggtree to Visualize Data on Tree-Like Structures // *Current protocols in bioinformatics*. – 2020. – Vol. 69, Issue 1. – P. e96.
- 114 Wang H., Gu Z., Eils R. et al. ggtreeExtra: exporting tree to other formats and utilities for tree manipulation // *Bioinformatics*. – 2021. – Vol. 37, Issue 24. – P. 4979-4980.
- 115 Barkalov K., Shtanyuk A., Sysoyev A. A Fast kNN Algorithm Using Multiple Space-Filling Curves // *Entropy (Basel)*. – 2022. – Vol. 24, Issue 6. – P. 767-1-767-18.
- 116 Zhang Z. Naïve Bayes classification in R // *Ann Transl Med*. – 2016. – Vol. 4, Issue 12. – P. 241-1-241-5.
- 117 Song X., Zhu J., Tan X. et al. XGBoost-Based Feature Learning Method for Mining COVID-19 Novel Diagnostic Markers // *Front Public Health*. – 2022. – Vol. 10. – P. 926069-1-926069-6.
- 118 Ozer M.E., Sarica P.O., Arga K.Y. New Machine Learning Applications to Accelerate Personalized Medicine in Breast Cancer: Rise of the Support Vector Machines // *OMICS*. – 2020. – Vol. 24, Issue 5. – P. 241-246.
- 119 ROC-кривая (в медицине) // <https://bigenc.ru/c/roc-krivaia-8c39b7>. 10.11.2023.
- 120 Rajagopala S.V., Bakhoun N.G., Pakala S.B. et al. Metatranscriptomics to characterize respiratory virome, microbiome, and host response directly from clinical samples // *Cell Rep Methods*. – 2021. – Vol. 1, Issue 6. – P. 100091.
- 121 Sandybayev N.T., Belousov V.Y., Stochkov V.M. et al. The nasopharyngeal virome in adults with acute respiratory infection //

<https://www.biorxiv.org/content/10.1101/2023.08.21.554191v1>. 10.11.2023.

122 Renuse S., Vanderboom P.M., Maus A.D. et al. A mass spectrometry-based targeted assay for detection of SARS-CoV-2 antigen from clinical specimens // *EBioMedicine*. – 2021. – Vol. 69. – P. 103465.

# ПРИЛОЖЕНИЕ А

## Акты внедрения



### АКТ

#### Внедрения результатов научно-исследовательской работы

В Клинику НАО МУК г. Караганды внедрена разработка докторанта PhD Султанбековой А.А. «Экспресс-метод определения вируса SARS-CoV-2 в материале, полученном из назального мазка при помощи время-пролетной массе-спектрометрии и методов машинного обучения».

Работа внедрена в инициативном порядке, при выполнении PhD диссертации Султанбековой А.А. «Детекция SARS-CoV-2 в назальных мазках на основе время-пролетной массе-спектрометрии (MALDI-MS)».

Форма внедрения: «Экспресс-метод определения вируса SARS-CoV-2 в материале, полученном из назального мазка при помощи время-пролетной массе-спектрометрии и методов машинного обучения» (Свидетельство о государственной регистрации прав на объект авторского права № 24868 от «7» апреля 2022 года, авторы Кадьрова И.А., Колесниченко С. И., Коршуков И.В., Барханская В.И., Султанбекова А. А., Колесникова Е.А., Бабенко Д.Б.).

Ответственный за внедрение и исполнитель: Докторант НАО МУК Султанбекова А.А.

Эффективность внедрения определяется способностью различать положительные и отрицательные образцы SARS-CoV-2 с помощью массе-спектрометрии и методов машинного обучения.

Предложения: для более широкого внедрения необходимо провести дальнейшую апробацию метода.

Сроки внедрения в течение 2023-2024 годы.

**Председатель комиссии:**  
Главный хирург КМУ МУК  
к.м.н., профессор

 Абатов И.Т.

**Члены комиссии**  
(ответственные за внедрение):  
С.п.с. НИЛ ИНОЖ НАО МУК  
PhD, асс. Профессор

 Кадьрова И.А.

Заведующий  
НИЛ ИНОЖ НАО МУК

 Лавриненко А.В.

Докторант Эгола  
Специальности «медицина»

 Султанбекова А.А.

Ф КГМУ 4/7-03/01  
ПП КГМУ 4/01

 «УТВЕРЖДАЮ»  
Ректор НАО МУК  
Турмухамбетова А.А.  
«\_\_» \_\_\_\_\_ 2023г.

**АКТ**

**Внедрения результатов научно-исследовательской работы  
НАО «Карагандинский медицинский университет»**

**Наименование предложения** Экспресс-метод определения вируса SARS-CoV-2 в материале, полученном из назального мазка при помощи время-пролетной масс-спектрометрии и методов машинного обучения

**Работа включена в перечень методик НИЛ ИНОЖ НАО МУК**

**Форма внедрения:** метод для научных исследований

**Ответственный за внедрение и исполнители** Кадырова И.А., Султанбекова А.А.

**Эффективность внедрения:** научно-исследовательская.

**Предложения, замечания учреждения, осуществляющего внедрение:** Рекомендуется для внедрения рабочий процесс НИЛ ИНОЖ НАО МУК

**Сроки внедрения в течение 2023 года** \_\_\_\_\_

**Председатель комиссии:**  
Директор ИНОЖ НАО МУК  
к.м.н., ассон. профессор

 Клоев Д.А.

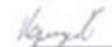
**Члены комиссии**  
(ответственные за внедрение):  
С.н.с. НИЛ ИНОЖ НАО МУК  
PhD, ассон. Профессор

 Кадырова И.А.

Докторант 3года  
Специальности «медицина»

 Султанбекова А.А.

Ассистент профессора  
кафедры информатики и  
биостатистики

 Коршуков И.В.

стажер исследователь НИЛ ИНОЖ НАО МУК

 Барханская В.И.

## ПРИЛОЖЕНИЕ Б

### Свидетельство об авторском праве

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН

АВТОРЛЫҚ ҚҰҚЫҚПЕН ҚОРҒАЛАТЫН ОБЪЕКТІЛЕРГЕ ҚҰҚЫҚТАРДЫҢ  
МЕМЛЕКЕТТІК ТІЗІЛІМГЕ МӘЛІМЕТТЕРДІ ЕНГІЗУ ТУРАЛЫ

КУӘЛІК

2022 жылғы «7» сәуір № 24868

Автордың (лардың) жөні, аты, әкесінің аты (егер ол жеке басым куәландыратын құжатта көрсетілсе):  
ҚАДЫРОВА ПРИНА АДІЛЬЕВНА, Колесниченко Светлана Игоревна, Коршукон Илья Васильевич,  
Барданская Валентина Ильяовна, Султанбекова Айлана Асқаровна, Колеснякова Евгения  
Александровна, Бабенко Дмитрий Борисович

Авторлық құқық объектісі: ғылыми тұжырым

Объектінің атауы: РАЗРАБОТКА ЭКСПРЕСС-МЕТОДА ОПРЕДЕЛЕНИЯ ВИРУСА SARS-COV-2 В  
МАТЕРИАЛЕ, ПОЛУЧЕННОМ ИЗ НАЗАЛЬНОГО МАЗКА ПРИ ПОМОЩИ ВРЕМЯПРОЛЕТНОЙ  
МАСС-СПЕКТРОМЕТРИИ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Объектіні жасаған күні: 15.03.2022



Құжат транскрипциялы <http://www.kazpatent.kz> сайтының  
"Авторлық құқық" бөлімінде тексеруге болады. <http://copyright.kazpatent.kz>  
Подлинность документа возможно проверить на сайте [kazpatent.kz](http://kazpatent.kz)  
в разделе «Авторское право» <http://copyright.kazpatent.kz>

ЭЦҚ қол қойылды

А. Естаев

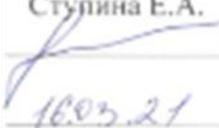
## ПРИЛОЖЕНИЕ В

### Стандартная операционная процедура «Правила анкетирования пациентов»

Правила анкетирования пациентов

СОП 06-73/01

|   |   |
|---|---|
| Наименование медицинской организации    | <b>НАО «МУК»</b><br>Лаборатория коллективного пользования НИЦ |
| Наименование структурного подразделения |   |
| <b>Название документа:</b>              | <b>Правила анкетирования пациентов</b>                        |

|  |   |  |
|--|---|--|
| <b>Разработчики:</b><br>Кадырова И.А.<br>Егоров С.В. | <b>Утверждено</b><br><br>                           | <b>Зав. ЛКП</b><br>Ахмалдинова Л.Л.<br><br><hr/>          |
|  | <b>Дата утверждения</b>                             | 15-03-2021 год   |
|  | <b>Согласовано</b>                                  | <b>Менеджер по качеству</b><br>Авдеев О.В.<br><br><hr/> |
| <b>Дата разработки:</b><br>12-03-21                  | <b>Дата согласования</b>                            | 15-03-2021 год   |
|  | <b>Ответственный за исполнение</b>                  | Ступина Е.А.<br><br>16.03.21<br><hr/> дата              |
|  | <b>Введен в действие</b>                            | 16-03- 2021 год  |
|  | <b>Сотрудник отвечающий за выполнение процедуры</b> | Ступина Е.А.   |

## **1. НАЗНАЧЕНИЕ**

Настоящая стандартная операционная процедура устанавливает порядок проведения анкетирования у пациента с новой коронавирусной инфекцией. Исполнителем является врач инфекционного центра.

## **2. ТЕХНИКА БЕЗОПАСНОСТИ И ОХРАНА ТРУДА**

Анкетирование пациентов с новой коронавирусной инфекцией выполняет медицинский персонал с использованием шапочек, халата, медицинского респиратора или маски, медицинских резиновых перчаток, очков, бахил.

## **3. Правила анкетирования персонала**

### **1. Определение:**

Опрос пациентов – это довольно эффективный инструмент, благодаря которому можно выявить симптоматику и собрать необходимую для исследования информацию.

### **2. Ресурсы:**

1) Бланки анкет / электронная форма Google forms.

### **3. Документирование:**

1) Анкета пациента;

2) Анализ результата анкетирования.

## **4. Процедура анкетирования пациентов:**

1. Анкетирование проводится врачом-инфекционистом (членом ВНК проекта).

2. Анкетирование может быть сплошным (при котором опрашиваются все пациенты) или выборочным (при котором опрашивается определенная категория разьяснением цели проведения анкетирования и правил заполнения анкеты).

3. Анкета предоставляется на государственном и/или русском языках.

4. Анкетирование проводится анонимно на добровольной основе путем заполнения бланка анкеты.

5. Пациентов необходимо заранее оповестить о предстоящем мероприятии. Кроме того необходимо донести с какой целью это делается, объяснить им, что мероприятие в их интересах и от них многое зависит.

6. Заполненные бланки анкет находятся у руководителя проекта.

7. В целях сохранения анонимности заполненные персоналом анкеты обезличены, имеют идентификационный номер и хранятся в закрытом месте.

#### **5. Процедура анализа анкет:**

1. Специалистом-аналитиком ВНК основании анкет проводится анализ.
2. Таблица анализа анкет ведется в программе Google tables.
3. Специалистом Отдела кадров результаты анкетирования оформляются в виде диаграмм из данных таблицы.
4. Анкеты хранятся у руководителя проекта до 3 лет, далее уничтожаются.
5. На основе полученной информации делается прогноз развития ситуации.
6. Анализ данных ассоциируется с данными лабораторной диагностики и приводится в отчете.

#### **6. Ссылки**

- 1) Приказ Министра здравоохранения Республики Казахстан от 2 октября 2012 года № 676 «Об утверждении стандартов аккредитации медицинских организаций» (с изменениями и дополнением от 5 июня 2018), Приложение 2, Глава 4: Лечение и уход за пациентом, Параграф 4: Лабораторные услуги и служба крови, Стандарт 90. Обращение с биоматериалом пациента;
- 2) WHO guidance for laboratories shipping specimens to WHO reference laboratories that provide confirmatory testing for COVID-19 virus, 2020;
- 3) Diagnostic testing for SARS-CoV-2, 2020.

#### **8. Формы записей**

Номера полученных анкет регистрируются в журнале регистрации биобразцов.

## Приложение

| <b>№</b> | <b>Дата забора и транспортировки</b> | <b>Ф.И.О. пациента</b> | <b>ИД образца/ анкеты</b> | <b>Направляющее материал учреждение</b> | <b>Дата</b> | <b>Подпись</b> |
|----------|--------------------------------------|------------------------|---------------------------|---|-------------|----------------|
| 1        |                                      |                        |                           |   |             |                |
| 2        |                                      |                        |                           |   |             |                |
| 3        |                                      |                        |                           |   |             |                |
| 4        |                                      |                        |                           |   |             |                |
| 5        |                                      |                        |                           |   |             |                |
| 6        |                                      |                        |                           |   |             |                |
| 7        |                                      |                        |                           |   |             |                |
| 8        |                                      |                        |                           |   |             |                |
| 9        |                                      |                        |                           |   |             |                |
| 10       |                                      |                        |                           |   |             |                |

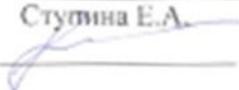
## ПРИЛОЖЕНИЕ Г

### Стандартная операционная процедура «Забор назального мазка»

Забор назального мазка

СОП

|   |  |
|---|--|
| Наименование медицинской организации    | НАО «МУК»<br>Лаборатория коллективного пользования НИЦ |
| Наименование структурного подразделения |  |
| Название документа:                     | Забор назального мазка                                 |

|   |  |  |
|---|--|--|
| Разработчики:<br>Кадырова И.А.<br>Егоров С.В. | Утверждено                                   | Зав. ЛКП<br>Ахматдинова Л.Л.<br>            |
|   | Дата утверждения                             | 15-03-2021 год   |
|   | Согласовано                                  | Менеджер по качеству<br>Авдиенко О.В.<br> |
| Дата разработки:<br>12-03-21                  | Дата согласования                            | 15-03-2021 год   |
|   | Ответственный за исполнение                  | Ступина Е.А.<br><br>16.03.21<br>дата      |
|   | Введен в действие                            | 16-03- 2021 год  |
|   | Сотрудник отвечающий за выполнение процедуры | Ступина Е.А.   |

## 1. НАЗНАЧЕНИЕ

Настоящая стандартная операционная процедура устанавливает порядок забора назального мазка у пациента, а также транспортировки биологического материала с клинической базы в Лабораторию коллективного пользования НИЦ  
Исполнитель врач инфекционного центра и персонал Лаборатории коллективного пользования НИЦ.

## 2. ТЕХНИКА БЕЗОПАСНОСТИ И ОХРАНА ТРУДА

Все процедуры по забору клинического материала выполняет медицинский персонал с использованием шапочек, халата, медицинского респиратора или маски, медицинских резиновых перчаток, очков, бахил.

## 3. ОБОРУДОВАНИЕ, МАТЕРИАЛЫ И РЕАКТИВЫ

Для транспортировки и хранения мазков из верхних дыхательных путей применять зарегистрированные транспортные среды для проведения молекулярно-генетических исследований. Допускается использование пробирок: типа «Эшпендорф» со специальной транспортной средой для хранения и транспортировки назальных мазков – 1 шт, Зонд-тампон для отбора, транспортировки и хранения биологических проб (ПС+Виск, стерильный) – 1 шт.

## 4. ПРОЦЕДУРА

1. Мазки берут сухими стерильными зондами.
2. Зонд вводят легким движением ПО НАРУЖНОЙ СТЕНКЕ полости носа на глубину 2-3 см до нижней носовой раковины.
3. Затем зонд слегка опускают книзу, вводят в нижний носовой ход под нижнюю носовую раковину и удаляют вдоль НАРУЖНОЙ СТЕНКИ полости носа, производя вращательные движения (3-4 см у детей и 5-6 см – у взрослых).



4. Конец зонда отламывают с расчетом, чтобы он позволил плотно закрыть крышку пробирки.
5. Пробирку с раствором и рабочей частью зонда закрывают.
6. В отдельный полиэтиленовый пакет вкладывают бланк направления с указанием следующих данных: наименование направляющего материал учреждения; ФИО больного; возраст больного; дата заболевания; место жительства; предварительный диагноз; эпидемиологический анамнез; страна прибытия; вид материала; дата и время отбора материала.

#### **5. УСЛОВИЯ ХРАНЕНИЯ И ТРАНСПОРТИРОВКИ МАТЕРИАЛА**

В специальном термоконтейнере с охлаждающими элементами или в термосе при температуре +2 до +8 °С – в течение 24 часов.

#### **6. КОНТРОЛЬ ПРОЦЕССА**

Контроль процесса осуществляет ответственный по биобезопасности.

#### **7. ССЫЛКИ**

- 1) Приказ Министра здравоохранения Республики Казахстан от 2 октября 2012 года № 676 «Об утверждении стандартов аккредитации медицинских организаций» (с изменениями и дополнением от 5 июня 2018), Приложение 2, Глава 4: Лечение и уход за пациентом, Параграф 4: Лабораторные услуги и служба крови, Стандарт 90. Обращение с биоматериалом пациента;
- 2) WHO guidance for laboratories shipping specimens to WHO reference laboratories that provide confirmatory testing for COVID-19 virus, 2020;
- 3) Diagnostic testing for SARS-CoV-2, 2020.

#### **8. ФОРМЫ ЗАПИСЕЙ**

Полученные образцы регистрируются в журнале регистрации биобразцов.

## Приложение

| <b>№</b>  | <b>Дата забора и транспортировки</b> | <b>Ф.И.О. пациента</b> | <b>ID образца</b> | <b>Направляющее материал учреждение</b> | <b>Дата</b> | <b>Подпись</b> |
|-----------|--------------------------------------|------------------------|-------------------|---|-------------|----------------|
| <b>1</b>  |                                      |                        |                   |   |             |                |
| <b>2</b>  |                                      |                        |                   |   |             |                |
| <b>3</b>  |                                      |                        |                   |   |             |                |
| <b>4</b>  |                                      |                        |                   |   |             |                |
| <b>5</b>  |                                      |                        |                   |   |             |                |
| <b>6</b>  |                                      |                        |                   |   |             |                |
| <b>7</b>  |                                      |                        |                   |   |             |                |
| <b>8</b>  |                                      |                        |                   |   |             |                |
| <b>9</b>  |                                      |                        |                   |   |             |                |
| <b>10</b> |                                      |                        |                   |   |             |                |

## ПРИЛОЖЕНИЕ Д

Таблица Д.1 – Масс-спектрометрия, выполненная с использованием стратегии Nachtigall и соавт. на образцах из Казахстана (Анализ I)

| Образцы из Казахстана |                            |                            |                            |                            |           |
|-----------------------|----------------------------|----------------------------|----------------------------|----------------------------|-----------|
| Peak (m/z)            | ОРИ, N = 98                | АС, N = 39                 | SARS-CoV-2/2020, N = 108   | SARS-CoV-2/2022, N = 7     | p-value   |
| 1                     | 2                          | 3                          | 4                          | 5                          | 6         |
| 3044                  | 4.5e-05 [2.3e-05; 8.1e-05] | 7.2e-05 [2.7e-05; 1.1e-04] | 8.4e-05 [4.8e-05; 1.1e-04] | 1.5e-04 [7.7e-05; 1.9e-04] | <0.001*** |
| 3095                  | 8.4e-05 [4.5e-05; 1.3e-04] | 1.1e-04 [7.6e-05; 1.5e-04] | 1.1e-04 [7.5e-05; 1.4e-04] | 1.1e-04 [5.7e-05; 2.0e-04] | 0.035*    |
| 3112                  | 5.9e-05 [2.4e-05; 1.2e-04] | 9.6e-05 [8.2e-05; 1.3e-04] | 9.0e-05 [5.2e-05; 1.3e-04] | 1.2e-05 [3.1e-06; 3.1e-05] | <0.001*** |
| 3139                  | 1.6e-04 [1.2e-04; 2.4e-04] | 1.1e-04 [7.7e-05; 1.4e-04] | 1.1e-04 [8.2e-05; 1.7e-04] | 1.8e-04 [1.2e-04; 2.3e-04] | <0.001*** |
| 3152                  | 5.4e-05 [1.6e-05; 1.5e-04] | 8.5e-05 [3.2e-05; 1.3e-04] | 7.5e-05 [4.4e-05; 1.1e-04] | 2.1e-05 [1.9e-06; 1.5e-04] | 0.7       |
| 3193                  | 1.2e-04 [7.2e-05; 1.9e-04] | 1.1e-04 [8.1e-05; 1.3e-04] | 1.2e-04 [9.5e-05; 1.9e-04] | 1.5e-04 [1.2e-04; 1.6e-04] | 0.046*    |
| 3242                  | 8.8e-05 [5.3e-05; 1.4e-04] | 1.1e-04 [7.9e-05; 1.4e-04] | 1.1e-04 [8.9e-05; 1.6e-04] | 7.5e-05 [5.6e-05; 1.7e-04] | 0.008**   |
| 3256                  | 9.7e-05 [5.2e-05; 1.5e-04] | 1.1e-04 [7.5e-05; 1.5e-04] | 7.1e-05 [4.0e-05; 1.2e-04] | 1.2e-04 [7.6e-05; 1.3e-04] | 0.005**   |
| 3297                  | 1.6e-04 [1.0e-04; 2.8e-04] | 1.1e-04 [6.9e-05; 1.3e-04] | 1.0e-04 [7.2e-05; 1.4e-04] | 1.9e-04 [7.3e-05; 2.8e-04] | <0.001*** |
| 3318                  | 1.1e-04 [4.4e-05; 1.6e-04] | 1.0e-04 [3.9e-05; 1.6e-04] | 6.8e-05 [1.9e-05; 1.0e-04] | 1.4e-04 [1.1e-04; 2.2e-04] | <0.001*** |
| 3337                  | 1.6e-04 [7.5e-05; 4.5e-04] | 3.3e-04 [1.6e-04; 4.3e-04] | 1.5e-04 [9.9e-05; 2.7e-04] | 2.3e-04 [1.8e-04; 4.0e-04] | 0.019*    |
| 3358                  | 1.4e-04 [7.6e-05; 2.0e-04] | 1.1e-04 [7.3e-05; 1.6e-04] | 8.9e-05 [5.3e-05; 1.3e-04] | 1.3e-04 [8.0e-05; 1.7e-04] | <0.001*** |
| 3372                  | 6.5e-04 [2.4e-04; 3.0e-03] | 4.8e-04 [2.5e-04; 1.8e-03] | 3.8e-04 [1.8e-04; 1.9e-03] | 1.0e-03 [2.0e-04; 1.8e-03] | 0.3       |
| 3392                  | 1.6e-04 [8.1e-05; 3.1e-04] | 9.5e-05 [6.6e-05; 1.3e-04] | 1.1e-04 [5.8e-05; 1.8e-04] | 1.9e-04 [1.1e-04; 2.8e-04] | 0.004**   |
| 3443                  | 7.2e-04 [2.2e-04; 3.3e-03] | 9.7e-04 [2.5e-04; 2.8e-03] | 5.7e-04 [2.7e-04; 2.9e-03] | 1.3e-03 [2.3e-04; 1.6e-03] | 0.9       |
| 3464                  | 3.9e-04 [2.0e-04; 8.3e-04] | 1.8e-04 [1.5e-04; 3.0e-04] | 1.9e-04 [1.3e-04; 5.6e-04] | 4.9e-04 [1.3e-04; 9.0e-04] | <0.001*** |
| 3476                  | 5.2e-04 [1.8e-04; 8.7e-04] | 3.5e-04 [2.3e-04; 5.1e-04] | 3.4e-04 [1.8e-04; 5.4e-04] | 2.7e-04 [1.5e-04; 9.3e-04] | 0.048*    |
| 3487                  | 7.3e-04 [3.0e-04; 2.0e-03] | 5.7e-04 [1.8e-04; 1.8e-03] | 3.8e-04 [1.6e-04; 1.1e-03] | 7.8e-04 [2.0e-04; 1.6e-03] | 0.027*    |
| 3516                  | 8.7e-05 [4.1e-05; 1.5e-04] | 1.1e-04 [7.6e-05; 1.4e-04] | 1.1e-04 [7.7e-05; 1.6e-04] | 2.4e-04 [1.6e-04; 3.1e-04] | 0.034*    |
| 3609                  | 1.5e-04 [1.1e-04; 2.1e-04] | 7.6e-05 [4.7e-05; 1.1e-04] | 1.0e-04 [5.4e-05; 1.3e-04] | 2.2e-04 [1.7e-04; 3.4e-04] | <0.001*** |
| 3651                  | 8.6e-05 [2.9e-05; 1.3e-04] | 1.1e-04 [8.1e-05; 1.3e-04] | 1.2e-04 [8.8e-05; 1.5e-04] | 1.1e-04 [7.9e-05; 1.3e-04] | 0.004**   |

Продолжение таблицы Д.1

| 1    | 2                             | 3                             | 4                          | 5                             | 6         |
|------|-------------------------------|-------------------------------|----------------------------|-------------------------------|-----------|
| 3710 | 3.5e-04 [1.5e-04;<br>5.1e-04] | 1.4e-04 [9.7e-05;<br>2.3e-04] | 1.4e-04 [8.2e-05; 2.3e-04] | 4.6e-04 [1.8e-04;<br>6.3e-04] | <0.001*** |
| 3732 | 1.6e-04 [1.1e-04;<br>3.0e-04] | 1.5e-04 [1.1e-04;<br>2.0e-04] | 9.1e-05 [6.0e-05; 1.3e-04] | 8.8e-05 [8.0e-05;<br>2.2e-04] | <0.001*** |
| 3754 | 1.2e-04 [6.1e-05;<br>1.8e-04] | 8.1e-05 [5.4e-05;<br>1.1e-04] | 1.5e-04 [1.1e-04; 2.0e-04] | 2.1e-04 [1.2e-04;<br>2.4e-04] | <0.001*** |
| 3779 | 6.6e-05 [2.0e-05;<br>1.2e-04] | 1.0e-04 [5.2e-05;<br>1.4e-04] | 1.1e-04 [7.1e-05; 1.7e-04] | 3.0e-05 [1.2e-06;<br>4.9e-05] | <0.001*** |
| 3792 | 4.7e-05 [1.2e-05;<br>1.1e-04] | 7.3e-05 [4.8e-05;<br>1.0e-04] | 6.5e-05 [2.9e-05; 1.0e-04] | 1.4e-05 [4.8e-06;<br>8.3e-05] | 0.3       |
| 3804 | 1.9e-04 [1.1e-04;<br>5.3e-04] | 1.9e-04 [1.3e-04;<br>2.9e-04] | 1.4e-04 [1.1e-04; 2.4e-04] | 1.3e-04 [6.7e-05;<br>2.8e-04] | 0.053     |
| 3827 | 3.5e-04 [1.4e-04;<br>7.2e-04] | 1.8e-04 [1.4e-04;<br>2.7e-04] | 1.6e-04 [1.0e-04; 2.5e-04] | 8.0e-04 [1.6e-04;<br>2.1e-03] | <0.001*** |
| 3915 | 1.1e-04 [6.9e-05;<br>1.5e-04] | 1.1e-04 [7.6e-05;<br>1.2e-04] | 8.4e-05 [5.4e-05; 1.3e-04] | 1.9e-04 [1.3e-04;<br>1.9e-04] | 0.033*    |
| 3934 | 8.5e-05 [1.7e-05;<br>1.5e-04] | 7.6e-05 [5.6e-05;<br>1.0e-04] | 9.4e-05 [6.6e-05; 1.4e-04] | 9.4e-05 [1.1e-05;<br>1.2e-04] | 0.10      |
| 3981 | 2.3e-04 [1.2e-04;<br>7.0e-04] | 4.8e-04 [3.5e-04;<br>6.1e-04] | 3.7e-04 [1.6e-04; 6.8e-04] | 6.6e-04 [1.6e-04;<br>1.9e-03] | 0.021*    |
| 4138 | 3.1e-04 [1.6e-04;<br>8.5e-04] | 7.5e-04 [4.1e-04;<br>1.1e-03] | 1.2e-04 [6.9e-05; 2.2e-04] | 5.7e-04 [2.2e-04;<br>1.4e-03] | <0.001*** |
| 4160 | 1.8e-04 [1.4e-04;<br>2.8e-04] | 1.5e-04 [1.2e-04;<br>2.0e-04] | 1.1e-04 [5.9e-05; 1.5e-04] | 2.5e-04 [1.9e-04;<br>3.6e-04] | <0.001*** |
| 4192 | 1.4e-04 [8.5e-05;<br>1.9e-04] | 1.2e-04 [1.1e-04;<br>1.5e-04] | 1.1e-04 [7.1e-05; 1.6e-04] | 1.8e-04 [1.5e-04;<br>4.2e-04] | 0.041*    |
| 4229 | 6.1e-05 [1.7e-05;<br>1.0e-04] | 6.7e-05 [4.3e-05;<br>1.1e-04] | 8.2e-05 [4.7e-05; 1.1e-04] | 2.6e-05 [5.4e-06;<br>1.0e-04] | 0.086     |
| 4356 | 1.2e-04 [6.0e-05;<br>1.7e-04] | 1.2e-04 [1.1e-04;<br>1.5e-04] | 1.4e-04 [9.8e-05; 2.5e-04] | 6.0e-05 [5.2e-05;<br>8.7e-05] | <0.001*** |
| 4374 | 1.1e-04 [6.6e-05;<br>1.4e-04] | 1.2e-04 [1.0e-04;<br>1.5e-04] | 4.4e-04 [1.3e-04; 2.4e-03] | 1.2e-04 [5.9e-05;<br>1.4e-04] | <0.001*** |
| 4393 | 9.8e-05 [5.0e-05;<br>1.3e-04] | 1.1e-04 [9.3e-05;<br>1.3e-04] | 1.5e-04 [1.1e-04; 2.6e-04] | 5.1e-05 [3.6e-05;<br>9.8e-05] | <0.001*** |
| 4428 | 1.7e-04 [1.2e-04;<br>2.2e-04] | 1.4e-04 [1.2e-04;<br>1.7e-04] | 1.4e-04 [9.8e-05; 1.8e-04] | 1.3e-04 [1.1e-04;<br>1.6e-04] | 0.014*    |
| 4473 | 9.5e-05 [4.9e-05;<br>1.3e-04] | 8.9e-05 [4.3e-05;<br>9.9e-05] | 8.6e-05 [5.0e-05; 1.1e-04] | 9.2e-05 [8.1e-05;<br>1.4e-04] | 0.3       |
| 4532 | 8.6e-05 [5.2e-05;<br>1.2e-04] | 1.0e-04 [7.4e-05;<br>1.2e-04] | 1.3e-04 [9.7e-05; 2.0e-04] | 7.5e-05 [6.6e-05;<br>1.1e-04] | <0.001*** |
| 4551 | 2.7e-04 [1.6e-04;<br>4.6e-04] | 1.2e-04 [1.1e-04;<br>1.4e-04] | 1.3e-04 [1.0e-04; 1.7e-04] | 3.3e-04 [1.3e-04;<br>4.6e-04] | <0.001*** |
| 4574 | 1.3e-04 [7.8e-05;<br>1.7e-04] | 1.0e-04 [8.1e-05;<br>1.3e-04] | 9.2e-05 [5.0e-05; 1.2e-04] | 1.5e-04 [1.1e-04;<br>1.9e-04] | <0.001*** |
| 4636 | 2.0e-04 [1.2e-04;<br>3.4e-04] | 1.0e-04 [7.4e-05;<br>1.3e-04] | 1.5e-04 [1.2e-04; 2.4e-04] | 2.5e-04 [2.1e-04;<br>3.2e-04] | <0.001*** |
| 4715 | 3.1e-04 [1.3e-04;<br>8.3e-04] | 1.3e-04 [1.0e-04;<br>1.6e-04] | 1.2e-04 [8.4e-05; 1.5e-04] | 4.9e-04 [1.7e-04;<br>9.0e-04] | <0.001*** |
| 4738 | 1.3e-04 [9.8e-05;<br>1.9e-04] | 9.3e-05 [5.7e-05;<br>1.2e-04] | 1.0e-04 [5.2e-05; 1.4e-04] | 1.6e-04 [1.2e-04;<br>1.9e-04] | <0.001*** |
| 4811 | 7.5e-05 [1.7e-05;<br>1.5e-04] | 7.7e-05 [4.2e-05;<br>1.2e-04] | 9.2e-05 [5.9e-05; 1.2e-04] | 5.3e-05 [1.1e-05;<br>1.0e-04] | 0.4       |
| 4842 | 2.6e-04 [1.3e-04;<br>5.9e-04] | 1.2e-04 [9.1e-05;<br>1.5e-04] | 8.9e-05 [5.5e-05; 1.2e-04] | 2.9e-04 [1.7e-04;<br>8.7e-04] | <0.001*** |

Продолжение таблицы Д.1

| 1    | 2                             | 3                             | 4                          | 5                             | 6         |
|------|-------------------------------|-------------------------------|----------------------------|-------------------------------|-----------|
| 4901 | 1.9e-04 [1.3e-04;<br>2.9e-04] | 7.4e-05 [5.4e-05;<br>1.0e-04] | 9.5e-05 [5.7e-05; 2.0e-04] | 8.2e-05 [2.4e-05;<br>1.5e-04] | <0.001*** |
| 4940 | 1.6e-04 [1.1e-04;<br>2.6e-04] | 3.2e-04 [1.8e-04;<br>4.2e-04] | 2.4e-04 [1.7e-04; 3.7e-04] | 1.4e-04 [1.0e-04;<br>1.8e-04] | <0.001*** |
| 4966 | 2.9e-04 [1.3e-04;<br>6.4e-04] | 3.4e-04 [1.8e-04;<br>8.5e-04] | 2.9e-04 [1.6e-04; 4.6e-04] | 4.2e-04 [1.3e-04;<br>5.2e-04] | 0.4       |
| 4986 | 1.3e-04 [7.4e-05;<br>2.1e-04] | 1.2e-04 [8.7e-05;<br>1.5e-04] | 9.5e-05 [3.8e-05; 1.3e-04] | 7.2e-05 [5.9e-05;<br>1.5e-04] | <0.001*** |
| 5005 | 7.1e-05 [1.9e-05;<br>1.3e-04] | 1.1e-04 [6.7e-05;<br>2.0e-04] | 1.0e-04 [7.4e-05; 1.5e-04] | 8.4e-05 [8.2e-06;<br>1.0e-04] | <0.001*** |
| 5047 | 1.3e-04 [7.2e-05;<br>1.7e-04] | 1.2e-04 [8.6e-05;<br>1.4e-04] | 7.3e-05 [3.3e-05; 1.2e-04] | 1.1e-04 [7.9e-05;<br>1.8e-04] | <0.001*** |
| 5147 | 3.8e-04 [1.6e-04;<br>7.9e-04] | 1.3e-04 [1.0e-04;<br>1.8e-04] | 6.8e-05 [3.9e-05; 1.0e-04] | 2.1e-04 [1.7e-04;<br>2.9e-04] | <0.001*** |
| 5218 | 2.5e-04 [1.3e-04;<br>5.0e-04] | 1.7e-04 [1.1e-04;<br>2.5e-04] | 1.7e-04 [1.0e-04; 2.4e-04] | 4.3e-04 [1.7e-04;<br>6.7e-04] | <0.001*** |
| 5236 | 1.7e-03 [2.5e-04;<br>3.2e-03] | 3.4e-04 [2.0e-04;<br>4.6e-04] | 2.4e-04 [1.5e-04; 3.3e-04] | 3.9e-03 [1.5e-04;<br>4.5e-03] | <0.001*** |
| 5256 | 2.3e-04 [1.3e-04;<br>5.1e-04] | 6.7e-05 [1.9e-05;<br>1.3e-04] | 9.8e-05 [6.8e-05; 1.4e-04] | 5.3e-04 [1.3e-04;<br>8.5e-04] | <0.001*** |
| 5285 | 2.4e-04 [1.3e-04;<br>8.5e-04] | 3.7e-04 [2.2e-04;<br>5.4e-04] | 1.7e-04 [9.7e-05; 3.3e-04] | 4.7e-04 [1.6e-04;<br>1.4e-03] | <0.001*** |
| 5382 | 1.9e-03 [2.8e-04;<br>5.8e-03] | 1.5e-03 [9.2e-04;<br>2.6e-03] | 7.1e-04 [1.9e-04; 1.6e-03] | 3.2e-03 [1.9e-04;<br>7.0e-03] | <0.001*** |
| 5402 | 3.8e-04 [1.5e-04;<br>1.1e-03] | 2.0e-04 [1.3e-04;<br>3.0e-04] | 1.6e-04 [9.0e-05; 4.5e-04] | 7.9e-04 [1.5e-04;<br>1.2e-03] | <0.001*** |
| 5423 | 3.9e-04 [1.6e-04;<br>7.0e-04] | 6.4e-05 [3.1e-05;<br>9.1e-05] | 7.9e-05 [4.0e-05; 1.1e-04] | 3.3e-04 [1.6e-04;<br>6.7e-04] | <0.001*** |
| 5530 | 1.8e-04 [9.5e-05;<br>4.0e-04] | 2.1e-04 [1.3e-04;<br>2.3e-04] | 1.7e-04 [1.3e-04; 2.5e-04] | 1.6e-04 [1.1e-04;<br>5.1e-04] | 0.9       |
| 5594 | 8.9e-05 [1.6e-05;<br>1.6e-04] | 1.3e-04 [9.3e-05;<br>2.1e-04] | 2.0e-04 [9.3e-05; 9.6e-04] | 3.0e-05 [0.0e+00;<br>3.8e-05] | <0.001*** |
| 5869 | 2.8e-04 [1.2e-04;<br>4.8e-04] | 3.1e-04 [2.1e-04;<br>3.9e-04] | 2.1e-04 [1.3e-04; 3.4e-04] | 2.5e-04 [9.3e-05;<br>3.2e-04] | 0.3       |
| 5950 | 1.0e-04 [4.6e-05;<br>1.4e-04] | 2.1e-04 [1.7e-04;<br>2.7e-04] | 3.2e-04 [2.1e-04; 6.2e-04] | 1.3e-04 [9.3e-05;<br>1.3e-04] | <0.001*** |
| 6192 | 1.0e-04 [5.4e-05;<br>1.5e-04] | 1.5e-04 [1.2e-04;<br>2.1e-04] | 1.5e-04 [1.1e-04; 2.0e-04] | 9.3e-05 [4.3e-05;<br>1.2e-04] | <0.001*** |
| 6361 | 1.1e-04 [4.4e-05;<br>1.9e-04] | 7.6e-05 [3.0e-05;<br>1.1e-04] | 1.2e-04 [9.4e-05; 1.5e-04] | 1.1e-04 [2.0e-05;<br>1.5e-04] | <0.001*** |
| 6639 | 8.2e-05 [2.1e-05;<br>1.4e-04] | 9.9e-05 [7.8e-05;<br>1.3e-04] | 1.2e-04 [7.9e-05; 1.6e-04] | 3.5e-06 [1.4e-06;<br>3.5e-05] | <0.001*** |
| 6964 | 5.5e-04 [1.6e-04;<br>1.5e-03] | 2.0e-04 [9.9e-05;<br>3.4e-04] | 1.5e-04 [7.5e-05; 2.6e-04] | 1.5e-04 [1.5e-04;<br>1.9e-03] | <0.001*** |
| 7349 | 4.6e-04 [1.8e-04;<br>1.1e-03] | 1.2e-04 [8.9e-05;<br>1.7e-04] | 1.3e-04 [8.4e-05; 2.1e-04] | 1.3e-04 [1.1e-04;<br>3.9e-04] | <0.001*** |
| 7612 | 2.9e-04 [9.9e-05;<br>8.2e-04] | 7.0e-04 [3.0e-04;<br>8.3e-04] | 3.4e-04 [1.7e-04; 7.3e-04] | 1.7e-04 [1.5e-04;<br>7.1e-04] | 0.073     |
| 7654 | 4.5e-05 [1.8e-05;<br>1.0e-04] | 4.5e-05 [2.5e-05;<br>8.2e-05] | 7.6e-05 [3.4e-05; 1.1e-04] | 8.8e-05 [2.1e-05;<br>1.2e-04] | 0.010**   |
| 7765 | 2.7e-05 [8.8e-06;<br>6.6e-05] | 8.0e-05 [6.6e-05;<br>1.1e-04] | 1.1e-04 [7.6e-05; 1.8e-04] | 1.7e-05 [7.8e-06;<br>1.2e-04] | <0.001*** |

Продолжение таблицы Д.1

| 1     | 2                             | 3                             | 4                              | 5                             | 6         |
|-------|-------------------------------|-------------------------------|--------------------------------|-------------------------------|-----------|
| 8215  | 1.1e-04 [4.6e-05;<br>1.7e-04] | 9.4e-05 [6.8e-05;<br>1.4e-04] | 9.9e-05 [5.7e-05; 1.3e-<br>04] | 7.6e-05 [5.9e-05;<br>1.1e-04] | 0.6       |
| 8452  | 1.0e-04 [2.4e-05;<br>1.4e-04] | 1.1e-04 [7.7e-05;<br>1.2e-04] | 9.5e-05 [7.2e-05; 1.2e-<br>04] | 5.5e-05 [2.6e-05;<br>8.4e-05] | 0.2       |
| 8469  | 4.5e-05 [2.2e-05;<br>9.8e-05] | 6.7e-05 [4.3e-05;<br>1.1e-04] | 6.0e-05 [3.6e-05; 9.4e-<br>05] | 4.6e-05 [1.5e-05;<br>1.0e-04] | 0.11      |
| 8568  | 1.3e-04 [7.1e-05;<br>2.1e-04] | 7.1e-05 [4.1e-05;<br>9.4e-05] | 6.8e-05 [4.0e-05; 9.4e-<br>05] | 1.0e-04 [8.6e-05;<br>1.3e-04] | <0.001*** |
| 8589  | 7.3e-05 [2.3e-05;<br>1.3e-04] | 9.3e-05 [5.9e-05;<br>1.1e-04] | 7.7e-05 [2.9e-05; 1.1e-<br>04] | 2.8e-05 [1.9e-05;<br>9.0e-05] | 0.3       |
| 8742  | 1.3e-04 [3.3e-05;<br>1.9e-04] | 7.7e-05 [5.2e-05;<br>1.0e-04] | 7.9e-05 [5.1e-05; 9.9e-<br>05] | 1.4e-04 [1.1e-04;<br>1.6e-04] | 0.001**   |
| 9956  | 5.3e-05 [2.5e-05;<br>1.1e-04] | 6.8e-05 [4.3e-05;<br>8.8e-05] | 9.1e-05 [6.8e-05; 1.1e-<br>04] | 5.6e-05 [2.5e-05;<br>1.2e-04] | <0.001*** |
| 10096 | 1.0e-04 [3.5e-05;<br>1.5e-04] | 7.0e-05 [4.6e-05;<br>1.1e-04] | 8.7e-05 [4.9e-05; 1.1e-<br>04] | 3.8e-05 [1.7e-05;<br>8.5e-05] | 0.2       |
| 10116 | 3.4e-05 [1.4e-05;<br>8.3e-05] | 7.2e-05 [5.0e-05;<br>8.7e-05] | 8.3e-05 [4.0e-05; 1.1e-<br>04] | 1.3e-05 [1.0e-05;<br>1.3e-04] | <0.001*** |
| 10444 | 1.7e-04 [1.0e-04;<br>3.0e-04] | 2.0e-04 [1.4e-04;<br>3.0e-04] | 2.1e-04 [1.4e-04; 3.4e-<br>04] | 1.4e-04 [1.1e-04;<br>2.5e-04] | 0.059     |
| 10837 | 1.3e-04 [4.6e-05;<br>1.8e-04] | 8.8e-05 [6.0e-05;<br>1.0e-04] | 9.1e-05 [6.0e-05; 1.2e-<br>04] | 8.9e-05 [3.6e-05;<br>1.3e-04] | 0.001**   |
| 11011 | 5.1e-05 [1.1e-05;<br>1.3e-04] | 7.3e-05 [4.6e-05;<br>9.8e-05] | 1.0e-04 [7.4e-05; 1.4e-<br>04] | 4.8e-05 [1.7e-05;<br>9.2e-05] | <0.001*** |
| 11735 | 2.3e-04 [1.4e-04;<br>3.3e-04] | 3.2e-04 [2.4e-04;<br>5.0e-04] | 1.6e-04 [9.9e-05; 2.5e-<br>04] | 2.0e-04 [8.5e-05;<br>3.1e-04] | <0.001*** |
| 14692 | 2.8e-04 [1.6e-04;<br>4.8e-04] | 1.1e-04 [8.4e-05;<br>1.4e-04] | 1.2e-04 [9.2e-05; 1.8e-<br>04] | 2.2e-04 [1.4e-04;<br>3.3e-04] | <0.001*** |

## ПРИЛОЖЕНИЕ Е

Таблица Е.1 – Пики MALDI-TOF MS, обнаруженные в объединенном наборе данных, содержащем образцы как из Казахстана, так и из Южной Америки и присутствующие хотя бы в 80% каждой подгруппы

| Peak<br>( <i>m/z</i> ) | Образцы из Южной Америки         |                                  |                                  |                                  |                                  |                                  | Образцы из Казахстана            |                                  |                                  |                                  | p-value   |
|------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
|                        | SARS-CoV-2+                      |                                  |                                  | ОРИ                              |                                  |                                  | ОРИ                              | AC                               | SARS-CoV-2                       |                                  |           |
|                        | Chile<br>N = 10                  | Argentina<br>N = 167             | Peru<br>N = 34                   | Chile,<br>N = 20                 | Argentina,<br>N = 114            | Peru,<br>N = 17                  | N = 71                           | N = 66                           | 2020,<br>N = 108                 | 2022,<br>N = 7                   |           |
| 1                      | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
| 3013                   | 8.6e-05<br>[3.4e-05;<br>1.8e-04] | 1.0e-04<br>[6.2e-05;<br>1.6e-04] | 1.6e-04<br>[1.1e-04;<br>2.0e-04] | 8.6e-05<br>[5.3e-05;<br>1.1e-04] | 1.6e-04<br>[6.6e-05;<br>3.6e-04] | 1.5e-04<br>[1.1e-04;<br>1.8e-04] | 9.8e-05<br>[5.3e-05;<br>1.8e-04] | 1.2e-04<br>[7.8e-05;<br>2.0e-04] | 9.0e-05<br>[6.0e-05;<br>1.4e-04] | 1.2e-04<br>[1.1e-04;<br>1.6e-04] | <0.001*** |
| 3037                   | 8.4e-05<br>[2.3e-05;<br>1.5e-04] | 1.2e-04<br>[6.7e-05;<br>2.2e-04] | 1.1e-04<br>[5.8e-05;<br>1.4e-04] | 9.8e-05<br>[5.7e-05;<br>1.2e-04] | 1.1e-04<br>[5.1e-05;<br>2.1e-04] | 2.4e-04<br>[1.5e-04;<br>4.6e-04] | 6.2e-05<br>[3.4e-05;<br>1.1e-04] | 5.7e-05<br>[4.1e-05;<br>9.0e-05] | 1.0e-04<br>[5.5e-05;<br>3.3e-04] | 1.5e-04<br>[7.7e-05;<br>1.9e-04] | <0.001*** |
| 3052                   | 1.5e-04<br>[9.9e-05;<br>2.3e-04] | 1.6e-04<br>[8.9e-05;<br>2.9e-04] | 1.4e-04<br>[7.3e-05;<br>2.0e-04] | 1.7e-04<br>[1.5e-04;<br>2.1e-04] | 2.1e-04<br>[8.9e-05;<br>4.3e-04] | 1.3e-04<br>[1.0e-04;<br>2.2e-04] | 9.4e-05<br>[6.4e-05;<br>1.6e-04] | 7.6e-05<br>[5.5e-05;<br>1.1e-04] | 6.5e-05<br>[4.3e-05;<br>1.1e-04] | 1.6e-04<br>[1.2e-04;<br>3.1e-04] | <0.001*** |
| 3095                   | 2.0e-04<br>[1.6e-04;<br>2.5e-04] | 1.2e-04<br>[5.5e-05;<br>2.2e-04] | 1.2e-04<br>[7.0e-05;<br>1.7e-04] | 1.2e-04<br>[8.2e-05;<br>1.5e-04] | 2.6e-04<br>[1.5e-04;<br>5.0e-04] | 2.4e-04<br>[1.5e-04;<br>2.8e-04] | 8.1e-05<br>[4.5e-05;<br>1.2e-04] | 8.8e-05<br>[6.7e-05;<br>1.3e-04] | 9.4e-05<br>[6.6e-05;<br>1.3e-04] | 5.9e-05<br>[4.0e-05;<br>1.1e-04] | <0.001*** |
| 3113                   | 2.2e-04<br>[1.3e-04;<br>3.6e-04] | 8.3e-05<br>[4.0e-05;<br>1.8e-04] | 7.8e-05<br>[3.3e-05;<br>1.3e-04] | 9.1e-05<br>[7.1e-05;<br>1.2e-04] | 2.1e-04<br>[1.3e-04;<br>2.8e-04] | 6.1e-05<br>[2.7e-05;<br>9.8e-05] | 5.6e-05<br>[2.1e-05;<br>9.6e-05] | 1.2e-04<br>[8.3e-05;<br>1.5e-04] | 4.9e-05<br>[3.1e-05;<br>8.4e-05] | 4.0e-05<br>[1.3e-05;<br>4.3e-05] | <0.001*** |
| 3128                   | 1.1e-04<br>[3.7e-05;<br>1.1e-04] | 1.1e-04<br>[3.8e-05;<br>2.1e-04] | 1.0e-04<br>[4.3e-05;<br>1.5e-04] | 8.1e-05<br>[4.3e-05;<br>1.3e-04] | 9.3e-05<br>[3.8e-05;<br>1.5e-04] | 5.8e-05<br>[2.4e-05;<br>9.9e-05] | 7.7e-05<br>[4.2e-05;<br>1.3e-04] | 7.5e-05<br>[4.2e-05;<br>1.0e-04] | 5.4e-05<br>[2.9e-05;<br>7.6e-05] | 1.1e-04<br>[5.6e-05;<br>2.1e-04] | <0.001*** |
| 3147                   | 7.9e-05<br>[1.9e-05;<br>1.6e-04] | 3.1e-04<br>[1.5e-04;<br>5.1e-04] | 2.6e-04<br>[1.7e-04;<br>3.5e-04] | 5.4e-05<br>[1.8e-05;<br>1.4e-04] | 8.3e-05<br>[3.1e-05;<br>1.5e-04] | 3.2e-04<br>[2.2e-04;<br>3.7e-04] | 1.7e-04<br>[1.3e-04;<br>2.4e-04] | 6.4e-05<br>[3.8e-05;<br>9.3e-05] | 7.4e-05<br>[5.0e-05;<br>1.1e-04] | 1.8e-04<br>[1.4e-04;<br>2.3e-04] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                             | 3                             | 4                             | 5                             | 6                             | 7                             | 8                             | 9                             | 10                            | 11                            | 12        |
|------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-----------|
| 3192 | 7.4e-05<br>[4.0e-05; 1.0e-04] | 4.0e-05<br>[8.8e-06; 1.1e-04] | 4.5e-05<br>[1.5e-05; 1.1e-04] | 8.9e-05<br>[3.1e-05; 1.2e-04] | 4.0e-05<br>[1.1e-06; 1.2e-04] | 2.1e-04<br>[1.5e-04; 3.1e-04] | 8.6e-05<br>[5.4e-05; 1.6e-04] | 7.0e-05<br>[5.2e-05; 1.0e-04] | 9.8e-05<br>[6.5e-05; 1.6e-04] | 1.5e-04<br>[1.2e-04; 1.6e-04] | <0.001*** |
| 3218 | 1.7e-04<br>[1.0e-04; 2.8e-04] | 4.9e-05<br>[6.5e-06; 1.3e-04] | 7.7e-05<br>[4.7e-05; 1.4e-04] | 2.0e-04<br>[1.3e-04; 2.4e-04] | 1.2e-04<br>[5.2e-05; 1.8e-04] | 7.5e-05<br>[4.3e-05; 1.5e-04] | 8.5e-05<br>[4.7e-05; 1.3e-04] | 8.0e-05<br>[6.8e-05; 9.6e-05] | 7.6e-05<br>[4.3e-05; 9.9e-05] | 1.0e-04<br>[7.4e-05; 1.3e-04] | <0.001*** |
| 3242 | 1.0e-04<br>[8.8e-05; 1.9e-04] | 1.0e-04<br>[5.2e-05; 1.8e-04] | 1.3e-04<br>[9.3e-05; 1.7e-04] | 9.7e-05<br>[6.9e-05; 1.5e-04] | 9.3e-05<br>[5.0e-05; 1.5e-04] | 1.4e-04<br>[1.1e-04; 1.6e-04] | 7.1e-05<br>[5.0e-05; 1.1e-04] | 6.7e-05<br>[5.3e-05; 1.1e-04] | 7.5e-05<br>[4.9e-05; 1.3e-04] | 7.5e-05<br>[5.6e-05; 1.7e-04] | <0.001*** |
| 3256 | 8.9e-05<br>[5.5e-05; 1.8e-04] | 1.1e-04<br>[4.0e-05; 2.4e-04] | 1.1e-04<br>[3.5e-05; 1.4e-04] | 5.3e-05<br>[3.2e-05; 1.1e-04] | 7.7e-05<br>[3.1e-05; 1.3e-04] | 1.5e-04<br>[1.1e-04; 1.8e-04] | 5.9e-05<br>[3.4e-05; 8.7e-05] | 7.1e-05<br>[5.5e-05; 1.2e-04] | 5.4e-05<br>[3.7e-05; 8.1e-05] | 6.4e-05<br>[5.7e-05; 9.3e-05] | <0.001*** |
| 3300 | 1.0e-04<br>[5.7e-05; 1.9e-04] | 1.3e-04<br>[2.4e-05; 3.1e-04] | 1.2e-04<br>[5.8e-05; 2.1e-04] | 8.4e-05<br>[1.7e-05; 1.3e-04] | 1.0e-04<br>[1.1e-05; 3.0e-04] | 2.3e-04<br>[1.6e-04; 3.3e-04] | 1.1e-04<br>[6.7e-05; 2.2e-04] | 7.8e-05<br>[3.9e-05; 1.1e-04] | 6.8e-05<br>[4.2e-05; 8.9e-05] | 2.0e-04<br>[1.9e-04; 3.7e-04] | <0.001*** |
| 3318 | 1.1e-04<br>[6.5e-05; 1.5e-04] | 1.6e-04<br>[1.0e-07; 3.6e-04] | 2.1e-04<br>[1.5e-04; 3.1e-04] | 6.5e-05<br>[3.0e-06; 2.0e-04] | 7.9e-05<br>[6.7e-06; 3.6e-04] | 2.4e-04<br>[1.8e-04; 3.5e-04] | 8.7e-05<br>[3.8e-05; 1.4e-04] | 5.7e-05<br>[2.0e-05; 1.0e-04] | 6.3e-05<br>[1.9e-05; 9.9e-05] | 1.3e-04<br>[6.5e-05; 1.5e-04] | <0.001*** |
| 3337 | 2.4e-04<br>[2.0e-04; 3.0e-04] | 1.3e-04<br>[7.2e-06; 3.3e-04] | 1.8e-04<br>[1.5e-04; 2.4e-04] | 4.0e-04<br>[3.2e-04; 6.1e-04] | 2.5e-04<br>[7.5e-05; 5.5e-04] | 2.6e-04<br>[2.1e-04; 4.8e-04] | 1.4e-04<br>[7.2e-05; 4.5e-04] | 3.3e-04<br>[1.6e-04; 4.3e-04] | 1.3e-04<br>[7.8e-05; 2.4e-04] | 2.3e-04<br>[1.8e-04; 4.0e-04] | <0.001*** |
| 3358 | 6.5e-05<br>[2.2e-06; 9.8e-05] | 1.1e-04<br>[4.0e-05; 1.9e-04] | 9.2e-05<br>[5.3e-05; 1.4e-04] | 6.6e-05<br>[1.7e-05; 1.8e-04] | 2.6e-04<br>[1.9e-04; 3.8e-04] | 1.1e-04<br>[8.6e-05; 1.4e-04] | 9.6e-05<br>[6.2e-05; 1.8e-04] | 8.4e-05<br>[6.3e-05; 1.5e-04] | 5.8e-05<br>[3.0e-05; 1.1e-04] | 9.3e-05<br>[6.2e-05; 1.7e-04] | <0.001*** |
| 3372 | 1.2e-03<br>[8.6e-04; 2.0e-03] | 1.2e-03<br>[5.7e-04; 3.0e-03] | 4.7e-04<br>[2.9e-04; 6.2e-04] | 1.2e-03<br>[4.4e-04; 2.3e-03] | 2.3e-03<br>[1.2e-03; 3.6e-03] | 5.6e-04<br>[4.2e-04; 6.2e-04] | 6.5e-04<br>[2.4e-04; 3.0e-03] | 4.8e-04<br>[2.5e-04; 1.8e-03] | 3.6e-04<br>[1.8e-04; 1.9e-03] | 1.0e-03<br>[2.0e-04; 1.8e-03] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 3393 | 2.7e-04<br>[2.0e-04;<br>5.1e-04] | 5.0e-04<br>[2.5e-04;<br>8.6e-04] | 1.1e-04<br>[7.8e-05;<br>1.4e-04] | 2.4e-04<br>[1.5e-04;<br>3.7e-04] | 4.3e-04<br>[2.6e-04;<br>7.4e-04] | 1.1e-04<br>[7.7e-05;<br>1.3e-04] | 1.1e-04<br>[5.1e-05;<br>2.9e-04] | 1.3e-04<br>[9.3e-05;<br>2.3e-04] | 8.2e-05<br>[3.5e-05;<br>1.3e-04] | 1.8e-04<br>[1.1e-04;<br>2.7e-04] | <0.001*** |
| 3443 | 1.5e-03<br>[7.5e-04;<br>2.3e-03] | 1.7e-03<br>[7.5e-04;<br>4.1e-03] | 5.0e-04<br>[3.1e-04;<br>6.5e-04] | 1.5e-03<br>[7.3e-04;<br>2.6e-03] | 3.0e-03<br>[1.5e-03;<br>4.5e-03] | 6.7e-04<br>[4.7e-04;<br>8.3e-04] | 7.2e-04<br>[2.2e-04;<br>3.3e-03] | 9.7e-04<br>[2.2e-04;<br>2.8e-03] | 5.7e-04<br>[2.7e-04;<br>2.9e-03] | 1.3e-03<br>[2.3e-04;<br>1.6e-03] | <0.001*** |
| 3463 | 5.6e-04<br>[3.7e-04;<br>7.1e-04] | 5.7e-04<br>[3.0e-04;<br>1.1e-03] | 2.0e-04<br>[1.5e-04;<br>2.6e-04] | 4.5e-04<br>[2.7e-04;<br>6.9e-04] | 9.0e-04<br>[4.2e-04;<br>1.5e-03] | 3.0e-04<br>[2.1e-04;<br>4.3e-04] | 3.9e-04<br>[2.0e-04;<br>8.3e-04] | 1.6e-04<br>[1.2e-04;<br>2.5e-04] | 1.9e-04<br>[1.3e-04;<br>5.7e-04] | 4.7e-04<br>[1.0e-04;<br>8.0e-04] | <0.001*** |
| 3476 | 2.7e-04<br>[2.1e-04;<br>3.2e-04] | 3.9e-04<br>[2.2e-04;<br>6.0e-04] | 2.3e-04<br>[1.7e-04;<br>3.2e-04] | 3.9e-04<br>[2.5e-04;<br>6.2e-04] | 3.9e-04<br>[2.4e-04;<br>6.8e-04] | 3.2e-04<br>[2.1e-04;<br>4.6e-04] | 5.2e-04<br>[1.4e-04;<br>8.7e-04] | 3.0e-04<br>[2.1e-04;<br>4.9e-04] | 3.3e-04<br>[1.5e-04;<br>5.1e-04] | 2.7e-04<br>[1.2e-04;<br>9.2e-04] | 0.001**   |
| 3487 | 1.1e-03<br>[7.8e-04;<br>1.7e-03] | 7.5e-04<br>[3.0e-04;<br>2.0e-03] | 2.8e-04<br>[1.6e-04;<br>5.1e-04] | 9.4e-04<br>[3.2e-04;<br>2.2e-03] | 1.5e-03<br>[7.4e-04;<br>3.4e-03] | 3.5e-04<br>[3.0e-04;<br>6.4e-04] | 7.3e-04<br>[2.9e-04;<br>2.0e-03] | 5.7e-04<br>[1.8e-04;<br>1.8e-03] | 3.4e-04<br>[1.4e-04;<br>1.1e-03] | 7.8e-04<br>[2.0e-04;<br>1.6e-03] | <0.001*** |
| 3516 | 2.1e-04<br>[1.2e-04;<br>4.6e-04] | 4.4e-05<br>[0.0e+00;<br>1.7e-04] | 1.1e-04<br>[5.2e-05;<br>1.4e-04] | 5.7e-05<br>[6.5e-06;<br>1.3e-04] | 1.9e-05<br>[0.0e+00;<br>3.7e-04] | 2.1e-04<br>[1.5e-04;<br>2.7e-04] | 8.5e-05<br>[4.1e-05;<br>1.5e-04] | 7.0e-05<br>[3.6e-05;<br>1.0e-04] | 6.6e-05<br>[3.8e-05;<br>1.0e-04] | 1.2e-04<br>[6.9e-05;<br>1.7e-04] | <0.001*** |
| 3560 | 1.6e-05<br>[4.9e-06;<br>3.4e-05] | 1.6e-06<br>[0.0e+00;<br>5.8e-05] | 7.1e-05<br>[3.0e-05;<br>1.2e-04] | 3.3e-05<br>[6.4e-06;<br>1.0e-04] | 2.1e-05<br>[0.0e+00;<br>1.2e-04] | 1.1e-04<br>[2.5e-05;<br>1.7e-04] | 1.1e-04<br>[7.1e-05;<br>1.5e-04] | 5.9e-05<br>[4.7e-05;<br>7.2e-05] | 5.2e-05<br>[1.9e-05;<br>9.0e-05] | 1.7e-04<br>[1.2e-04;<br>2.2e-04] | <0.001*** |
| 3589 | 1.3e-04<br>[6.2e-05;<br>1.7e-04] | 7.9e-05<br>[2.1e-05;<br>1.6e-04] | 8.1e-05<br>[6.2e-05;<br>1.2e-04] | 1.4e-04<br>[8.8e-05;<br>2.2e-04] | 1.2e-04<br>[5.0e-05;<br>2.2e-04] | 1.7e-04<br>[8.0e-05;<br>2.7e-04] | 1.3e-04<br>[8.4e-05;<br>2.5e-04] | 1.6e-04<br>[1.2e-04;<br>2.1e-04] | 9.0e-05<br>[4.9e-05;<br>1.4e-04] | 2.1e-04<br>[1.6e-04;<br>4.0e-04] | <0.001*** |
| 3610 | 7.2e-05<br>[5.8e-05;<br>1.4e-04] | 1.1e-04<br>[3.8e-05;<br>1.8e-04] | 8.7e-05<br>[3.8e-05;<br>1.5e-04] | 1.4e-04<br>[6.1e-05;<br>2.3e-04] | 1.0e-04<br>[5.0e-05;<br>2.0e-04] | 1.3e-04<br>[1.2e-04;<br>1.8e-04] | 1.2e-04<br>[6.6e-05;<br>1.7e-04] | 6.7e-05<br>[4.6e-05;<br>8.9e-05] | 6.1e-05<br>[2.8e-05;<br>9.1e-05] | 2.2e-04<br>[1.7e-04;<br>3.4e-04] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                             | 3                             | 4                             | 5                             | 6                             | 7                             | 8                             | 9                             | 10                            | 11                            | 12        |
|------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-----------|
| 3633 | 1.1e-04<br>[7.8e-05; 1.5e-04] | 6.9e-05<br>[1.5e-05; 1.3e-04] | 1.1e-04<br>[4.8e-05; 2.1e-04] | 1.2e-04<br>[5.3e-05; 1.8e-04] | 1.7e-04<br>[8.3e-05; 2.8e-04] | 1.1e-04<br>[6.2e-05; 1.4e-04] | 1.3e-04<br>[8.7e-05; 2.6e-04] | 1.3e-04<br>[1.0e-04; 1.7e-04] | 9.2e-05<br>[6.2e-05; 1.3e-04] | 1.7e-04<br>[1.2e-04; 3.2e-04] | <0.001*** |
| 3658 | 1.3e-04<br>[1.0e-04; 1.4e-04] | 2.3e-05<br>[3.5e-06; 7.4e-05] | 6.2e-05<br>[2.1e-05; 1.1e-04] | 4.0e-05<br>[4.5e-07; 1.0e-04] | 6.5e-05<br>[1.5e-05; 1.4e-04] | 7.5e-05<br>[4.9e-05; 1.1e-04] | 4.5e-05<br>[2.1e-05; 8.3e-05] | 6.8e-05<br>[2.9e-05; 8.7e-05] | 7.0e-05<br>[5.1e-05; 1.1e-04] | 1.6e-04<br>[1.5e-04; 2.7e-04] | <0.001*** |
| 3711 | 2.4e-04<br>[1.2e-04; 4.8e-04] | 3.3e-04<br>[1.3e-04; 7.4e-04] | 1.4e-04<br>[6.5e-05; 2.3e-04] | 3.1e-04<br>[2.1e-04; 5.5e-04] | 4.2e-04<br>[2.2e-04; 8.7e-04] | 1.1e-04<br>[6.1e-05; 2.4e-04] | 2.2e-04<br>[1.0e-04; 4.2e-04] | 1.1e-04<br>[8.2e-05; 2.0e-04] | 1.0e-04<br>[6.2e-05; 1.9e-04] | 4.6e-04<br>[1.8e-04; 6.3e-04] | <0.001*** |
| 3732 | 1.4e-04<br>[1.2e-04; 3.4e-04] | 8.8e-05<br>[3.4e-05; 1.7e-04] | 8.0e-05<br>[4.7e-05; 1.0e-04] | 1.1e-04<br>[4.1e-05; 1.8e-04] | 1.1e-04<br>[2.2e-05; 2.0e-04] | 6.0e-05<br>[1.5e-05; 9.6e-05] | 9.0e-05<br>[4.5e-05; 1.4e-04] | 9.8e-05<br>[7.2e-05; 1.7e-04] | 9.1e-05<br>[6.0e-05; 1.3e-04] | 8.8e-05<br>[8.0e-05; 2.2e-04] | 0.034*    |
| 3752 | 6.2e-05<br>[4.5e-06; 8.9e-05] | 1.2e-04<br>[2.1e-05; 2.5e-04] | 1.7e-04<br>[8.9e-05; 3.4e-04] | 8.5e-05<br>[4.3e-05; 1.3e-04] | 8.2e-05<br>[4.1e-06; 1.6e-04] | 2.6e-04<br>[1.5e-04; 3.7e-04] | 8.7e-05<br>[5.3e-05; 1.4e-04] | 8.1e-05<br>[5.6e-05; 1.2e-04] | 1.0e-04<br>[6.0e-05; 1.7e-04] | 2.1e-04<br>[1.2e-04; 2.4e-04] | <0.001*** |
| 3776 | 1.1e-04<br>[5.0e-05; 2.3e-04] | 1.3e-04<br>[7.0e-05; 2.6e-04] | 1.3e-04<br>[7.8e-05; 2.6e-04] | 1.2e-04<br>[8.0e-05; 1.8e-04] | 1.2e-04<br>[5.6e-05; 2.5e-04] | 2.1e-04<br>[1.6e-04; 3.6e-04] | 4.4e-05<br>[2.0e-05; 7.8e-05] | 7.4e-05<br>[4.5e-05; 9.0e-05] | 8.3e-05<br>[4.9e-05; 1.4e-04] | 1.3e-04<br>[1.1e-04; 2.1e-04] | <0.001*** |
| 3792 | 2.8e-04<br>[1.2e-04; 1.1e-03] | 1.6e-04<br>[5.2e-05; 3.5e-04] | 6.7e-05<br>[2.6e-05; 9.0e-05] | 2.2e-04<br>[1.3e-04; 5.7e-04] | 1.6e-04<br>[6.5e-05; 3.7e-04] | 5.7e-05<br>[1.3e-05; 8.0e-05] | 3.5e-05<br>[9.2e-06; 8.0e-05] | 6.2e-05<br>[4.0e-05; 8.3e-05] | 5.3e-05<br>[2.6e-05; 7.7e-05] | 1.4e-05<br>[4.8e-06; 5.0e-05] | <0.001*** |
| 3805 | 1.4e-04<br>[1.2e-04; 1.5e-04] | 1.4e-04<br>[2.3e-05; 2.6e-04] | 2.5e-04<br>[1.8e-04; 4.2e-04] | 2.3e-04<br>[1.9e-04; 3.0e-04] | 9.5e-05<br>[2.2e-05; 3.0e-04] | 2.6e-04<br>[1.3e-04; 4.9e-04] | 1.7e-04<br>[8.4e-05; 4.4e-04] | 1.9e-04<br>[1.2e-04; 2.9e-04] | 1.2e-04<br>[7.7e-05; 2.3e-04] | 1.7e-04<br>[1.3e-04; 2.8e-04] | <0.001*** |
| 3827 | 7.7e-05<br>[3.2e-05; 1.1e-04] | 1.2e-04<br>[4.4e-05; 2.7e-04] | 1.8e-04<br>[1.0e-04; 3.4e-04] | 9.9e-05<br>[5.2e-05; 2.4e-04] | 1.5e-04<br>[5.5e-05; 3.6e-04] | 3.8e-04<br>[1.6e-04; 5.8e-04] | 3.5e-04<br>[1.4e-04; 7.2e-04] | 1.7e-04<br>[1.3e-04; 2.6e-04] | 1.3e-04<br>[6.5e-05; 2.1e-04] | 8.0e-04<br>[1.5e-04; 2.1e-03] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 3848 | 3.7e-05<br>[1.7e-05;<br>5.8e-05] | 5.6e-05<br>[1.3e-05;<br>1.3e-04] | 8.2e-05<br>[3.0e-05;<br>1.5e-04] | 9.7e-06<br>[0.0e+00;<br>6.2e-05] | 9.1e-05<br>[3.4e-05;<br>1.7e-04] | 9.7e-05<br>[5.7e-05;<br>2.2e-04] | 8.7e-05<br>[5.1e-05;<br>1.5e-04] | 6.5e-05<br>[5.4e-05;<br>1.0e-04] | 4.6e-05<br>[1.7e-05;<br>8.6e-05] | 1.4e-04<br>[1.0e-04;<br>2.9e-04] | <0.001*** |
| 3862 | 3.2e-05<br>[1.9e-06;<br>5.8e-05] | 3.3e-05<br>[7.6e-07;<br>7.2e-05] | 4.8e-05<br>[1.7e-05;<br>1.2e-04] | 5.6e-05<br>[3.4e-05;<br>1.4e-04] | 7.8e-05<br>[1.6e-05;<br>1.3e-04] | 9.8e-05<br>[6.7e-05;<br>1.6e-04] | 1.3e-04<br>[8.4e-05;<br>3.5e-04] | 1.2e-04<br>[8.8e-05;<br>2.0e-04] | 9.0e-05<br>[4.6e-05;<br>1.5e-04] | 2.7e-04<br>[1.4e-04;<br>6.0e-04] | <0.001*** |
| 3877 | 4.1e-05<br>[2.5e-05;<br>9.0e-05] | 5.5e-05<br>[2.3e-05;<br>1.2e-04] | 2.9e-05<br>[4.8e-06;<br>8.4e-05] | 7.9e-05<br>[5.4e-05;<br>1.4e-04] | 6.9e-05<br>[2.0e-05;<br>1.3e-04] | 8.6e-05<br>[4.4e-05;<br>1.3e-04] | 8.5e-05<br>[4.2e-05;<br>1.2e-04] | 1.5e-04<br>[9.7e-05;<br>2.5e-04] | 1.1e-04<br>[7.2e-05;<br>1.9e-04] | 2.1e-04<br>[2.0e-04;<br>8.9e-04] | <0.001*** |
| 3901 | 2.6e-05<br>[9.3e-06;<br>4.4e-05] | 2.6e-05<br>[2.7e-06;<br>7.1e-05] | 4.2e-05<br>[1.5e-05;<br>6.1e-05] | 7.0e-05<br>[1.6e-05;<br>9.3e-05] | 1.5e-05<br>[0.0e+00;<br>6.9e-05] | 1.9e-05<br>[7.4e-06;<br>4.1e-05] | 4.5e-05<br>[1.3e-05;<br>1.0e-04] | 7.6e-05<br>[4.4e-05;<br>9.8e-05] | 5.4e-05<br>[1.9e-05;<br>7.6e-05] | 1.7e-04<br>[1.0e-04;<br>1.9e-04] | <0.001*** |
| 3917 | 1.5e-04<br>[7.9e-05;<br>2.0e-04] | 1.2e-04<br>[5.3e-05;<br>1.8e-04] | 9.1e-05<br>[4.8e-05;<br>1.5e-04] | 2.2e-04<br>[1.5e-04;<br>2.7e-04] | 1.6e-04<br>[8.9e-05;<br>2.4e-04] | 9.6e-05<br>[3.7e-05;<br>1.2e-04] | 7.5e-05<br>[4.6e-05;<br>1.2e-04] | 7.4e-05<br>[5.1e-05;<br>1.0e-04] | 7.9e-05<br>[5.0e-05;<br>1.1e-04] | 1.9e-04<br>[1.5e-04;<br>1.9e-04] | <0.001*** |
| 3934 | 1.8e-04<br>[1.4e-04;<br>3.2e-04] | 1.4e-04<br>[5.6e-05;<br>2.3e-04] | 1.6e-04<br>[1.2e-04;<br>2.2e-04] | 1.1e-04<br>[6.9e-05;<br>1.6e-04] | 1.4e-04<br>[4.4e-05;<br>2.8e-04] | 1.6e-04<br>[8.5e-05;<br>2.3e-04] | 4.0e-05<br>[6.6e-06;<br>1.0e-04] | 7.3e-05<br>[5.6e-05;<br>9.8e-05] | 9.4e-05<br>[6.6e-05;<br>1.4e-04] | 1.2e-05<br>[1.0e-05;<br>7.6e-05] | <0.001*** |
| 3980 | 6.8e-05<br>[1.6e-05;<br>1.1e-04] | 1.9e-04<br>[1.2e-04;<br>2.7e-04] | 1.7e-04<br>[1.1e-04;<br>2.7e-04] | 1.4e-05<br>[1.0e-07;<br>1.9e-04] | 1.1e-04<br>[5.2e-05;<br>1.8e-04] | 3.0e-04<br>[1.4e-04;<br>4.8e-04] | 2.0e-04<br>[8.8e-05;<br>5.8e-04] | 4.4e-04<br>[2.9e-04;<br>5.9e-04] | 2.5e-04<br>[1.1e-04;<br>4.8e-04] | 6.6e-04<br>[1.3e-04;<br>1.8e-03] | <0.001*** |
| 4000 | 1.0e-04<br>[4.2e-05;<br>1.8e-04] | 5.9e-05<br>[2.9e-05;<br>1.1e-04] | 4.8e-05<br>[2.5e-05;<br>1.0e-04] | 5.4e-05<br>[5.5e-06;<br>1.2e-04] | 5.9e-05<br>[2.5e-05;<br>1.0e-04] | 1.2e-04<br>[6.8e-05;<br>3.2e-04] | 8.0e-05<br>[4.8e-05;<br>1.3e-04] | 1.0e-04<br>[7.6e-05;<br>1.3e-04] | 8.4e-05<br>[4.7e-05;<br>1.3e-04] | 1.8e-04<br>[1.3e-04;<br>2.8e-04] | <0.001*** |
| 4136 | 2.7e-04<br>[1.9e-04;<br>5.4e-04] | 1.9e-04<br>[1.1e-04;<br>4.2e-04] | 1.3e-04<br>[6.5e-05;<br>2.0e-04] | 1.1e-04<br>[5.9e-05;<br>2.7e-04] | 4.7e-04<br>[2.2e-04;<br>9.3e-04] | 2.1e-04<br>[1.3e-04;<br>3.3e-04] | 2.5e-04<br>[1.1e-04;<br>4.7e-04] | 7.5e-04<br>[4.1e-04;<br>1.1e-03] | 1.5e-04<br>[7.2e-05;<br>3.1e-04] | 5.7e-04<br>[2.2e-04;<br>1.4e-03] |           |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 4157 | 2.3e-04<br>[1.8e-04;<br>4.1e-04] | 1.4e-04<br>[6.8e-05;<br>3.2e-04] | 8.3e-05<br>[2.2e-05;<br>1.3e-04] | 1.4e-04<br>[3.5e-05;<br>2.4e-04] | 3.6e-04<br>[1.0e-04;<br>9.0e-04] | 1.6e-04<br>[5.6e-05;<br>2.4e-04] | 1.1e-04<br>[6.0e-05;<br>1.9e-04] | 1.0e-04<br>[8.1e-05;<br>1.3e-04] | 6.2e-05<br>[3.0e-05;<br>1.0e-04] | 2.5e-04<br>[1.9e-04;<br>3.6e-04] | <0.001*** |
| 4171 | 9.9e-05<br>[6.2e-05;<br>1.5e-04] | 5.1e-05<br>[8.2e-06;<br>1.1e-04] | 6.7e-05<br>[3.6e-05;<br>1.2e-04] | 8.6e-05<br>[4.8e-05;<br>1.2e-04] | 1.3e-04<br>[3.3e-05;<br>4.0e-04] | 9.7e-05<br>[5.0e-05;<br>2.0e-04] | 1.2e-04<br>[6.9e-05;<br>2.2e-04] | 1.2e-04<br>[9.7e-05;<br>1.7e-04] | 7.5e-05<br>[4.0e-05;<br>1.2e-04] | 1.9e-04<br>[1.6e-04;<br>3.4e-04] | <0.001*** |
| 4193 | 6.2e-05<br>[1.8e-05;<br>8.0e-05] | 1.4e-04<br>[2.2e-05;<br>5.8e-04] | 6.4e-04<br>[3.8e-04;<br>8.5e-04] | 7.3e-05<br>[2.4e-05;<br>1.7e-04] | 1.7e-04<br>[5.5e-05;<br>3.5e-04] | 3.7e-04<br>[2.6e-04;<br>5.8e-04] | 8.7e-05<br>[5.1e-05;<br>1.3e-04] | 9.2e-05<br>[6.7e-05;<br>1.2e-04] | 8.5e-05<br>[5.7e-05;<br>1.4e-04] | 1.8e-04<br>[1.5e-04;<br>4.2e-04] | <0.001*** |
| 4229 | 3.7e-04<br>[1.7e-04;<br>6.8e-04] | 1.5e-04<br>[8.7e-05;<br>2.9e-04] | 1.3e-04<br>[6.6e-05;<br>2.1e-04] | 2.9e-04<br>[2.0e-04;<br>5.9e-04] | 1.3e-04<br>[4.8e-05;<br>3.1e-04] | 7.8e-05<br>[2.6e-05;<br>1.0e-04] | 4.8e-05<br>[1.6e-05;<br>8.8e-05] | 6.4e-05<br>[4.3e-05;<br>8.2e-05] | 6.6e-05<br>[4.0e-05;<br>1.0e-04] | 2.6e-05<br>[5.4e-06;<br>4.3e-05] | <0.001*** |
| 4305 | 1.4e-05<br>[6.3e-07;<br>4.2e-05] | 3.2e-05<br>[8.1e-06;<br>6.9e-05] | 6.4e-05<br>[1.2e-05;<br>9.1e-05] | 4.0e-05<br>[1.6e-05;<br>1.0e-04] | 2.0e-05<br>[1.2e-06;<br>1.0e-04] | 1.7e-04<br>[1.2e-05;<br>2.8e-04] | 7.7e-05<br>[4.1e-05;<br>1.1e-04] | 8.3e-05<br>[5.4e-05;<br>1.1e-04] | 5.9e-05<br>[2.5e-05;<br>8.7e-05] | 7.7e-05<br>[6.7e-05;<br>1.1e-04] | <0.001*** |
| 4356 | 1.8e-04<br>[1.3e-04;<br>2.2e-04] | 3.0e-04<br>[1.9e-04;<br>4.3e-04] | 2.3e-04<br>[1.9e-04;<br>2.8e-04] | 1.8e-04<br>[4.9e-05;<br>2.2e-04] | 3.4e-04<br>[2.1e-04;<br>4.5e-04] | 1.8e-04<br>[1.5e-04;<br>2.3e-04] | 7.0e-05<br>[3.6e-05;<br>1.2e-04] | 8.8e-05<br>[6.1e-05;<br>1.2e-04] | 1.2e-04<br>[7.6e-05;<br>2.3e-04] | 5.9e-05<br>[3.9e-05;<br>5.9e-05] | <0.001*** |
| 4374 | 1.6e-04<br>[1.4e-04;<br>4.2e-04] | 1.8e-04<br>[1.2e-04;<br>5.7e-04] | 3.6e-04<br>[1.3e-04;<br>6.8e-04] | 2.3e-04<br>[1.4e-04;<br>3.9e-04] | 1.9e-04<br>[8.9e-05;<br>3.5e-04] | 6.2e-04<br>[2.3e-04;<br>1.0e-03] | 7.7e-05<br>[4.9e-05;<br>1.0e-04] | 8.7e-05<br>[6.4e-05;<br>1.3e-04] | 4.4e-04<br>[1.2e-04;<br>2.3e-03] | 4.7e-05<br>[3.8e-05;<br>6.0e-05] | <0.001*** |
| 4393 | 1.8e-04<br>[1.7e-04;<br>2.2e-04] | 9.5e-05<br>[4.7e-05;<br>2.4e-04] | 1.9e-04<br>[1.4e-04;<br>3.3e-04] | 1.6e-04<br>[1.1e-04;<br>2.6e-04] | 1.4e-04<br>[8.3e-05;<br>2.3e-04] | 2.9e-04<br>[4.4e-05;<br>5.3e-04] | 6.2e-05<br>[3.6e-05;<br>9.8e-05] | 8.6e-05<br>[6.2e-05;<br>1.0e-04] | 1.1e-04<br>[6.8e-05;<br>2.0e-04] | 4.6e-05<br>[3.6e-05;<br>6.5e-05] | <0.001*** |
| 4428 | 1.4e-04<br>[1.1e-04;<br>2.2e-04] | 6.0e-05<br>[1.7e-05;<br>1.3e-04] | 7.3e-05<br>[4.0e-05;<br>9.8e-05] | 6.1e-05<br>[7.4e-06;<br>1.1e-04] | 6.3e-05<br>[1.8e-05;<br>1.5e-04] | 2.9e-04<br>[2.0e-04;<br>3.5e-04] | 7.7e-05<br>[4.7e-05;<br>1.2e-04] | 8.7e-05<br>[5.8e-05;<br>1.2e-04] | 9.7e-05<br>[6.6e-05;<br>1.4e-04] | 7.4e-05<br>[4.1e-05;<br>8.4e-05] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 4471 | 1.7e-04<br>[1.6e-04;<br>1.9e-04] | 9.2e-05<br>[4.3e-05;<br>1.4e-04] | 1.2e-04<br>[6.1e-05;<br>1.8e-04] | 1.0e-04<br>[6.8e-05;<br>1.2e-04] | 1.1e-04<br>[4.8e-05;<br>2.0e-04] | 1.1e-04<br>[3.1e-05;<br>1.3e-04] | 8.6e-05<br>[4.7e-05;<br>1.1e-04] | 6.7e-05<br>[4.0e-05;<br>9.2e-05] | 6.7e-05<br>[3.6e-05;<br>9.4e-05] | 9.2e-05<br>[8.1e-05;<br>1.4e-04] | <0.001*** |
| 4532 | 5.9e-05<br>[3.1e-05;<br>1.6e-04] | 1.2e-04<br>[8.3e-05;<br>1.8e-04] | 2.0e-04<br>[1.6e-04;<br>2.6e-04] | 7.7e-05<br>[1.4e-05;<br>1.2e-04] | 3.8e-05<br>[1.2e-05;<br>9.6e-05] | 2.5e-04<br>[1.5e-04;<br>3.7e-04] | 5.7e-05<br>[2.9e-05;<br>9.1e-05] | 7.5e-05<br>[5.9e-05;<br>1.0e-04] | 1.1e-04<br>[7.2e-05;<br>1.7e-04] | 7.4e-05<br>[6.6e-05;<br>8.2e-05] | <0.001*** |
| 4551 | 2.5e-04<br>[6.6e-05;<br>2.9e-04] | 2.1e-04<br>[8.2e-05;<br>3.2e-04] | 5.2e-04<br>[2.6e-04;<br>8.0e-04] | 3.2e-04<br>[2.1e-04;<br>5.6e-04] | 1.3e-04<br>[6.1e-05;<br>3.7e-04] | 5.9e-04<br>[4.8e-04;<br>6.6e-04] | 2.7e-04<br>[1.6e-04;<br>4.6e-04] | 8.5e-05<br>[7.0e-05;<br>1.1e-04] | 9.4e-05<br>[6.7e-05;<br>1.3e-04] | 3.3e-04<br>[1.3e-04;<br>4.6e-04] | <0.001*** |
| 4574 | 1.6e-04<br>[1.3e-04;<br>2.2e-04] | 1.4e-04<br>[4.2e-05;<br>2.3e-04] | 1.5e-04<br>[9.6e-05;<br>2.5e-04] | 1.4e-04<br>[9.6e-05;<br>2.6e-04] | 9.6e-05<br>[3.5e-05;<br>1.5e-04] | 1.7e-04<br>[1.4e-04;<br>2.3e-04] | 8.6e-05<br>[5.1e-05;<br>1.2e-04] | 6.7e-05<br>[5.4e-05;<br>1.0e-04] | 5.1e-05<br>[3.5e-05;<br>8.4e-05] | 1.5e-04<br>[1.1e-04;<br>1.9e-04] | <0.001*** |
| 4635 | 1.3e-04<br>[7.2e-05;<br>2.4e-04] | 1.2e-04<br>[7.5e-05;<br>1.8e-04] | 1.7e-04<br>[1.0e-04;<br>3.5e-04] | 1.8e-04<br>[1.3e-04;<br>2.4e-04] | 7.9e-05<br>[4.3e-05;<br>1.2e-04] | 3.0e-04<br>[1.5e-04;<br>4.3e-04] | 1.4e-04<br>[8.6e-05;<br>2.5e-04] | 9.6e-05<br>[7.8e-05;<br>1.2e-04] | 1.1e-04<br>[7.9e-05;<br>1.4e-04] | 2.5e-04<br>[2.1e-04;<br>3.2e-04] | <0.001*** |
| 4714 | 2.0e-04<br>[1.6e-04;<br>4.3e-04] | 1.1e-04<br>[5.7e-05;<br>1.8e-04] | 1.4e-04<br>[8.2e-05;<br>2.9e-04] | 2.3e-04<br>[7.7e-05;<br>5.1e-04] | 8.6e-05<br>[4.0e-05;<br>1.5e-04] | 1.3e-04<br>[7.9e-05;<br>4.5e-04] | 2.9e-04<br>[1.2e-04;<br>8.1e-04] | 1.0e-04<br>[8.2e-05;<br>1.3e-04] | 8.4e-05<br>[4.8e-05;<br>1.1e-04] | 4.9e-04<br>[1.7e-04;<br>9.0e-04] | <0.001*** |
| 4738 | 1.7e-04<br>[1.2e-04;<br>3.1e-04] | 6.0e-05<br>[2.0e-05;<br>1.4e-04] | 3.7e-05<br>[1.3e-05;<br>1.0e-04] | 2.5e-04<br>[1.3e-04;<br>2.6e-04] | 4.1e-05<br>[1.0e-05;<br>1.0e-04] | 9.6e-05<br>[7.4e-05;<br>1.2e-04] | 9.8e-05<br>[5.8e-05;<br>1.5e-04] | 6.7e-05<br>[5.2e-05;<br>9.8e-05] | 5.9e-05<br>[2.6e-05;<br>9.6e-05] | 1.1e-04<br>[7.1e-05;<br>1.2e-04] | <0.001*** |
| 4763 | 8.9e-05<br>[3.1e-05;<br>2.0e-04] | 7.5e-05<br>[3.8e-05;<br>1.2e-04] | 5.6e-05<br>[3.5e-05;<br>9.2e-05] | 1.2e-04<br>[3.9e-05;<br>1.7e-04] | 9.6e-05<br>[2.3e-05;<br>1.9e-04] | 4.3e-05<br>[3.2e-05;<br>1.1e-04] | 7.6e-05<br>[3.7e-05;<br>1.2e-04] | 5.2e-05<br>[3.7e-05;<br>8.2e-05] | 5.1e-05<br>[2.2e-05;<br>8.4e-05] | 1.4e-04<br>[9.7e-05;<br>1.6e-04] | <0.001*** |
| 4811 | 1.4e-04<br>[8.5e-05;<br>1.7e-04] | 4.6e-05<br>[1.7e-05;<br>8.0e-05] | 8.3e-05<br>[4.9e-05;<br>1.1e-04] | 5.3e-05<br>[2.8e-05;<br>9.9e-05] | 4.8e-05<br>[1.4e-05;<br>1.3e-04] | 7.1e-05<br>[4.0e-05;<br>1.2e-04] | 4.7e-05<br>[1.5e-05;<br>1.1e-04] | 6.2e-05<br>[3.4e-05;<br>8.5e-05] | 6.6e-05<br>[4.4e-05;<br>9.4e-05] | 3.9e-05<br>[1.1e-05;<br>9.5e-05] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 4842 | 1.8e-04<br>[1.1e-04;<br>3.0e-04] | 1.0e-04<br>[5.3e-05;<br>2.3e-04] | 1.5e-04<br>[6.7e-05;<br>2.4e-04] | 2.0e-04<br>[1.3e-04;<br>3.1e-04] | 9.9e-05<br>[4.2e-05;<br>2.4e-04] | 1.9e-04<br>[1.8e-04;<br>2.3e-04] | 2.6e-04<br>[1.2e-04;<br>5.9e-04] | 1.2e-04<br>[9.1e-05;<br>1.5e-04] | 9.1e-05<br>[5.3e-05;<br>1.3e-04] | 2.9e-04<br>[1.7e-04;<br>8.7e-04] | <0.001*** |
| 4901 | 1.8e-04<br>[7.3e-05;<br>2.8e-04] | 9.9e-05<br>[4.3e-05;<br>1.9e-04] | 3.3e-04<br>[1.8e-04;<br>5.1e-04] | 2.0e-04<br>[1.3e-04;<br>2.6e-04] | 5.4e-05<br>[2.1e-05;<br>1.2e-04] | 2.8e-04<br>[2.0e-04;<br>4.4e-04] | 1.3e-04<br>[7.3e-05;<br>2.6e-04] | 6.9e-05<br>[5.2e-05;<br>9.5e-05] | 9.2e-05<br>[5.3e-05;<br>1.9e-04] | 4.4e-05<br>[1.7e-05;<br>8.9e-05] | <0.001*** |
| 4941 | 2.7e-04<br>[1.2e-04;<br>8.4e-04] | 1.9e-04<br>[1.3e-04;<br>3.8e-04] | 2.4e-04<br>[1.5e-04;<br>3.5e-04] | 3.1e-04<br>[1.8e-04;<br>4.8e-04] | 2.5e-04<br>[1.1e-04;<br>4.9e-04] | 2.6e-04<br>[1.3e-04;<br>3.0e-04] | 1.2e-04<br>[6.5e-05;<br>2.3e-04] | 3.2e-04<br>[1.8e-04;<br>4.2e-04] | 2.2e-04<br>[1.4e-04;<br>3.4e-04] | 1.2e-04<br>[8.6e-05;<br>1.7e-04] | <0.001*** |
| 4966 | 1.4e-03<br>[3.7e-04;<br>2.4e-03] | 6.8e-04<br>[3.8e-04;<br>1.4e-03] | 5.1e-04<br>[4.1e-04;<br>6.9e-04] | 1.3e-03<br>[9.5e-04;<br>1.8e-03] | 1.2e-03<br>[6.4e-04;<br>2.4e-03] | 3.6e-04<br>[2.5e-04;<br>8.0e-04] | 1.9e-04<br>[7.2e-05;<br>5.3e-04] | 3.4e-04<br>[1.4e-04;<br>8.5e-04] | 2.0e-04<br>[9.8e-05;<br>4.0e-04] | 4.2e-04<br>[1.3e-04;<br>5.2e-04] | <0.001*** |
| 4986 | 9.7e-04<br>[6.2e-04;<br>1.5e-03] | 8.7e-04<br>[4.7e-04;<br>1.4e-03] | 3.1e-04<br>[2.3e-04;<br>3.7e-04] | 9.1e-04<br>[6.0e-04;<br>1.1e-03] | 9.2e-04<br>[5.1e-04;<br>1.6e-03] | 3.4e-04<br>[1.6e-04;<br>4.2e-04] | 1.1e-04<br>[5.0e-05;<br>2.0e-04] | 8.6e-05<br>[6.9e-05;<br>1.4e-04] | 4.7e-05<br>[2.2e-05;<br>9.5e-05] | 7.1e-05<br>[5.9e-05;<br>1.2e-04] | <0.001*** |
| 5005 | 3.5e-04<br>[2.1e-04;<br>6.6e-04] | 4.7e-04<br>[2.0e-04;<br>7.5e-04] | 1.7e-04<br>[1.2e-04;<br>2.3e-04] | 3.7e-04<br>[2.8e-04;<br>4.5e-04] | 5.2e-04<br>[2.4e-04;<br>9.0e-04] | 1.9e-04<br>[3.4e-05;<br>2.5e-04] | 4.5e-05<br>[1.8e-05;<br>9.1e-05] | 9.4e-05<br>[4.1e-05;<br>1.7e-04] | 7.1e-05<br>[3.3e-05;<br>1.0e-04] | 1.6e-05<br>[2.1e-07;<br>7.9e-05] | <0.001*** |
| 5043 | 1.5e-04<br>[1.2e-04;<br>2.9e-04] | 8.0e-05<br>[2.8e-05;<br>1.6e-04] | 5.2e-05<br>[2.4e-05;<br>8.2e-05] | 4.4e-05<br>[9.3e-06;<br>1.6e-04] | 7.8e-05<br>[2.8e-05;<br>2.1e-04] | 2.3e-05<br>[1.1e-05;<br>6.4e-05] | 6.7e-05<br>[2.6e-05;<br>1.0e-04] | 7.3e-05<br>[5.4e-05;<br>1.0e-04] | 5.9e-05<br>[2.3e-05;<br>8.8e-05] | 1.1e-04<br>[7.9e-05;<br>1.8e-04] | <0.001*** |
| 5059 | 5.0e-05<br>[3.5e-05;<br>8.4e-05] | 4.0e-05<br>[9.2e-06;<br>7.2e-05] | 6.3e-05<br>[3.6e-05;<br>1.2e-04] | 7.4e-05<br>[4.0e-05;<br>1.6e-04] | 5.4e-05<br>[5.5e-06;<br>1.9e-04] | 6.2e-05<br>[9.3e-06;<br>1.3e-04] | 9.4e-05<br>[5.7e-05;<br>1.4e-04] | 5.6e-05<br>[2.6e-05;<br>8.5e-05] | 5.6e-05<br>[3.0e-05;<br>8.4e-05] | 1.4e-04<br>[1.2e-04;<br>3.1e-04] | <0.001*** |
| 5136 | 1.6e-04<br>[4.7e-05;<br>2.6e-04] | 1.4e-04<br>[4.2e-05;<br>4.1e-04] | 2.9e-04<br>[1.6e-04;<br>3.7e-04] | 1.9e-04<br>[1.1e-04;<br>2.6e-04] | 8.5e-05<br>[2.2e-05;<br>3.4e-04] | 2.3e-04<br>[1.6e-04;<br>3.9e-04] | 5.4e-04<br>[1.6e-04;<br>1.2e-03] | 8.9e-05<br>[6.6e-05;<br>1.2e-04] | 7.4e-05<br>[4.7e-05;<br>1.1e-04] | 1.4e-03<br>[1.3e-04;<br>1.6e-03] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 5156 | 1.2e-04<br>[7.7e-05;<br>2.4e-04] | 1.6e-04<br>[4.4e-05;<br>3.3e-04] | 1.6e-04<br>[6.2e-05;<br>2.9e-04] | 3.2e-04<br>[1.6e-04;<br>4.0e-04] | 1.2e-04<br>[5.4e-05;<br>4.6e-04] | 3.0e-04<br>[1.9e-04;<br>4.2e-04] | 3.8e-04<br>[1.5e-04;<br>7.9e-04] | 1.3e-04<br>[8.3e-05;<br>1.6e-04] | 9.2e-05<br>[7.0e-05;<br>1.3e-04] | 6.4e-04<br>[1.7e-04;<br>1.4e-03] | <0.001*** |
| 5177 | 2.9e-05<br>[1.4e-05;<br>9.7e-05] | 8.2e-05<br>[4.0e-05;<br>1.6e-04] | 8.5e-05<br>[5.0e-05;<br>1.3e-04] | 9.2e-05<br>[4.4e-05;<br>1.4e-04] | 3.7e-05<br>[1.1e-05;<br>2.0e-04] | 1.2e-04<br>[6.7e-05;<br>1.7e-04] | 7.1e-05<br>[4.0e-05;<br>1.3e-04] | 2.6e-05<br>[4.8e-06;<br>6.0e-05] | 6.1e-05<br>[3.1e-05;<br>1.0e-04] | 1.4e-04<br>[1.3e-04;<br>1.7e-04] | <0.001*** |
| 5191 | 2.8e-06<br>[0.0e+00;<br>6.2e-05] | 2.6e-05<br>[4.9e-06;<br>6.8e-05] | 6.2e-05<br>[3.2e-05;<br>9.0e-05] | 3.0e-05<br>[2.1e-06;<br>8.3e-05] | 2.0e-05<br>[9.0e-07;<br>5.6e-05] | 5.1e-05<br>[3.1e-05;<br>6.7e-05] | 7.6e-05<br>[2.9e-05;<br>1.0e-04] | 4.9e-05<br>[3.3e-05;<br>7.1e-05] | 5.9e-05<br>[2.9e-05;<br>9.3e-05] | 1.1e-04<br>[7.7e-05;<br>1.2e-04] | <0.001*** |
| 5218 | 7.7e-05<br>[3.7e-05;<br>1.3e-04] | 9.9e-05<br>[4.1e-05;<br>2.7e-04] | 2.7e-04<br>[1.6e-04;<br>3.3e-04] | 1.7e-04<br>[9.7e-05;<br>2.0e-04] | 1.7e-04<br>[7.6e-05;<br>3.2e-04] | 1.6e-04<br>[7.6e-05;<br>2.7e-04] | 2.2e-04<br>[1.0e-04;<br>4.3e-04] | 1.3e-04<br>[7.8e-05;<br>1.8e-04] | 1.2e-04<br>[7.3e-05;<br>1.7e-04] | 3.9e-04<br>[1.4e-04;<br>6.7e-04] | <0.001*** |
| 5235 | 1.3e-04<br>[6.4e-05;<br>4.9e-04] | 1.0e-04<br>[3.8e-05;<br>5.9e-04] | 3.2e-04<br>[1.7e-04;<br>6.3e-04] | 5.3e-04<br>[2.1e-04;<br>9.9e-04] | 2.1e-04<br>[6.8e-05;<br>6.1e-04] | 4.2e-04<br>[3.0e-04;<br>5.9e-04] | 1.7e-03<br>[2.5e-04;<br>3.2e-03] | 3.4e-04<br>[2.0e-04;<br>4.6e-04] | 2.1e-04<br>[1.3e-04;<br>3.1e-04] | 3.9e-03<br>[1.5e-04;<br>4.5e-03] | <0.001*** |
| 5256 | 6.2e-05<br>[2.6e-05;<br>1.9e-04] | 7.4e-05<br>[2.3e-05;<br>2.2e-04] | 1.4e-04<br>[5.4e-05;<br>2.3e-04] | 1.8e-04<br>[8.5e-05;<br>2.9e-04] | 7.6e-05<br>[2.0e-05;<br>2.8e-04] | 1.8e-04<br>[1.3e-04;<br>2.1e-04] | 1.7e-04<br>[8.0e-05;<br>4.3e-04] | 5.1e-05<br>[1.4e-05;<br>8.4e-05] | 6.3e-05<br>[2.9e-05;<br>9.1e-05] | 5.3e-04<br>[1.3e-04;<br>8.5e-04] | <0.001*** |
| 5286 | 9.8e-05<br>[5.5e-05;<br>1.2e-04] | 3.7e-05<br>[7.8e-06;<br>9.1e-05] | 1.0e-04<br>[5.7e-05;<br>2.6e-04] | 7.6e-05<br>[4.0e-05;<br>1.5e-04] | 2.3e-05<br>[2.3e-09;<br>9.6e-05] | 1.8e-04<br>[1.5e-04;<br>3.0e-04] | 2.3e-04<br>[1.1e-04;<br>7.9e-04] | 3.7e-04<br>[2.2e-04;<br>5.4e-04] | 1.6e-04<br>[7.1e-05;<br>3.2e-04] | 4.7e-04<br>[1.6e-04;<br>1.4e-03] | <0.001*** |
| 5303 | 4.5e-05<br>[8.7e-06;<br>9.2e-05] | 5.0e-05<br>[2.1e-05;<br>9.7e-05] | 9.3e-05<br>[5.1e-05;<br>1.6e-04] | 6.9e-05<br>[1.4e-05;<br>1.3e-04] | 3.2e-05<br>[4.9e-06;<br>1.3e-04] | 1.8e-04<br>[8.8e-05;<br>3.0e-04] | 1.7e-04<br>[8.0e-05;<br>9.3e-04] | 4.3e-04<br>[2.6e-04;<br>6.6e-04] | 1.7e-04<br>[6.4e-05;<br>4.1e-04] | 2.4e-04<br>[1.8e-04;<br>2.0e-03] | <0.001*** |
| 5382 | 1.8e-04<br>[1.0e-04;<br>4.9e-04] | 1.3e-04<br>[5.4e-05;<br>2.3e-04] | 4.5e-04<br>[2.1e-04;<br>7.6e-04] | 3.8e-04<br>[2.1e-04;<br>5.7e-04] | 1.7e-04<br>[9.4e-05;<br>2.7e-04] | 8.7e-04<br>[6.3e-04;<br>1.0e-03] | 1.9e-03<br>[2.8e-04;<br>5.8e-03] | 1.5e-03<br>[9.2e-04;<br>2.6e-03] | 7.0e-04<br>[1.9e-04;<br>1.6e-03] | 3.2e-03<br>[1.9e-04;<br>7.0e-03] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 5403 | 2.0e-04<br>[8.5e-05;<br>2.9e-04] | 1.3e-04<br>[6.8e-05;<br>2.5e-04] | 2.3e-04<br>[8.4e-05;<br>3.4e-04] | 1.8e-04<br>[1.3e-04;<br>2.1e-04] | 1.2e-04<br>[5.3e-05;<br>2.3e-04] | 3.7e-04<br>[2.5e-04;<br>4.8e-04] | 3.3e-04<br>[9.1e-05;<br>9.5e-04] | 1.6e-04<br>[1.0e-04;<br>2.8e-04] | 1.3e-04<br>[5.0e-05;<br>3.5e-04] | 7.9e-04<br>[1.5e-04;<br>1.2e-03] | <0.001*** |
| 5423 | 3.7e-04<br>[1.9e-04;<br>6.1e-04] | 2.1e-04<br>[1.0e-04;<br>4.6e-04] | 2.0e-04<br>[1.4e-04;<br>2.8e-04] | 3.6e-04<br>[2.0e-04;<br>8.4e-04] | 2.7e-04<br>[8.7e-05;<br>5.4e-04] | 2.1e-04<br>[1.7e-04;<br>3.3e-04] | 3.6e-04<br>[1.3e-04;<br>6.0e-04] | 6.4e-05<br>[3.3e-05;<br>9.1e-05] | 7.9e-05<br>[4.0e-05;<br>1.1e-04] | 3.3e-04<br>[1.5e-04;<br>6.7e-04] | <0.001*** |
| 5510 | 1.7e-05<br>[4.8e-06;<br>3.4e-05] | 6.0e-05<br>[1.6e-05;<br>1.2e-04] | 1.1e-04<br>[7.3e-05;<br>1.8e-04] | 3.4e-05<br>[4.1e-06;<br>9.5e-05] | 2.2e-05<br>[3.9e-06;<br>6.1e-05] | 1.2e-04<br>[3.1e-05;<br>1.8e-04] | 5.0e-05<br>[1.1e-05;<br>8.2e-05] | 5.5e-05<br>[3.0e-05;<br>7.3e-05] | 5.5e-05<br>[2.5e-05;<br>8.7e-05] | 1.1e-04<br>[7.2e-05;<br>1.4e-04] | <0.001*** |
| 5529 | 3.7e-05<br>[3.7e-07;<br>1.1e-04] | 5.1e-05<br>[4.9e-06;<br>1.4e-04] | 1.5e-04<br>[7.6e-05;<br>2.3e-04] | 6.4e-05<br>[2.7e-05;<br>1.6e-04] | 5.2e-05<br>[9.9e-06;<br>1.5e-04] | 2.0e-04<br>[1.3e-04;<br>3.4e-04] | 1.4e-04<br>[6.9e-05;<br>3.1e-04] | 1.8e-04<br>[1.3e-04;<br>2.2e-04] | 1.3e-04<br>[9.4e-05;<br>1.9e-04] | 1.6e-04<br>[1.1e-04;<br>5.1e-04] | <0.001*** |
| 5594 | 8.6e-05<br>[5.0e-05;<br>1.9e-04] | 8.6e-05<br>[4.5e-05;<br>3.2e-04] | 2.0e-04<br>[9.3e-05;<br>4.6e-04] | 1.5e-04<br>[9.6e-05;<br>2.4e-04] | 6.6e-05<br>[2.3e-05;<br>1.6e-04] | 2.3e-04<br>[1.7e-04;<br>5.3e-04] | 5.0e-05<br>[1.3e-05;<br>1.3e-04] | 7.4e-05<br>[5.3e-05;<br>1.6e-04] | 1.9e-04<br>[7.9e-05;<br>8.8e-04] | 3.0e-05<br>[0.0e+00;<br>3.8e-05] | <0.001*** |
| 5687 | 1.3e-04<br>[4.7e-05;<br>2.1e-04] | 1.2e-04<br>[5.6e-05;<br>2.1e-04] | 2.2e-04<br>[1.2e-04;<br>4.3e-04] | 8.9e-05<br>[7.2e-05;<br>1.7e-04] | 7.7e-05<br>[3.1e-05;<br>1.5e-04] | 1.1e-04<br>[4.9e-05;<br>2.5e-04] | 9.2e-05<br>[5.2e-05;<br>1.6e-04] | 3.1e-04<br>[2.1e-04;<br>4.2e-04] | 3.2e-04<br>[1.8e-04;<br>4.8e-04] | 9.7e-05<br>[6.2e-05;<br>1.2e-04] | <0.001*** |
| 5763 | 5.0e-05<br>[1.9e-05;<br>8.0e-05] | 3.7e-05<br>[5.9e-06;<br>8.1e-05] | 7.9e-05<br>[4.6e-05;<br>1.3e-04] | 6.5e-05<br>[4.4e-05;<br>1.0e-04] | 4.4e-05<br>[9.8e-06;<br>9.3e-05] | 5.6e-05<br>[2.3e-05;<br>8.9e-05] | 7.0e-05<br>[3.3e-05;<br>1.6e-04] | 1.5e-04<br>[1.1e-04;<br>2.2e-04] | 9.1e-05<br>[5.4e-05;<br>1.6e-04] | 4.1e-05<br>[6.7e-06;<br>9.8e-05] | <0.001*** |
| 5864 | 3.4e-04<br>[1.5e-04;<br>5.2e-04] | 1.9e-04<br>[1.0e-04;<br>5.7e-04] | 5.0e-04<br>[3.5e-04;<br>6.6e-04] | 3.8e-04<br>[2.9e-04;<br>5.2e-04] | 1.3e-04<br>[7.5e-05;<br>3.7e-04] | 4.8e-04<br>[2.2e-04;<br>5.6e-04] | 1.7e-04<br>[6.3e-05;<br>3.4e-04] | 3.1e-04<br>[2.1e-04;<br>3.9e-04] | 1.9e-04<br>[1.0e-04;<br>3.0e-04] | 1.3e-04<br>[6.6e-05;<br>3.0e-04] | <0.001*** |
| 5948 | 5.9e-05<br>[5.1e-05;<br>2.1e-04] | 6.9e-05<br>[3.0e-05;<br>1.3e-04] | 1.9e-04<br>[1.0e-04;<br>3.1e-04] | 6.3e-05<br>[3.6e-05;<br>8.1e-05] | 4.5e-05<br>[8.9e-06;<br>8.2e-05] | 2.3e-04<br>[1.3e-04;<br>3.2e-04] | 6.8e-05<br>[3.8e-05;<br>1.1e-04] | 1.6e-04<br>[1.1e-04;<br>2.1e-04] | 2.6e-04<br>[1.6e-04;<br>5.0e-04] | 9.0e-05<br>[6.7e-05;<br>1.2e-04] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 5971 | 6.6e-05<br>[6.8e-06;<br>2.3e-04] | 4.8e-05<br>[3.3e-06;<br>1.4e-04] | 1.1e-04<br>[6.0e-05;<br>1.5e-04] | 6.4e-05<br>[9.0e-06;<br>3.0e-04] | 1.7e-05<br>[4.4e-07;<br>6.8e-05] | 9.9e-05<br>[8.3e-05;<br>2.2e-04] | 4.4e-05<br>[1.1e-05;<br>0.4]     | 1.5e-04<br>[8.3e-05;<br>2.0e-04] | 9.5e-05<br>[5.2e-05;<br>1.4e-04] | 1.1e-04<br>[8.5e-05;<br>2.5e-04] | <0.001*** |
| 6192 | 2.1e-04<br>[1.1e-04;<br>2.5e-04] | 2.3e-04<br>[9.8e-05;<br>4.1e-04] | 1.8e-04<br>[1.3e-04;<br>2.8e-04] | 1.4e-04<br>[1.3e-04;<br>1.8e-04] | 1.8e-04<br>[1.2e-04;<br>2.5e-04] | 1.2e-04<br>[5.3e-05;<br>1.4e-04] | 8.1e-05<br>[4.7e-05;<br>1.2e-04] | 1.1e-04<br>[7.8e-05;<br>1.5e-04] | 9.6e-05<br>[5.9e-05;<br>1.4e-04] | 6.5e-05<br>[2.3e-05;<br>9.1e-05] | <0.001*** |
| 6227 | 4.2e-05<br>[2.2e-05;<br>9.4e-05] | 3.1e-05<br>[1.1e-05;<br>7.5e-05] | 5.9e-05<br>[3.8e-05;<br>9.0e-05] | 4.2e-05<br>[2.1e-05;<br>9.0e-05] | 3.4e-05<br>[8.3e-06;<br>7.3e-05] | 5.1e-05<br>[1.8e-05;<br>8.4e-05] | 2.8e-05<br>[7.4e-06;<br>5.2e-05] | 7.9e-05<br>[5.2e-05;<br>1.0e-04] | 6.5e-05<br>[3.7e-05;<br>1.0e-04] | 7.9e-05<br>[6.0e-05;<br>1.3e-04] | <0.001*** |
| 6278 | 6.0e-05<br>[2.8e-05;<br>8.8e-05] | 3.7e-05<br>[9.5e-06;<br>6.5e-05] | 6.5e-05<br>[2.8e-05;<br>9.5e-05] | 4.6e-05<br>[1.3e-05;<br>7.1e-05] | 2.3e-05<br>[5.9e-06;<br>4.7e-05] | 8.3e-05<br>[6.4e-05;<br>1.0e-04] | 3.6e-05<br>[1.9e-05;<br>8.3e-05] | 1.3e-04<br>[9.9e-05;<br>1.7e-04] | 8.2e-05<br>[5.6e-05;<br>1.1e-04] | 2.7e-05<br>[1.5e-05;<br>4.8e-05] | <0.001*** |
| 6361 | 2.2e-04<br>[1.7e-04;<br>2.5e-04] | 8.9e-05<br>[3.4e-05;<br>1.6e-04] | 6.5e-05<br>[2.5e-05;<br>1.1e-04] | 1.4e-04<br>[5.1e-05;<br>2.7e-04] | 1.4e-04<br>[4.7e-05;<br>2.4e-04] | 8.4e-05<br>[1.7e-05;<br>1.2e-04] | 6.5e-05<br>[2.1e-05;<br>1.5e-04] | 5.7e-05<br>[2.6e-05;<br>7.3e-05] | 6.9e-05<br>[4.7e-05;<br>1.0e-04] | 1.6e-05<br>[4.3e-06;<br>8.7e-05] | <0.001*** |
| 6639 | 7.1e-05<br>[3.3e-05;<br>1.0e-04] | 8.2e-05<br>[4.7e-05;<br>1.2e-04] | 2.3e-04<br>[1.5e-04;<br>3.5e-04] | 1.3e-04<br>[1.1e-04;<br>1.5e-04] | 7.1e-05<br>[3.9e-05;<br>1.6e-04] | 9.5e-05<br>[6.0e-05;<br>2.0e-04] | 3.8e-05<br>[1.3e-05;<br>1.0e-04] | 6.2e-05<br>[4.6e-05;<br>9.7e-05] | 8.4e-05<br>[6.2e-05;<br>1.4e-04] | 3.5e-06<br>[1.4e-06;<br>3.1e-05] | <0.001*** |
| 6953 | 4.5e-04<br>[1.7e-04;<br>6.5e-04] | 1.4e-04<br>[3.4e-05;<br>8.2e-04] | 2.8e-04<br>[1.2e-04;<br>8.4e-04] | 4.5e-04<br>[1.8e-04;<br>1.5e-03] | 1.9e-04<br>[5.2e-05;<br>7.3e-04] | 1.4e-04<br>[7.0e-05;<br>1.2e-03] | 5.5e-04<br>[1.5e-04;<br>1.5e-03] | 1.2e-03<br>[4.9e-04;<br>1.6e-03] | 7.2e-04<br>[2.9e-04;<br>1.4e-03] | 1.5e-04<br>[1.5e-04;<br>1.9e-03] | <0.001*** |
| 6974 | 2.5e-04<br>[1.6e-04;<br>4.4e-04] | 8.3e-05<br>[2.4e-05;<br>2.9e-04] | 1.4e-04<br>[9.5e-05;<br>3.3e-04] | 1.9e-04<br>[9.3e-05;<br>6.3e-04] | 9.8e-05<br>[4.2e-05;<br>4.4e-04] | 1.9e-04<br>[1.1e-04;<br>6.4e-04] | 1.5e-04<br>[7.2e-05;<br>4.1e-04] | 1.2e-04<br>[7.7e-05;<br>1.9e-04] | 9.6e-05<br>[6.4e-05;<br>1.6e-04] | 1.1e-04<br>[7.4e-05;<br>3.4e-04] | <0.001*** |
| 7093 | 3.6e-05<br>[1.3e-05;<br>7.0e-05] | 4.0e-05<br>[1.1e-05;<br>7.9e-05] | 7.0e-05<br>[3.9e-05;<br>9.8e-05] | 8.5e-05<br>[3.6e-05;<br>1.5e-04] | 2.3e-05<br>[5.6e-06;<br>4.4e-05] | 4.8e-05<br>[2.5e-05;<br>5.7e-05] | 7.9e-05<br>[3.3e-05;<br>1.3e-04] | 7.2e-05<br>[4.6e-05;<br>1.0e-04] | 9.9e-05<br>[7.2e-05;<br>1.4e-04] | 8.3e-05<br>[4.1e-05;<br>1.3e-04] | <0.001*** |

Продолжение таблицы Е.1

| 1    | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 7207 | 2.2e-05<br>[1.4e-05;<br>3.5e-05] | 2.5e-05<br>[3.8e-06;<br>5.3e-05] | 2.2e-05<br>[4.5e-06;<br>5.2e-05] | 5.7e-05<br>[3.3e-06;<br>8.7e-05] | 7.8e-06<br>[1.2e-07;<br>2.8e-05] | 1.5e-05<br>[4.3e-06;<br>4.9e-05] | 1.4e-05<br>[1.9e-06;<br>5.1e-05] | 1.4e-04<br>[1.0e-04;<br>1.9e-04] | 1.1e-04<br>[6.5e-05;<br>1.7e-04] | 2.5e-05<br>[1.2e-05;<br>2.9e-05] | <0.001*** |
| 7270 | 1.3e-05<br>[0.0e+00;<br>3.1e-05] | 6.1e-05<br>[2.1e-05;<br>1.1e-04] | 9.4e-05<br>[1.9e-05;<br>1.6e-04] | 2.9e-05<br>[0.0e+00;<br>6.2e-05] | 2.1e-05<br>[3.0e-08;<br>5.7e-05] | 6.5e-05<br>[3.0e-05;<br>1.1e-04] | 5.6e-05<br>[2.3e-05;<br>1.1e-04] | 2.8e-04<br>[1.5e-04;<br>5.6e-04] | 2.1e-04<br>[1.1e-04;<br>3.7e-04] | 5.0e-05<br>[1.6e-05;<br>1.8e-04] | <0.001*** |
| 7349 | 1.3e-03<br>[8.0e-04;<br>1.7e-03] | 4.8e-04<br>[2.2e-04;<br>1.0e-03] | 6.7e-04<br>[4.1e-04;<br>1.0e-03] | 8.5e-04<br>[6.9e-04;<br>1.1e-03] | 4.6e-04<br>[1.8e-04;<br>8.5e-04] | 5.5e-04<br>[2.5e-04;<br>7.6e-04] | 4.1e-04<br>[1.3e-04;<br>1.0e-03] | 7.3e-05<br>[4.7e-05;<br>1.6e-04] | 1.1e-04<br>[5.7e-05;<br>1.7e-04] | 1.2e-04<br>[1.0e-04;<br>3.9e-04] | <0.001*** |
| 7531 | 9.6e-05<br>[4.8e-05;<br>1.5e-04] | 4.4e-05<br>[3.0e-06;<br>1.6e-04] | 9.3e-05<br>[3.6e-05;<br>2.2e-04] | 1.6e-04<br>[7.8e-05;<br>3.2e-04] | 8.8e-05<br>[2.6e-05;<br>1.8e-04] | 1.3e-04<br>[7.8e-05;<br>2.4e-04] | 9.4e-05<br>[3.5e-05;<br>2.1e-04] | 2.5e-04<br>[1.7e-04;<br>3.4e-04] | 1.4e-04<br>[8.7e-05;<br>2.2e-04] | 8.5e-05<br>[4.2e-05;<br>2.3e-04] | <0.001*** |
| 7611 | 9.8e-05<br>[7.3e-05;<br>1.9e-04] | 5.8e-05<br>[1.8e-05;<br>1.5e-04] | 2.1e-04<br>[9.8e-05;<br>6.9e-04] | 3.1e-04<br>[2.5e-04;<br>6.7e-04] | 9.1e-05<br>[3.5e-05;<br>3.5e-04] | 6.0e-04<br>[1.5e-04;<br>1.1e-03] | 2.8e-04<br>[8.1e-05;<br>7.3e-04] | 7.0e-04<br>[3.0e-04;<br>8.3e-04] | 3.4e-04<br>[1.7e-04;<br>7.3e-04] | 1.6e-04<br>[1.1e-04;<br>7.1e-04] | <0.001*** |
| 7654 | 2.1e-05<br>[0.0e+00;<br>7.0e-05] | 3.0e-05<br>[9.3e-06;<br>6.1e-05] | 8.2e-05<br>[4.7e-05;<br>1.2e-04] | 6.9e-05<br>[2.8e-05;<br>1.3e-04] | 4.1e-05<br>[1.5e-05;<br>8.9e-05] | 1.6e-04<br>[1.1e-04;<br>2.4e-04] | 4.3e-05<br>[1.8e-05;<br>7.2e-05] | 3.4e-05<br>[2.3e-05;<br>5.5e-05] | 4.5e-05<br>[3.1e-05;<br>7.9e-05] | 2.3e-05<br>[1.8e-05;<br>8.0e-05] | <0.001*** |
| 7765 | 2.1e-05<br>[6.7e-06;<br>4.2e-05] | 2.8e-05<br>[5.9e-06;<br>6.6e-05] | 9.8e-05<br>[5.3e-05;<br>1.5e-04] | 5.4e-05<br>[1.3e-05;<br>1.0e-04] | 1.9e-05<br>[3.7e-06;<br>3.7e-05] | 1.3e-04<br>[1.0e-04;<br>3.4e-04] | 2.1e-05<br>[7.9e-06;<br>5.3e-05] | 4.5e-05<br>[2.9e-05;<br>6.9e-05] | 5.6e-05<br>[2.7e-05;<br>9.3e-05] | 1.6e-05<br>[7.8e-06;<br>2.8e-05] | <0.001*** |
| 7932 | 1.9e-04<br>[1.1e-04;<br>2.0e-04] | 4.8e-05<br>[1.5e-05;<br>8.9e-05] | 6.9e-05<br>[3.6e-05;<br>1.1e-04] | 1.5e-04<br>[5.3e-05;<br>2.3e-04] | 5.8e-05<br>[2.6e-05;<br>2.1e-04] | 7.9e-05<br>[1.1e-05;<br>1.1e-04] | 4.2e-05<br>[1.7e-05;<br>6.9e-05] | 1.8e-04<br>[1.2e-04;<br>2.8e-04] | 1.2e-04<br>[8.0e-05;<br>2.1e-04] | 4.1e-05<br>[2.5e-05;<br>8.2e-05] | <0.001*** |
| 8215 | 1.3e-04<br>[1.1e-04;<br>1.9e-04] | 5.5e-05<br>[2.7e-05;<br>1.0e-04] | 6.9e-05<br>[4.4e-05;<br>1.0e-04] | 7.0e-05<br>[5.0e-05;<br>1.1e-04] | 2.8e-05<br>[1.1e-05;<br>6.7e-05] | 1.8e-04<br>[3.3e-05;<br>6.6e-04] | 6.7e-05<br>[3.6e-05;<br>1.3e-04] | 7.9e-05<br>[6.2e-05;<br>1.1e-04] | 7.2e-05<br>[5.2e-05;<br>1.0e-04] | 6.2e-05<br>[4.5e-05;<br>7.6e-05] | <0.001*** |

Продолжение таблицы Е.1

| 1     | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|-------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 8452  | 2.5e-04<br>[1.4e-04;<br>7.2e-04] | 1.5e-04<br>[7.5e-05;<br>2.6e-04] | 8.1e-05<br>[5.6e-05;<br>1.2e-04] | 2.9e-04<br>[1.5e-04;<br>5.2e-04] | 9.3e-05<br>[5.4e-05;<br>1.5e-04] | 5.1e-05<br>[2.7e-05;<br>7.2e-05] | 5.8e-05<br>[2.2e-05;<br>1.1e-04] | 6.9e-05<br>[5.7e-05;<br>1.0e-04] | 6.6e-05<br>[4.0e-05;<br>9.4e-05] | 4.1e-05<br>[1.9e-05;<br>4.9e-05] | <0.001*** |
| 8469  | 1.8e-04<br>[1.3e-04;<br>3.9e-04] | 6.1e-05<br>[2.4e-05;<br>1.1e-04] | 5.4e-05<br>[1.8e-05;<br>8.4e-05] | 1.2e-04<br>[5.0e-05;<br>2.0e-04] | 4.2e-05<br>[2.2e-05;<br>8.4e-05] | 5.3e-05<br>[2.7e-05;<br>7.4e-05] | 4.4e-05<br>[2.2e-05;<br>9.2e-05] | 6.1e-05<br>[4.2e-05;<br>8.1e-05] | 5.4e-05<br>[3.4e-05;<br>7.2e-05] | 2.2e-05<br>[1.5e-05;<br>5.6e-05] | <0.001*** |
| 8568  | 1.0e-04<br>[7.6e-05;<br>1.3e-04] | 1.4e-04<br>[6.4e-05;<br>2.3e-04] | 3.1e-04<br>[2.6e-04;<br>4.9e-04] | 1.5e-04<br>[6.9e-05;<br>3.1e-04] | 1.6e-04<br>[9.6e-05;<br>3.1e-04] | 2.4e-04<br>[1.4e-04;<br>4.1e-04] | 1.1e-04<br>[5.3e-05;<br>1.8e-04] | 4.4e-05<br>[3.4e-05;<br>6.7e-05] | 6.4e-05<br>[4.0e-05;<br>8.3e-05] | 1.5e-04<br>[1.0e-04;<br>1.5e-04] | <0.001*** |
| 8589  | 3.5e-05<br>[1.0e-05;<br>4.3e-05] | 4.0e-05<br>[6.7e-06;<br>8.6e-05] | 1.4e-04<br>[1.1e-04;<br>1.7e-04] | 1.1e-04<br>[5.6e-05;<br>1.8e-04] | 4.6e-05<br>[1.7e-05;<br>7.8e-05] | 9.7e-05<br>[5.5e-05;<br>1.5e-04] | 4.8e-05<br>[2.1e-05;<br>8.3e-05] | 4.0e-05<br>[2.4e-05;<br>6.6e-05] | 4.3e-05<br>[2.4e-05;<br>6.9e-05] | 2.0e-05<br>[1.3e-05;<br>3.0e-05] | <0.001*** |
| 8742  | 1.1e-04<br>[8.7e-05;<br>1.5e-04] | 3.9e-05<br>[1.3e-05;<br>6.6e-05] | 4.7e-05<br>[2.9e-05;<br>8.4e-05] | 7.2e-05<br>[4.4e-05;<br>1.0e-04] | 2.6e-05<br>[9.3e-06;<br>5.0e-05] | 5.9e-05<br>[1.7e-05;<br>9.0e-05] | 5.7e-05<br>[2.1e-05;<br>1.5e-04] | 6.2e-05<br>[4.4e-05;<br>8.1e-05] | 4.9e-05<br>[3.3e-05;<br>6.9e-05] | 5.7e-05<br>[1.5e-05;<br>1.5e-04] | <0.001*** |
| 9956  | 1.4e-04<br>[1.2e-04;<br>1.7e-04] | 7.4e-05<br>[4.0e-05;<br>1.1e-04] | 9.9e-05<br>[7.8e-05;<br>1.3e-04] | 9.3e-05<br>[7.4e-05;<br>1.1e-04] | 5.2e-05<br>[3.1e-05;<br>7.9e-05] | 6.6e-05<br>[3.1e-05;<br>1.1e-04] | 4.4e-05<br>[2.3e-05;<br>7.0e-05] | 6.0e-05<br>[4.3e-05;<br>8.2e-05] | 6.7e-05<br>[5.2e-05;<br>8.8e-05] | 5.0e-05<br>[2.5e-05;<br>7.6e-05] | <0.001*** |
| 10096 | 2.3e-04<br>[5.2e-05;<br>4.0e-04] | 9.1e-05<br>[4.2e-05;<br>1.7e-04] | 8.2e-05<br>[4.6e-05;<br>9.8e-05] | 2.6e-04<br>[1.7e-04;<br>4.8e-04] | 6.1e-05<br>[3.3e-05;<br>1.2e-04] | 9.1e-05<br>[5.4e-05;<br>1.1e-04] | 5.8e-05<br>[3.1e-05;<br>1.0e-04] | 5.5e-05<br>[4.2e-05;<br>8.8e-05] | 5.5e-05<br>[3.7e-05;<br>7.4e-05] | 3.5e-05<br>[1.7e-05;<br>4.2e-05] | <0.001*** |
| 10116 | 1.5e-04<br>[6.1e-05;<br>1.9e-04] | 6.7e-05<br>[3.6e-05;<br>1.3e-04] | 6.8e-05<br>[2.8e-05;<br>8.9e-05] | 1.5e-04<br>[1.2e-04;<br>2.2e-04] | 3.6e-05<br>[1.6e-05;<br>7.2e-05] | 4.4e-05<br>[2.0e-05;<br>7.5e-05] | 3.4e-05<br>[1.4e-05;<br>7.2e-05] | 6.5e-05<br>[4.6e-05;<br>7.8e-05] | 5.7e-05<br>[3.0e-05;<br>8.2e-05] | 1.0e-05<br>[9.4e-06;<br>2.0e-05] | <0.001*** |
| 10437 | 7.1e-05<br>[3.1e-05;<br>1.3e-04] | 6.8e-05<br>[3.6e-05;<br>1.3e-04] | 9.4e-05<br>[5.8e-05;<br>1.8e-04] | 1.7e-04<br>[1.3e-04;<br>2.6e-04] | 9.5e-05<br>[4.8e-05;<br>1.8e-04] | 1.3e-04<br>[7.8e-05;<br>2.7e-04] | 1.1e-04<br>[5.5e-05;<br>2.1e-04] | 2.0e-04<br>[1.4e-04;<br>3.0e-04] | 1.6e-04<br>[1.0e-04;<br>2.6e-04] | 1.3e-04<br>[1.0e-04;<br>2.2e-04] | <0.001*** |

Продолжение таблицы Е.1

| 1     | 2                                | 3                                | 4                                | 5                                | 6                                | 7                                | 8                                | 9                                | 10                               | 11                               | 12        |
|-------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|-----------|
| 10760 | 8.3e-05<br>[3.0e-05;<br>2.0e-04] | 9.2e-05<br>[1.6e-05;<br>1.7e-04] | 1.4e-04<br>[7.4e-05;<br>2.0e-04] | 1.5e-04<br>[2.9e-05;<br>2.2e-04] | 8.0e-05<br>[1.8e-05;<br>3.6e-04] | 9.5e-05<br>[5.2e-05;<br>1.8e-04] | 7.1e-05<br>[2.9e-05;<br>1.3e-04] | 1.6e-04<br>[1.2e-04;<br>2.1e-04] | 5.5e-05<br>[2.5e-05;<br>1.0e-04] | 1.5e-04<br>[1.2e-04;<br>2.8e-04] | <0.001*** |
| 10837 | 2.2e-04<br>[2.0e-04;<br>4.7e-04] | 1.8e-04<br>[1.1e-04;<br>5.2e-04] | 1.7e-04<br>[7.8e-05;<br>2.7e-04] | 1.8e-04<br>[1.2e-04;<br>4.9e-04] | 3.0e-04<br>[1.7e-04;<br>6.8e-04] | 7.7e-05<br>[6.1e-05;<br>9.7e-05] | 8.5e-05<br>[3.8e-05;<br>1.3e-04] | 5.0e-05<br>[3.1e-05;<br>6.8e-05] | 5.8e-05<br>[3.5e-05;<br>7.8e-05] | 3.2e-05<br>[1.5e-05;<br>5.9e-05] | <0.001*** |
| 11011 | 5.5e-05<br>[5.7e-06;<br>7.7e-05] | 8.6e-05<br>[2.7e-05;<br>1.3e-04] | 2.4e-04<br>[1.6e-04;<br>3.1e-04] | 3.3e-05<br>[1.3e-05;<br>7.7e-05] | 1.3e-05<br>[4.3e-08;<br>6.6e-05] | 1.5e-04<br>[1.1e-04;<br>2.4e-04] | 2.7e-05<br>[9.2e-06;<br>6.5e-05] | 4.7e-05<br>[3.1e-05;<br>6.6e-05] | 7.2e-05<br>[4.7e-05;<br>1.0e-04] | 4.8e-05<br>[1.7e-05;<br>6.5e-05] | <0.001*** |
| 11725 | 2.9e-04<br>[1.2e-04;<br>4.1e-04] | 1.2e-04<br>[5.2e-05;<br>2.5e-04] | 3.5e-04<br>[1.8e-04;<br>4.1e-04] | 2.3e-04<br>[1.2e-04;<br>2.8e-04] | 1.2e-04<br>[5.9e-05;<br>2.0e-04] | 2.9e-04<br>[2.0e-04;<br>3.6e-04] | 1.3e-04<br>[7.6e-05;<br>2.3e-04] | 3.1e-04<br>[2.2e-04;<br>4.7e-04] | 1.6e-04<br>[8.8e-05;<br>2.4e-04] | 9.3e-05<br>[6.7e-05;<br>2.6e-04] | <0.001*** |
| 14692 | 6.0e-04<br>[3.1e-04;<br>6.9e-04] | 3.3e-04<br>[1.6e-04;<br>6.6e-04] | 3.8e-04<br>[1.9e-04;<br>5.5e-04] | 3.0e-04<br>[2.5e-04;<br>4.0e-04] | 2.6e-04<br>[1.8e-04;<br>5.3e-04] | 2.6e-04<br>[1.2e-04;<br>3.4e-04] | 1.8e-04<br>[9.9e-05;<br>4.1e-04] | 7.7e-05<br>[5.5e-05;<br>1.0e-04] | 8.1e-05<br>[5.4e-05;<br>1.3e-04] | 1.4e-04<br>[1.3e-04;<br>2.6e-04] | <0.001*** |