

Получена: 28 ноябрь 2022 / Принята: 20 январь 2023 / Опубликовано online: 15 марта 2023 г.
 УДК 616.9:578.834.1-084-071
 DOI 10.53511/PHARMKAZ.2023.96.89.041

А.А. Султанбекова¹, И.А. Кадырова¹, Д.Б. Бабенко¹

¹Ғылыми орталық, Қарағанды медицина университеті, Қарағанды, Қазақстан

MALDI-TOF МАСС-СПЕКТРОМЕТРИЯ ЖӘНЕ МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ АРҚЫЛЫ SARS-COV-2 ВИРУСЫН ДИАГНОСТИКАЛАУДЫҢ ЭКСПРЕСС ӘДІСІН БАҒАЛАУ

Түйін: COVID-19 адамзат үшін, әсіресе денсаулық сақтау жүйесі үшін ауыртпалығын көрсетті. Пандемия емдік-диагностикалық процестерді ұйымдастырудағы көптеген олқылықтарды анықтады және бұл аурушаңдық ең қиын кезеңде стационарлық және зертханалық қызметтерінің құлдырауына әкелді.

Мақсаты: MALDI-MS комбинациясын және машиналық оқыту әдісін қолдана отырып, мұрын жағындыларында SARS-CoV-2 вирусын жылдам анықтау әдісін бағалау.

Әдістер: SARS-CoV-2 КТ-ПТР және MALDI-TOF MS сынағы үшін COVID-19 оң (n=203) және теріс (n=101) субъектілерінен мұрын-жұтқыншақ жағындылары жиналды. Алынған MALDI спектрлері сапаны бақылау, түрлендіру және тегістеу, базалық түзету, қарқындылықты калибрлеу және т.б. қоса, әдеттегі жұмыс үрдісіне сәйкес алдын ала өңделді. PCA өлшемділікті азайту әдісі ретінде шыңдардың алдын ала өңделген матрицасында оқытылған және сыналған төрт машиналық оқыту әдісі (GLM, CART, RF және XGBoost) арқылы орындалды.

Нәтижелер: RF, CART және XGBoost машиналық оқыту үлгілері SARS-CoV-2 оң/теріс үлгілерін 100% дәлдікпен, 100% сезімталдықпен және 100% ерекшелігімен ажырату қабілетінде ең жақсы нәтижелерді көрсетті, ал GLM моделі сыналған деректер жиынтығы бойынша 71,4% F- score және 88,8% дәлдікке қол жеткізді.

Қорытындылар: MALDI-TOF MS мен машиналық оқыту талдауын пайдалана отырып, мұрын жағынды материалында SARS-CoV-2 анықтаудың ұсынылған әдісі қолжетімді, арзан және жоғары өткізу қабілеті бар сынақ болып табылады. Биоматериал үлгілерін алдын ала дайындаудың қажетсіздігі және қымбат реагенттерді пайдаланбауы бұл әдісті скринингтік әдіс ретінде жіктейді.

Түйінді сөздер: COVID-19; SARS-CoV-2; MALDI-TOF MS; машиналық оқыту; зертханалық диагностика; скрининг; экспресс әдіс

А.А. Султанбекова¹, И.А. Кадырова¹, Д.Б. Бабенко¹

¹Научно-исследовательский Центр, Медицинский Университет Караганды, г. Караганда, Казахстан

ОЦЕНКА ЭКСПРЕСС-МЕТОДА ОПРЕДЕЛЕНИЯ ВИРУСА SARS-COV-2 ПРИ ПОМОЩИ MALDI-TOF МАСС-СПЕКТРОМЕТРИИ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Резюме: COVID-19 стал тяжелым испытанием для человечества и, особенно, для системы здравоохранения. Пандемия выявила множество пробелов в организации лечебно-диагностических процессов и привела к коллапсу стационарной и лабораторной службы в периоды наибольшей заболеваемости.

Цель: Оценить экспресс-метод обнаружения вируса SARS-CoV-2 в мазках из носа с использованием комбинации MALDI-MS и подхода машинного обучения.

A.A. Sultanbekova¹, I.A. Kadyrova¹, D.B. Babenko¹

¹Research Center, Medical University of Karaganda, Karaganda, Kazakhstan

EVALUATION OF THE EXPRESS METHOD FOR THE DETECTION OF THE SARS-COV-2 VIRUS USING MALDI-TOF MASS SPECTROMETRY AND MACHINE LEARNING APPROACH

Resume: COVID-19 became a travail for humanity and, especially, for the healthcare system. The pandemic revealed many gaps in the organization of medical and diagnostic processes and it led to the collapse of the inpatient and laboratory services during periods of the highest incidence.

Aim: To evaluate a rapid detection method for the SARS-CoV-2 virus in nasal swabs using a combination of MALDI-MS and a machine

Методы. Мазки из носоглотки были собраны для тестирования SARS-CoV-2 ОТ-ПЦР и MALDI-TOF MS у положительных на COVID-19 (n = 203) и отрицательных (n = 101) субъектов. Полученные спектры MALDI были предварительно обработаны в соответствии с обычным рабочим процессом, включая контроль качества, преобразование и сглаживание, коррекцию базовой линии, калибровку интенсивности и т. д. PCA как метод уменьшения размерности был выполнен с использованием четырех обученных методов машинного обучения (GLM, CART, RF и XGBoost) и протестирован на предварительно обработанной матрице пиков.

Результаты. Модели машинного обучения RF, CART и XGBoost продемонстрировали наилучшие результаты в способности различать положительные и отрицательные образцы SARS-CoV-2 с точностью 100 %, чувствительностью 100 % и специфичностью 100 %, в то время как модель GLM достигли 71,4% F- score и 88,8% точности на протестированном наборе данных.

Выводы: Предлагаемый метод выявления SARS-CoV-2 в материале, полученном из мазка из носа, с использованием MALDI-TOF MS и анализа с помощью машинного обучения, является доступным, малозатратным и высокопроизводительным тестом. Отсутствие предварительной подготовки образцов биоматериала и использование дорогостоящих реагентов относят этот метод к скрининговым.

Ключевые слова: COVID-19; SARS-CoV-2; MALDI-TOF MS; машинное обучение; лабораторная диагностика; скрининг; экспресс-метод

learning approach.

Methods: Nasopharyngeal swabs were collected for SARS-CoV-2 RT-qPCR and MALDI-TOF MS testing from COVID-19 positive (n=203) and negative (n=101) subjects. Obtained MALDI spectra were preprocessed according to the conventional workflow including quality control, transformation, and smoothing, baseline correction, intensity calibration, etc. PCA as a dimensionality reduction technique was performed following four machine learning methods (GLM, CART, RF and XGBoost) trained and tested on preprocessed peak matrix.

Results: Machine learning models RF, CART and XGBoost demonstrated the best results in the ability to distinct SARS-CoV-2 positive/negative samples with an accuracy of 100%, sensitivity of 100%, and specificity of 100%, while the GLM model achieved 71.4 % F-score and 88.8% in accuracy on the tested dataset.

Conclusions: The proposed method for detecting SARS-CoV-2 in the material obtained from a nasal swab using MALDI-TOF MS and the machine learning analysis is an accessible, low time-consuming and high throughput test. The absence of the preliminary preparation of biomaterial samples and the usage of expensive reagents qualify this technique as a screening method.

Keywords: COVID-19; SARS-CoV-2; MALDI-TOF MS; machine learning; laboratory diagnostics; screening; express method

Кіріспе. 2019 жылы пайда болған жаңа коронавирустық инфекция адамзат үшін, әсіресе денсаулық сақтау жүйесі үшін ауыртпалық болды. COVID-19 тез таралуы және жаңа штаммдардың пайда болуына байланысты денсаулық сақтау мәселесі болып қала береді. Әсіресе, эпидемиологиялық қадағалаудың жеткіліксіздігінен зардап шегетін қоғамдық денсаулық сақтау инфрақұрылымы шектеулі аймақтарға зиян тигізеді [1]. Пандемия емдік-диагностикалық процестерді ұйымдастырудағы көптеген олқылықтарды анықтады және бұл аурушаңдық ең қиын кезеңдердегі стационарлық және зертханалық қызметтің құлдырауына әкелді [2].

Заманауи диагностикалық әдістерді әзірлеу мен қолданудағы елеулі прогреске қарамастан, SARS-CoV-2 анықтау эпидемиологиялық қызмет, емхана дәрігерлері және қоғам үшін ең өзекті мәселе болып қала береді. Бұл кезде аурудың нәтижесі дер кезінде диагностика мен емдеуге тікелей байланысты. Клиникалық зертханалық диагностикада SARS-CoV-2 вирусының элементтерін анықтауға немесе оның ағзада болуының жанама белгілерін анықтауға арналған диагностикалық әдістердің кең спектрі бар [4, 5, 6]. Қазіргі уақытта коронавирустық инфекция диагностикасының «алтын стандарты» кері транскрипциялық ПТР (КТ-ПТР) арқылы биоматериалдағы вирустық РНҚ-ны анықтау болып табылады. КТ-ПТР – жоғары дәлдікте-

гі әдіс; дегенмен, осы диагностикалық әдісті қолдану стандартталған зертханалық хаттамаларды, күшейту алдындағы, күшейту және кейінгі күшейту аймақтары анық бөлінген, жабдықталған зертханаларды қажет етеді. [7, 8, 9]. Сонымен қатар, зертханалық диагностикалық сынақтар көптеген салалар үшін қажет: симптомсыз популяция скринингі, жоғары қауіп тобының мақсатты скринингі, клиникалық диагностика, инфекциялық мониторинг. КТ-ПТР жоғарыда аталған барлық тапсырмалардағы талапты орындай алмайды. Скринингтік тапсырмалар үшін COVID-19 инфекциясын анықтаудың жылдам және дәл зертханалық әдістері әзірленуі керек. Қолайлы нұсқалардың бірі – MALDI-TOF масс-спектрометриясы.

MALDI-TOF MS клиникалық микробиологияда микробтық изолаттарды скрининг үшін кеңінен қолданылады, дегенмен диагностикада, болжамда, емдік препараттарды бақылауда және көптеген ауруларды биопсиялық бейнелеуде MALDI-TOF MS қолданылуы оның баламалы әдістермен салыстырғанда жоғары сезімталдығы мен ерекшелігіне байланысты орасан зор әлеует пен мүмкіндіктер беріледі. Алынған MALDI-TOF MS үлгілері күрделі сондықтан машиналық оқыту (ML) сияқты қосымша аналитикалық тәсілдерді қажет етеді.

Fabiane M. Nachtigall және бірлескен авторлар осы салада бірінші болып MALDI-TOF MS және ML (0,93 дәл-

дік) перспективті комбинациясын көрсетті. Осыған ұқсас зерттеулерді Meritxell Deulofeu мен басқалары (0,98 дәлдік), Nam K. Tran мен басқа (0,983 дәлдік) авторлары жүргізді. Өртүрлі тасымалдау орталары, ML алгоритмдері мен түрлі жабдықтар ұсынылды. Біздің ойымызша, массалық спектрлердің үлгілерін анықтау үшін жалпылама сызықтық модельдер (GLM) [16], классификация және регрессия ағашы (CART) [17], кездейсоқ орман (RF) [18] және экстремалды градиентті күшейту (XGBoost) ML алгоритмдері жоғары дәлдікке қол жеткізуге мүмкіндік береді, сондай-ақ R ортасында алдын ала өңдеу бұл әдісті зерттеушілер мен зертхана қызметкерлері үшін қолжетімді етеді. Мұндай сценарий әдістің дәлдігі мен қолжетімділігін арттырады. Осылайша, бұл зерттеудің мақсаты COVID-19-ға күдікті науқастарды жылдам, тиімді және жоғары дәлдікпен сынау үшін SARS-CoV-2 инфекциясының оң және теріс жағынды үлгілерінің масс-спектрометриялық үлгілері бойынша машиналық оқыту моделін әзірлеу болды.

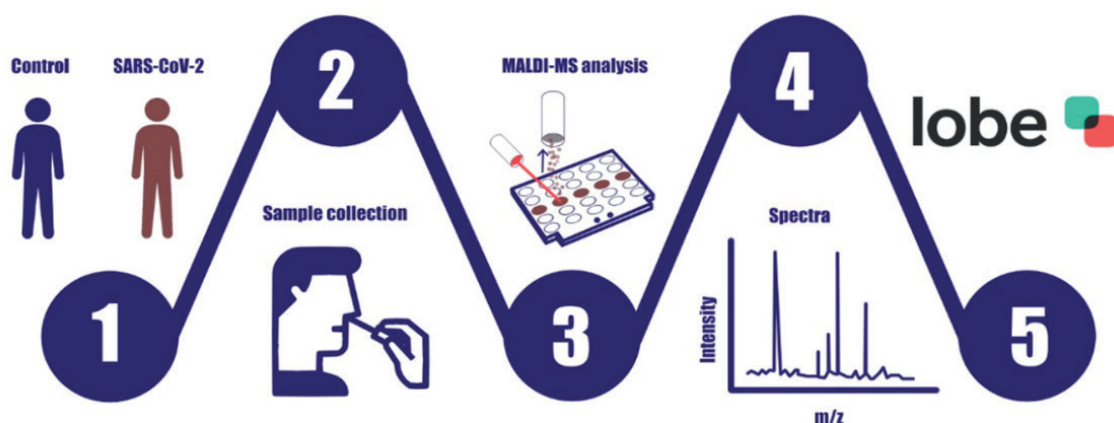
Материалдар мен тәсілдер. Проспективті зерттеу КЕАҚ «Қарағанды қаласының медициналық университеті», КМК «Облыстық клиникалық аурухана» инфекциялық орталығы, КЕАҚ «Қарағанды қаласының медициналық университеті» еңбек гигиенасы клиникасы, Қалалық алғашқы медициналық-санитарлық көмек орталығы негізінде жүргізілді. Жұмыс КЕАҚ «Қарағанды қаласының медициналық университеті» биоэтика комиссиясының 06.04.2020 жылғы № 45 қаулысымен бекітілген этикалық хаттамаларға сәйкес жүргізілді. Ешбір үлгінің идентификациясы пациенттің атымен немесе қатысушының сәйкестендіруіне әкелетін басқа ақпаратпен байланысты емес. Науқастың нақты демографиялық ақпараты (арудың ауырлығы, жасы және ілеспелі ауру) емделушілерден хабардар етілген келісімге қол қойылғаннан кейін ауруханада сұрау кезінде алынды. Барлық үлгілер MALDI пластинасында байқалды және Ортақ ресурстар зертханасының хаттамаларына сәйкес пайдаланылды. үлгілерді биоқорғау қорабына салғаннан кейін масс-спектрометрдің ластануын бол-

дырмау үшін MALDI пластиналар 20 минут бойы ультракүлгін шаммен сәулелендірілді. Жұмыс барысы 1-суретте көрсетілген.

2.1 Үлгі жинағы
 COVID-19 пациенттері мен сау қатысушылардан барлығы 304 (203 КТ-ПТР оң және 101 КТ-ПТР теріс, Ct мәндері жоқ немесе Ct 40-тан жоғары) мұрын шырышты қабатының клиникалық үлгілері алынды. (1-сурет). Коллекция 2021 жылдың қаңтар-қыркүйек айларында пандемия кезінде қалыптасты. Дүниежүзілік денсаулық сақтау ұйымының ПТР клиникалық хаттамаларына сәйкес, барлық COVID-19 пациенттерінің COVID-19 белгілерін тудыратын және КТ-ПТР талдаулары үшін мұрын-жұтқыншақ жағынды үлгісі алынған. Биоматериал тампонды таңдайға параллель мұрын жолдарына мұрын тесігі мен құлақтың сыртқы тесіктерінен бірдей қашықтықта енгізу арқылы жүзеге асырылды және секрецияны сіңіру үшін бірнеше секундқа қалдырылды. Содан кейін жағынды дереу 3 мл тасымалдау ортасы (RNA/DNA Shield, Zymo Research, Irvine, US) бар стерильді түтікке салынды. Мұрын жағындылары талдау алдында -80°C температурада DNA/RNA Shield тасымалдау ортасында сақталды.

2.2 Мұрын жағындысындағы SARS-CoV-2 ПТР растау
 Биоматериалда SARS-CoV-2 бар екенін растау үшін молекулалық-генетикалық талдау жасалды. Жалпы РНҚ мұрын жағынды үлгілерінен өндірушілердің хаттамасын (RealBest Sorbitus, Vector-Best, Новосибирск, Ресей) пайдаланып, содан кейін Real-Best RNA SARS-CoV-2 жинақтарын (Vector-Best, Новосибирск, Ресей) пайдалана отырып, нақты уақыттағы КТ-ПТР жүргізілді.

2.3 Дақтарды талдау
 Үлгілер бөлме температурасында ерітілді және ақырын шайқалды. Барлық үлгілер Milli-Q суымен 1:10 мкл сұйылтылды. Нысанды пластинаға 0,5 мкл ерітінді 5 қайталаумен жағылды. Бөлме температурасында мұқият кептіруден кейін үлгілер 1 мкл HCCA матрицасымен (құрамында ацетонитрилде сұйылтылған α-циано-4-гидроксидиннамин қышқылы, 2,5% трифторсірке

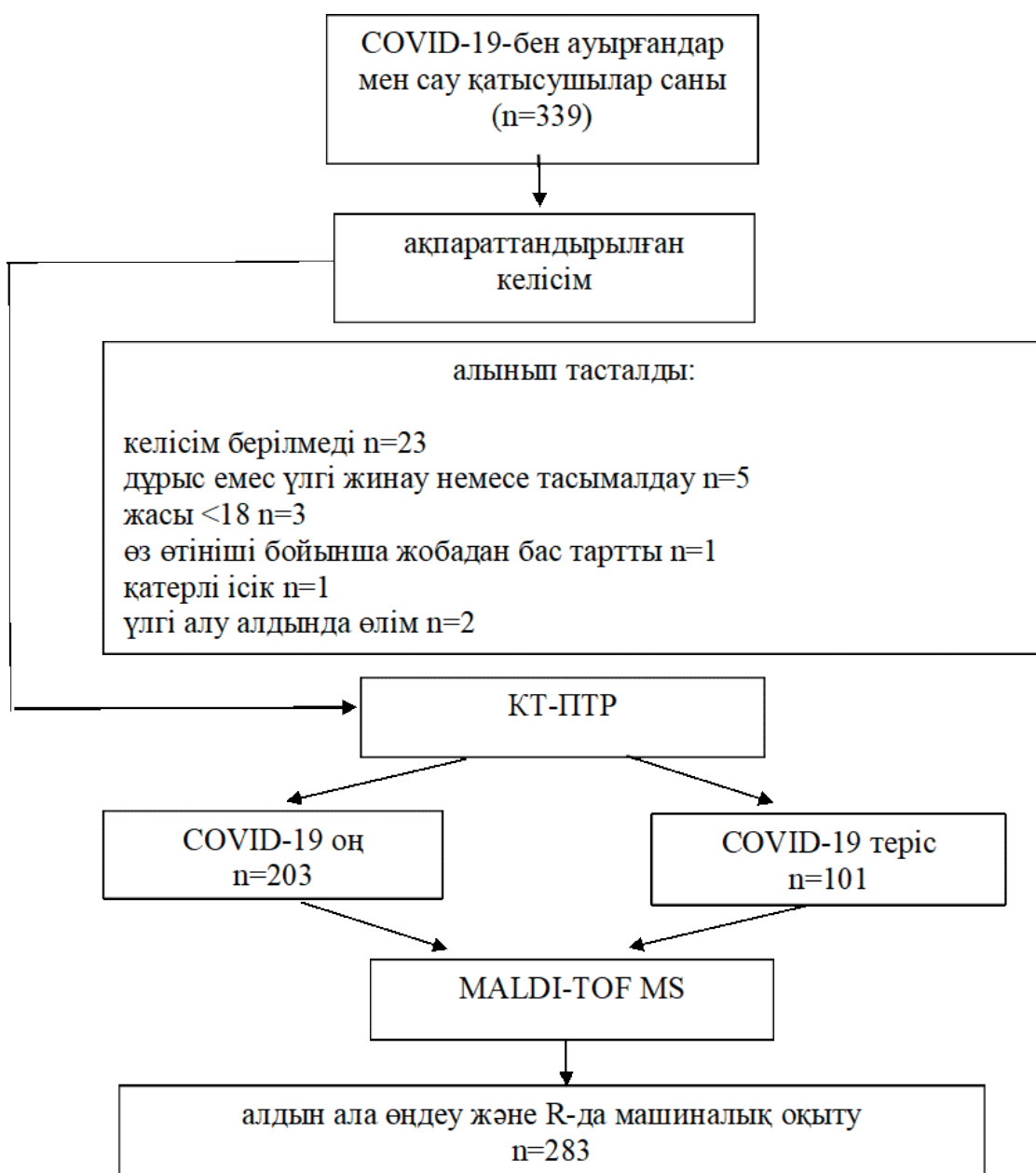


1-сурет - Жұмыс барысы

қышқылы және нуклеазасыз су бар ерітінді) қапталды. Ерітіндіннің үстіңгі қабаты мұқият кептіріліп, қосымша 1 мкл HCCA матрицасы қолданылды. Шұңқырлар кепкен кезде пластина MicroFlex LT 3.4 нұсқасының құралына жүктелді (Bruker Daltonics, Бремен, Германия). Спектрлер лазерді нысанды нүктеде кездейсоқ позицияларда түсіру арқылы 3 және 20 қДа аралығындағы 500 жалғыз спектрді (10 * 50 түсірілім) қосу арқылы жасалды [10].

2.4 Спектралды алдын ала өңдеу
Алынған MALDI спектрлері R пакеттерін пайдалана отырып, әдеттегі жұмыс үрдісіне сәйкес алдын ала өңделген.

MALDI спектрлерінің өңделмеген файлдары MALDIquantForeign бумасы [11] арқылы R ішіне импортталды. Негізгі сапаны бақылау үшін біз MALDIquant R бумасын пайдаланып график құру арқылы барлық спектрлерде деректер нүктелерінің бірдей саны бар-жоғын және бос еместігін тексердік. Квадрат түбірді түрлендіру графикалық визуализацияны жеңілдету және дисперсияның орташа мәннен ықтимал тәуелділігін жеңу үшін қолданылды. Сигналдарды өшіру немесе тегістеу MALDIrppa жүйесінде wavSmoothing функциясын қолданып жойылмаған дискретті толқындық түрлендіру (UDWT) арқылы жүргізіледі [12]. Базалық сызықты жою



2-сурет - Зерттеуге қатысушыларды тартуды көрсететін блок-схема

30 итерациямен Ryan мен басқалар [13] жұмысында сипатталған статистикалық сезімтал сызықты емес итеративті шыңды кесу (SNIP) алгоритмі арқылы орындалды. Қарқынды калибрлеу спектрлерді масштабтау үшін сенімді әдіс болып табылатын PQN нормалауында (Ықтималдық үлесті нормалау) орындалды [14]. Пиктерді анықтау үшін сигнал-шу қатынасы (SNR) 4 және массалық терезе өлшемінің мәндері 30 болатын медиана абсолютті ауытқу (MAD) функциясы орнатылды. Терістеу және біріктіру процесі 70% минималды жиілікпен және 0,5% ауытқумен профилдер бойынша ең жоғары экстракциядан кейін болды. Соңғы қарқындылық матрицасы MALDIquant бумасындағы intensityMatrix функциясы арқылы машиналық оқыту әдістерін пайдалана отырып, әрі қарай талдау үшін алынды.

2.5 Машиналық оқыту классификациясының моделі Алынған деректердің сапасын бақылау MALDIquant R пакетінде бағаланды. Машиналық оқытуды талдау үшін R бағдарламалау тіліне арналған Rtemis кітапханасы пайдаланылды [15]. Бірнеше алгоритмдер, соның ішінде жалпылама сызықтық модельдер (GLM) [16], жіктеу және регрессия ағашы (CART) [17], кездейсоқ орман (RF) [18] және экстремалды градиентті күшейту (XGBoost) [19], MALDI-TOF-MS арқылы алынған оң және теріс COVID-19 үлгілерін жіктеу үшін қарқындылық матрицасында сыналған. Модельді оқыту алдында бүкіл деректер жинағы кездейсоқ түрде 2 ішкі жиынға бөлінді – оқыту үшін 75% және модельді тестілеу үшін 25%. Нәтижелерде ауытқуды болдырмау үшін модель 20 стратифицирленген ішкі үлгілер бойынша оқытылды, өйткені болжау үшін модельдің орташа мәні пайдаланылды.

Дәлдік, сезімталдық, ерекшелік және F1-score алынған модельдердің жалған оң (FP), ақиқат оң (TP), жалған теріс (FN) және шынайы теріс (TN) нәтижелерін ескере отырып, белгілі теңдеулерді пайдалана отырып әзірленген модельдің өнімділік өлшемі ретінде есептелді (1,2,3-теңдеу).

$F1-score = 2 \cdot \text{дәлдік} \cdot \text{жауап} / (\text{дәлдік} + \text{жауап})$ (1-теңдеу)

Дәлдік = $TP / (TP + FP)$ (2-теңдеу)

Жауап = $TP / (TP + FN)$ (3-теңдеу)

2 және 3 өлшемді кеңістіктегі спектрлерді зерттеу және

салыстыру үшін rtemis пакеттерін пайдалану арқылы негізгі компоненттерді талдау (PCA) жүргізілді.

Нәтижелер. 2021 жылдың қаңтар-қыркүйек айларында SARS-CoV-2 пандемиясы кезінде Қарағанды облысында 304 сынама жиналды, олар MALDI-TOF MS-пен де, ПТП арқылы да тексерілді. (2-сурет). 203 сынама SARS-CoV-2 оң және 101 сынама теріс деп анықталды. 21 үлгі полимер ластануына байланысты жарамсыз болды және MALDI-TOF MS талдауынан шығарылды. КТ-ПТП циклінің шекті мәні (Ct) мәндерімен өлшенген орташа (SD) вирустық жүктеме 14,5–36,8 цикл диапазоны бар 25,7 (10,9) цикл болды.

Содан кейін талдау COVID-19 позитивті жағдайларды теріс жағдайлардан ажырату тапсырмасы үшін ең жақсы жұмыс істейтін модельді құру және анықтау үшін пайдаланылды. MALDI-MS көмегімен алынған алдын ала өңделген шыңдық матрицада оқытылған және сыналған төрт ML әдісінің (GLM, CART, RF және XGBoost) негізгі мәндері (сезімталдық, ерекшелік және дәлдік) нәтижелері 1-кестеде берілген. Оған қоса кестеде Fabiane M. Nachtigall және авторлардың [10] жұмысында жарияланған көрсеткіштер берілген.

10. Fabiane M. Nachtigall, Alfredo Pereira, Oleksandra S. Trofymchuk & Leonardo S. Santos: Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. Nature Biotechnology 38, бет 1168–1173 (Springer, 2020)

1-кестеде көрсетілгендей, RF, CART және XGBoost SARS-CoV-2 оң/теріс үлгілерін ажырату мүмкіндігінде ең жақсы нәтижелерді көрсетті, ал GLM моделі сыналған деректер жиынында 88,8% дәлдікке қол жеткізді, ал сезімталдық көрсеткіші бойынша тек 55,5% болды. Талқылау. COVID-19 пандемиясы денсаулық сақтау жүйесіне сын болды және жоғары өнімділік, жоғары тиімділік, спецификалық және сезімтал әдістерді дамыту үшін маңызды олқылық туғызды. [20]. Уақыт факторы да маңызды болды, алайда бір сағаттан аз уақыт ішінде нәтиже беретін жылдам шешімдер зертханалық аналогтарымен салыстырғанда төмен клиникалық ерекшелік пен сезімталдықты көрсетті. [21].

Біздің зерттеуімізде MALDI-TOF MS шыңының үлгілері COVID-позитивті және COVID-теріс жағынды үлгілері үшін анық ерекшеленеді. Зерттеуде көрсетілгендей,

Кесте 1 - Машиналық оқыту үлгілерінің сипаттамалары

	Сезімталдық	Ерекшелік	Дәлдік
GLM	0.555	1.000	0.888
Random Forest	1.000	1.000	1.000
CART	1.000	1.000	1.000
XGBoost	1.000	1.000	1.000
DT ¹⁰	0.455	0.917	0.966
Random Forest ¹⁰	0.867	0.923	0.866
KNN ¹⁰	0.947	0.475	0.966
NB ¹⁰	0.923	0.873	0.899
SVM-L ¹⁰	0.894	0.837	0.872
SVM-R ¹⁰	0.889	0.857	0.881

қарқындылықтың шыңы матрицасында жасалған негізгі компоненттік талдау COVID жағдайларын COVID емес жағдайлардан жақсы ажырата алды. Бір қызығы, массалық спектрдің шыңдары вирустық РНҚ емес, мұрын жағындысынан жиналған иондалатын құрамды білдіреді. CART, RF және XGBoost сияқты бақыланатын машиналық оқыту алгоритмдері бар жасырын үлгілерді пайдалана отырып, күрделі MALDI-TOF MS спектрлерін ажыратуда жоғары дәл нәтижелерге қол жеткізе алады. Тасымалдаушы ортада (мысалы, ViCUM, DeltaSwab вирусы және т.б.) жиналған COVID-19 мұрын жағындыларын MALDI-TOF MS-пен сынау үшін ML қолдану басқа топпен зерттелді және жақында жарияланды. Олардың нәтижесі 96,4% дәлдікпен, 100% сезімталдықпен және ерекшелігі 92,3% [22], сондай-ақ Cary Blair микробиологиялық тасымалдау ортасы үшін 93,9% ерекшелігі, PPA 95% және NPA 93% шықты [23]. Сондай-ақ, Meritxell Deulofeu және басқалар [22] PCA түрлендірілген деректерде SVM және XBoost алгоритмдерін қолданатын зерттеулерінде жоғары дәлдік, сезімталдық пен ерекшелікті көрсетті, барлық жағдайларда ең оқытылған модельдер үшін 90% жоғары мәндерге жетті. Біздің жұмысымызда да, басқа басылымдарда да [22-24] ұсынылған машиналық оқыту әдістеріне негізделген кейінгі талдаумен MALDI-TOF MS қолдану молекулалық-генетикалық әдістерге әлеуетті балама болып табылады. Бұл тәсілдің бірқатар артықшылықтары бар: біріншіден, MALDI-TOF MS талдауында оңай қол жетімді реагенттер (CHCA, этанол, ацетонитрил, су және трифторсірке қышқылы) пайдаланылады, ал ПТР РНҚ оқшаулау және ПТР-КТ өзі үшін жинақтар мен шығын материалдарын қажет етеді. MALDI-TOF MS-ң екінші сөзсіз артықшылығы - талдау уақыты. Шың үлгісін алады, шындық өңдеуді және оқытылған үлгі негізінде үлгі диагностикасын ескере отырып, толық процедура үшін 1 сағаттан аз уақыт қажет. Ақырында, MALDI-TOF MS бір құралда күніне жүздеген үлгілерді талдауға мүмкіндік беретін жоғары өнімділік әдісі болып табылады [23]. Барлық талдаулар, соның ішінде пиктерді өңдеу қадамдары, негізгі компоненттерді талдау, машиналық

оқыту алгоритмдерін қолданатын оқыту және тестілеу үлгілері және барлық графиктер зерттеушілер арасында ең көп қолданылатын бағдарламалау тілдерінің бірі болып табылатын R тілінде орындалды [25]. Бұл зерттеу ML және MALDI-TOF MS комбинациясы теріс сынақтан өткендермен салыстырғанда COVID-19 бар оң пациенттерді анықтай алатынын анықтауға арналған, бірақ COVID-19-ды басқа жедел респираторлық инфекциялардан (тұмау, аденовирустық инфекция және т.б.) ажыратуға емес. Сонымен қатар, осы зерттеудің шектеулері вирустық ақуыздар мен шырышты қабықтың пептидтерін бөлу үшін анықталған шыңдардың протеомдық профилін жасау қажеттілігін қамтиды. Полимердің ластануы 21 үлгіде анықталды, бұл иондануға жол бермеді және жарамсыз нәтижеге әкелді. Масс-спектрометриядағы полимердің ластану көзі тұзды ерітіндінің өзі, жинау түтігі немесе жағынды болды, өкінішке орай, бұл кең таралған мәселе [24]. Қорытынды. Осы жұмыста біз мұрын-жұтқыншақ жағындыларында SARS-CoV-2 анықтаудың экспресс әдісін ұсынамыз. Ұсынылған әдістің артықшылығы - биоматериал үлгісін алдын ала дайындаудың болмауы. Қымбат реагенттерді пайдаланудың қажеті жоқ және клиникалық диагностикалық зертханаларда жүйелі түрде қолданылатын жабдықтың болуы. Мұрын жағындысынан алынған материалда SARS-CoV-2 анықтаудың ұсынылған әдісі КТ-ПТР-ге қарағанда қол жетімді және аз уақытты қажет ететін сынақ болады. Дегенмен, бұл тәсіл ПТР-ға жақсы қосымша екенін және ол көптеген сайттар бойынша нәтижелерді кеңірек тексеруді қажет ететінін түсіну қажет. Сонымен қатар, жоғарыда аталған сипаттамаларға байланысты MALDI-MS және ML талдауын қолдану арқылы оң және теріс SARS-CoV-2 жағынды үлгілерін анықтау әдісі жоғары өткізу қабілетіне ие әрі оның құнының төмендігін ескере отырып, оны скрининг әдісі ретінде қабылдау мүмкін.

Қаржыландыру. Зерттеуді Қазақстан Республикасы Білім және ғылым министрлігі (AP09259123) қаржыландырды

ӘДЕБИЕТТЕР ТІЗІМІ

- 1 Kadyrova I., Yegorov S., Negmetzhanov B. et al. Seroprevalence of anti-SARS-CoV-2 IgG and IgA antibodies before the launch of COVID-19 vaccination in Kazakhstan. The preprint server medRxiv. September 10, 2021. doi: <https://doi.org/10.1101/2021.09.03.21262885>
- 2 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;4(100096):1-10. <https://doi.org/10.1016/j.lanep.2021.100096>
- 3 <https://who.maps.arcgis.com/apps/dashboards/a19d5d1f86ee4d99b013eed5f637232d>
- 4 Yüce M., Filiztekin E., Özkaya K.G. COVID-19 diagnosis -A review of current methods. Biosens Bioelectron. 2021;172(112752):1-15. DOI: 10.1016/j.bios.2020.112752
- 5 Taleghani N., Taghipour F. Diagnosis of COVID-19 for controlling the pandemic: A review of the state-of-the-art. Biosens Bioelectron. 2021;174(112830). DOI: 10.1016/j.bios.2020.112830 (дата обращения 20-10-2021)
- 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;49(1):21-29. DOI: 10.1016/j.ajic.2020.07.011
- 7 Dramé M., Tabue Teguo M., Proye E. et al. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? J Med Virol. 2020;92(11):2312-2313. DOI: 10.1002/jmv.25996
- 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;93(4):2281–2286. doi: 10.1002/jmv.26691
- 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;100(2_suppl):131S-138S. doi: 10.1177/0145561320953231
- 10 Fabiane M. Nachtigall, Alfredo Pereira, Oleksandra S. Trofymchuk, Leonardo S. Santos Detection of SARS-CoV-2 in nasal swabs using MALDI-MS.

- Nature biotechnology.2020;38:1168–1173. <https://doi.org/10.1038/s41587-020-0644-7>
- 11 Sebastian Gibb (2021). MALDIquantForeign: Import/Export Routines for 'MALDIquant'. R package version 0.13. <https://strimmerlab.github.io/software/malDIquant/> <https://github.com/sgibb/MALDIquantForeign/>
- 12 Coombes K.R., et al. Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform. *Proteomics*. 2005; 5:4107–4117.
- 13 Ryan, C., Clayton, E., Griffin, W., Sie, S., and Cousens, D. Snip, a statistics-sensitive background treatment for the quantitative analysis of pxe spectra in geoscience applications. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. 1988;34(3):396 – 402.
- 14 Dieterie, Frank, Alfred Ross, Götz Schlotterbeck, and Hans Senn. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry* 2006;78(13):4281–90. <https://doi.org/10.1021/ac051632c>
- 15 Gennatas, E. D. Towards Precision Psychiatry: Gray Matter Development and Cognition in Adolescence (Doctoral dissertation, University of Pennsylvania). 2017
- 16 McCullagh, P., Nelder, J. A. *Generalized Linear Models*. London: Chapman & Hall / CRC. 1989
- 17 Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. New York: Chapman and Hall; 1984
- 18 Breiman, L. Random Forests. *Machine Learning*. 2001;45:5–32. <https://doi.org/10.1023/A:1010933404324>
- 19 Chen, T., & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM. 2016:785–794. <https://doi.org/10.1145/2939672.2939785>
- 20 Vandenberg, O. et al. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* 2020;14:1–13.
- 21 Mogling, R. et al. Delayed laboratory response to COVID-19 caused by molecular diagnostic contamination. *Emerg. Infect. Dis.* 2020;26:1944–1946.
- 22 Deulofeu M., García-Cuesta E., Peña-Méndez E. M et al. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining and Artificial Intelligence. *Frontiers in Medicine*. 2021; 8:1-12 doi: 10.3389/fmed.2021.661358
- 23 Nam K. Tran et al. Novel application of automated machine learning with MALDI-TOF-MS for rapid high-throughput screening of COVID-19: a proof of concept. *Scientific Reports*. 2021;11(8219). <https://doi.org/10.1038/s41598-021-87463-w>
- 24 Veloso, A., Leiza, H. R. Cross-contamination from vial caps for SEC analysis determined by MALDI-TOF mass spectrometry. *J. Chromatogr. Sci.* 2017; 55:137–141.
- 25 Tippmann, Sylvia. *Programming tools: Adventures with R*. Nature. 2014;517 (7532): 109–110. doi:10.1038/517109a. PMID 25557714

REFERENCES

- 26 Kadyrova I., Yegorov S., Negmetzhanov B. et al. Seroprevalence of anti-SARS-CoV-2 IgG and IgA antibodies before the launch of COVID-19 vaccination in Kazakhstan. The preprint server medRxiv. September 10, 2021. doi: <https://doi.org/10.1101/2021.09.03.21262885>
- 27 Yegorov S., Goremykina M., Ivanova R. et al. Epidemiology, clinical characteristics, and virologic features of COVID-19 patients in Kazakhstan: A nationwide retrospective cohort study. *The Lancet Regional Health – Europe*. 2021;4(100096):1-10. <https://doi.org/10.1016/j.lanep.2021.100096>
- 28 <https://who.maps.arcgis.com/apps/dashboards/a19d5d1f86ee4d99b013eed5f637232d>
- 29 Yüce M., Filiztekin E., Özkaya K.G. COVID-19 diagnosis -A review of current methods. *Biosens Bioelectron.* 2021;172(112752):1-15. DOI: 10.1016/j.bios.2020.112752
- 30 Taleghani N., Taghipour F. Diagnosis of COVID-19 for controlling the pandemic: A review of the state-of-the-art. *Biosens Bioelectron.* 2021;174(112830). DOI: 10.1016/j.bios.2020.112830 (дата обращения 20-10-2021)
- 31 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;49(1):21-29. DOI: 10.1016/j.ajic.2020.07.011
- 32 Dramé M., Tabue Teguo M., Proye E. et al. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? *J Med Virol.* 2020;92(11):2312-2313. DOI: 10.1002/jmv.25996
- 33 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;93(4):2281–2286. doi: 10.1002/jmv.26691
- 34 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;100(2_suppl):131S-138S. doi: 10.1177/0145561320953231
- 35 Fabiane M. Nachtigall, Alfredo Pereira, Oleksandra S. Trofymchuk, Leonardo S. Santos Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. *Nature biotechnology*.2020;38:1168–1173. <https://doi.org/10.1038/s41587-020-0644-7>
- 36 Sebastian Gibb (2021). MALDIquantForeign: Import/Export Routines for 'MALDIquant'. R package version 0.13. <https://strimmerlab.github.io/software/malDIquant/> <https://github.com/sgibb/MALDIquantForeign/>
- 37 Coombes K.R., et al. Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform. *Proteomics*. 2005; 5:4107–4117.
- 38 Ryan, C., Clayton, E., Griffin, W., Sie, S., and Cousens, D. Snip, a statistics-sensitive background treatment for the quantitative analysis of pxe spectra in geoscience applications. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. 1988;34(3):396 – 402.
- 39 Dieterie, Frank, Alfred Ross, Götz Schlotterbeck, and Hans Senn. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry* 2006;78(13):4281–90. <https://doi.org/10.1021/ac051632c>
- 40 Gennatas, E. D. Towards Precision Psychiatry: Gray Matter Development and Cognition in Adolescence (Doctoral dissertation, University of Pennsylvania). 2017
- 41 McCullagh, P., Nelder, J. A. *Generalized Linear Models*. London: Chapman & Hall / CRC. 1989
- 42 Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. New York: Chapman and Hall; 1984
- 43 Breiman, L. Random Forests. *Machine Learning*. 2001;45:5–32. <https://doi.org/10.1023/A:1010933404324>
- 44 Chen, T., & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM. 2016:785–794. <https://doi.org/10.1145/2939672.2939785>
- 45 Vandenberg, O. et al. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* 2020;14:1–13.
- 46 Mogling, R. et al. Delayed laboratory response to COVID-19 caused by molecular diagnostic contamination. *Emerg. Infect. Dis.* 2020;26:1944–1946.
- 47 Deulofeu M., García-Cuesta E., Peña-Méndez E. M et al. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining and Artificial Intelligence. *Frontiers in Medicine*. 2021; 8:1-12 doi: 10.3389/fmed.2021.661358
- 48 Nam K. Tran et al. Novel application of automated machine learning with MALDI-TOF-MS for rapid high-throughput screening of COVID-19: a proof of concept. *Scientific Reports*. 2021;11(8219). <https://doi.org/10.1038/s41598-021-87463-w>
- 49 Veloso, A., Leiza, H. R. Cross-contamination from vial caps for SEC analysis determined by MALDI-TOF mass spectrometry. *J. Chromatogr. Sci.* 2017; 55:137–141.
- 50 Tippmann, Sylvia. *Programming tools: Adventures with R*. Nature. 2014;517 (7532): 109–110. doi:10.1038/517109a. PMID 25557714

Вклад авторов. Все авторы принимали равное участие при написании данной статьи.

Конфликт интересов – не заявлен.

Данный материал не был заявлен ранее, для публикации в других изданиях и не находится на рассмотрении другими издательствами. При проведении данной работы не было финансирования сторонними организациями и медицинскими представителями.

Финансирование – Исследование финансировалось Министерством образования и науки Республики Казахстан (AP09259123).

Авторлардың үлесі. Барлық авторлар осы мақаланы жазуға тең дәрежеде қатысты.

Мүдделер қақтығысы – мәлімделген жоқ.

Бұл материал басқа басылымдарда жариялау үшін бұрын мәлімделмеген және басқа басылымдардың қарауына ұсынылмаған. Осы жұмысты жүргізу кезінде сыртқы ұйымдар мен медициналық өкілдіктердің қаржыландыруы жасалған жоқ.

Қаржыландыру Зерттеуді Қазақстан Республикасы Білім және ғылым министрлігі (AP09259123) қаржыландырды..

Authors' Contributions. All authors participated equally in the writing of this article.

No conflicts of interest have been declared.

This material has not been previously submitted for publication in other publications and is not under consideration by other publishers.

There was no third-party funding or medical representation in the conduct of this work.

Funding - The research was funded by the Ministry of Education and Science of the Republic of Kazakhstan (AP09259123).

Сведения об авторах:

Султанбекова Айдана Аскарвна. 8D110100 «Медицина» мамандығының 3 жыл докторанты. ҚеАҚ «Қарағанды медицина университеті», қ. Қарағанды, Қазақстан Қарағанды, Гоголь 40 к-сі. A.Sultanbekova@qmu.kz. ORCID: 0000-0001-5835-3958

Кадырова Ирина Адильевна. PhD, асс. профессор., аға ғылыми қызметкер. ҚеАҚ «Қарағанды медицина университетінің» ғылыми орталығы, қ. Қарағанды, Қазақстан Қарағанды, Гоголь 40 к-сі. ikadyrova@qmu.kz. ORCID: 0000-0001-7173-3138

Бабенко Дмитрий Борисович. PhD, асс. профессор. ғылыми қызметкер, ҚеАҚ «Қарағанды медицина университетінің» ғылыми орталығы, қ. Қарағанды, Қазақстан Қарағанды, Гоголь 40 к-сі. dmitriy.b.babenko@gmail.com. ORCID:0000-0003-2280-2146