MAYO CLINIC

ORIGINAL ARTICLE

Personalized Prediction of Hospital Mortality in COVID-19–Positive Patients

Daniel Rozenbaum, MD; Jacob Shreve, MD, MS; Nathan Radakovich, MD; Abhijit Duggal, MD, MPH, MSc; Lara Jehi, MD, MHCDS; and Aziz Nazha, MD

Abstract

Objective: To develop predictive models for in-hospital mortality and length of stay (LOS) for coronavirus disease 2019 (COVID-19)–positive patients.

Patients and Methods: We performed a multicenter retrospective cohort study of hospitalized COVID-19–positive patients. A total of 764 patients admitted to 14 different hospitals within the Cleveland Clinic from March 9, 2020, to May 20, 2020, who had reverse transcriptase-polymerase chain reaction–proven coronavirus infection were included. We used LightGBM, a machine learning algorithm, to predict in-hospital mortality at different time points (after 7, 14, and 30 days of hospitalization) and in-hospital LOS. Our final cohort was composed of 764 patients admitted to 14 different hospitals within our system.

Results: The median LOS was 5 (range, 1-44) days for patients admitted to the regular nursing floor and 10 (range, 1-38) days for patients admitted to the intensive care unit. Patients who died during hospitalization were older, initially admitted to the intensive care unit, and more likely to be white and have worse organ dysfunction compared with patients who survived their hospitalization. Using the 10 most important variables only, the final model's area under the receiver operating characteristics curve was 0.86 for 7-day, 0.88 for 14-day, and 0.85 for 30-day mortality in the validation cohort.

Conclusion: We developed a decision tool that can provide explainable and patient-specific prediction of in-hospital mortality and LOS for COVID-19–positive patients. The model can aid health care systems in bed allocation and distribution of vital resources.

 ■ Mayo Clin Proc Inn Qual Out 2021;5(4):795-801

Despite several international and local efforts, the coronavirus pandemic caused by the severe acute respiratory syndrome coronavirus 2 has infected more than 122 million individuals worldwide, and more than 2.7 million people have died to date.[1] The pandemic is far from over, with increasing new cases in several parts of the United States and the world. Consequently, health care systems continue to face several challenges regarding bed availabilities/allocations and resource use.

Whereas some infected patients can be asymptomatic, others can experience severe respiratory distress syndrome, multiorgan failure, and death. Thus, identifying patients with higher risk for early mortality during their hospitalization could aid hospitals and health care providers in predicting the disease trajectory, distributing vital resources efficiently, and consequently improving patients' outcomes.

We developed a clinical decision tool that uses clinical and demographic variables within 24 hours of hospitalization to provide personalized predictions of patient mortality and length of stay (LOS) that are specific for a given patient.

PATIENTS AND METHODS

All patients admitted to our health care system from March 9, 2020, to May 20, 2020, who had reverse transcriptase-polymerase chain reaction–proven coronavirus disease 2019 (COVID-19) infection were included in our database (n=962). We excluded patients: (1) who had not been discharged or died by May 21, 2020 (n=103), (2) for whom discharge disposition was unknown (due to missing information or transfer to another hospital; n=89), and (3) who were younger than 18 years (n=6). Our final cohort was composed of 764 patients admitted to 14 different hospitals within our system. The

From the Department of Hematology and Medical Oncology, Cleveland Clinic (D.R., J.S., A.N.); Lerner College of Medicine (N.R., A.N.); Department of Critical Care, Cleveland Clinic (A.D.); Neurological Institute and Lerner College of Medicine (L.J.); and Center for Clinical Artificial Intelligence, Cleveland Clinic, Cleveland, OH (A.N.).

TABLE. Patients' Characteristics[a,b]

Characteristic	All Patients (n = 764)[c]	Death or Hospice (n=116)	Survived (n=648)	*P*
Demographic characteristic				
Race, no. (%)				
White	433 (56.7)	82 (70.7)	351 (54.2)	.001
African American	277 (36.3)	30 (25.9)	247 (38.1)	.02
Asian	10 (1.3)	1 (0.9)	9 (1.4)	>.99
Multiracial	28 (3.7)	1 (0.9)	27 (4.2)	.11
Ethnicity, no. (%)				
Non-Hispanic	705 (94.8)	109 (96.5)	596 (94.5)	.51
Hispanic	39 (5.2)	4 (3.5)	35 (5.5)	
Age (y), median (Q1, Q3)	64 (53, 76)	80 (72, 84)	62 (52, 73)	<.001
Sex, no. (%)				
Female	366 (47.9)	57 (49.1)	309 (47.7)	.85
Male	398 (52.1)	59 (50.9)	339 (52.3)	.85
Body mass index (kg/m^2), median (Q1, Q3)	30.1 (25.9, 35.4)	30.3 (26.5, 35.6)	28.6 (22.9, 32.7)	<.001
Previous medical history, no. (%)				
Chronic obstructive pulmonary disease	95 (13.5)	17 (16.2)	78 (13.0)	.47
Asthma	156 (22.1)	16 (15.1)	140 (23.3)	.08
Diabetes	284 (39.9)	50 (46.3)	234 (38.7)	.17
Hypertension	528 (72.3)	96 (83.5)	432 (70.2)	.005
Coronary artery disease	152 (21.6)	44 (40.4)	108 (18.1)	<.001
Heart failure	139 (19.6)	44 (40.0)	95 (15.9)	<.001
Any cancer	142 (19.4)	35 (31.5)	107 (17.3)	.001
Laboratory parameters, median (Q1, Q3)				
Metabolic indexes				
Sodium (mEq/L)	137.0 (134.0, 139.0)	138.0 (134.0, 141.0)	137.0 (134.0, 139.0)	.02
Potassium (mEq/L)	4.0 (3.7, 4.4)	4.2 (3.8, 4.5)	4.0 (3.7, 4.3)	<.001
Creatinine (mg/dL)	1.0 (0.8, 1.4)	1.6 (1.1, 2.3)	1.0 (0.8, 1.3)	<.001
Lactate (mg/dL)	1.4 (1.0, 1.8)	1.5 (1.2, 2.1)	1.3 (1.0, 1.8)	.02
Hepatic indexes				
Alanine aminotransferase (U/L)	24.0 (15.0, 40.0)	27.0 (15.0, 41.0)	23.0 (15.0, 39.0)	0.38
Aspartate aminotransferase (U/L)	34.0 (24.0, 52.0)	43.0 (32.0, 79.0)	32.0 (23.0, 49.0)	<.001
Total bilirubin (mg/dL)	0.4 (0.3, 0.6)	0.5 (0.3, 0.7)	0.4 (0.3, 0.6)	.05
Alkaline phosphatase (U/L)	72.0 (57.5, 94.5)	82.0 (63.5, 104.0)	71.0 (57.0, 93.2)	.01
Albumin (g/dL)	3.7 (3.4, 4.0)	3.4 (3.0, 3.8)	3.7 (3.4, 4.0)	<.001
Hematologic indexes				
Hemoglobin (g/dL)	13.1 (11.6, 14.5)	11.9 (9.9, 13.8)	13.3 (11.9, 14.6)	<.001
White blood cell count (k/μL)	6.4 (4.8, 8.5)	7.7 (5.4, 10.9)	6.3 (4.8, 8.2)	<.001
Platelet count (k/μL)	207.0 (160.0, 267.0)	198.5 (144.2, 245.2)	209.0 (163.0, 268.0)	.04
Coagulation indexes				
International normalized ratio	1.0 (1.0, 1.1)	1.1 (1.0, 1.2)	1.0 (1.0, 1.1)	.04
Partial thromboplastin time (s)	29.6 (27.1, 33.4)	30.8 (27.0, 33.7)	29.4 (27.1, 33.2)	.50
D-Dimer (ng/mL)	840.0 (490.0, 1615.0)	1470.0 (825.0, 3380.0)	780.0 (470.0, 1390.0)	<.001
Inflammatory indexes				
Lactate dehydrogenase (U/L)	299.0 (229.8, 401.0)	400.0 (308.0, 531.0)	288.0 (223.5, 366.5)	<.001
C-Reactive protein (mg/dL)	6.5 (3.0, 12.2)	11.9 (5.7, 17.5)	5.9 (2.5, 11.3)	<.001
Procalcitonin (ng/mL)	0.1 (0.1, 0.4)	0.3 (0.2, 1.4)	0.1 (0.1, 0.3)	<.001
Ferritin (ng/mL)	511.4 (255.3, 1009.2)	852.9 (351.9, 1747.5)	485.5 (235.1, 893.2)	<.001
Cardiac enzymes				
Troponin T (ng/mL)	0.0 (0.0, 0.1)	0.1 (0.0, 0.2)	0.0 (0.0, 0.1)	.06
Creatine kinase (U/L)	135.0 (69.5, 297.0)	242.0 (105.0, 753.0)	115.0 (65.8, 228.2)	.001
Treatment-related variables, no. (%)				
Intensive care unit on admission	147 (19.2)	48 (41.4)	99 (15.3)	<.001

Continued on next page

Mayo Clin Proc Inn Qual Out ■ August 2021;5(4):795-801 ■ https://doi.org/10.1016/j.mayocpiqo.2021.05.001
www.mcpiqojournal.org

TABLE. Continued

Characteristic	All Patients (n = 764)[c]	Death or Hospice (n=116)	Survived (n=648)	*P*
Laboratory parameters, median (Q1, Q3), continued				
Need for noninvasive mechanical ventilation	96 (12.6)	34 (29.3)	62 (9.6)	<.001
Mechanical ventilation on d 1	74 (9.7)	34 (29.3)	40 (6.2)	<.001
Mechanical ventilation during stay	133 (17.4)	59 (74.7)	74 (27.4)	<.001
Use of hydroxychloroquine	293 (52.6)	39 (48.1)	254 (53.4)	.45
Use of tocilizumab	50 (9.0)	8 (9.9)	42 (8.8)	.92
New use of steroids	94 (12.3)	32 (27.6)	62 (9.6)	<.001

[a]Q, quartile.

[b]SI conversion factors: To convert sodium and potassium values to mmol/L, multiply by 1.0; to convert creatinine values to μmol/L, multiply by 88.4; to convert lactate values to mmol/L, multiply by 0.111; to convert total bilirubin values to μmol/L, multiply by 17.104; to convert albumin and hemoglobin values to g/L, multiply by 10; to convert white blood cell values to $\times 10^9$/L, multiply by 1; to convert platelet values to $\times 10^9$/L, multiply by 1; to convert D-dimer values to nmol/L, multiply by 5.476; to convert C-reactive protein values to mg/L, multiply by 10; to convert ferritin values to μg/L, multiply by 1; to convert troponin T values to μg/L, multiply by 1.0.

[c]For the categorical variables, percentages are calculated out of non-missing data points instead of out of 764.

study was approved by the Cleveland Clinic Institutional Review Board and conducted in accordance with the Declaration of Helsinki.

Data Set and Outcomes Definition

For each patient, demographic, clinical, and laboratory variables (109 variables; Supplemental Table, available online at https://mcpiqojournal.org) were included and structured from the electronic health care record. All variables were collected within the first 24 hours of hospitalization. Twenty-two percent of our data was missing, mostly because some laboratory tests were not ordered on the day of admission or the test was not ordered at all for the patient. Missing data were handled by the built-in algorithm from the machine learning model used in our analysis.

The main outcomes evaluated were mortality at 7, 14, and 30 days of hospitalization and hospital LOS, which was defined as the time between hospitalization and death or discharge from the hospital. We also built a model for prediction of intensive care unit (ICU) transfer (or death before ICU transfer) among patients admitted to the regular nursing floor.

Statistical Analyses

To ensure that all variables are treated equally regardless of their significance in univariate analysis and to account for the variables that can be significant only in the context of other variables, we used a machine learning model, Light Gradient Boosting Machine (LightGBM),[2] in our analysis. LightGBM is a model based on the gradient boosting framework. In gradient boosting, models with weak predictive capability, such as decision trees, are used together to achieve high predictive performance. During training of a gradient boosting model, decision trees are created using the available variables to separate instances belonging to different classes (eg, survivors vs nonsurvivors). These decision trees are created in a sequential fashion to minimize the prediction errors made by the previous trees. When facing a new case, the model will use the framework of decision trees created during training to classify the new example.

The data set was divided randomly into training (80.0%) and test (20.0%) sets and the models were initially trained with all our variables. The most influential 10 variables as determined by the values originated by the SHapley Additive exPlanations (SHAP) algorithm[3] (an algorithm that is widely used to determine the most important variables that affected a model's decision) were extracted. These were ranked from the most to the least important variable and used to fit reduced clinically usable versions of our models.

Hyperparameter optimization using a Bayesian optimization algorithm was obtained to ensure that the most robust models were used, and 10-fold cross-validation was also used to ensure the reproducibility of the final models. Model performance in the validation

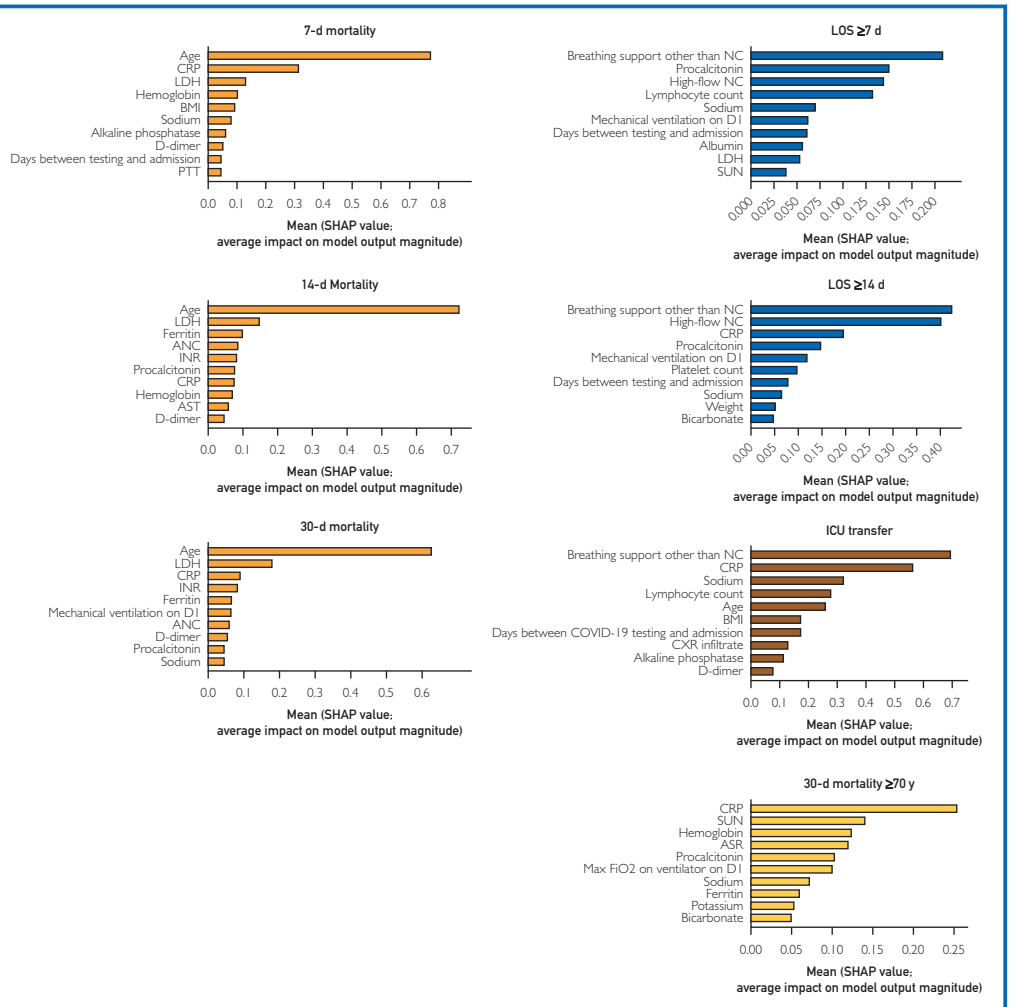

FIGURE1. Ten most important variables for each model. Bar plots show the 10 most important variables for each model based on their SHapley Additive exPlanations (SHAP) values (values generated using the SHAP algorithm indicating how much a variable contributed to the model's decisions). ANC, absolute neutrophil count; AST, aspartate aminotransferase; BMI, body mass index; CK, creatinine kinase; COVID-19, coronavirus disease 2019; CRP, C-reactive protein; CXR, chest radiograph; D1, day 1; ICU, intensive care unit; INR, international normalized ratio; LDH, lactate dehydrogenase; LOS, length of stay; NC, nasal cannula; PTT, partial thromboplastin time; SUN, serum urea nitrogen.

set is reported using the area under the receiver operating characteristics curve (ROC AUC).

RESULTS

Patient Population

Among the 764 patients included in the analysis, 116 (15.2%) either died (n=87) or were transitioned to hospice care (n=29). The median age was 64 (range, 19-98) years and 147 patients (19.2%) were admitted directly to the ICU. The median LOS was 5 (range, 1-44) days for patients admitted to the regular nursing floor and 10 (range, 1-38) days for patients admitted to the ICU. The Table summarizes the clinical characteristics of our cohort. As expected, patients who died during their hospitalization were older, were more likely to be initially admitted to the ICU, and had worse organ dysfunction and inflammatory biomarker levels compared with patients who survived their

Mayo Clin Proc Inn Qual Out ■ August 2021;5(4):795-801 ■ https://doi.org/10.1016/j.mayocpiqo.2021.05.001
www.mcpiqojournal.org

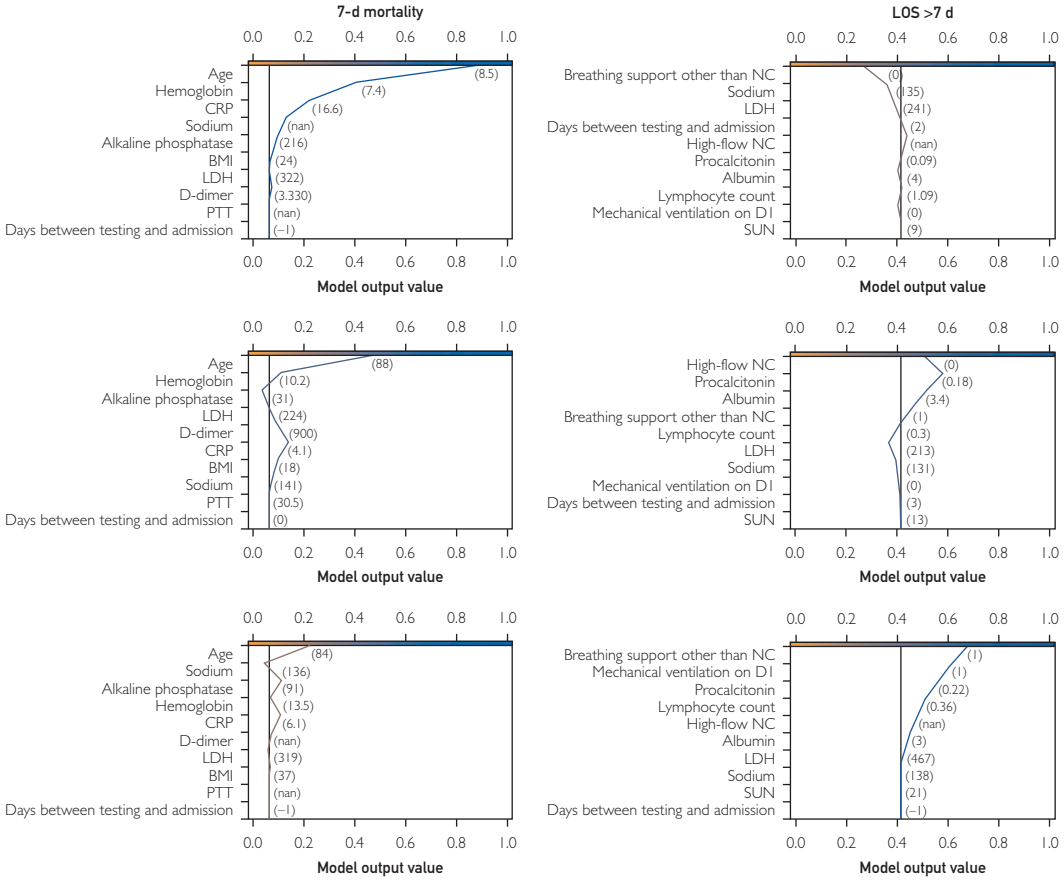

FIGURE 2. Personalized prediction of mortality and length of stay (LOS). Decision plots show how the probability of the outcome (7-day mortality on the left and LOS >7 days on the right) shifts as each variable is considered for 3 different patients on each side. The starting point in the bottom of each graph is the pre-test probability (ie, overall percentage of patients who died within 7 days or whose LOS was >7 days). For instance, in the top panel left, the probability of dying goes from about 40% to 90% as the patient's age (of 85 years) is considered by the algorithm. On the left, the 3 patients depicted had similar ages but different outcomes (top 1 died and the other 2 survived), all of which were correctly predicted by the model. On the right, from top to bottom, LOS was 5, 8, and 24 days. BMI, body mass index; CRP, C-reactive protein; D1, day 1; LDH, lactate dehydrogenase; Nan, missing value; NC, nasal cannula; PTT, partial thromboplastin time; SUN, serum urea nitrogen.

hospitalization (Table). Interestingly, men did not have worse outcomes compared with women and African American patients had a lower mortality rate compared with whites in our cohort (Table).

Mortality Models

A total of 109 clinical variables (Supplemental Table) were included in the algorithm to predict mortality after 7, 14, and 30 days of hospitalization. A feature extraction algorithm was used to identify the top 10 variables that affected mortality at each time point. Although variables such as age and lactate dehydrogenase, ferritin, and C-reactive protein levels were shown as important at each time point but at a different level of importance, others such as being treated with a mechanical ventilator in the first 24 hours only affected mortality at 30 days (Figure 1).

Using the top 10 variables only, the final model ROC AUC when applied to the validation cohort was 0.86 for 7-day mortality, 0.88 for 14-day mortality, and 0.85 for 30-day

mortality. The model can provide personalized and explainable prediction for an individual patient (Figure 2).

LOS Model

Using similar methodology, the top 10 variables that affected hospital LOS longer than 7 days and longer than 14 days are shown in Figure 1. Using these variables, the final model ROC AUC was 0.80 for LOS longer than 7 days and 0.82 for LOS longer than 14 days.

Other Outcomes

We also used the same methodology to build a model to predict ICU transfer (or death before ICU transfer) among patients admitted initially to the regular nursing floor. The final model ROC AUC with only the top 10 variables was 0.80. The top 10 clinical variables that affected the risk for ICU transfer as well as 30-day mortality in patients older than 70 years can also be found in Figure 1.

DISCUSSION

In this study, we developed personalized prediction models that use clinical variables within 24 hours of admission to predict mortality and LOS that are specific for COVID-19—infected patients. The proposed models showed robust AUCs in predicting mortality and LOS at different time points during hospitalization. Our models' predictions could alert physicians regarding adverse outcomes for hospitalized patients with COVID-19 infection such as hospital mortality and transfer to the ICU. It can also help hospitals manage a COVID-19 surge by identifying the expected LOS in the hospital and ICU. We also explored the clinical variables that affected these outcomes during hospitalization and showed that although some variables such as age, lactate dehydrogenase level, and ferritin level have a significant impact on mortality at each time point, others such as procalcitonin level can only affect mortality after 14 and 30 days. More importantly, our models can provide an explainable prediction that is specific for a given patient. This explainability will allow physicians to understand the significant clinical variables that affected their patients' outcomes.

Several studies have evaluated the impact of clinical variables on mortality during hospitalization for patients with COVID-19 infection.[4-7] Although all showed that age and comorbid conditions could affect the outcome, the effect of other clinical variables varies. These differences in the outcomes could be related to the difference in the methodology of conducting the multivariate analyses. Our machine learning model included all clinical variables initially to ensure that all variables are treated equally regardless of their significance in univariate analyses. We then focused on the analysis of the top 10 variables that affected the overall outcomes. Although machine learning models are often viewed as a "black box," our model can provide an explainable output that is specific for a given patient.

Our study has important limitations. First, as a retrospective study importing data from the electronic medical record, a high proportion of missing data is expected. Although missing data will worsen the performance of a prediction algorithm, empirically we were able to verify that the model still had robust performance on our test set (ie, validation cohort). Second, given that each surge may have its own specific characteristics and that all patients came from hospitals within the same health care system, our ability to generalize our findings may be limited to some extent.

CONCLUSION

We built personalized prediction models to predict outcomes for hospitalized patients with COVID-19 infection. The models can aid physicians and health care systems in understating the disease trajectory and expected outcomes for a given patient.

ACKNOWLEDGMENTS

We acknowledge the COVID-19 (coronavirus disease 2019) Enterprise Research Registry Workstream for assistance with data collection. This research was supported by the National Institutes of Health under award number R01NS097719.

SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at https://mcpiqojournal.org. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Abbreviations and Acronyms: **ANC** = absolute neutrophil count; **AST** = aspartate aminotransferase; **BMI** = body mass index; **CK** = creatinine kinase; **COVID-19** = coronavirus disease 2019; **CRP** = C-reactive protein; **CXR** = chest radiograph; **D1** = day 1; **ICU** = intensive care unit; **INR** = international normalized ratio; **LDH** = lactate dehydrogenase; **LightGBM** = Light Gradient Boosting Machine; **LOS** = length of stay; **Nan** = missing value; **NC** = nasal cannula; **PTT** = partial thromboplastin time; **Q** = quartile; **ROC AUC** = area under the receiver operating characteristics curve; **SHAP** = SHapley Additive exPlanations; **SUN** = serum urea nitrogen

Potential Competing Interests: The authors report no competing interests.

Correspondence: Address to Aziz Nazha, MD, Cleveland Clinic Center for Clinical Artificial Intelligence, Lerner College of Medicine/Case Western Reserve University, Department of Hematology and Medical Oncology, Taussig Cancer Institute, Cleveland Clinic, Desk R35, 9500 Euclid Ave, Cleveland, OH 44195 (nazhaa@ccf.org; Twitter: @AzizNazhaMD).

ORCID

Daniel Rozenbaum: https://orcid.org/0000-0002-9478-8585; Lara Jehi: https://orcid.org/0000-0002-8041-6377

REFERENCES

1. World Health Organization. Weekly epidemiological update on COVID-19 - 23 March 2021. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—23-march-2021. Accessed March 28, 2021.
2. Ke G, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree. Paper presented at: Advances in Neural Information Processing Systems 30 (NIPS 2017); December 4-9, 2017; Long Beach, CA. https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree. Accessed October 22, 2020.
3. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. Paper presented at: Advances in Neural Information Processing Systems 30 (NIPS 2017); December 4-9, 2017; Long Beach, CA. https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions. Accessed October 22, 2020.
4. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062.
5. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19 death using OpenSAFELY. *Nature*. 2020;584(7821):430-436.
6. Zhang J, Wang X, Jia X, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. *Clin Microbiol Infect*. 2020;26(6):767-772.
7. Yu C, Lei Q, Li W, et al. Clinical characteristics, associated factors, and predicting COVID-19 mortality risk: a retrospective study in Wuhan, China. *Am J Prev Med*. 2020;59(2):168-175.