

Novel Application of Automated Machine Learning With Maldi-Tof-Ms For Rapid High-Throughput Screening of Covid-19: A Proof Of Concept

Nam Tran (✉ nktran@ucdavis.edu)

University of California, Davis

Taylor Howard

University of California, Davis

Ryan Walsh

Shimadzu Scientific Instruments

John Pepper

Allegiant Airlines

Julia Loegering

University of California, Davis

Brett Phinney

University of California, Davis

Michelle Salemi

University of California, Davis

Hooman Rashidi

University of California, Davis

Research Article

Keywords: Algorithm, artificial intelligence, infection, k-nearest neighbor, logistic regression, mass spectrometry, naïve Bayes, neural network, polymerase chain reaction, principal component analysis, random forest, SARS-CoV-2, support vector machine, and gradient boosting machine

Posted Date: January 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-141397/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on April 15th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-87463-w>.

Abstract

The 2019 novel coronavirus infectious disease (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has created an unsustainable need for molecular diagnostic testing. Molecular approaches such as reverse transcription (RT) polymerase chain reaction (PCR) offers highly sensitive and specific means to detect SARS-CoV-2 RNA, however, despite it being the accepted “gold standard”, molecular platforms often require a tradeoff between speed versus throughput. Matrix assisted laser desorption ionization (MALDI) – time of flight (TOF) – mass spectrometry (MS) has been proposed as a potential solution for COVID-19 testing and finding a balance between analytical performance, speed, and throughput, without relying on impacted supply chains. Combined with machine learning (ML), this MALDI-TOF-MS approach could overcome logistical barriers encountered by current testing paradigms. We evaluated the analytical performance of an ML-enhanced MALDI-TOF-MS method for screening COVID-19. Residual nasal swab samples from adult volunteers were used for testing and compared against RT-PCR. Two optimized ML models were identified, exhibiting accuracy of 98.3%, positive percent agreement (PPA) of 100%, negative percent agreement (NPA) of 96%, and accuracy of 96.6%, PPA of 98.5%, and NPA of 94% respectively. Machine learning enhanced MALDI-TOF-MS for COVID-19 testing exhibited performance comparable to existing commercial SARS-CoV-2 tests.

Background

The 2019 novel coronavirus infectious disease (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has created a significant demand for testing.^{1,2} In the United States, delays in establishing high throughput testing capacity early in the pandemic and subsequent supply shortages limited the nation’s ability to control the spread of COVID-19.^{3,4}

SARS-CoV-2 diagnostic testing relies on molecular or antigen platforms (**Table 1**).⁵ Molecular methods such as reverse transcription (RT) polymerase chain reaction (PCR) offers highly sensitive and specific means to detect SARS-CoV-2 RNA. Unfortunately, despite molecular technologies serving as the accepted “gold standard” for SARS-CoV-2 diagnostics, these techniques are often dependent on constrained supplies chains (*e.g.*, molecular grade reagents and consumables, processing plates/pipettes, extraction kits, etc).⁶ Moreover, molecular platforms typically tradeoff between speed versus throughput.^{6,7} Rapid (<20 minute) point-of-care molecular platforms, for example, are often being limited to testing one sample at a time, while high throughput laboratory-base instruments tests in batches every few hours – realistically producing results within 24 to 48 hours. Antigen testing offers a unique alternative to molecular diagnostics by detecting SARS-CoV-2 proteins rather than RNA.^{8,9} However, current data suggests antigen methods exhibit lower sensitivity and specificity especially when testing asymptomatic populations. *To this end, there is a critical need for a highly sensitive and specific high throughput method that is cost-effective, and rapid, for screening COVID-19.*

Matrix assisted laser desorption ionization (MALDI) – time of flight (TOF) – mass spectrometry (MS) has been proposed as a potential solution for COVID-19 testing.¹⁰ Briefly, MALDI-TOF-MS has been employed

in clinical microbiology over the last decade to accelerate identification of bacterial and fungal species from positive culture samples.^{11,12} The technique produces mass spectra that represents ionizable protein components found in the sample that may correspond with a pathogen and/or disease state. For any given sample, there could be numerous mass spectra peaks perhaps hundreds or thousands – making analysis of complex samples or diseases challenging. Current advances in machine learning approaches complements and enhances performance of these MS-based technologies to analyze these complex samples. In this paper, we describe a proof-of-concept novel automated machine learning (ML) enhanced MALDI-TOF-MS approach for testing nasal swabs from patients with suspected COVID-19. An automated ML platform was used for data analysis.

Methods

We conducted a bench analytic study to evaluate the performance of the MALDI-TOF-MS COVID-19 testing method using SARS-CoV-2 RNA PCR positive and negative samples. The goal of this study was to determine the accuracy along with the positive percent agreement (PPA) and negative percent agreement (NPA) of the MALDI-TOF-MS method to the PCR method that was used as a comparative approach.

Study Population / Samples:

The study was approved by the UC Davis Institutional Review Board. Informed consent was obtained for 226 nasal swab samples (anterior nares) preserved in saline transport media were obtained from the UC Davis Clinical Laboratory Biorepository. Patients included asymptomatic and symptomatic populations including those meeting COVID-19 testing criteria (*i.e.*, patients who presented with or without symptoms at the time of collection) as well as asymptomatic volunteers as part of workplace screening. Saline viral transport media was used due to its widespread availability and compatibility with MALDI-TOF-MS techniques. Commercially available swabs (Copan, Murrieta, CA) were used for collection. All samples were stored at -70°C prior to testing.

MALDI-TOF-MS Method:

The study testing workflow is illustrated by Figure 1. Mass spectrometry testing was performed on a Shimadzu 8020 (Shimadzu Scientific Instruments, Columbia, MD) MALDI-TOF-MS analyzer. Sample processing was conducted under a Class II Biosafety Cabinet. Nasal swabs were first plated directly onto the MALDI-TOF-MS target plate, followed by addition of α -cyano-4-hydroxycinnamic acid (CHCA), ethanol, acetonitrile, and water solution with 3% trifluoroacetic acid (TFA). Plated samples were then inactivated by ultraviolet (UV) irradiation for 10 minutes. Thereafter, the target plate was transferred to the MALDI-TOF-MS analyzer for testing. MALDI-TOF-MS settings included a mass range of 2,000 to 20,000 Daltons. Ten laser shots were fired for each profile at a frequency of 100 Hz using a dithering pattern (total of 1,000 shots per well) and Gaussian smoothing method. Post-acquisition baseline subtraction and smoothing was performed using MALDI Solutions software (Shimadzu Scientific Instruments, Columbia,

MD) (parameters: Baseline Filter Width = 250, Smoothing = Gaussian, and Smoothing Width = 50, Peak Width = 5). Peak picking was also performed by MALDIQuant software (Shimadzu Scientific Instruments, Columbia, MD). Threshold Apex algorithm was used for peak selection where the peak mass is assigned by selecting the highest point on the peak. Based on this protocol, the MALDI-TOF-MS would complete 48 runs (samples and quality control) every 20 minutes. Mass spectra were then standardized prior to analysis by ML with peak selection / alignment performed using MALDIQuant software.

Comparative Method:

Residual saline transport media was tested by RT-PCR using Food and Drug Administration (FDA) emergency use authorized (EUA) assays (Table 1).⁵ These EU assays included the cobas 6800 (Roche Molecular Systems, Pleasanton, CA), and digital droplet RT-PCR (Bio-Rad, Hercules, CA). Briefly, the cobas 6800 SARS-CoV-2 EUA assay targets open reading frame 1ab (ORF1ab) and envelope protein (E) gene regions, while the digital droplet RT-PCR method targeted two regions within the nucleocapsid (N) protein region. Both assays report sensitivity and specificity of >99% based on their FDA EUA documentation. The use of two different assays was due to supply constraints during the pandemic.

Machine Learning:

The Machine Learning (ML) aspects of this study were carried out through the Machine Intelligence Learning Optimizer (MILO) automated ML platform (MILO ML, LLC, Sacramento, CA) which has been published in several recent papers.¹³⁻¹⁶ Briefly, MILO includes an automated data processor, a data feature selector (ANOVA F select percentile feature selector and RF Feature Importances Selector) and feature set transformer (*e.g.*, principal component analysis), followed by its custom supervised ML model builder using its custom hyperparameter search tools (*i.e.*, its custom grid search along with its random search tools) to help find the optimal hyperparameter combinations within the variety of its embedded supervised algorithms/methods (*i.e.*, deep neural network [DNN], logistic regression [LR], naïve Bayes [NB], k-nearest neighbor [*k*-NN], support vector machine [SVM], random forest [RF], and XGBoost gradient boosting machine GBM]). Ultimately, MILO employs a combination of unsupervised and supervised ML platforms from a large set of algorithms, scalars, scorers and feature selectors/ transformers to create thousands of unique ML pipelines (Figure 2) that generates over a hundred thousand models that are then statistically assessed to ultimately identify the best performing model for one's given task.

For this study, we imported the trial data into MILO using COVID-19 status as the outcome measure for analysis. The aforementioned functions were then performed automatically by MILO. Information is assessed to ensure model training and the initial validation step is based on a balanced dataset. Initially in the build phase of MILO, the first balanced Dataset A is split into training and validation test sets in an 80-20 split with a 10k-fold cross validation step, respectively. Since many algorithms benefit from scaling, in addition to using the unscaled data, the training dataset also underwent two possible scaling

transformations (*i.e.*, standard scaler and minmax scaler). To evaluate the effect of different features within the datasets on model performance, a combination of various statistically significant feature subsets (*i.e.*, various MS peaks) or transformed feature sets were also selected to build new datasets with less features or transformed feature sets to feed into the various aforementioned supervised algorithms. The features selected in this step are derived from several well-established unsupervised ML/statistical techniques including ANOVA F-statistic value select percentile, RF Feature Importances or transformed using its principal component analysis approach.⁹ A large number of supervised machine learning models are then built through this approach from these datasets with optimal parameters through MILO's various supervised algorithms (*i.e.*, DNN, SVM, NB, LR, *k*-NN, RF, and GBM), scalers, hyper-parameters, and feature sets. Notably, the final validation of each model within MILO is not based on the 20% test set mentioned earlier that generated from the initial training dataset (*i.e.*, Dataset A) but rather each ML model's true performance is based on its predictive capability on the independent secondary dataset (Dataset B). Ultimately, for final model validation, MILO's thousands of generated models are then individually passed onto this next phase of the MILO engine generalization assessment phase (Figure 2). This secondary testing approach markedly reduces the possibility of overfitted ML models since the model's final performance measures are based on an independent secondary dataset only (Dataset B) as noted above. The final machine learning model performance data results are then tabulated by MILO's interface and reported as clinical sensitivity, specificity, accuracy, negative predictive value (NPV), positive predictive value (PPV), F1 score, receiver operator characteristic (ROC) curves, and brier scores with reliability curves.

Statistical Analysis:

Statistical analysis was performed using JMP Software (SAS Institute, Cary, NC). Area under the ROC curve analysis was also performed, as well as calculating PPA and NPA which served as surrogates for sensitivity and specificity. The use of PPA and NPA is recommended by the FDA due to not having a proven "gold standard" for SARS-CoV-2 detection at this time.^{5,17} An independent Principal component analysis (PCA) within scikit learn was also performed on the greater than 600 MS peaks evaluated here and its PC1, 2 and 3 components (results not shown) highlighted many of the shared peaks noted within the MILO feature selector approach (*i.e.*, RF Importances features [25%]) that found one of the best performing ML models for this study.

Results

Study Population:

A total of 226 samples were collected with 199 tested by both MALDI-TOF-MS and RT-PCR. Twenty-seven were invalid due to polymer contamination of the sample. For the remaining 199 samples tested, 107 samples were COVID-19 positive with 92 determined to be negative (Figure 3). Mean (SD) viral load as measured by RT-PCR cycle threshold (Ct) values was 25.7 (10.9) cycles with a range of 14.5 to 36.8

cycles. Cycle threshold values were similar between Datasets A and B. Examples of MALDI spectra for COVID-19 PCR positive versus PCR negative patient samples area shown in Figure 4A and 4B.

Data Analysis:

Prior to ML analysis, an independent principal component analysis was performed (Figure 4B and 4C). Thereafter, ML analysis was employed to build and identify the best performing model for the task of distinguishing the COVID positive cases from the negative cases. MILO's automated ML engine was initially trained on a selected subset of the aforementioned data with 82 cases (42 COVID-19 negative and 40 COVID-19 positive) known as Dataset A used for generating the large number of the ML models with initial validation followed by testing each of these models on a secondary generalization Dataset B comprising the remaining 117 cases (50 negative cases and 67 positive cases). MILO produced a total of 379,269 models and identified two models with high performance characteristics (Figure 5). The first is a DNN model with 75% of the total features/MS peaks (487 peaks [range: 1,993.91 – 19,590.89 m/z]) that exhibited an accuracy of 98.3% (95% CI: 94.0 – 99.8%), PPA of 100% (95% CI: 94.6 – 100.0%), NPA of 96% (95% CI: 86.3 – 99.5%), with an area under the ROC of 99.9 (95% CI: 65.6 – 100.0). The second model is a GBM model with 25% of the total features/MS peaks (166 peaks [range: 2,002.72 – 19,590.89 m/z]) that exhibited an accuracy of 96.6% (95% CI: 91.5 – 99.1%), PPA of 98.5% (95% CI: 92.0– 100.0%), NPA of 94% (95% CI: 83.5–98.8%), with an area under the ROC of 99.0 (95% CI: 86.7– 100).

Discussion

The COVID-19 pandemic has created a critical shortage of high performance, high throughput, and rapid solutions for detecting SARS-CoV-2 infection.^{1,2} Often, many of these attributes are mutually exclusive, with platforms producing results in < 1 hour having lower throughput and/or exhibiting lower clinical sensitivity and specificity compared to their laboratory counterparts.^{3,4} Although, as of this paper, supply chains have improved to support testing of symptomatic individuals in a hospital setting, resources remain limited to facilitate and enable widespread rapid screening of asymptomatic individuals necessary for reopening businesses, schools, and other non-hospital settings. Novel molecular solutions have been devised using automated RT-PCR and sequencing platforms. Unfortunately, these mass testing platforms trade high throughput with speed – exhibiting real-world TATs of 24 to 48 hours or even longer.

Our MALDI-TOF-MS approach attempts to address both throughput and speed limitations exhibited by molecular platforms, while maintaining high positive and negative percent agreement. Uniquely, MALDI-TOF-MS does not detect viral RNA and the mass spectra peaks visualized in this technique represents ionizable contents collected from a nasal swab. As illustrated by our study, MALDI-TOF-MS spectra is highly complex, but patterns exist that is differentiable by PCA, but more importantly, by ML techniques that can then generate models than can distinguish COVID from non-COVID cases with a great degree of confidence, matching or even outperforming many of the current screening test performances including antigen-based methods. The use of ML for COVID-19 MALDI-TOF-MS testing of nasal swabs collected in

microbiological transport media (*i.e.*, Cary Blair media) has been studied by another group and recently published ¹⁰ with a reported accuracy of 93.9%, and PPA of 95% and NPA of 93%.

Our study differs from Nachigall *et al.*¹⁰ in that we directly tested swabs rather than the transport media itself. Additionally, we used readily available saline as a preservative rather than Cary Blair media. Moreover, we accelerated development of ML models with improved percent agreement and eliminated programmer bias by utilizing a clinically validated auto-ML platform to identify optimized models. As discussed in this study, as well as others, the manual programming of ML models is both laborious and prone to user bias. These data scientists would need to assess the performance of every feature combination, scaler, and other parameters across all types of ML techniques. Since it is not possible to accomplish this in a reasonable amount of time – especially during the pandemic, data scientists must base their ML development on their experience. The use of MILO auto-ML enables stakeholders to evaluate the performance of every feature combination, scaler, and other parameters across a very large number of ML techniques in about 24 hours. In the case of this study, MILO identified two very promising ML models (DNN and GBM), both offering enhanced performance compared to the SVM based model proposed by Nachigall *et al.*¹⁰ Notably, MILO's best SVM (accuracy of 96.6%, and PPA of 98.5% and NPA of 94%), as a comparison, also outperformed the SVM model proposed by Nachigall *et al.* (accuracy of 93.9%, and PPA of 95% and NPA of 93%) which further supports the need of use of such auto-ML platforms within this arena.

Frequent COVID-19 testing is key to reopening schools and businesses until herd immunity is achieved.¹⁸ However, this is not presently sustainable with molecular techniques and rapid antigen tests are still limited by reagent availability, and more importantly, false negative and false positive rates.^{8,9} MALDI-TOF-MS combined with ML offers several unique advantages over SARS-CoV-2 molecular and antigen testing. Firstly, the MALDI-TOF-MS technique is rapid, with an analysis time of 20 minutes. Total turnaround time would be < 1 hour and could be accelerated if multiple instruments are available to support random-access testing. Secondly, MALDI-TOF-MS can provide high throughput, with up to 46 samples per run with plus two levels of controls, it has the potential to perform up to 1,104 analyses per day per instrument – limited by instrument down time (*e.g.*, preventative maintenance, repairs, etc) and incoming test volume. Lastly, this technique is less dependent on complex supply chains, using readily available bulk chemicals such as acetonitrile, CHCA, ethanol, water, and TFA, whereas molecular assays require a long list of reagents including RNA extraction kits, master mixes, and molecular grade processing plates and pipettes which remain in short supply.⁶ Therefore, our approach provides an opportunity to be exploited for re-opening schools and businesses, where testing can be performed with both speed and scale. In order to operationalize this approach, patient registration, specimen collection, testing by MALDI-TOF-MS, and analysis by ML must work as a system rather than be deployed piecemeal. **Figure 6** provides a conceptual model for ML-enhanced MALDI-TOF-MS integrated with secure mobile software.

Limitations of this study include the use of frozen clinical specimens as a proof of concept. Proteomic profiling of identified peaks would also be needed to characterize the presence of viral proteins and host response factors. This study was intended to determine if ML-enhanced MALDI-TOF-MS could differentiate between PCR positive COVID-19 patients versus those who tested negative. The study did not evaluate the detection of other coronavirus or influenza like illnesses in the community. Notably, local “shelter in place” have greatly reduced influenza prevalence in the community.¹⁹ Polymer contamination of samples prevented ionization of 27 samples – resulting in an invalid result. Sources of polymer include, but are not limited to the saline itself, specimen collection tube, and the swab. Polymer contamination is unfortunately a common challenge in mass spectrometry.²⁰

Conclusions

Machine learning – enhanced screening of COVID-19 in symptomatic and asymptomatic patients by MALDI-TOF-MS exhibits acceptable positive and negative percent agreement for screening applications. This approach may have great value for testing at satellite laboratories to rapidly screen large numbers of individuals requiring access to businesses, schools and other large facilities. Larger multicenter studies are needed to determine the feasibility of large-scale MALDI-TOF-MS-based COVID-19 detection for workplace screening and further refine ML models that encompass a range of non-COVID-19 respiratory infectious diseases.

Declarations

Acknowledgements

We thank University of California, Davis Health and Mr. Scott Bainbridge for his support of this study. The project was developed jointly between UC Davis, SpectraPass, and Shimadzu Scientific Instruments.

Competing Interests

Dr. Tran is a co-inventor the MILO software and co-owner of MILO-ML, LLC. He is also a consultant for Roche Diagnostics and Roche Molecular Systems. Dr. Tran is the Principal Investigator of this SpectraPass, LLC funded study. Dr. Howard is a co-investigator for this SpectraPass, LLC funded study. Dr. Walsh is employed by Shimadzu Scientific Instrumentations who manufacturers the MALDI-TOF-MS instrument used in the study.

Mr. Pepper is employed by Allegiant Airlines which collaborators with SpectraPass, LLC.

Ms. Loegering is a co-investigator for this SpectraPass, LLC funded study. Dr. Phinney is a co-investigator for this SpectraPass, LLC funded study. Ms. Salemi is a co-investigator for this SpectraPass, LLC funded study. Dr. Rashidi is a co-inventor of MILO, and co-owner of MILO-ML, LLC. He is a co-investigator for this SpectraPass, LLC funded study.

AUTHOR CONTRIBUTIONS

Nam Tran (N.T.): N.T. serves as study PI for the study and wrote the manuscript in its entirety. Performed basic statistical analysis of the study data as well as supporting the ML algorithm development with H.R. Significantly contributed to the review and editing of the manuscript with co-authors. Generated Figures and Tables for the manuscript.

Taylor Howard (T.H.): T.H is co-investigator for the study who significantly provided input on the operation of the MALDI-TOF-MS platform, performed testing described in the study, and reviewed/edited the paper.

Ryan Walsh (R.W.): R.W. is co-investigator for the study and provided MALDI-TOF-MS expertise in the study and provided technical support for the testing platform. He reviewed and edited the paper.

John Pepper (J.P.): J.P. is co-investigator for the study and for Allegiant Airlines which is a study collaborator. He helped performed the initial analysis of the study using PCA and early ML proof of concept work. Significantly contributed to the development of this paper including review and editing, as well as art resources to enhance figures.

Julia Loegering (J.L.): J.L. is co-investigator for the study and helped with testing of MALDI-TOF-MS samples and specimen biobanking. Significantly contributed to writing and reviewing the manuscript.

Brett Phinney (B.P.): B.P. is co-investigator on the study and provided proteomics expertise to the project as well as the manuscript. Significantly contributed to the writing, review, and editing of the manuscript.

Michelle R. Salemi (M.R.S.): M.R.S. mass spectrometry and proteomics expertise to the study. Significantly contributed to the writing, review, and editing of the manuscript.

Hooman Rashidi (H.R.): H.R. is co-investigator and developed and validated the ML algorithms used in the manuscript. Significantly contributed to the writing, review, and editing of the manuscript. All authors reviewed the manuscript.

References

1. Medical Laboratory Observer (MLO) Article: <https://www.mlo-online.com/disease/infectious-disease/article/21159031/data-collection-tool-shows-shortages-of-covid19-and-other-testing-supplies>, Accessed on December 20, 2020.
2. Vandenberg O, Martiny D, Rochas O, et al. Considerations for diagnostic COVID-19 tests. *Nat Rev Microbiol* 2020;14:1-13.
3. Washington Post Article: https://www.washingtonpost.com/investigations/contamination-at-cdc-lab-delayed-rollout-of-coronavirustests/2020/04/18/fd7d3824-7139-11ea-aa80-c2470c6b2034_story.html, Accessed on December 20, 2020.

4. Mogling R, Meijer A, Berginc N, et al. Delayed laboratory response to COVID-19 caused by molecular diagnostic contamination. *Emerg Infect Dis* 2020;26:1944-1946.
5. United States Food and Drug Administration Emergency Use Authorization website: <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/vitro-diagnostics-euas>. Accessed on December 27, 2020.
6. Tang YW, Schmidtz JE, Persing DH, et al. Laboratory diagnosis of COVID-19: current issues and challenges. *J Clin Microbiol* 2020;58:e00512-20.
7. Hanson G, Marino J, Wang ZX, et al. Clinical performance of the point-of-care cobas Liat for detection of SARS-CoV-2 in 20 minutes: A multicenter study. *J Clin Microbiol* 2020 [in press].
8. Fitzpatrick MC, Pandey, Wells CR, et al. Buyer beware: inflated claims of sensitivity for rapid COVID-19 tests. *Lancet* 2020 [in press].
9. Scohy A, Anantharaajah A, Bodeus M, et al. Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. *J Clin Virol* 2020;129:104455.
10. Nachtigall FM, Pereira A, Trofymchuk OS, et al. Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. *Nat Biotech* 2020;38:1168-1173.
11. Patel R. MALDI-TOF-MS for the diagnosis of infectious diseases. *Clin Chem* 2015;61:100-111.
12. Croxatto A, Prod'hom G, Greub G, et al. Application of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* 2012;36:380-407.
13. Tran NK, Albahra S, Pham TN, et al. Novel application of an automated-machine learning development tool for predicting burn sepsis: A proof of concept. *Sci Rep* 2020;10:12354.
14. Rashidi HH, Makley A, Palmieri TL, et al. Enhancing military burn- and trauma-related acute kidney injury prediction through an automated machine learning platform and point-of-care testing. *Arch Pathol Lab Med* 2020 [in press].
15. Rashidi HR, Tran NK, Vali Betts E, et al. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Acad Pathol* 2019;6:2374289519873088.
16. Jen KY, Albahra S, Yen F, et al. Automated en masse machine learning model generation shows comparable performance as classic regression models for predicting delayed graft function in renal allografts. *Transplantation* 2020 [accepted].
17. United States Food and Drug Administration Guidance on Statistical Methods for Evaluating In Vitro Diagnostic Tests: <https://www.fda.gov/media/71147/download>. Accessed on December 27, 2020.
18. Mina M, Larremore DB. COVID-19 test sensitivity – a strategy for containment. *N Engl J Med* 2020;383:e120.
19. Olson SJ, Azziz-Baumgarner E, Budd AP, et al. Decreased influenza activity during the COVID-19 pandemic – United States, Australia, Chile, and South Africa, 2020. *MMWR* 2020;69:1305-1309.
20. Veloso A, Leiza HR. Cross-contamination from vial caps for SEC analysis determined by MALDI-TOF mass spectrometry. *J Chromatography Sci* 2017;55:137-141.

Table

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

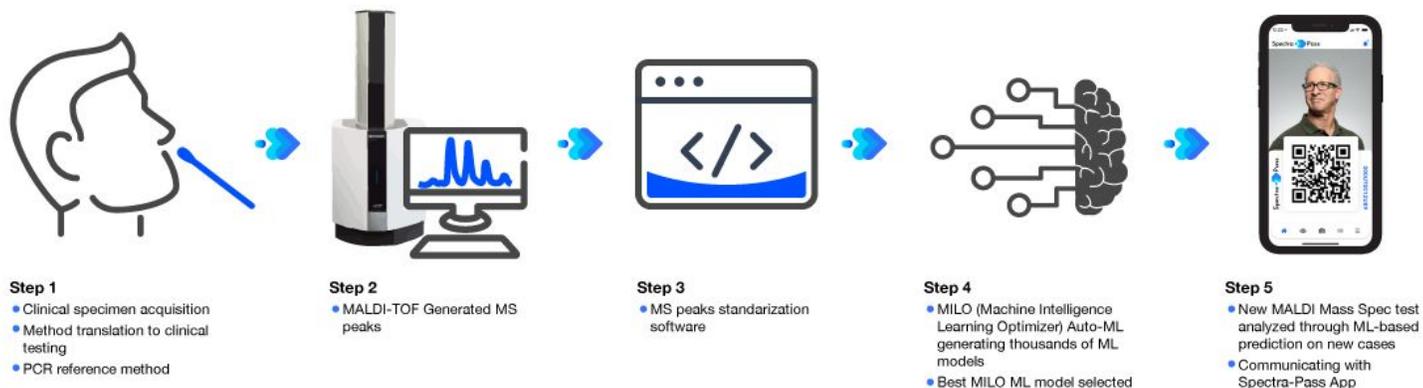


Figure 1

Conceptual Drawing of Study Workflow. The study workflow consisted of patients providing a nasal swab specimen preserved in saline transport media. Media was tested by RT-PCR (Step 1) and swabs plated onto the MALDI-TOF-MS platform (Step 2). Mass spectra were standardized (Step 3) and then analyzed using machine learning via the Auto-ML MILO platform (Step 4). COVID-19 status is then exported to a smart device app (Step 5).

Figure 2

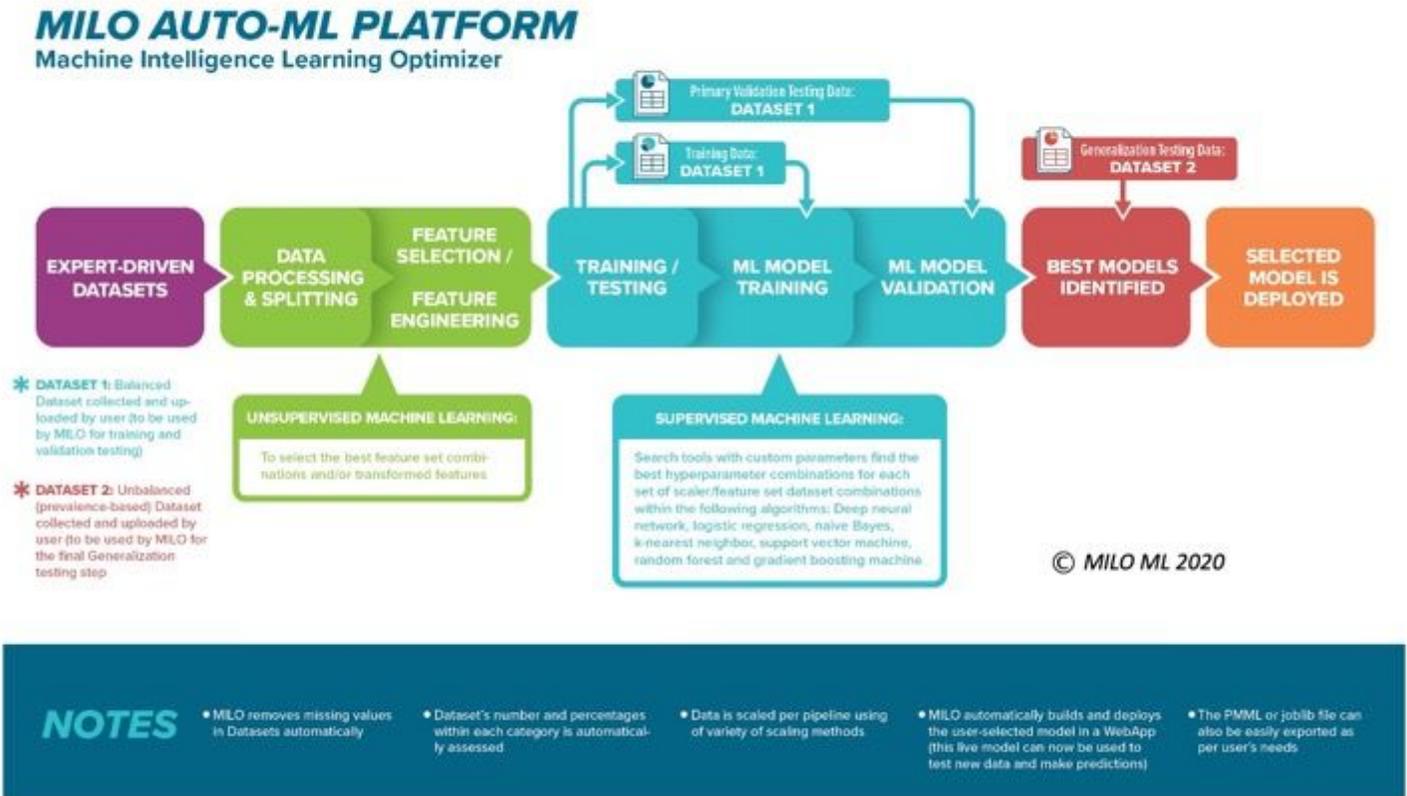


Figure 2

Machine Intelligence Learning Optimizer Figure 1. The MILO auto-machine learning (ML) infrastructure consists of beginning with two datasets: (a) balanced data (Dataset A) set used for training and initial validation, and (b) an unbalanced dataset (Dataset B) for generalization/ secondary testing. MILO initially removes the missing values followed by providing several scaling options for the given dataset which is then assessed by the software. Unsupervised ML is then used for feature selection and feature engineering. The generated models are then trained on a subset (80%) of dataset A (depicted as Dataset 1 in the image above) and then initially tested with the remaining subset (20%) of Data Set A during its supervised ML stage. Following this training/initial validation stage, each of the ML models generated in this stage are then secondarily tested on Dataset B (depicted as Dataset 2 in the image above) for generalization testing. Selected models can then be deployed thereafter as predictive model markup language (PMML) or joblib files. For this study, we imported the study data into MILO using COVID-19 status as the outcome measure for analysis. The following functions are then performed automatically by MILO.

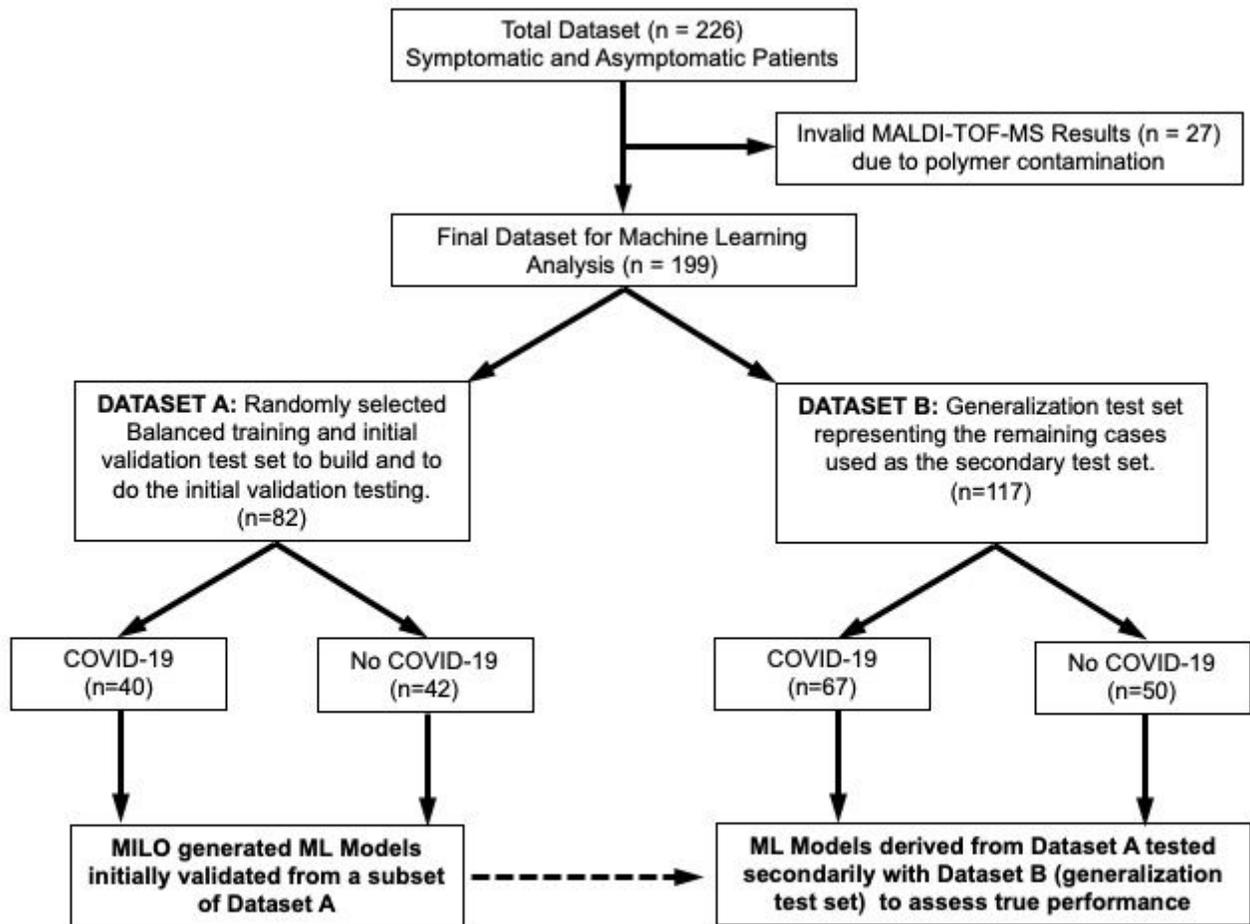


Figure 3

Study Datasets. A total of 226 asymptomatic and symptomatic patients were enrolled. Twenty-seven samples were invalid due to polymer contamination, preventing MALDI-TOF-MS analysis. The remaining 199 were successfully tested by MALDI-TOF-MS and produced spectra. These data were divided into Datasets A and B, with Dataset A serving as the training/ initial validation dataset. Optimized models produced from Dataset A were then secondarily tested with Dataset B for generalization to assess their true performance.

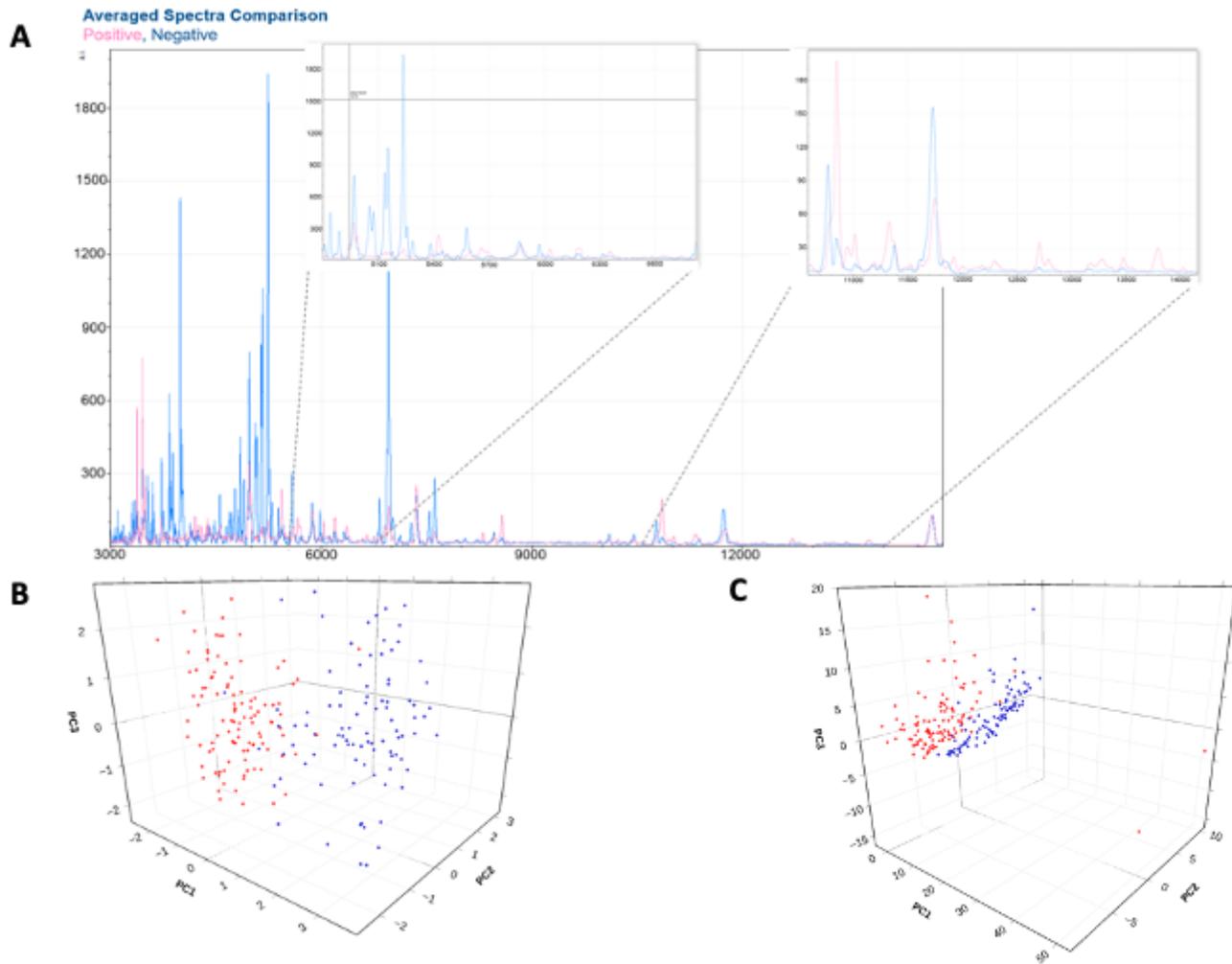


Figure 4

Example MALDI-TOF-MS Spectra and PCA Analysis of COVID-19 Positive vs. Negative Samples. Panels A illustrates the MALDI-TOF-MS spectra for patients that were SARS-CoV-2 RNA PCR positive (pink) versus PCR negative (blue). Zoomed in regions of interested are also shown. X-axis is mass to charge (m/z) ratio and Y-axis is relative abundance. Panels B and C show unscaled and scaled PCA, respectively, for the 199 samples (red = positive, blue = negative) tested by the MALDI-TOF-MS method.

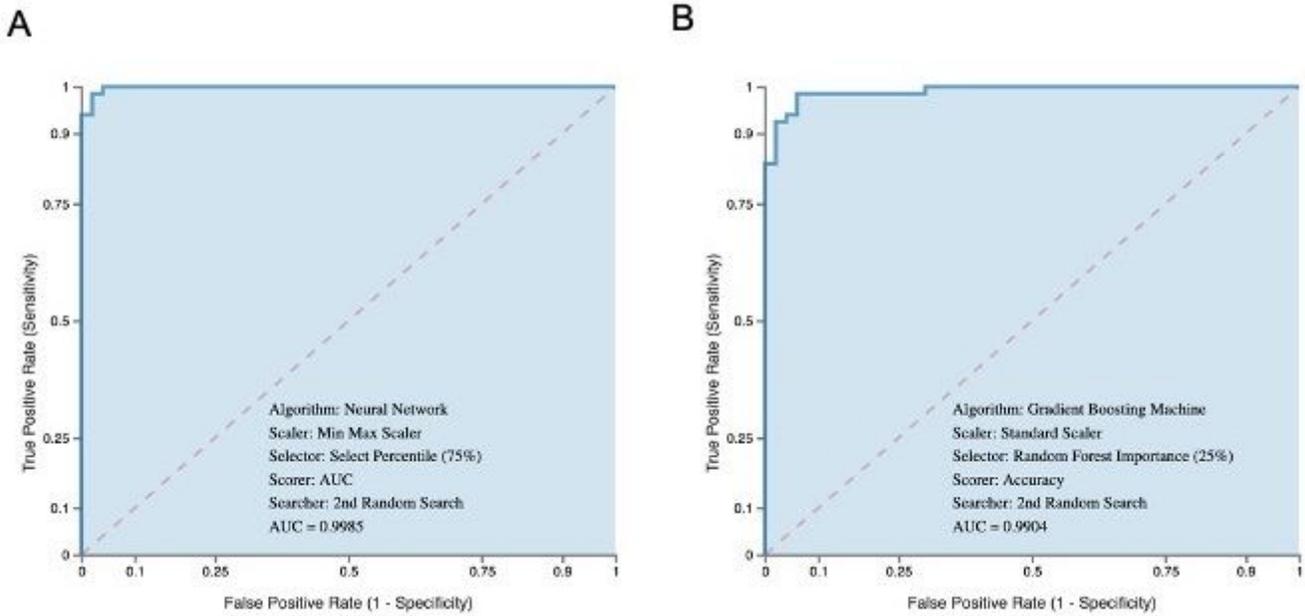


Figure 5

Receiver Operator Characteristic Curves Top Performing ML Models. The figure illustrates optimized deep neural network (Panel A) and gradient boosting machine (Panel B) ML models secondarily tested by Dataset B. For the deep neural network, the ML model used 75% of MS peaks (features) to yield a positive percent agreement of 100% (95% CI 95-100%), and negative percent agreement of 96% (95% CI 86-99%), with an area under the receiver operator curve of 0.9985. In contrast, the Gradient Boosting Machine ML model used only 25% of the MS peaks (features) to yield a positive percent agreement of 99% (95% CI 92-100%) and negative percent agreement of 94% (95% CI 84-99%) with an area under the receiver operator curve of 0.9904.

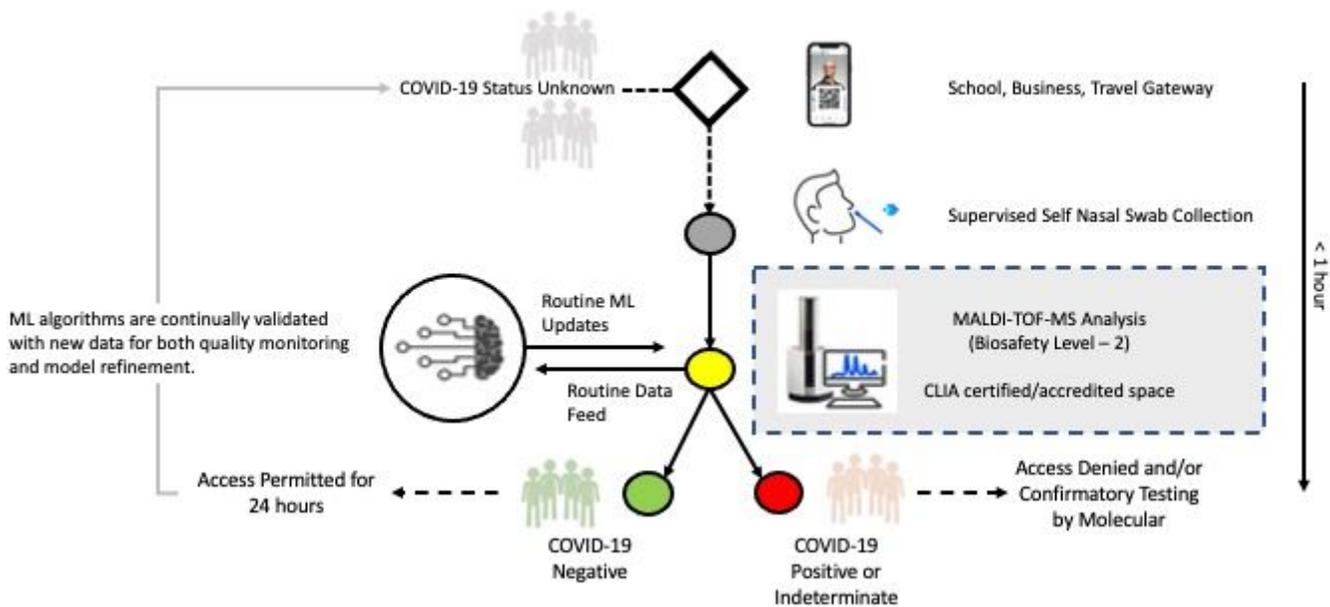


Figure 6

Conceptual Model for Near Patient ML-Enhanced MALDI-TOF-MS COVID-19 Testing. The Figure outlines the conceptual workflow for our ML-enhanced MALDI-TOF-MS COVID-19 testing method when performed near patient. Individuals with unknown COVID-19 status register via a smart device app which links their identity with a unique quick-response (Q-R) barcode. The Q-R code is paired to the nasal swab specimen which is self-collected under supervision. The sample is tested by MALDI-TOF-MS and mass spectra analyzed by the ML algorithm to report out a COVID-19 result. COVID-19 individuals are allowed entry for 24 hours until. COVID-19 positive / indeterminant individuals will be denied entry and/or require follow-up testing by molecular methods. Data from MALDI-TOF-MS is fed routinely to the automated ML platform for both quality assurance and continual refinement of models. Total time from sample collection to result is < 1 hour.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)