A Time-Series Feature-Based Recursive Classification Model to Optimize Treatment Strategies for Improving Outcomes and Resource Allocations of COVID-19 Patients

Lin Wang, Zheng Yin, Mamta Puppala, Chika F. Ezeana, Kelvin K. Wong, Tiancheng He, Deepa B. Gotur, and Stephen T. C. Wong, Fellow, IEEE

Abstract—This paper presents a novel Lasso Logistic Regression model based on feature-based time series data to determine disease severity and when to administer drugs or escalate intervention procedures in patients with coronavirus disease 2019 (COVID-19). Advanced features were extracted from highly enriched and time series vital sign data of hospitalized COVID-19 patients, including oxygen saturation readings, and with a combination of patient demographic and comorbidity information, as inputs into the dynamic feature-based classification model. Such dynamic combinations brought deep insights to guide clinical decision-making of complex COVID-19 cases, including prediction progression, timing of drug administration, admission to intensive care units, and application of intervention procedures like ventilation and intubation. The COVID-19 patient classification model was developed utilizing 900 hospitalized COVID-19 patients in a leading multi-hospital system in Texas, United States. By providing mortality prediction based on time-series physiologic data, demographics, and clinical records of individual COVID-19 patients, the dynamic feature-based classification model can be used to improve efficacy of the COVID-19 patient treatment, prioritize medical resources, and reduce casualties. The uniqueness of our model is that it is based on just the first 24 hours of vital sign data such that clinical interventions can be decided early and applied effectively.

Such a strategy could be extended to prioritize resource allocations and drug treatment for future pandemic events.

Index Terms—COVID-19, clinical decision making, time-series feature extraction, machine learning, SpO2.

I. INTRODUCTION

CORONAVIRUS disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). First reported in December 2019, it has since spread around the globe. It was announced as a Public Health Emergency of International Concern by World Health Organization (WHO) on January 30, 2020, and was officially declared as a pandemic on March 11, 2020 [1].

COVID-19 presents enormous challenges to public health and health systems and significantly changes human behavior around the globe. As a result of the high incidence of the disease that is spread from human-to-human transmission, there is undue burden on even the most sophisticated facilities, even those with the highest level of disaster preparedness. Clinical decisions and outcomes are influenced by availability of rapid diagnostic tests, treatment options, scarcity of resources, risks to staff, the timing of intubation and use of supplemental oxygen therapy, availability of rescue measures like extra-corporeal membrane oxygenation, and patient characteristics and demographics.

A comprehensive plan based on surge projections is essential. It is imperative to maintain, augment, and extend the hospital workforce and allocate limited healthcare resources in an ethical and organized manner to maximize positive patient outcomes. Although the prevention and testing strategy for COVID-19 is outside the scope of this work, it is worth noting that the public health policies and plans of many countries, especially during early stages of the pandemic, calculate the fear of overwhelming available health systems. This paper addresses the following questions that are accounted for in many countries’ health policies and approaches during the early stages of a pandemic:

1) How to accurately evaluate and stratify patient status?
2) How to effectively distribute valuable and limited medical resources, e.g., ICU beds, ventilators, or effective treatments for certain symptoms?
The biomedical and health informatics community is actively involved in the fight against the COVID-19 pandemic on multiple fronts. From systems biology modeling of multi-omics networks involved in the infection process [2]–[4] to automated classification and interpretation of computer tomography (CT) chest images [5]–[7], from quantification of clinical trial results [8]–[10] to the prediction of virus transmission trend and hospital capacity in a neighborhood, a country, or around the world [11]–[18], machine learning technologies are contributing significantly on every possible scale from molecules to organs to human populations. As the disease spreads from its local epicenter to across the globe, the characteristics of the disease depend on the substrate population with variations in demographic extent and disease severity. The ability to accurately predict potential outcomes, including disease severity, is paramount in effectively managing resources and triage care. Specifically, variant machine learning models have been employed to sort COVID-19 patients according to the outcomes of their hospital stays (discharged or expired) [19]. Other salient features from laboratory tests, especially those from blood tests, have been incorporated into the models [20].

Our research aims to improve the classification models by finding more useful features just from information available at early stage, like at patient’s admission, to go beyond clinical outcomes and tackle the multiple layers of complex decision-making in the care of in-hospital COVID-19 patients. Readings from blood tests take hours, if not days to process. Meanwhile, hospitals are taking vital signs like temperature, blood pressure, and oxygen saturation levels for COVID-19 patients, even measuring as frequently as at intervals of ten or fifteen minutes, only to make decisions based on the latest measurement or the few most recently sampled readings, while neglecting indication to outcomes hidden in the trend of vital sign changes over extended time. In contrast, we took advantage of millions of vital sign data points of each of the COVID-19 patients recorded in the electronic medical record (EMR) to derive an evidence-based strategy to improve allocation of medical resources and optimize assignment of ICU and hospital wards by understanding the patterns and trends of changes in time-series vital sign data.

One prominent vital sign is oxygen saturation, SpO2, the fraction of oxygen-saturated hemoglobin relative to total hemoglobin in the blood. The human body requires a precise balance of oxygen in the blood that is largely disrupted in COVID-19 patients as the disease affects oxygen exchange in the pulmonary system [21], [22]. Our research indicates that daily differences between the maximum and minimum SpO2 serve as powerful factors within classification models for patient outcomes. Increased daily range of SpO2 levels or fluctuations of heart rate and diastolic blood pressure (DBP) may be an early sign of cytokine release syndrome, which is a hyperinflammatory state, predominantly affecting the lungs, causing acute respiratory distress syndrome, a critical condition that often leads to the sudden death of COVID-19 patients.

II. OVERVIEW OF DATASET

The data was obtained from the Methodist Environment for Translational enhancement and Outcomes Research (METEOR) [23] clinical data warehouse that integrates electronic medical record databases, picture archiving and communication system image archives, outpatient care databases, hospital administrative databases, and specialty care research databases across the eight hospitals of Houston Methodist in the greater Houston area. This study was approved as a quality improvement investigation by the hospital administration and the study was under the IRB protocol (Pro00011135) of Houston Methodist Hospital.

The dataset covered all patients who tested positive for SARS-Cov-2 and stayed as in-patients in the flagship hospital of Houston Methodist at Texas Medical Center between February 26 (the day the first such patient checked into Houston Methodist) to June 3, 2020. Altogether there were 1027 COVID-19 patients during this period, including 86 expired, 814 discharged, and 127 still hospitalized as of June 3, 2020. Fig. 1 summarizes the daily numbers of COVID-19 in-patients and shows a trend of hospitalized patients starting from late May, dragging though June 3, and ultimately eclipsing the previous peak in mid-April.

The dataset for each patient includes eight demographic features, status of 19 comorbidities (pre-existing conditions that affect the treatment received), 15 complications (conditions that arise during the hospital stay), lab test results, time series profiles for nine vital signs, and the administration of 49 drugs during the hospital stay, resulting in more than 1.8 million data points. Lab test results were organized into categories like blood gases, hematology, biochemistry, urinalysis, immunology, endocrine, and coagulation. The timing for allocation of key medical resources, including intensive care unit (ICU) admission, ventilation, intubation, extracorporeal membrane oxygenation (ECMO), and pressor medications form a complex web of decision-making points that affects the length of hospital stay and clinical outcomes.

Our COVID-19 patient database contains highly enriched time-series profiles of vital tests and lab signs. All the patient records are standardized and derived from the eight hospitals within the same health system, which greatly facilitates data harmonization and analysis.

A. Comorbidities and Complications Diseases

Use one space after periods and colons. Hyphenate complex modifiers: “zero-field-cooled magnetization.” Avoid dangling participles, such as, “Using (1), the potential was calculated.” [It is not clear who or what used (1).] Write instead, “The potential was calculated by using (1),” or “Using (1), we calculated the potential.”

For each patient, we collected the status of 19 comorbidities and 15 complications; some of these were found to possibly worsen conditions of COVID-19 patients, ending in patient death. We checked the top-ranking comorbidities, including heart disease, cancer, Alzheimer’s disease, diabetes, lung and kidney chronic diseases, and smoking history of patients. We also checked the frequency of all the complications transpired during hospital stay for COVID-19.
For each disease, we used a one-way ANOVA analysis between the deceased patient group and the discharged patient group to observe any significant differences between these two patient groups. To rephrase, if a disease appears more often in deceased patients statistically, then that disease may be considered as a potential risk factor.

Table I lists all the 19 comorbidities and 15 complications ordered by the P value of the one-way ANOVA analysis. There are many diseases enriched in the deceased patients, including comorbidities such as chronic renal disease, diabetes, and dementia that the patients already had before being affected with COVID-19. Severe complications were observed to have stronger correlation with the deterioration of the condition of the COVID-19 patients. Since salient disease features carry prediction power for patient death, we will train a prediction model using certain identified disease features and compared the model with other models using different features.

B. Demographic Features

Of the 814 discharged patients, the average age is 57.6, while the average age of the 86 deceased patients is 74.6 with P value less than 10-15. We emphasize our data highlights that ‘age’ is a key feature for deceased patients. Of the 86 deceased patients, males accounted for 58.1%, while only accounting for 47.4% of the discharged patient population, demonstrating a sex bias on the deceased patients with P value of 0.075. African American patients have a lower death rate compared to other races. Table II compares demographic features between discharged patients and expired patients. We compared the differences of sex and race between the discharged group and the expired group using Chi-Square test, alongside using t-test to compare the difference of age means between discharged and expired patients.

III. VITAL SIGN FEATURE ABSTRACTION

An urgent need passed back from the clinical front line is to identify the key features that lead to patient death or severe conditions in order to guide resource allocation like ICU or ventilators to the most severely ill patients before it is too late. Vital signs are intuitive and real-time measurements are better than the lab signals because even though lab tests were processed multiple times, the results of the tests are not an accurate representation of the situation. The dynamic vital sign data may contain more information than just one single vital value. Our goal is to extract the features of time-series profile of vital signs that have significant impact on patient death. In addition to the demographic and disease features, we studied the time-series vitals data of the first day and extracted features relevant to predict risk of patient death.

The vitals mostly include SpO2, fraction of inspired oxygen (FiO2), diastolic blood pressure (DBP), systolic blood pressure (SBP), respiratory rate, heart rate, temperature, and mean arterial pressure (MAP). Once the patients have been admitted in our hospital system, their vitals are measured every 3-4 hours, whereas in ICU the vitals are measured in a 15-minute interval. Just for the first 24 hours, we have already acquired a large set of data points for each vital signal. The statistical mean and standard deviation are the first two features obtained from the time-series vital sign data. We extracted the minimum and maximum values of the vitals in the first 24 hours as well as the range between them.
Considering the stability of time series vital signs, we were curious about how many ups and downs or fluctuations of the vital signals occurred within the first 24 hours. More fluctuation indicates instability of the patient’s condition. The number of the fluctuations is calculated by counting the number of -2 (indicating a local minimum) in the equation: diff(sgn(diff(X)), where X is a vector which represents the time series data of a vital sign, and diff function computes the lagged differences between the previous and next value; sgn function returns a vector with the signs of the corresponding elements of X. The sign of a real number is 1, 0, or -1 if the number is positive, zero, or negative, respectively.

For each vital sign measurement, we calculated the mean, variance, minimum value, maximum value, range, and number of fluctuations of the time-series data in the first 24 hours. Multiple variable ANOVA analyses were implemented to test if any feature is significant for each vital sign. Table III shows the p-value of the two-way ANOVA analysis for the features from each specific vital sign. The number of fluctuations of respiratory rate, heart rate, DBP, and MAP all have significant differences between expired and discharged patients. The prediction based on such features could provide clinicians timely information about a patient’s condition for decision making.

### IV. Recursive Model Training for Vital Feature Selection

When Houston Methodist started to accept COVID-19 patients in late February 2020, we obtained the patients’ dataset in a timely manner and updated our patient’s dataset every week, during peak period every one or two days. As we were augmented with more and more data, and after all the feature analysis from which we discovered many significant features that may impact patients’ outcomes, we initiated development of a feature-based classification model for death prediction. This is an urgent need from the clinical frontline to help clinical decision making, including assessment of the severity, death likelihood, and resource allocation. This model serves to assist clinicians on the frontline in the battle against COVID-19 to make important clinical decisions in assessment of severity and resource allocations.

We trained a predicting model inserting all the extracted vital sign features and basic demographics features for deceased patients since mid-April 2020. Specifically, these features correspond to the significant features shown in Tables II and III. Every week when we get the latest week’s patients’ data, we retrain our model.

| TABLE I | P VALUES OF ONE-WAY ANOVA ANALYSIS FOR COMORBIDITIES AND COMPILATIONS SORTED BY THEIR P VALUES IN THE ASCENDING ORDER |
|---|---|
| No. | Discharged (n = 814) | Expired (n = 86) | P Value |
| Comorbidities | | | |
| Dementia | 71 (8.72%) | 30 (34.88%) | 1.25E-13 |
| Chronic Renal Disease | 174 (21.38%) | 43 (50.00%) | 2.65E-09 |
| Coronary Artery Disease | 190 (23.34%) | 45 (52.33%) | 4.43E-09 |
| Chronic Heart Failure | 138 (16.95%) | 36 (41.86%) | 2.10E-08 |
| COPD | 121 (14.86%) | 33 (38.37%) | 2.95E-08 |
| Diabetes Type 2 | 341 (41.89%) | 58 (67.44%) | 5.22E-06 |
| Diabetes Mellitus | 341 (41.89%) | 57 (66.28%) | 1.38E-05 |
| Hypertension | 536 (65.85%) | 75 (87.21%) | 5.17E-05 |
| History of Flu | 29 (3.56%) | 11 (12.79%) | 7.46E-05 |
| Former Smoker | 146 (17.94%) | 29 (33.72%) | 0.000424 |
| Peripheral Vascular Disease | 79 (9.71%) | 19 (22.09%) | 0.000441 |
| Diabetes Type 1 | 29 (3.56%) | 9 (10.47%) | 0.002440 |
| Chronic Lung Disease | 203 (24.94%) | 34 (39.53%) | 0.003437 |
| Hypertensive heart disease | 88 (10.81%) | 18 (20.93%) | 0.005593 |
| End Stage Renal Failure | 51 (6.27%) | 12 (13.95%) | 0.007837 |
| Cancer | 70 (8.60%) | 12 (13.95%) | 0.101042 |
| Current Smoker | 74 (9.09%) | 4 (4.65%) | 0.164361 |
| Atrial Flutter | 20 (2.46%) | 3 (3.49%) | 0.564842 |
| Chronic Liver Disease | 13 (1.60%) | 1 (1.16%) | 0.757267 |

| Complications No. (%) | | | |
| --- | --- | --- | |
| Shock | 97 (11.92%) | 60 (69.77%) | 1.13E-45 |
| Septic Shock | 82 (10.07%) | 54 (62.79%) | 2.29E-42 |
| Acute Respiratory failure | 298 (36.61%) | 78 (90.70%) | 3.04E-23 |
| Atrial fibrillation | 89 (10.93%) | 36 (41.86%) | 1.09E-15 |
| Acute Respiratory Distress Syndrome | 171 (21.01%) | 46 (53.49%) | 1.26E-11 |
| Disseminated Intravascular Coagulopathy | 7 (0.86%) | 8 (9.30%) | 4.51E-09 |
| Immunocompromised | 394 (48.40%) | 66 (76.74%) | 4.98E-07 |
| Myocardial Infarction | 71 (8.72%) | 21 (24.42%) | 4.43E-06 |
| Acute Respiratory Failure | 37 (4.55%) | 12 (13.95%) | 0.000247 |
| Hyperlipidemia | 64 (7.86%) | 16 (18.26%) | 0.008294 |
| Other Arrhythmia | 71 (8.72%) | 15 (17.44%) | 0.008865 |
| Cerebrovascular Accident Stroke | 67 (8.23%) | 14 (16.28%) | 0.013100 |
| Severe Sepsis | 64 (7.86%) | 10 (11.63%) | 0.227149 |
| Transient Ischemic Attack | 20 (2.46%) | 1 (1.16%) | 0.450143 |
| Renal Insufficiency | 35 (4.30%) | 5 (5.81%) | 0.517508 |

| TABLE II | DEMOGRAPHICS OF 900 PATIENTS COMPARED TO THOSE DISCHARGED WITH THOSE DECEASED PATIENTS |
|---|---|
| No. | Discharged (n = 814) | Expired (n = 86) | P Value |
| Characteristic | | | |
| Age, mean (SD) | 57.6 (17.1) | 74.6 (12.6) | < 10^-15 |
| Sex, No. (%) | | | 0.075 |
| Male (n = 436) | 386 (47.4) | 50 (58.1) | |
| Female (n = 464) | 428 (52.6) | 36 (41.9) | |
| Race, No. (%) | | | |
| Asian | 48 (5.9) | 9 (10.5) | 0.04523 |
| African American | 271 (33.3) | 17 (19.8) | |
| Caucasian | 456 (56.0) | 55 (64.0) | |
| Others | 39 (4.8) | 5 (5.8) | |

| TABLE III | P VALUE OF TWO-WAY ANOVA ANALYSIS FOR EACH VITAL SIGN |
|---|---|
| Respiratory Rate | Mean | Standard | No. | Range (max - min) | Min. |
| | | Variation | Fluctuations | | |
| Heart Rate | 0.06765 | 0.004 | 2.31*10^-4 | 6.21*10^-12 | 0.69524 |
| DBP | 3.11*10^-6 | 9.41*10^-6 | 2.11*10^-4 | < 10^-12 | 0.94326 |
| SBP | 9.91*10^-4 | 0.003 | 0.05168 | < 10^-16 | 0.07552 |
| SpO2 | 2.61*10^-6 | 0.006 | 0.12821 | < 10^-16 | 0.20961 |
| MAP | 4.61*10^-10 | 5.11*10^-10 | 3.31*10^-12 | < 10^-16 | 2.45*10^-8 |
| Temperature | 0.012 | 0.47454 | 0.41488 | 5.91*10^-4 | 0.34221 |
model using all the previous data and test the retrained model on the latest week’s patients’ data. That is, we used the data from February to April 9, 2020 to train the first iteration model and then tested it using the data of patients discharged or expired between April 10th to April 16th, and so on. The model was trained and tested recursively eight times through June 3, 2020.

A Lasso regression model was employed to develop the classification model, as the Lasso regression model helped to find the most contributive variables resulting in optimal model performance and shrinking the coefficients of the other variables toward zero. Cross-validation method was used to establish the penalization parameter Lambda. We choose the value of lambda that minimizes the cross-validation error of the model and obtain the model coefficients at that value of lambda. Table S.1 in the supplement shows the Lambda chosen and the performance with Area Under the Receiver Operating Characteristics (AUC) of each recursive training on the testing data. During the eight rounds of recursive training, we investigated the trend of every feature’s coefficient value. If the coefficient of a feature had a non-zero value in one iteration, it means that this feature has been kept within the Lasso model for this iteration. We realize that if the coefficients were kept relatively stable, then insight into which vital feature is important and stable for patient death prediction can be obtained. If a feature has a non-zero coefficient in more than four iterations, we consider that it is important. Fig. 2 shows the trends of the coefficients of the eight most important vital features through the recursive training. In addition, after two rounds of recursive training, the trends of their coefficients maintained the relative stability until the end.

V. RECURSIVE MODEL COMPARISON WITH DIFFERENT GROUPS OF FEATURES

Once we have patients’ demographic data, comorbidities conditions prior to admission, complications during stay in hospital for COVID-19, and the extracted vital features within 24 hours, we want to know the prediction power of each feature group for deceased patients. Five Lasso logistic regression models were trained and employed by using different feature groups to compare their prediction performance. The first baseline model (model 1) was trained applying demographic features of age, race, and sex shown in Table II. Model 2 employed the previous demographic features in conjunction with significant comorbidity features shown in Table I. Model 3 included demographic features and comorbidities in Model 2, plus significant complication features shown in Table I. Model 3 incorporated features unavailable at the first 24 hours. Nevertheless, this model can be treated as a benchmark and compared with other models using the first 24 hours vital features to validate their prediction performance. In the fourth model, we used significant demographic features and complication features like previous models, along with significant extracted vital sign features shown in Table III. The last model (Model 5) added one more parameter to the fourth model: the number of SpO2 bad days. The definition for SpO2 “Bad Days” was set as “daily range of SpO2 equal to or greater than 10%.”

As described in the previous section, we used Lasso Logistic Regression model instead of Logistic Regression as there is abundant information in the features and only the most important features are used for prediction. As in the previous section, Lambda in the Lasso model selects the value that minimizes the cross-validation error of the model. The model tells us which group of features were critical for deceased patients. AUC with 95% confident interval of each model was calculated. The 95% CI of AUC was computed with 2000 stratified bootstrap using R package ‘pROC’.

As shown in the Fig. 3, the baseline model (Model 1) shown in the black curve using the demographic feature obtains an AUC of 0.776 (95% CI, 0.694-0.857), the performance would improve to 0.792 (95% CI, 0.704-0.881) after adding the comorbidities (model 2, brown curve). Model 3 demonstrates that the performance of using patient demographics and all the comorbidities and complications information achieved an AUC of 0.907 (95% CI: 0.861-0.953).

Table S.2 in the supplement shows the key features selected by the Lasso model. Complications that contribute most to the Model 3 are seven acute diseases including kidney failure and shock. However, even though a complication demonstrates strong prediction power in the Lasso logistic regression model, the prediction may not have practical implication, as it is often
COVID-19 P..distress syndrome (ARDS), has been described in up to 20%
manifested by fever and pneumonia, leading to acute respiratory
are still trying to understand emerging aspects of the disease,
repurposed to fight the coronavirus disease. The physicians
no available vaccine, known drugs for other indications are
Syndrome (CRS)
Treatment for Treating Cytokine Release
the 99 days covered by our dataset, a tolerable increase according
and 7 of these 8 patients had at least one SpO2 “Bad Day,”
in our dataset, only 8 expired without being admitted to ICU,
the hospital would avoid situations where patients die without
ventilation and intubation. As long as there are enough ICU beds,
care. Severely affected patients would receive intensive moni-
tions in predicting patient death. With one more feature
of SpO2 in Model 5, the performance is significantly better than
Model 3. Vital sign features that can predict patient death before
the signs of complications is beneficial and critical for reducing
mortality risk of severely affected patients.

VI. SpO2 PATTERN RECOGNITION CAN IMPROVE DECISION MAKING IN COVID-19 PATIENT CARE

A. Most Expired Patients Not Admitted Into ICU Had SpO2 “Bad Days”

ICU admission is a critical decision during COVID-19 patient
care. Severely affected patients would receive intensive moni-
toring and access to intervention procedures such as mechanical
ventilation and intubation. As long as there are enough ICU beds,
the hospital would avoid situations where patients die without
ever being admitted to the ICU. Among the 86 expired patients
in our dataset, only 8 expired without being admitted to ICU,
and 7 of these 8 patients had at least one SpO2 “Bad Day,”
including the one who expired on the same day of admission.
If every patient who had at least one SpO2 “Bad Day” were to be
admitted to ICU, there would be 82 more ICU admissions over
the 99 days covered by our dataset, a tolerable increase according
to the ICU capacity of the hospitals in our health system.

B. SpO2 Patterns Had the Potential of Helping Improve the Timing of Administering Monoclonal Antibody Treatment for Treating Cytokine Release Syndrome (CRS)

In the portfolio of all the countermeasures, in the setting of
no available vaccine, known drugs for other indications are
repurposed to fight the coronavirus disease. The physicians
are still trying to understand emerging aspects of the disease,
like the cytokine release syndrome[24]–[26]. Severe COVID-19
manifested by fever and pneumonia, leading to acute respiratory
distress syndrome (ARDS), has been described in up to 20%
of COVID-19 cases. CRS is an exaggerated immune response
to the SARS-CoV2 virus. It is characterized by the release of
high levels of pro-inflammatory cytokines such as Interleukin-
6 (IL-6) that instigates an amplification cascade resulting in
lymphocytic changes and differentiation, enhanced vascular
permeability, and increased neutrophilic and monocytic recruit-
ment in lung parenchyma. Inflammatory injury to the alveolar
capillary barrier, with extravasation of protein-rich edema fluid
into the airspace contributes to the pathophysiology of ARDS
in COVID-19 patients with resultant hypoxemia and drop in
SpO2. CRS affects other organ systems leading to hypotension
and shock. Given this experience, urgently needed therapeutics
based on profound immunosuppression, such as tocilizumab,
a monoclonal antibody to IL-6 that inhibits cellular signal trans-
duction and CRS, has been approved by FDA for treatment of
COVID-19 hospitalized adults and pediatric patients (2 years of
age and older), receiving systemic corticosteroids and requiring
supplemental oxygen, mechanical ventilation, or extracorporeal
membrane oxygenation (ECMO).

One challenge for applying tocilizumab is the timing of
starting the drug. CRS is diagnosed based on routinely checked
inflammatory markers from blood tests, and there have been
extensive discussions regarding modifying the treatment algo-
rithms so that the drug can be applied to a reasonably larger
population at an earlier stage of CRS. Table S.4 shows that
among the users of tocilizumab, those who had never been to ICU
and those starting the drug before ICU admission had more than a
two-fold lower ratio of death compared to those starting the drug
within one week of ICU admission, while both patients starting
the drug more than 10 days after ICU admissions expired.

VII. DISCUSSION AND CONCLUSION

Taking advantage of the comprehensive clinical datasets of
COVID-19 patients available in the Houston Methodist ME-
TEOR data warehouse, we extracted powerful features from
time series vital signs and developed a time-series feature-based
classification Lasso Logistic model to differentiate 86 expired
and 814 discharged COVID-19 patients during the period be-
tween February 26, 2020 and June 3, 2020. The time-series
feature-based classification model accurately evaluated and
stratified patient status of a cohort of 900 COVID-19 in-patients
and provided evidence to help distribute valuable medical re-
sources, e.g., ICU beds, ventilators, or treatments for certain
symptoms.

In the presented Lasso logistic regression model, advanced
features concerning the pattern changes were extracted from
time-lapse profile of vital signs and the concept of SpO2 “Bad
Days” was introduced based on the daily range of SpO2 read-
ings. Our research findings indicate that SpO2 patterns not
only aided to differentiate expired and discharged patients, but
also indicated the potential of the utilization of the model in
clinical decision-making regarding ICU admissions and timing
of starting certain known drugs such as tocilizumab.

In the presented Lasso logistic regression model, all other
features kept in the model are the extracted features from
time-series vital signs. Most of the extracted vital sign features
except “SpO2 bad days” are from the first 24 hours of patient
admission, which leads us to hypothesize that the time series
vital signs can give us an early warning signal that the patient’s condition may deteriorate downwards to a severely ill state. Such information is currently hidden behind the trend of the vital signs extended over time. The earlier that the warning signal we can capture, the more resources and treatment can be deployed to the COVID-19 patient who needs them the most.

Several machine learning models have been proposed to predict the outcome of COVID-19 patients [27]–[31]. However, most of the key features of these models are lab values and may not have the ability to provide early COVID-19 mortality prediction. Instead, the uniqueness of our model is that it is based on just the first 24 hours of vital data such that clinical interventions can be decided early and applied effectively.

In terms of the performance of these models, most of their AUCs are between 90% and 95%, to which the performance of our model is comparable. Karthikeyan et al. [31] proposed a model that performed with AUC of 99% when predicted at the day of outcome but performed worse when it predicted far from the day of outcome. This demonstrated the difficulty of early COVID-19 mortality prediction. The novelty of our model is that it achieves comparable predictive performance with published models but as an early mortality indicator by focusing on the vital signs signal instead of lab test results.

Our ongoing work is evaluating the performance of the model to include “SpO2 bad days” earlier to give an early indicator; studying the hidden relationship between the extracted features of time-series vital signs and the critical complications like Cytokine release syndrome to optimize clinical decisions, including medical resource allocations, ICU admission, ventilation and intubation, timing for drug administration, and length of stay, with the ultimate goal of reducing the mortality rate of hospitalized COVID-19 patients.

ACKNOWLEDGMENT

The authors would like to thank Rebecca L. Danforth for her suggestions and proofreading work on this paper and the ICU clinical staff at Houston Methodist.

REFERENCES

[1] “WHO timeline for COVID-19,” 2020. [Online]. Available: https://www.who.int/news-room/detail/27-04-2020-who-timeline-covid-19
[2] W. Sungnak et al., “SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes,” Nature Med., vol. 26, no. 5, pp. 681–687, May 2020, doi: 10.1038/s41591-020-0868-6.
[3] D. Bojkova et al., “Proteomics of SARS-CoV-2-infected host cells reveals therapy targets,” Nature, vol. 583, pp. 469–472, May 2020, doi: 10.1038/s41551-020-0458-9.
[4] D. E. Gordon et al., “A SARS-CoV-2 protein interaction map reveals targets for drug repurposing,” Nature, vol. 583, pp. 459–468, Apr. 2020, doi: 10.1038/s41586-020-2286-9.
[5] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, “Automated detection of COVID-19 cases using deep neural networks with X-ray images,” Comput. Biol. Med., vol. 121, pp. 103792–103792, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.
[6] S. Wang et al., “A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis,” Eur Respir J., vol. 56, no. 2, Aug. 2020, Art. no. 2000775, doi: 10.1183/13993003.00775-2020.
[7] M. A. Elaziz, K. M. Hosny, A. Salah, M. M. Darwish, S. Lu, and A. T. Sahlo, “New machine learning method for image-based diagnosis of COVID-19,” PLoS One, vol. 15, no. 6, Jun. 2020, Art. no. e0235187, doi: 10.1371/journal.pone.0235187.
[8] J. H. Beigel et al., “Remdesivir for the treatment of COVID-19—Preliminary report,” N. Engl. J. Med., vol. 383, no. 19, pp. 1813–1836, May 2020, doi: 10.1056/NEJMoa2007640.
[9] J. Grein et al., “Compassionate use of remdesivir for patients with severe COVID-19,” N. Engl. J. Med., vol. 382, no. 24, pp. 2327–2346, Apr. 2020, doi: 10.1056/NEJMoa2007016.
[10] J. Geleris et al., “Observational study of hydroxychloroquine in hospitalized patients with COVID-19,” N. Engl. J. Med., vol. 382, no. 25, 2411–2418, May 2020, doi: 10.1056/NEJMoa2012410.
[11] “Regional hospital capacity calculator by Harvard Global Health Institute and the Harvard TH Chan School of Public Health,” 2020. [Online]. Available: https://globalepidemics.org/hospital-capacity/
[12] “Qventus: Localized COVID-19 model and scenario planner,” 2020. [Online]. Available: https://covid19.qventus.com/modeling/hospitals
[13] “Hospital resource calculator for COVID-19 by rush university center for quality, safety & value analytics,” 2020. [Online]. Available: http://covid19forecast.rush.edu
[14] “CHIME: The COVID-19 hospital impact model for epidemics from penn medicine,” 2020. [Online]. Available: http://predictivehealthcare.pennmedicine.org/2020/03/14/accounting-chime.html
[15] “Health catalyst the capacity planning tool,” 2020. [Online]. Available:https://www.healthcatalyst.com/covid-19-cappland
[16] “Hospital bed projections tool from Stanford Medicine,” 2020. [Online]. Available: https://surf.stanford.edu/covid-19-tools/covid-19-hospital-projections/
[17] “Critical care surge response strategies tool from RAND corporation,” 2020. [Online]. Available: https://www.rand.org/pubs/tools/TLA164-1.html
[18] “COVID-19 forecasting tool from the University of Washington’s Institute for health metrics and evaluation,” 2020. [Online]. Available: https://covid19.healthdata.org/united-states-of-america
[19] J. Friedman et al., “Predictive performance of international COVID-19 mortality forecasting models,” Nat. Commun., vol. 12, 2021, Art. no. 2609, doi: 10.1038/s41467-020-0311-8.
[20] A. S. Yadaw et al., “Proteomics of SARS-CoV-2-infected host cells reveals therapy targets,” Nature, vol. 583, pp. 469–472, May 2020, doi: 10.1038/s41551-020-0458-9.
[21] W. Guan et al., “Clinical characteristics of coronavirus disease 2019 in China,” N. Engl. J. Med., vol. 382, no. 8, pp. 1708–1720, Feb. 2020, doi: 10.1056/NEJMoa200232.
[22] M. Z. Tay, C. M. Poh, L. Reina, P. A. MacAry, and L. F. P. Ng, “The tragedy of COVID-19: Immunity, inflammation and intervention,” Nature Rev. Immunol., vol. 20, no. 6, pp. 363–374, Jun. 2020, doi: 10.1038/s41577-020-0311-8.
[23] J. Friedman et al., “Covid-19 mortality risk assessment: An international multi-center study,” PLoS One, vol. 15, 2020, Art. no. e0243262.
[24] A. S. Yadaw et al., “Clinical features of COVID-19 mortality: A multi-center study,” Lancet Digit. Health, vol. 2, no. 10, pp. e516–e525, 2020, doi: 10.1016/S2589-7500(20)30217-X.
[25] N. Barda et al., “Developing a COVID-19 mortality risk prediction model when individual-level data are not available,” Nat. Commun., vol. 11, 2020, Art. no. 4439, doi: 10.1038/s41467-020-18297-9.
[26] D. Bertsimas et al., “From predictions to prescriptions: A data-driven decision support system for early COVID-19 mortality prediction,” Front. Public Health, vol. 9, May 2021, Art. no. 26697, doi: 10.3389/fpubh.2021.626697.