Machine Learning for Prediction of Hemodialysis Patients with an Undetected SARS-CoV-2 Infection

*Caitlin K. Monaghan¹, *John W. Larkin¹, Sheetal Chaudhuri¹⁺⁻⁷, Hao Han¹, Yue Jiao¹, Kristine M. Bermudez², Eric D. Weinhandl², Ines A. Dahne-Steuber², Kathleen Belmonte³, Luca Neri⁴, Peter Kotanko⁵⁺⁶, Jeroen P. Kooman⁷, Jeffrey L. Hymes², Robert J. Kossmann², Len A. Usvyat¹, Franklin W. Maddux⁸

*Co-first authors

¹ Fresenius Medical Care, Global Medical Office, Waltham, United States; ² Fresenius Medical Care North America, Medical Office, Waltham, United States; ³ Fresenius Kidney Care, Waltham, United States; ⁴ Fresenius Medical Care Deutschland GmbH, EMEA Medical Office, Bad Homburg, Germany; ⁵ Renal Research Institute, New York, United States; ⁶ Icahn School of Medicine at Mount Sinai, New York, United States; ⁷ Maastricht University Medical Center, Maastricht, The Netherlands; ⁸ Fresenius Medical Care AG & Co. KGaA, Global Medical Office, Bad Homburg, Germany

**Corresponding Author:** Len A. Usvyat, PhD; Fresenius Medical Care, Global Medical Office; 920 Winter Street, Waltham, MA, 02451 USA; Phone: (+1) 781-699-2535; E-Mail: len.usvyat@fmc-na.com
Key Points

- We developed a machine learning predictive model to detect dialysis patients with a SARS-CoV-2 infection 3 days before symptoms onset.
- Changes in physiological markers were subtle independently; model appeared to detect important combinations for each patient’s prediction.
- We proposed a conceptual workflow for application of model directed mitigation and testing within the standard practices of a provider.

Abstract

**Background:**
We developed a machine learning (ML) model that predicts the risk of a hemodialysis (HD) patient having an undetected SARS-CoV-2 infection that is identified after the following 3 or more days.

**Methods:**
As part of a healthcare operations effort we used patient data from a national network of dialysis clinics (February-September 2020) to develop a ML model (XGBoost) that uses 81 variables to predict the likelihood of an adult HD patient having an undetected SARS-CoV-2 infection that is identified in the subsequent ≥3 days. We used a 60:20:20% randomized split of COVID-19 positive samples for the training, validation, and testing datasets.

**Results:**
We used a select cohort of 40,490 HD patients to build the ML model (11,166 COVID-19 positive cases and 29,324 unaffected (control) patients). The prevalence of COVID-19 in the cohort (28% COVID-19 positive) was by design higher than the HD population. The prevalence of COVID-19 was set to 10% in the testing dataset to estimate the prevalence observed in the national HD population. The threshold for classifying observations as positive or negative was set at 0.80 to minimize false positives. Precision for the model was 0.52, the recall was 0.07, and the lift was 5.3 in the testing dataset. Area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) for
the model was 0.68 and 0.24 in the testing dataset, respectively. Top predictors of an HD patient having a SARS-CoV-2 infection were the change in interdialytic weight gain from the previous month, mean pre-HD body temperature in the prior week, and the change in post-HD heart rate from the previous month.

**Conclusions:**

The developed ML model appears suitable for predicting HD patients at risk of having COVID-19 at least three days before there would be a clinical suspicion of the disease.
Introduction:

The 2019 coronavirus disease (COVID-19) pandemic is challenging the world’s healthcare systems, including bringing complexities to the maintenance of dialysis in people with end stage kidney disease (ESKD) (1-5). In the United States, most ESKD patients are treated by outpatient hemodialysis (HD) where social distancing can be difficult and heightened infection control measures are required (e.g. temperature screenings, universal masking, isolation treatments/shifts/clinics) (1-5). ESKD patients are typically older and have multiple comorbidities, placing the population at higher risk for requiring intensive care and dying if affected by COVID-19 (6-12).

Early reports from the United States show an 11% COVID-19 mortality in ESKD (13), which is higher than the 3% COVID-19 mortality shown in the national population (14, 15). This is not unexpected with reports from Asia and Europe suggesting a 16% to 23% COVID-19 mortality in ESKD (16-19). Albeit the high mortality rate, an impaired immune response may render dialysis patients more frequently asymptomatic when infected by SARS-CoV-2 (16, 17). In both the general and ESKD populations, the most prevalent symptoms of COVID-19 at presentation are fever (11%-66% in dialysis; 82% in general population) and cough (37%-57% in dialysis; 62% in general population) (16, 20-22). The less frequent occurrence of signs and symptoms indicative of COVID-19 in dialysis patients could be making the outbreak even more challenging to manage.

Dialysis providers routinely capture patient/clinical data during care. The robust data collected during HD treatments (generally thrice weekly) provide unique opportunities to leverage artificial intelligence (AI) in predicting COVID-19 outcomes. AI modeling helped identify onset of the outbreak in China (23, 24) and is currently being used to help with early detection of areas and individuals in the general population at risk for COVID-19 (25-27).
As part of a healthcare operations effort in response to the COVID-19 outbreak, an integrated kidney disease healthcare company aimed to develop a machine learning (ML) prediction model that identifies the risk of an HD patient having an undetected severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infection. We analyzed the model performance to determine the possible utility for testing in the HD population.

**Materials and Methods:**

**General:**

An integrated kidney disease healthcare company (Fresenius Medical Care, Waltham, MA, United States) used retrospective real world data from its national network of dialysis clinics to develop a ML model that predicts the risk of an adult HD patient having an undetected SARS-CoV-2 infection that is identified after the following ≥3 days.

This analysis was performed in adherence with the Declaration of Helsinki under an initial and revised protocol reviewed by New England Independent Review Board (NEIRB). This retrospective analysis was determined to be exempt and did not require patient consent (Needham Heights, MA, United States; Protocol version 1.0 NEIRB#1-17-1302368-1; Protocol revision version 1.1 NEIRB#17-1348994-1).

**COVID-19 Mitigation and Testing Practices:**

The national network of dialysis clinics (Fresenius Kidney Care, Waltham, MA, United States) started implementing modified infection control measures in late Feb 2020 in response to the COVID-19 outbreak in the general population. Universal mitigation efforts at the provider included, screening patients/staff before entry into the dialysis facility for high body temperature, signs or symptoms of flu-like illness, exposure to others with COVID-19, or a known infection diagnosed elsewhere (28). Patients
and staff were required to thoroughly wash their hands upon entering and leaving the facility. Patients were provided surgical masks and were required to wear them while in any area of the facility. Staff were required to wear enhanced personal protective equipment including use of masks, face shields, gowns, and gloves while being in the proximity of patients in any area. The first dialysis patients (n=2) at the provider were identified as COVID-19 positive on 03 Mar 2020.

All patients and staff with an elevated body temperature or symptoms of a flu-like illness were considered under investigation and had reverse transcription polymerase chain reaction (RT-PCR) laboratory testing for SARS-CoV-2 performed at a laboratory contracted by the dialysis provider. Patients under laboratory investigation for a SARS-CoV-2 infection were treated in dedicated isolation areas (rooms, shifts, or clinics) for suspected patients until confirmed negative by two RT-PCR tests that were more than 24 hours apart. Patients who had been exposed to others with COVID-19 were moved to unique isolation areas for exposed patients under investigation for 14 days and received RT-PCR testing if they presented with signs or symptoms of a flu-like illness. Patients with RT-PCR confirmed COVID-19 were treated in dedicated isolation areas for infected patients until two negative RT-PCR tests more than 24 hours apart were documented.

Population and Outcome:

We considered data from adult (age ≥18 years) HD patients treated throughout the national network for development of a model to predict individuals with an undetected SARS-CoV-2 infection. The observation period started on 27 Feb 2020. Positive arm included data from patients who had ≥1 confirmed positive RT-PCR COVID-19 test as of the end of the observation period (08 Sep 2020, n=11,166). Negative arm included data from patients who: 1) were found COVID-19 negative (n=7,959), or 2) were randomly sampled from all active patients at the dialysis provider without a reported
suspicion of COVID-19 as of the end of the observation period (n=21,365). The random sampling was performed using the ‘sample’ function from the ‘pandas’ Python package.

We defined the index date of a HD patient having a SARS-CoV-2 infection as the date of the COVID-19+ test. In control patients with a negative COVID-19 test result, the test date was used as the index date. In controls without a test, the index date was randomly sampled from the positive cases’ index dates occurring before 25 Aug 2020, two weeks before the end of the observation period. This cutoff was chosen to minimize the possibility that control patients were infected but had not displayed signs or symptoms leading to testing before the end of the observation period. We included data from patients with 1) ≥1 hemoglobin sample collected both 1-14 days and 31-60 days before the individual’s prediction date (3 days prior to index date, further defined below), and 2) ≥1 HD treatment both 1-7 days and 31-60 days preceding the prediction date. This was done to ensure we included only active patients as hemoglobin draws are conducted weekly for in-center HD (typically thrice weekly treatments). We excluded data from patients suspected to have COVID-19 who were pending laboratory testing or were classified as person under investigation (PUI) where no laboratory testing was performed or documented.

AI Model Development:

Software and ML Model Logic:

We used Python version 3.7.7 (Python Software Foundation, Delaware, United States) to build the ML model utilizing the XGBoost package (29). The XGBoost Python package used input variables from the training dataset to construct multiple decision trees, giving each a random sample, and established a series of thresholds that split variables to maximize the information gain. Decision trees were constructed iteratively, and new decision trees were added to predict prior errors. The decision trees
made by the XGBoost ML model are inherently able to handle missing values without imputation by including their presence when determining the splits (e.g. splitting observations with temperatures \( \geq 98.0^\circ F \geq 36.7^\circ C \) from temperatures <98.0°F (<36.7°C) or missing temperatures). After no further improvements in performance were achieved using the validation dataset (also used for hyperparameter tuning), the ensemble of decision trees produced the final ML model that was assessed with the testing dataset.

Undetected SARS-CoV-2 Prediction Model:

We used 81 *a priori* selected treatment/laboratory variables up to the individually defined prediction date (3 days prior to the index date defined above) to predict the risk of a SARS-CoV-2 infection being identified in the following \( \geq 3 \) days (*Figure 1*). This is intended to yield individual predictions at least 3 days in advance of symptoms that warranted testing. We used a 60:20:20% randomized split of COVID-19+ samples for the training, validation, and testing datasets, and added the same number of COVID-19 negative patients to only the training and validation datasets. The testing dataset used to evaluate final model performance had a higher number of COVID-19 negative samples added to more closely match the prevalence observed in the overall national HD population (30, 31).

Statistical Methods:

Descriptive Statistics:

Descriptive statistics for HD patients were tabulated for demographics and variables at the time of the prediction for an undetected SARS-CoV-2 infection. Data are stratified by HD patients who did, or did not, have laboratory confirmation of COVID-19 after the date of prediction.

Analysis of ML Model Feature Importance:
Shapley values (32, 33) were calculated using the SHAP python package to determine the influence of each variable on the predictions (34, 35). SHAP values are calculated for each variable and each observation, representing a measure of impact (positive or negative value) of the observed value on each individual prediction. SHAP methods withhold and include individual inputs in all possible combinations, and compare differences between withheld and included data, to compute the mean value of all possible differences for attributing the feature importance. SHAP values are output as log odds (i.e. the logarithm of the odds ratio), meaning they are additive explanations of feature importance. SHAP values for each variable are summed for each set of observations (in this case, for each patient) and converted from log odds to probability, which is then output by the model as the prediction. Thus, the more positive SHAP values increase the predicted probability, while more negative SHAP values decrease it. Overall feature importance for individual variables in the model were calculated from the SHAP values using the mean absolute values for each variable across all observations.

Analysis of ML Model Performance:

Performance of ML model was measured by the area under the receiver operating characteristic curve (AUROC) in the training, validation, and testing datasets, as well as the recall, precision, and lift in the testing datasets. Additionally, we evaluated the area under the precision-recall curve (AUPRC) in the testing dataset.

AUROC measures the rate of true and false positives classified by the prediction model across probability thresholds. The definition of true/false positives and negatives is shown in Table 1.

Recall (sensitivity) measures the rate of true positives classified by the model at a specified threshold and is calculated as follows:
Recall = \frac{\text{number of true positives classified by model}}{\text{number of true positives classified by model} + \text{number of false negatives classified by model}}

Precision measures the positive predictive value for the model at a specified threshold and is calculated as follows:

\text{Precision} = \frac{\text{number of true positives classified by model}}{\text{number of true positives classified by model} + \text{number of false positives classified by model}}

Lift measures the effectiveness of the model compared to random sampling and is calculated as follows:

\text{Lift} = \frac{\text{model precision}}{\text{proportion of positives in dataset}}

AUPRC measures the ratio of precision for corresponding recall values across probability thresholds (36).

AUROC, AUPRC, recall, and precision metrics yield scores on a scale of 0 (lowest) to 1 (highest). A model performing at chance would yield an AUROC of 0.5, an AUPRC equal to the proportion of positives in the dataset, and a lift value of 1. The cutoff threshold for classifying predictions were selected to optimize recall, precision, and lift according to the use case.

\text{Results:}

\text{Patient Characteristics:}

We identified data from a select cohort of 40,490 HD patients meeting eligibility criteria (11,166 COVID-19+ cases and 29,324 unaffected (control) patients). The prevalence of COVID-19 in the cohort (28% COVID-19 positive) was by design higher than the HD population. The prevalence of COVID-19+ cases
(about 28% COVID-19+) in the training and validation datasets was consistent within the cohort. For the testing dataset used to evaluate final model performance, there was a 10% prevalence of COVID-19+ cases based on the designed data split that was made to estimate the prevalence observed in the national HD population (30, 31).

In the cohort, there was a higher proportion of HD patients with a SARS-CoV-2 infection of black race, Hispanic ethnicity, and with diabetes (Table 2). Mean values for the 81 treatment and laboratory variables before a SARS-CoV-2 infection being identified in the subsequent ≥3 days (or concurrent index date in controls) are shown in Tables 3 & 4.

HD patients who contracted COVID-19 had only subtle, clinically unremarkable distinctions in treatment and laboratory characteristics before being suspected to have a SARS-CoV-2 infection compared to unaffected patients. Mean pre-/post-HD body temperatures (Table 3) and inflammatory markers (white blood cell (WBC) count and differential) (Table 4) before a SARS-CoV-2 infection being identified did not show a clinically relevant difference between groups. HD patients who had a SARS-CoV-2 infection identified in the following 3 days did appear to have somewhat higher ferritin levels compared to unaffected patients.

Prediction Model Feature Importance:

Calculation of variable feature importance with SHAP values found the top three predictors of HD patients having a SARS-CoV-2 infection were the change in interdialytic weight gain (IDWG) from the previous month, mean pre-HD body temperature in the prior week, and the change in post-HD pulse from the previous month (Figure 2A).

The SHAP value plot in Figure 2B further shows the degree of positive or negative impact of each individual measurement for each individual prediction. Each dot corresponds to an individual patient,
where the dot’s position on the x-axis represents that feature’s impact on the model prediction while the color indicates how high or low that feature’s value was. Features with missing values are indicated in gray.

For the top predictor of the change in interdialytic weight gain in the week before compared to the month before a SARS-CoV-2 infection, smaller (negative) values (cooler colors) were associated with a positive SHAP value, while larger values (warmer colors) were associated with a negative SHAP value. These results showed for each individual prediction, the model generally considered decreases in interdialytic weight gain from the previous month to be associated with a greater probability of an undetected SARS-CoV-2 infection and an increase in interdialytic weight gain to be associated with a lower likelihood of an undetected SARS-CoV-2 infection. In other words, patients who do not gain as much weight as usual in between dialysis treatments are deemed more likely to have an undetected SARS-CoV-2 infection by the model.

Along with highlighting directional effects as previously stated, Figure 2B also highlights different distributions of effects that might not be apparent when viewing the mean absolute values as in Figure 2A. For example, the eighth most important variable, change in monocytes from the previous month, produces the largest (most positive) SHAP values out of all the variables shown. This long, rightward tail along the x-axis indicates that despite having a lower mean absolute value in comparison to other variables, for some individuals this is very important. Specifically, the model assessed that patients with increased monocyte levels from the previous month are deemed more likely to have a SARS-CoV-2 infection, whereas the SHAP values for those with similar or lower levels of monocytes do not significantly decrease the prediction.

*Prediction Model Performance:*
The ML model had adequate performance in prediction of the 3-day risk for having an undetected SARS-CoV-2 infection. The ML model had an AUROC of 0.77, 0.67, and 0.68 in the training, validation, and testing datasets respectively (Figure 3). The ML model had an AUPRC of 0.24 in the testing dataset (Figure 4).

Setting the threshold for classifying observations as positive or negative at 0.80 to minimize false positives, the precision for the ML model in the testing dataset was 0.52 showing 52% of patients predicted to have a SARS-CoV-2 infection actually had symptoms in the subsequent ≥3 days and were confirmed to have COVID-19. Given the high threshold, recall was 0.07 showing the model correctly predicted true positives for a SARS-CoV-2 infection in 7% of positive HD patients. The lift was 5.3, suggesting model use is 5.3 times more effective in predicting a HD patient who contracts COVID-19, as compared to not having a model (Figure 5).

Discussion:

We successfully developed a ML prediction model using retrospective data that appears to have suitable performance in identifying HD patients at risk of having an undetected SARS-CoV-2 infection that is identified in the following ≥3 days. The top predictors of a patient having a SARS-CoV-2 infection were the change in interdialytic weight gain from the previous month, mean pre-HD body temperature in the prior week, and the change in post-HD pulse from the previous month.

Albeit some top predictors are not surprising, the observed distinctions were subtle. Without insights from the model considering an array of variables, it would not be clear where one should classify a higher or lower risk for an individual patient that is meaningful. For instance, assessing for a decrease in weekly IDWG of about 0.3 kg alone may not be considered actionable, and the same is true for assessing for an increase of about 0.2°F (0.1°C) in weekly pre-HD body temperature, or an increase in pulse of
about 1 beat per min (BPM). Notably, the average pre-HD body temperature was 97.6°F (36.4°C) (primarily oral measurements) in our analysis and has been previously reported as 98.2°F (36.7°C) (37).

Given 98.6°F (37°C) is the expected average in healthy populations, the lower body temperature of HD patients is of importance with the rather low incidence of fever presenting in dialysis patients with COVID-19 (11%-to-66% with fever (16, 20, 22)). Overall, the small changes observed for each individual variable suggest any one parameter alone has minimal value for detecting a patient’s risk of having COVID-19, especially since every affected patient will not have every symptom of COVID-19 consistently. However, the combinations of minor changes appear to be meaningful in the individualized ML model we developed, with each small change being one piece of the puzzle for each patient’s unique prediction.

Individual predictions can be further used to identify the risk level for dialysis clinics through the proportion of patients classified with an undetected SARS-CoV-2 infection. We anticipate using a combination of individual predictions along with reporting of the percent of patients at risk in each clinic may yield the greatest early insights on: 1) what otherwise asymptomatic HD patients might be most appropriate for enhanced screening, COVID-19 testing, and triage to an isolation area, and 2) where providers can focus additional resource allocations to combat COVID-19. Furthermore, flagging patients as potentially infectious may cut through some of the ‘COVID fatigue’ occurring during this prolonged pandemic. By adding this additional novelty and warning, the hope is additional care may be given in identifying of potential symptoms during screening. Prospective evaluation of ML model directed mitigation is currently being piloted at the national network of dialysis clinics.

The authors propose a conceptual workflow for the application of the ML model predictions to assist with directing care to individual patients and assisting with directing resource allocations to clinics (Figure 6). The model was trained using a target date of three days prior to patients presenting with
COVID-19 symptoms to alert clinicians at least one dialysis treatment earlier. Given this timeline, we believe it is prudent to run the prediction model on a per treatment basis. The delivery of reports on individual patient predictions to clinic staff would optimally be delivered on interdialytic days to provide the care team time to prepare for a more comprehensive screening by an advanced clinician at the next encounter and potential isolation of subsequent HD treatments. The delivery of reports on the percent of patients in each clinic at risk can be performed on a weekly basis to allow leadership and regional managers to meet with clinical managers and prepare for allocation of resources including additional staff, protective equipment, and isolation areas. We propose categorizing clinic-level reports to detail facilities with more than 5% of patients at risk for undetected SARS-CoV-2 infection.

Mitigation efforts at the national dialysis network include universal RT-PCR testing of patients with symptoms of a flu-like illness along with distinct isolation areas (rooms, shifts, clinics) for suspected patients under investigation and COVID-19 positive patients. We propose patients predicted to be at risk receive a comprehensive screening for signs and symptoms of a flu-like illness by an advanced practitioner (e.g. physician, physician assistant, nurse practitioner, experienced dialysis nurse) since there is a possibility of false positives. However, the comprehensive assessments should consider any minor sign or symptoms of a flu-like illness that may otherwise be considered normal based on the patient’s uremia and medical history (38, 39) to be a reason for suspicion of COVID-19. Given the predictions are derived for each individual patient, the reasons underlying the risk predictions can be provided for each patient, as well as the global importance of features for the model (Figure 2). This may help to provide additional insight into what the what a the more comprehensive screening assessment should focus on for each individual patient. For example, if a patient is classified by the model at risk with the top reason being related to a decrease in intradialytic weight gain, the next screening before entry to the clinic could include assessment of any change in appetite or fluid intake. High risk patients suspected with any mild sign of a flu-like illness could be triaged to unique isolation areas for patients
under investigation and receive RT-PCR testing. HD would be continued in a distinct isolation area until diagnosis of COVID-19 or not (determined by two negative RT-PCR tests >24 hours apart), whereby laboratory positive patients would be triaged to unique isolation areas for COVID-19, and negative patients would return to be treated with the general HD population (Figure 6), which is consistent with the providers' practices without the model. Patients diagnosed with COVID-19 at the provider are treated in distinct isolation areas until they have two negative RT-PCR tests >24 hours apart, after which recovered patients are transferred back to receive HD with the unaffected HD population.

The developed model has potential to provide a data-driven way for providers to identify individuals with undetected SARS-CoV-2 infections. The conceptual workflow provides a hypothetical strategy that can be adapted within the practice patterns of other providers, which may not include universal testing and require periods of isolation. Different strategies could utilize different thresholds for flagging patients depending on the intervention and implications of false positives and false negatives. Considering the possibility of prolonged viral shedding observed in the general and dialysis populations (40-42), the optimal period for isolation of dialysis patients affected by COVID-19 appears to be longer than 14 days (42). In countries or areas with testing limitations, especially those with a high positive-to-negative testing ratio (e.g. >25% positive test rate), it may be reasonable to consider having separate isolation areas for patients predicted at risk in addition to isolation areas for patients with symptoms of a flu-like illness. In this scenario, the 14-day timeframe for isolation of patients predicted to be at risk is anticipated to be appropriate if no signs or symptoms of a flu-like illness arise.

As more data is captured in the COVID-19 outbreak, further prediction models that can classify the risk of morbid/mortal outcomes in dialysis patients affected by COVID-19 need to be developed. The potential applications of AI for COVID-19 have been previously detailed (43); the first priority was suggested as “early detection and diagnosis of the infection”. The robustness of data and an a priori
selection of variables to be included in our ML model bring value through assessment of feature importance; this allows for interpretation of meaningfulness of predictors, albeit it does not determine causality. The selection of input variables was focused on biological changes reflected in clinical presentations and biomarkers allowing the model to be generalizable to all individual HD patients in the overall population, and not specific to the characteristics of outbreaks or the local population where patients reside. Although this approach yields more generalizability for the model to be used in the HD populations worldwide, external factors such as local incidence rates or social determinants of health are anticipated to impact the likelihood of a patient contracting COVID-19 and can be considered as appropriate. Ultimately, this strategy has the potential to allow for COVID-19 to be detected sooner than HD patients show symptoms, and for a localized HD population, earlier than it would be reported by national authorities.

A systematic review identified several models developed using data from China for early detection of COVID-19 in suspected individuals in the general population (27). One is an externally validated ML model that predicts COVID-19 in suspected asymptomatic patients (AUROC validation=0.872) (44). Another effort used a prediction model (AUROC validation=0.966) to develop logic for an 8 variable COVID-19 risk chart (45). A further model with an AUROC of 0.938 was created to detect COVID-19 pneumonia in patients admitting to a fever clinic (46). Other models used genomic/computed tomography data to diagnose COVID-19 (27). An effort using data from China not included in prior reviews developed various ML models to predict (AUROC testing=0.87 to 0.95) and identify features indicative of COVID-19 status across age categories among people in the general population presenting to a clinic/hospital (47). This model found the most important features for prediction of COVID-19 at presentation were lung infection, cough, pneumonia. Consistent variables used across models for predictions included age, body temperature, and flu-like illness symptoms (27, 47). Another distinct effort reported in the literature included the development of ML and traditional models using only full
blood count data to predict the likelihood of a COVID-19 among people in the general population presenting to the emergency department (AUROC training=0.80 to 0.86) of, or patients admitted at (AUROC training=0.94 to 0.95), a large hospital in Brazil (48). Although these models were all reported to have suitable performance, all were subject to bias due to non-generalizable sampling of controls without COVID-19 and possible overfitting. We cannot rule out that our ML model may have similar bias, although it included a large sample and the testing dataset had relatively generalizable sampling for the dialysis population with respect to positives and negatives (30, 31). Also, since we randomly selected a subset of patients for the negative arm who never had symptoms of COVID-19 and did not receive PCR testing, it is possible that we might have unintentionally included a small number of patients who were asymptomatic. However, this would have required patients to have had an asymptomatic SARS-CoV-2 infection that aligned with the randomly sampled time window. Given the balanced class design of the training and validation data splits, it is unlikely this introduced a remarkable bias in the model during training and validation. Yet, there is a possibly this could have introduced a minimal bias in evaluation of performance in the testing data since there were fewer positive cases to identify to offset any impact of an incorrectly labeled negative patient as positive. Additionally, the reported model performance may be on the conservative side when considering the constraints of the “ground truth” labels as they relate to how positive patients are identified by conventional screening. The extent of this depends on how well the model identifies individuals not included in the training sample but might show similar patterns and on the intervention. In any case, our model is unique in its ability to identify the risk of SARS-CoV-2 infection in patients without any suspicion of being affected with the disease.

The developed model holds promise to help providers through the COVID-19 pandemic and subsequent wave(s) of outbreak (49, 50). We recommend model use as augmentation and not replacement of symptom screening, as AI modeling is never 100% accurate and model risk classifications need to be interpreted within the extent of the model’s performance.
Conclusions:

The developed AI model showed a clinically meaningful performance in prediction of individual HD patients at risk of having an undetected SARS-CoV-2 infection at least three days before there would be any suspicion of the disease. Prospective testing is needed and underway at the national network of dialysis clinics. We proposed a conceptual workflow for application of ML model directed mitigation and testing. These efforts should provide key insights for consideration by healthcare providers.

Disclosures:

CKM, JWL, SC, HH, YJ, LAU, FWM are employees of Fresenius Medical Care in the Global Medical Office. KMB, EDW, IADS, KB, JLH, RJK are employees of Fresenius Medical Care North America. LN is an employee of Fresenius Medical Care Deutschland GmbH in the EMEA Medical Office. PK is an employee of Renal Research Institute, a wholly owned subsidiary of Fresenius Medical Care. IADS, KB, PK, JLH, RJK, LAU, FWM have share options/ownership in Fresenius Medical Care. PK receives honorarium from Up-To-Date and is on the Editorial Board of Blood Purification and Kidney and Blood Pressure Research. JLH has directorships in the Renal Physicians Association Board of Directors and Nephroceuticals LLC Scientific Advisory Board. FWM has directorships in the Fresenius Medical Care Management Board, Goldfinch Bio, and Vifor Fresenius Medical Care Renal Pharma. All remaining authors have nothing to disclose.

Funding:

Project and manuscript composition were supported internally by Fresenius Medical Care.

Acknowledgements:

We would like to acknowledge Vladimir M Rigodon for assistance with the composition of the regulatory protocol for this analysis.
Previous version of this manuscript appeared on preprint server MedRxiv, https://www.medrxiv.org/content/10.1101/2020.06.15.20131680v1.

Authors' contributions:

C Monaghan: Conceptualization; Data curation; Formal analysis; Methodology; Validation; Visualization; Writing - original draft; Writing - review and editing
J Larkin: Data curation; Resources; Supervision; Visualization; Writing - original draft; Writing - review and editing
S Chaudhuri: Data curation; Resources; Supervision; Writing - review and editing
H Han: Data curation; Writing - review and editing
Y Jiao: Data curation; Writing - review and editing
K Bermudez: Writing - review and editing
E Weinhandl: Writing - review and editing
I Dahne-Steuber: Conceptualization; Methodology; Writing - review and editing
K Belmonte: Conceptualization; Methodology; Writing - review and editing
L Neri: Writing - review and editing
P Kotanko: Conceptualization; Methodology; Supervision; Writing - review and editing
J Kooman: Supervision; Writing - review and editing
J Hymes: Conceptualization; Methodology; Supervision; Writing - review and editing
R Kossmann: Conceptualization; Methodology; Resources; Supervision; Writing - review and editing
L Usvyat: Conceptualization; Data curation; Formal analysis; Methodology; Resources; Supervision; Visualization; Writing - original draft; Writing - review and editing
F Maddux: Conceptualization; Methodology; Supervision; Writing - review and editing

The interpretation, drafting and revision of this manuscript was conducted by all authors. The decision to submit this manuscript for publication was jointly made by all authors and the manuscript was confirmed to be accurate and approved by all authors.
References:

EDTA Registry indicate a high mortality due to COVID-19 in dialysis patients and kidney transplant recipients across Europe. *Kidney Int*, 2020


Tables:

Table 1: Definition of true/false positive and negative predictions classified by the model in the assessment of performance in the testing dataset

<table>
<thead>
<tr>
<th>Prediction Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positives</td>
<td>Patients classified as COVID-19 positive by the model who were in the COVID-19 positive group</td>
</tr>
<tr>
<td>False positives</td>
<td>Patients classified as COVID-19 positive by the model who were in the COVID-19 negative group</td>
</tr>
<tr>
<td>True negatives</td>
<td>Patients classified as COVID-19 negative by the model who were in the COVID-19 negative group</td>
</tr>
<tr>
<td>False negatives</td>
<td>Patients classified as COVID-19 negative by the model who were in the COVID-19 positive group</td>
</tr>
</tbody>
</table>

Table 2: Demographics and Comorbidities of HD Patients with and without an Undetected SARS-CoV-2 Infection Identified in the Subsequent ≥3 Days

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unaffected Patients N (%) or Mean±SD</th>
<th>COVID-19+ Patients N (%) or Mean±SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HD patients</td>
<td>29,324</td>
<td>11,166</td>
</tr>
<tr>
<td>Age (Years)</td>
<td>62.66±14.25</td>
<td>62.62±13.92</td>
</tr>
<tr>
<td>Male</td>
<td>16,614 (56.66%)</td>
<td>6,149 (55.07%)</td>
</tr>
<tr>
<td>White Race</td>
<td>12,021 (40.99%)</td>
<td>4,338 (38.85%)</td>
</tr>
<tr>
<td>Black Race</td>
<td>7,838 (26.73%)</td>
<td>3,354 (30.04%)</td>
</tr>
<tr>
<td>Other Race</td>
<td>1,223 (4.17%)</td>
<td>372 (3.33%)</td>
</tr>
<tr>
<td>Unknown Race</td>
<td>8,242 (28.11%)</td>
<td>3,102 (27.78%)</td>
</tr>
<tr>
<td>Hispanic Ethnicity</td>
<td>2,849 (14.03%)</td>
<td>1,831 (23.34%)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>29.26±7.71</td>
<td>29.45±7.83</td>
</tr>
<tr>
<td>Dialysis Vintage (Years)</td>
<td>3.75±4.11</td>
<td>3.96±4.09</td>
</tr>
<tr>
<td>Diabetes</td>
<td>19,186 (65.58%)</td>
<td>8,085 (73.11%)</td>
</tr>
<tr>
<td>CHF</td>
<td>6,710 (22.93%)</td>
<td>2,595 (23.47%)</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>7,647 (26.14%)</td>
<td>2,830 (25.59%)</td>
</tr>
<tr>
<td>Central Venous Catheter Access</td>
<td>6,799 (23.19%)</td>
<td>2,738 (24.52%)</td>
</tr>
</tbody>
</table>

Age, gender, and catheter access variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days.

HD: hemodialysis; CHF: congestive heart failure; BMI: body mass index; N: patient count; SD: standard deviation
Table 3: Clinical and Treatment Characteristics of HD Patients with and without an Undetected SARS-CoV-2 Infection Identified in the Subsequent ≥3 Days

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unaffected Patients Mean±SD; N</th>
<th>COVID-19+ Patients Mean±SD; N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HD patients</td>
<td>29,324</td>
<td>11,166</td>
</tr>
<tr>
<td>Pre-HD Sitting SBP (mmHg)†</td>
<td>148.31±22.83; 29,324</td>
<td>146.03±23.03; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Sitting SBP (mmHg)‡</td>
<td>-0.40±15.78; 29,324</td>
<td>-1.95±16.72; 11,166</td>
</tr>
<tr>
<td>Pre-HD Sitting DBP (mmHg)†</td>
<td>76.87±13.86; 29,322</td>
<td>75.44±13.58; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Sitting DBP (mmHg)‡</td>
<td>-0.32±9.03; 29,322</td>
<td>-0.88±9.48; 11,166</td>
</tr>
<tr>
<td>Pre-HD Weight (kg)†</td>
<td>85.71±24.51; 29,323</td>
<td>85.09±24.51; 11,165</td>
</tr>
<tr>
<td>Change in Pre-HD Weight (kg)‡</td>
<td>-0.17±2.24; 29,323</td>
<td>-0.66±2.73; 11,165</td>
</tr>
<tr>
<td>Pre-HD Body Temperature (°F)†</td>
<td>97.56±0.61; 29,324</td>
<td>97.76±0.66; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Body Temperature (°F)‡</td>
<td>0.07±0.56; 29,324</td>
<td>0.22±0.65; 11,166</td>
</tr>
<tr>
<td>Post-HD Sitting SBP (mmHg)†</td>
<td>140.40±21.60; 29,321</td>
<td>144.44±21.62; 11,166</td>
</tr>
<tr>
<td>Change in Post-HD Sitting SBP (mmHg)‡</td>
<td>0.43±14.98; 29,320</td>
<td>1.55±15.74; 11,166</td>
</tr>
<tr>
<td>Post-HD Sitting DBP (mmHg)†</td>
<td>73.91±12.58; 29,319</td>
<td>73.56±12.33; 11,166</td>
</tr>
<tr>
<td>Change in Post-HD Sitting DBP (mmHg)‡</td>
<td>0.15±8.49; 29,318</td>
<td>0.41±8.79; 11,166</td>
</tr>
<tr>
<td>Pre-HD Body Temperature (°F)‡</td>
<td>97.58±0.56; 29,318</td>
<td>97.70±0.62; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Body Temperature (°F)‡</td>
<td>0.03±0.50; 29,317</td>
<td>0.14±0.57; 11,165</td>
</tr>
<tr>
<td>Pre-HD Respirations per Minute‡</td>
<td>17.64±1.16; 29,324</td>
<td>17.72±1.15; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Respirations per Minute‡</td>
<td>-0.001±0.97; 29,324</td>
<td>0.01±1.02; 11,166</td>
</tr>
<tr>
<td>Pre-HD Pulse (BPM)†</td>
<td>79.00±12.11; 29,324</td>
<td>79.02±11.90; 11,166</td>
</tr>
<tr>
<td>Change in Pre-HD Pulse (BPM)‡</td>
<td>0.11±7.26; 29,324</td>
<td>1.06±7.56; 11,166</td>
</tr>
<tr>
<td>Post-HD Respirations per Minute‡</td>
<td>17.56±1.15; 29,320</td>
<td>17.65±1.13; 11,165</td>
</tr>
<tr>
<td>Change in Post-HD Respirations per Minute‡</td>
<td>-0.007±0.95; 29,319</td>
<td>0.000±0.99; 11,165</td>
</tr>
<tr>
<td>Post-HD Pulse (BPM)†</td>
<td>75.80±11.23; 29,321</td>
<td>77.23±11.16; 11,166</td>
</tr>
<tr>
<td>Change in Post-HD Pulse (BPM)‡</td>
<td>-0.32±7.16; 29,320</td>
<td>1.30±7.87; 11,166</td>
</tr>
<tr>
<td>IDWG (kg)†</td>
<td>2.24±1.21; 29,083</td>
<td>1.95±1.29; 11,039</td>
</tr>
<tr>
<td>Change in IDWG (kg)†</td>
<td>0.01±0.90; 29,004</td>
<td>-0.26±1.09; 10,991</td>
</tr>
<tr>
<td>Post-HD Weight Loss (kg)†</td>
<td>-2.26±1.07; 29,317</td>
<td>-2.06±1.07; 11,160</td>
</tr>
<tr>
<td>Change in Post-HD Weight Loss (kg)‡</td>
<td>-0.01±0.68; 29,316</td>
<td>0.18±0.77; 11,159</td>
</tr>
<tr>
<td>Post-HD Body Temperature Change‡</td>
<td>0.01±0.66; 29,318</td>
<td>-0.06±0.70; 11,165</td>
</tr>
<tr>
<td>Change in Post-HD Body Temperature Change‡</td>
<td>-0.04±0.66; 29,317</td>
<td>-0.07±0.71; 11,165</td>
</tr>
<tr>
<td>Post-HD Respirations per Minute Change‡</td>
<td>-0.08±0.97; 29,320</td>
<td>-0.07±0.97; 11,165</td>
</tr>
<tr>
<td>Change in Post-HD Respirations per Minute Change‡</td>
<td>-0.01±1.04; 29,319</td>
<td>-0.01±1.07; 11,165</td>
</tr>
<tr>
<td>Post-HD Pulse Change (BPM)‡</td>
<td>-3.20±8.86; 29,321</td>
<td>-1.79±8.77; 11,166</td>
</tr>
<tr>
<td>Change in Post-HD Pulse Change (BPM)‡</td>
<td>-0.43±7.75; 29,320</td>
<td>0.24±8.06; 11,166</td>
</tr>
<tr>
<td>% HD Treatments with Nasal Oxygen Administered†</td>
<td>5.23±18.52; 29,324</td>
<td>5.67±19.16; 11,166</td>
</tr>
<tr>
<td>Change in % HD Treatments with Nasal Oxygen Administered‡</td>
<td>0.37±13.40; 29,324</td>
<td>0.72±14.12; 11,166</td>
</tr>
</tbody>
</table>

All variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days.

† Mean values of HD treatment variables 1-7 days before the prediction date (i.e. 3 days before suspicion of SARS-CoV-2 infection in standard clinical practice).

‡ Mean values of the difference in HD treatment variables 31-60 days to 1-7 days before the prediction date.

HD: hemodialysis; SBP: systolic blood pressure; DBP: diastolic blood pressure; IDWG: interdialytic weight gain; Post-HD Weight Loss: post-HD minus pre-HD weight (kg); N: patient count; SD: standard deviation. (100°F – 32) × 5/9 = 37.8°C
<table>
<thead>
<tr>
<th>Variable</th>
<th>Unaffected Patients Mean±SD, N</th>
<th>COVID-19+ Patients Mean±SD, N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HD patients</td>
<td>29,324</td>
<td>11,166</td>
</tr>
<tr>
<td>Albumin (g/dL)†</td>
<td>3.79±0.40; 13,723</td>
<td>3.69±0.46; 5,252</td>
</tr>
<tr>
<td>Change in Albumin (g/dL)‡</td>
<td>-0.002±0.25; 13,139</td>
<td>-0.03±0.27; 5,012</td>
</tr>
<tr>
<td>Creatinine (mg/dL)†</td>
<td>8.42±3.06; 13,323</td>
<td>8.41±3.14; 5,113</td>
</tr>
<tr>
<td>Change in Creatinine (mg/dL)‡</td>
<td>0.08±1.40; 12,711</td>
<td>0.16±1.52; 4,860</td>
</tr>
<tr>
<td>Bicarbonate (mmol/L)†</td>
<td>24.24±3.05; 13,395</td>
<td>24.22±3.22; 5,137</td>
</tr>
<tr>
<td>Change in Bicarbonate (mmol/L)‡</td>
<td>0.02±2.97; 12,772</td>
<td>-0.16±3.12; 4,864</td>
</tr>
<tr>
<td>BUN (mg/dL)‡</td>
<td>56.21±18.53; 14,941</td>
<td>56.17±19.27; 5,631</td>
</tr>
<tr>
<td>Change in BUN (mg/dL)‡</td>
<td>-0.21±15.50; 14,400</td>
<td>-0.13±16.55; 5,416</td>
</tr>
<tr>
<td>URR†</td>
<td>74.92±6.52; 14,273</td>
<td>75.05±6.61; 5,348</td>
</tr>
<tr>
<td>Change in URR‡</td>
<td>0.09±5.89; 13,548</td>
<td>0.07±6.08; 5,054</td>
</tr>
<tr>
<td>Sodium (mmol/L)†</td>
<td>137.50±3.37; 13,139</td>
<td>137.08±3.52; 5,046</td>
</tr>
<tr>
<td>Change in Sodium (mmol/L)‡</td>
<td>-0.10±2.83; 29,324</td>
<td>-0.25±3.10; 4,772</td>
</tr>
<tr>
<td>Potassium (mmol/L)†</td>
<td>4.80±0.68; 16,051</td>
<td>4.78±0.70; 6,217</td>
</tr>
<tr>
<td>Change in Potassium (mmol/L)‡</td>
<td>0.01±0.60; 15,499</td>
<td>-0.01±0.63; 6,003</td>
</tr>
<tr>
<td>Phosphate (mg/dL)†</td>
<td>5.55±1.74; 14,489</td>
<td>5.37±1.71; 5,913</td>
</tr>
<tr>
<td>Change in Phosphate (mg/dL)‡</td>
<td>0.01±1.48; 14,918</td>
<td>-0.03±1.46; 5,692</td>
</tr>
<tr>
<td>Chloride (meq/L)†</td>
<td>98.66±4.14; 12,602</td>
<td>98.33±4.13; 4,702</td>
</tr>
<tr>
<td>Change in Chloride (meq/L)‡</td>
<td>-0.19±3.35; 11,708</td>
<td>-0.24±3.50; 4,572</td>
</tr>
<tr>
<td>Calcium (mg/dL)†</td>
<td>8.89±0.69; 16,051</td>
<td>8.78±0.70; 6,217</td>
</tr>
<tr>
<td>Change in Calcium (mg/dL)‡</td>
<td>0.02±0.58; 14,882</td>
<td>-0.07±0.60; 5,659</td>
</tr>
<tr>
<td>Corrected Calcium (mg/dL)†</td>
<td>9.06±0.66; 12,865</td>
<td>9.04±0.71; 4,903</td>
</tr>
<tr>
<td>Change in Corrected Calcium (mg/dL)‡</td>
<td>0.01±0.54; 12,148</td>
<td>-0.03±0.59; 4,608</td>
</tr>
<tr>
<td>iPTH (pg/mL)†</td>
<td>489.46±454.13; 10,900</td>
<td>497.22±490.12; 3,801</td>
</tr>
<tr>
<td>Change in iPTH (pg/mL)‡</td>
<td>-21.39±280.41; 7,245</td>
<td>-21.84±296.12; 2,734</td>
</tr>
<tr>
<td>Ferritin (ng/mL)†</td>
<td>1029.94±576.07; 8,229</td>
<td>1197.32±900.22; 3,138</td>
</tr>
<tr>
<td>Change in Ferritin (ng/mL)‡</td>
<td>52.90±505.99; 4,400</td>
<td>142.00±739.89; 1,589</td>
</tr>
<tr>
<td>TSAT (%)†</td>
<td>33.07±14.10; 13,051</td>
<td>31.29±14.42; 10,090</td>
</tr>
<tr>
<td>Change in TSAT (%)‡</td>
<td>0.17±5.33; 12,310</td>
<td>-1.59±16.54; 4,689</td>
</tr>
<tr>
<td>Hgb (g/dL)†</td>
<td>10.76±1.24; 29,324</td>
<td>10.61±1.26; 11,166</td>
</tr>
<tr>
<td>Change in Hgb (g/dL)‡</td>
<td>0.05±1.07; 29,324</td>
<td>0.01±1.13; 11,166</td>
</tr>
<tr>
<td>Platelet Count (x 10⁹/L)†</td>
<td>195.49±72.47; 11,378</td>
<td>192.35±77.10; 4,293</td>
</tr>
<tr>
<td>Change in Platelet Count (x 10⁹/L)‡</td>
<td>-1.93±49.23; 10,959</td>
<td>-7.82±55.06; 3,963</td>
</tr>
<tr>
<td>WBC Count (x 10⁹/L)†</td>
<td>6.93±2.36; 13,043</td>
<td>6.55±2.39; 5,027</td>
</tr>
<tr>
<td>Change in WBC Count (x 10⁹/L)‡</td>
<td>0.03±1.76; 12,344</td>
<td>-0.36±1.93; 4,733</td>
</tr>
<tr>
<td>% of Neutrophils‡</td>
<td>66.11±9.50; 17,215</td>
<td>66.59±9.53; 6,941</td>
</tr>
<tr>
<td>Change in % of Neutrophils‡</td>
<td>0.06±1.07; 14,931</td>
<td>0.47±1.07; 5,997</td>
</tr>
<tr>
<td>% of Lymphocytes‡</td>
<td>20.22±7.98; 17,215</td>
<td>19.76±7.96; 6,941</td>
</tr>
<tr>
<td>Change in % of Lymphocytes‡</td>
<td>-0.04±4.99; 14,931</td>
<td>-0.53±4.57; 5,997</td>
</tr>
<tr>
<td>% of Monocytes‡</td>
<td>6.38±1.90; 17,215</td>
<td>6.69±2.14; 6,941</td>
</tr>
<tr>
<td>Change in % of Monocytes‡</td>
<td>0.02±1.48; 14,931</td>
<td>0.37±1.82; 5,997</td>
</tr>
<tr>
<td>% of Eosinophils‡</td>
<td>4.29±2.88; 17,212</td>
<td>3.95±2.84; 6,939</td>
</tr>
<tr>
<td>Change in % of Eosinophils‡</td>
<td>-0.10±2.03; 14,927</td>
<td>-0.40±2.28; 5,995</td>
</tr>
<tr>
<td>% of Basophils‡</td>
<td>0.75±0.47; 17,206</td>
<td>0.73±0.45; 6,934</td>
</tr>
<tr>
<td>Change in % of Basophils‡</td>
<td>0.05±0.54; 14,917</td>
<td>0.03±0.52; 5,988</td>
</tr>
</tbody>
</table>

All variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days.

† Mean values of laboratory variables 1-14 days before the prediction date (i.e. 3 days before suspicion of SARS-CoV-2 infection in standard clinical practice).
‡ Mean values of the difference in laboratory variables 31-60 days to 1-14 days before the prediction date.
HD: hemodialysis; Hgb: hemoglobin; WBC: white blood cell; TSAT: transferrin saturation; URR: urea reduction ratio; iPTH: intact parathyroid hormone; BUN: blood urea nitrogen; N: patient count; SD: standard deviation.
Figure Legends:

Figure 1: Prediction timeline for data ascertainment and prediction of HD patients with and without SARS-CoV-2 infection identified in the subsequent ≥3 days. ML model used HD treatment variables († mean values 1-7 days before the prediction date; ‡ difference in mean values 31-60 days to 1-7 days before the prediction date) and laboratory variables (◊ mean values 1-14 days before the prediction date; ○ difference in mean values 31-60 days to 1-14 days before the prediction date) for prediction of SARS-CoV-2 infection.

Figure 2: SHAP value plots for the ML model showing the extent each predictor contributes (positively or negatively) to each individual prediction. A) Bar plot of the mean absolute SHAP values for the top 10 predictors in descending order. B) SHAP value plot for the degree of the positive or negative impact of each individual measurement on the prediction (x-axis), with warmer colors representing higher observed values for that measurement, cooler colors indicating lower values for that measurement, and gray color representing a missing value for that measurement.

Figure 3: Area under the receiver operating characteristic curve (AUROC) plot for the ML model showing the rate of true and false positives classified by the prediction model across probability thresholds.

Figure 4: Area under the precision-recall curve (AUPRC) plot for the ML model showing the ratio of precision for corresponding recall values across probability thresholds.

Figure 5: Lift curve for the ML model showing the lift value (y-axis) by the proportion of the population predicted to have an undetected SARS-CoV-2 infection (x-axis).

Figure 6: Conceptual workflow for application of ML model predictions within current mitigation and testing practices at the provider.
Figure 1

Ascertainment period of data to make prediction

30-day mean values of HD treatment variables

30-day mean values of laboratory variables

3-day window before predicted outcome

Prediction period

SARS-CoV-2 infection identified in standard clinical practice

Prediction date

Index date

‡ Difference in mean values

† 7-day mean values of HD treatment variables

○ Difference in mean values

◊ 14-day mean values of laboratory variables

Days to Index Date

-60 -31 -17 -10 -4 -3 0 +14

+14

-60
Figure 2A

A

**Interdialytic Weight Gain: Change from previous month**

- Pre-HD Temperature: 1-week average
- Post-HD Pulse: Change from previous month
- Pre-to Post-HD pulse difference: 1-week average
- Post-HD Temperature: Change from previous month
- Pre-HD Pulse: Change from previous month
- Pre-HD Weight: Change from previous month
- Monocytes: Change from previous month
- Interdialytic Weight Gain: 1-week average
- Hemoglobin: 2-week average

*mean(|SHAP value|) (average impact on model output magnitude)*
Figure 2B

- **Interdialytic Weight Gain:** Change from previous month
- **Pre-HD Temperature:** 1-week average
- **Post-HD Pulse:** Change from previous month
- **Pre- to Post-HD pulse difference:** 1-week average
- **Post-HD Temperature:** Change from previous month
- **Pre-HD Pulse:** Change from previous month
- **Pre-HD Weight:** Change from previous month
- **Monocytes:** Change from previous month
- **Interdialytic Weight Gain:** 1-week average
- **Hemoglobin:** 2-week average
Figure 3

The graph shows the Receiver Operating Characteristic (ROC) curves for different datasets: Train, Val, and Test. The curves are color-coded as follows:

- *Train*: Red, AUC: 0.768
- *Val*: Purple, AUC: 0.666
- *Test*: Blue, AUC: 0.682

A dashed line represents the chance level, indicating the performance of a classifier that makes random predictions. The curves above this line indicate better performance.
Figure 5

Lift Curve

- Chance
- Threshold = 0.80
Figure 6

- Routine clinical & HD data captured at each session (typically thrice weekly)
- Routine laboratory data collected weekly, monthly, and quarterly
- AI model classifies risk for undetected SARS-CoV-2 infection for each patient after each HD treatment (typically three predictions per patient per week)
- Weekly predictions of clinics with >5% of patients at risk for undetected SARS-CoV-2 infection delivered to leadership and regional managers
- Leadership and regional managers meet with clinics at risk each week to review mitigation efforts and plan resource allocations needed
- Regional managers direct resources to at risk clinics as appropriate (e.g., excess staff/supplies, creation of additional isolation areas)
- Individual predictions delivered to clinic staff e-mail
- Predicted not at risk
- Routine screening by staff before entry to facility
- No symptoms for flu like illness
- Unaffected/recovered patients receive HD treatments in normal clinic environment with mitigation efforts
- rRT-PCR test COVID-19 Negative (two negative tests >24 hours apart)
- Predicted to be at risk for an undetected SARS-CoV-2 infection
- Comprehensive screening by dialysis nurse before entry to facility
- Symptomatic for flu like illness
- For patients predicted to be at risk, inclusive of any minor sign even if related to medical history
- Move HD to a unique isolation area for suspected symptomatic patients & perform SARS-CoV-2 rRT-PCR test, and if negative, perform second test >24 hours later
- rRT-PCR test COVID-19 Positive
- Move HD to a unique isolation area for COVID-19 positive patients & perform SARS-CoV-2 rRT-PCR test 14 days later and again >24 hours later if negative, or every 7 days thereafter if positive until patient has two negative tests >24 hours apart