Machine Learning Prediction of Death in Critically Ill Patients with COVID-19

Matthew M Churpek, MD, MPH, PhD,1* Shruti Gupta, MD, MPH,2* Alexandra B. Spicer, MS,1 Salim S. Hayek, MD,3 Anand Srivastava, MD, MPH, 4 Lili Chan, MD, MSCR,5 Michal L. Melamed, MD, MHS,6 Samantha K. Brenner, MD, MPH,7 Jared Radbel, MD,8 Farah Madhani-Lovely, MD,9 Pavan K. Bhatraju, MD, MSc,10 Anip Bansal, MD,11 Adam Green, MD, MBA,12 Nitender Goyal, MD,13 Shahzad Shaefi, MD, MPH,14 Chirag R. Parikh, MD, PhD,15 Matthew W. Semler, MD,16 David E. Leaf, MD, MMSc,2 for the STOP-COVID Investigators

*These authors contributed equally.

Online Data Supplement

STOP-COVID INVESTIGATORS

Baylor College of Medicine: Carol P. Walther*, Samaya J. Anumudu

Baylor University Medical Center: Justin Arunthamakun*, Kathleen F. Kopecky, Gregory P. Milligan, Peter A. McCullough, Thuy-Duyen Nguyen

Beth Israel Deaconess Medical Center: Shahzad Shaefi*, Megan L. Krajewski, Sidharth Shankar, Ameeka Pannu, Juan D. Valencia

Boston Medical Center: Sushrut S. Waikar*, Zoe A. Kibbelaar

Cook County Health: Ambarish M. Athavale*, Peter Hart, Oyintayo Ajiboye, Matthew Itteera

Cooper University Health Care: Adam Green*, Jean-Sebastien Rachoin, Christa A. Schorr, Lisa Shea

Duke University Medical Center: Daniel L. Edmonston*, Christopher L. Mosher

Hackensack Meridian Health Mountainside Medical Center: Alexandre M. Shehata*, Zaza Cohen, Valerie Allusson, Gabriela Bambrick-Santoyo, Noor ul aain Bhatti, Bijal Metha, Aquino Williams

Hackensack University Medical Center: Samantha K. Brenner*, Patricia Walters, Ronaldo C. Go, Keith M. Rose

Harvard T.H. Chan School of Public Health: Miguel A. Hernán

Harvard University: Amy M. Zhou, Ethan C. Kim, Rebecca Lisk

Icahn School of Medicine at Mount Sinai: Lili Chan*, Kusum S. Mathews*, Steven G. Coca, Deena R. Altman, Aparna Saha, Howard Soh, Huei Hsun Wen, Sonali Bose, Emily Leven, Jing G. Wang, Gohar Mosoyan, Girish N. Nadkarni

Indiana University School of Medicine/Indiana University Health: Allon N. Friedman*, John Guirguis, Rajat Kapoor, Christopher Meshberger

Johns Hopkins Hospital: Chirag R. Parikh*, Brian T. Garibaldi, Celia P. Corona-Villalobos, Yumeng Wen, Steven Menez, Rubab F. Malik, Carmen Elena Cervantes, Samir C. Gautam, Crystal Chang

Loma Linda University: H. Bryant Nguyen*, Afshin Ahoubim

Mayo Clinical, Arizona: Leslie F. Thomas*

Mayo Clinic, Florida: Pramod K. Guru*

Medical College of Wisconsin: Paul A. Bergl*, Yan Zhou, Jesus Rodriguez, Jatan A. Shah, Mrigank S. Gupta

MedStar Georgetown University Hospital: Princy N. Kumar*, Deepa G. Lazarous, Seble G. Kassaye

Montefiore Medical Center/Albert Einstein College of Medicine: Michal L. Melamed*, Tanya S. Johns. Ryan Mocerino, Kalyan Prudhvi, Denzel Zhu, Rebecca V. Levy, Yorg Azzi, Molly Fisher, Milagros Yunes, Kaltrina Sedaliu, Ladan Golestaneh, Maureen Brogan, Jyotsana Thakkar, Neelja Kumar, Michael J. Ross, Michael Chang

New York-Presbyterian Queens Hospital: Ritesh Raichoudhury*

New York-Presbyterian/Weill Cornell Medical Center: Edward J. Schenck*, Soo Jung Cho, Maria Plataki, Sergio L. Alvarez-Mulett, Luis G. Gomez-Escobar, Di Pan, Stefi Lee, Jamuna Krishnan, William Whalen

New York University Langone Hospital: David Charytan*, Ashley Macina

Northwell Health: Daniel W. Ross

Northwestern Memorial Hospital: Northwestern University Feinberg School of Medicine - Anand Srivastava*, Alexander S. Leidner, Carlos Martinez, Jacqueline M. Kruser, Richard G. Wunderink, Alexander J. Hodakowski

Ochsner Medical Center: Juan Carlos Q. Velez*, Eboni G. Price-Haywood, Luis A. Matute-Trochez, Anna E. Hasty, Muner MB. Mohamed

Oregon Health and Science University Hospital: Rupali S. Avasare*, David Zonies*

Partners Healthcare: Brigham and Women's Hospital, Brigham and Women's Faulkner Hospital, Massachusetts General Hospital, and Newton Wellesley Hospital - David E. Leaf*, Shruti Gupta*, Rebecca M. Baron, Meghan E. Sise, Erik T. Newman, Samah Abu

Omar, Kapil K. Pokharel, Shreyak Sharma, Harkarandeep Singh, Simon Correa Gaviria, Tanveer Shaukat, Omer Kamal, Wei Wang, Heather Yang, Jeffery O. Boateng, Meghan Lee, Ian A. Strohbehn, Jiahua Li, Saif A. Muhsin, Ernest I. Mandel, Ariel L. Mueller

ProMedica Health System: Promedical Toledo Hospital – Nicholas S. Cairl

Renown Health: Farah Madhani-Lovely, Chris Rowan*, Farah Madhai-Lovely*

Rush University Medical Center: Vasil Peev*, Jochen Reiser, John J. Byun, Andrew Vissing, Esha M. Kapania, Zoe Post, Nilam P. Patel, Joy-Marie Hermes

Rutgers/New Jersey Medical School: Anne K. Sutherland*, Amee Patrawalla, Diana G. Finkel, Barbara A. Danek, Sowminya Arikapudi, Jeffrey M. Paer

Rutgers/Robert Wood Johnson Medical School: Jared Radbel*, Sonika Puri, Jag Sunderram, Matthew T. Scharf, Ayesha Ahmed, Ilya Berim, Jayanth Vatson

Stanford Healthcare: Stanford University School of Medicine – Shuchi Anand*, Joseph E. Levitt, Pablo Garcia

Temple University Hospital: Suzanne M. Boyle*, Rui Song

Thomas Jefferson University Hospital: Jingjing Zhang*

Tulane Medical Center: Moh'd A. Sharshir*, Vadym V. Rusnak

University of Colorado Anschutz Medical Campus: Anip Bansal*, Amber S. Podoll, Michel Chonchol, Sunita Sharma, Ellen L. Burnham

University Hospitals Cleveland Medical Center: Arash Rashidi*, Rana Hejal

University of Alabama-Birmingham Hospital: Eric Judd*, Laura Latta, Ashita Tolwani

University of California-Davis Medical Center: Timothy E. Albertson*, Jason Y. Adams

University of California-Los Angeles Medical Center: Ronald Reagan-UCLA Medical Center - Steven Y. Chang*, Rebecca M. Beutler; UCLA Medical Center, Santa Monica – Carl E. Schulze

University of California-San Diego Medical Center: Etienne Macedo*, Harin Rhee

University of California-San Francisco Medical Center: Kathleen D. Liu*, Vasantha K. Jotwani

University of Chicago Medical Center: Jay L. Koyner*

University of Florida Health-Gainesville: Chintan V. Shah*

University of Florida-Health-Jacksonville: Vishal Jaikaransingh*

University of Illinois Hospital and Health Sciences System: Stephanie M. Toth-Manikowski*, Min J. Joo*, James P. Lash

University of Kentucky Medical Center: Javier A. Neyra*, Nourhan Chaaban

University Medical Center of Southern Nevada: Alfredo Iardino, Elizabeth H. Au, Jill H. Sharma

University of Miami Health System: Marie Anne Sosa*, Sabrina Taldone, Gabriel Contreras, David De La Zerda, Hayley B. Gershengorn

University of Michigan: Salim S. Hayek*, Pennelope Blakely, Hanna Berlin, Tariq U. Azam, Husam Shadid, Michael Pan, Patrick O' Hayer, Chelsea Meloche, Rafey Feroze, Kishan J. Padalia, Jeff Leya, John P. Donnelly, Andrew J. Admon

University of North Carolina Hospitals: Jennifer E. Flythe*, Matthew J. Tugman

University of Oklahoma Health Sciences Center: Brent R. Brown*

University of Pennsylvania Health System: Amanda K. Leonberg-Yoo*, Ryan C. Spiardi, Todd A. Miano, Meaghan S. Roche, Charles R. Vasquez

University of Pittsburgh Medical Center: Amar D. Bansal*, Natalie C. Ernecoff

University of Tennessee Health Science Center and Memphis VA Medical Center/Methodist University Hospital – Csaba P. Kovesdy*, Miklos Z. Molnar*

University of Texas Southwestern Medical Center and Parkland Health and Hospital System: S. Susan Hedayati*, Mridula V. Nadamuni, Sadaf S. Khan, Duwayne L. Willett

University of Vermont Larner College of Medicine: Samuel A.P. Short

University of Virginia Health System: Amanda D. Renaghan*

University of Washington Medical Center: Pavan Bhatraju*, A. Bilal Malik

Vanderbilt University Medical Center: Matthew W. Semler

Washington University in St. Louis/Barnes Jewish Hospital: Anitha Vijayan*, Christina Mariyam Joy, Tingting Li, Seth Goldberg, Patricia F. Kao

Wellforce Health System: Lowell General Hospital **-** Greg L. Schumaker*, Tufts Medical Center - Nitender Goyal*, Anthony J. Faugno, Greg L. Schumaker, Caroline M. Hsu, Asma Tariq, Leah Meyer

Westchester Medical Center: Marta Christov*

Yale School of Medicine: Francis P. Wilson*, Tanima Arora, Ugochukwu Ugwuowo

*Site Principal Investigator

eMethods

Definition of an intensive care unit (ICU) patient

Intensive care unit (ICU) admission was defined as admission to a usual ICU room or to a non-ICU room that was functioning as an ICU room due to surge capacity. Non-ICU rooms were considered to be functioning as an ICU room if: 1) the patient was being treated by an ICU team; 2) the patient was receiving extracorporeal membrane oxygenation or invasive mechanical ventilation; 3) the patient was receiving continuous renal replacement therapy; or 4) the patient was receiving vasopressors, inotropes, or mechanical cardiac support (e.g., a ventricular assist device) in a room where this usually would not be allowed.

Data collection and validation

REDCap, a secure web-based platform for building and managing databases and surveys, was used for data collection. The data were quality checked, with queries provided to each site enabling sites to double check questionable values and correct any input errors. The directions requested that PEEP and PaO2/FiO2 ratio should only be entered if a patient was on a mechanical ventilator. Thus, when this was not the case this variable was set to missing. PEEP was set to 0 if not a on a mechanical ventilator and missing if on BiPaP/CPAP/High flow Nasal Cannula.

Machine learning methods

Several machine learning methods were compared in this study. Each model type has its own set of hyperparameters that control model building. These hyperparameters can be thought of as a series of dials that can be optimized. All hyperparameters were selected in the training data using ten-fold cross validation to maximize the area under the receiver operating characteristic curve (AUC) from the out-of-sample folds (i.e., the training data was separated into ten parts to perform ten-fold cross-validation). Missing values were imputed using bagged trees, and the imputation models were developed in the training data, and then applied to the test datasets to impute missing values.

Elastic net logistic regression: This approach combines multivariable logistic regression with lasso and ridge regression penalty terms. These penalty terms shrink the model coefficients to decrease overfitting and consequently improve performance while also providing variable selection. Ten-fold cross-validation in the training data was used to determine the values of the penalty terms that maximized the area under the receiver operating characteristic curve (AUC). To account for potential non-linearity of the continuous predictor variables, restricted cubic splines were used. This allows the risk of mortality to vary for both low and high values of a variable in non-linear fashion, which can improve model accuracy.

eXtreme Gradient Boosting (XGBoost): Gradient boosted machines (GBM) is based on simple decision trees that separate patients with and without the outcome of interest using simple yes-no splits, which can be visualized in the form of a tree. GBM builds many trees sequentially such that each tree attempts to improve the model fit by weighing the difficult-to-predict cases to a greater degree, which results in a tree ensemble model that is more accurate than any one individual tree. An improved version of GBM, called eXtreme Gradient Boosting (XGBoost), was used in this work which increases the speed of the algorithm and includes penalty terms to avoid model overfitting. The number of trees, depth of trees, learning rate as new trees were added, and the minimum size of the terminal leaves were determined using ten-fold cross-validation in the training data.

Random forests: The random forests algorithm is similar to XGBoost in that it builds an ensemble of decision trees, but instead of building them sequentially it builds each tree separately based on a random sample of the training data. Within each tree, only a random

number of predictor variables are available for each yes-no split, which results in trees that are different from each other. The final random forest model is therefore a tree ensemble containing hundreds and sometimes thousands of individual decision trees, with each of them combining to make predictions on new patients. The number of trees and the number of predictor variables available at each split were determined using ten-fold cross-validation in the training data.

Neural networks: Neural networks are flexible, non-linear models that were initially inspired by how the brain works. These models are composed of a combination of individual neuron-like units that take the predictor variables as inputs, combine them in hidden layers, transform them through activation functions, and then output predictions. They can be shallow, with few hidden layers, or deep, with multiple hidden layers. Neural networks have revolutionized the private sector and have demonstrated high accuracy in large datasets in clinical medicine. In this study, a feed-forward multi-layer perceptron neural network was used. The number of hidden layers, size of each hidden layer, learning rate, and decay were determined using ten-fold cross-validation in the training data.

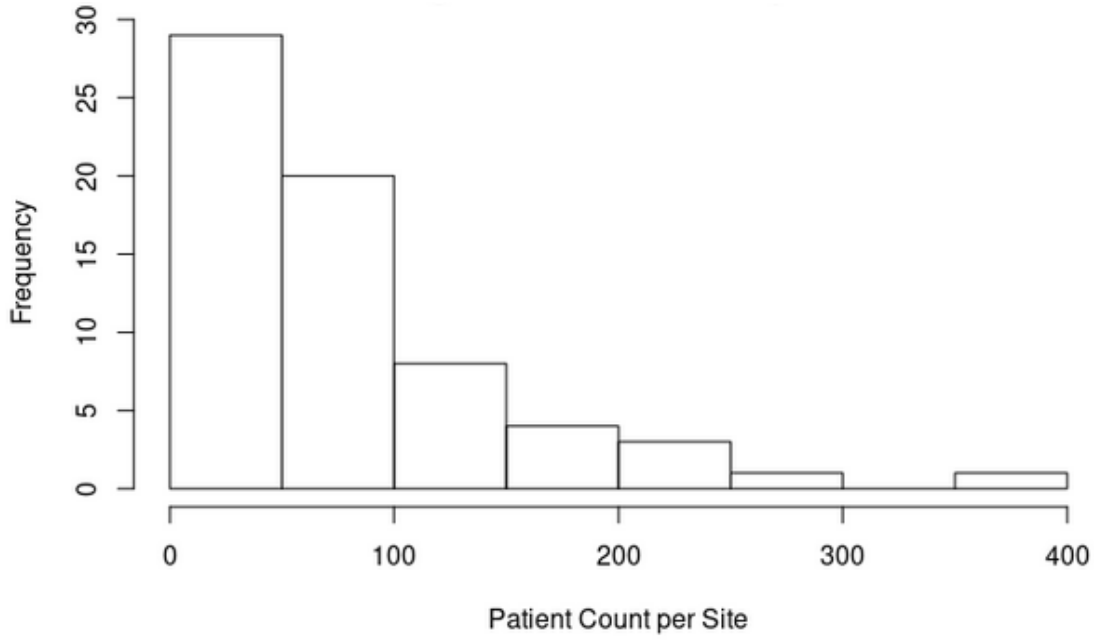
Support vector machines (SVMs): SVMs project the patient data onto a higher dimensional space and create a linear decision boundary in that space that attempts to maximize the margin between the patients categorized by outcome. This decision boundary maps to a nonlinear decision boundary in the original space of patient data. In this study, we utilized the radial basis kernel, which allows the decision boundary to be non-linear in the input space, potentially improving accuracy. The optimal value of the cost penalty, which penalizes the model for misclassified points, was determined using ten-fold cross-validation in the training data.

K-nearest neighbors (KNN): KNN is an approach that also projects the data into multidimensional space, but assigns the outcome of a new patient based on the majority outcome of the K closest training points (i.e., its neighbors). In this study, the number of nearest neighbors and the distance metric used were determined using ten-fold cross-validation in the training data.

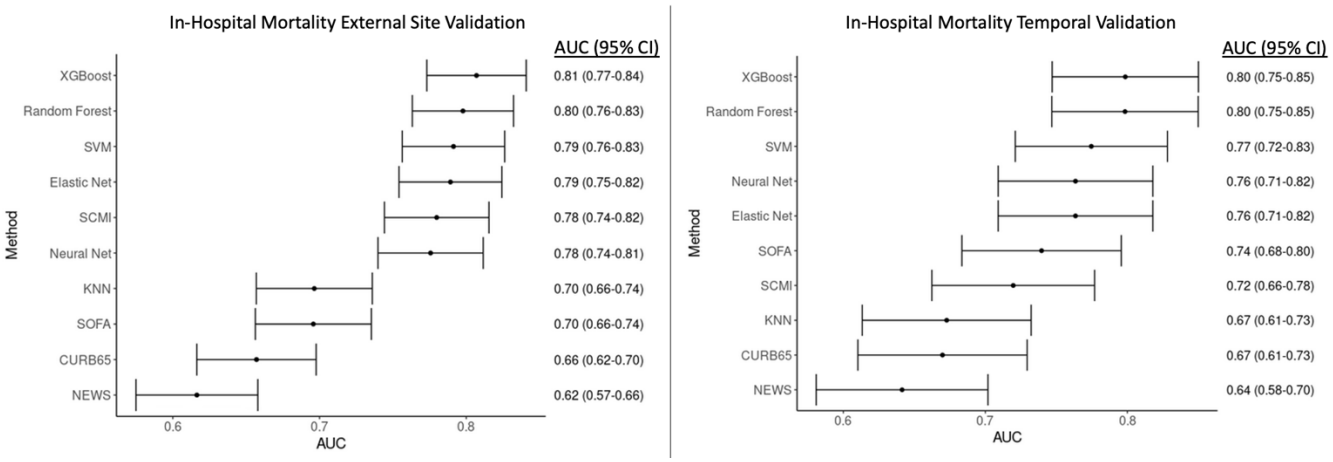
Sensitivity analysis of 28-day mortality in patients discharged from the hospital prior to 28 days

In a subset of patients admitted to six hospitals in Boston, MA who had been discharged from the hospital prior to 28 days, we called them or reviewed their charts to ascertain their 28-day survival status. All of the 50 discharged patients reviewed remained alive at 28 days.

eFigure 1. Distribution of the number of patients included in the study by site. Histogram displaying the number of patients enrolled in the study at each site.

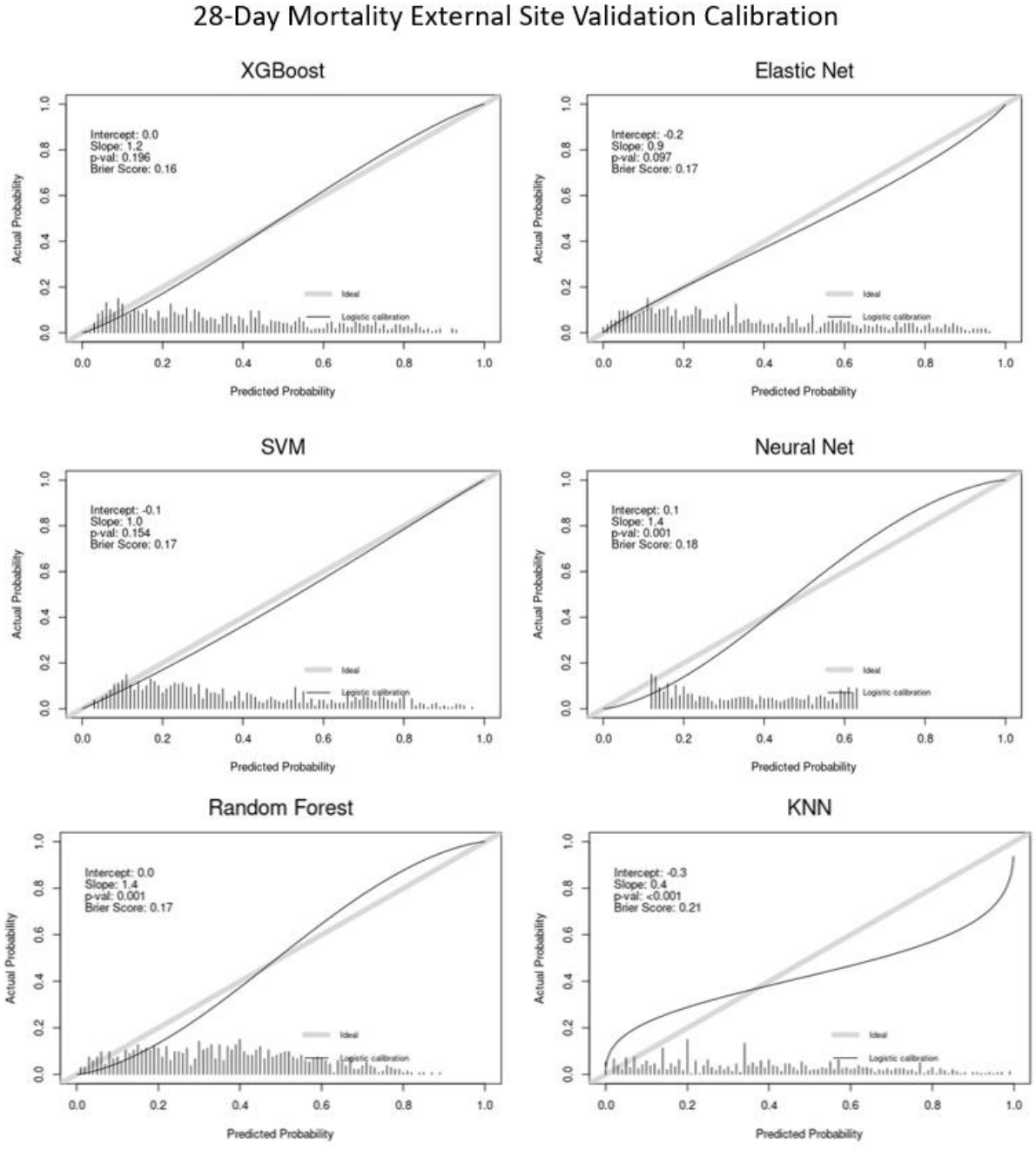


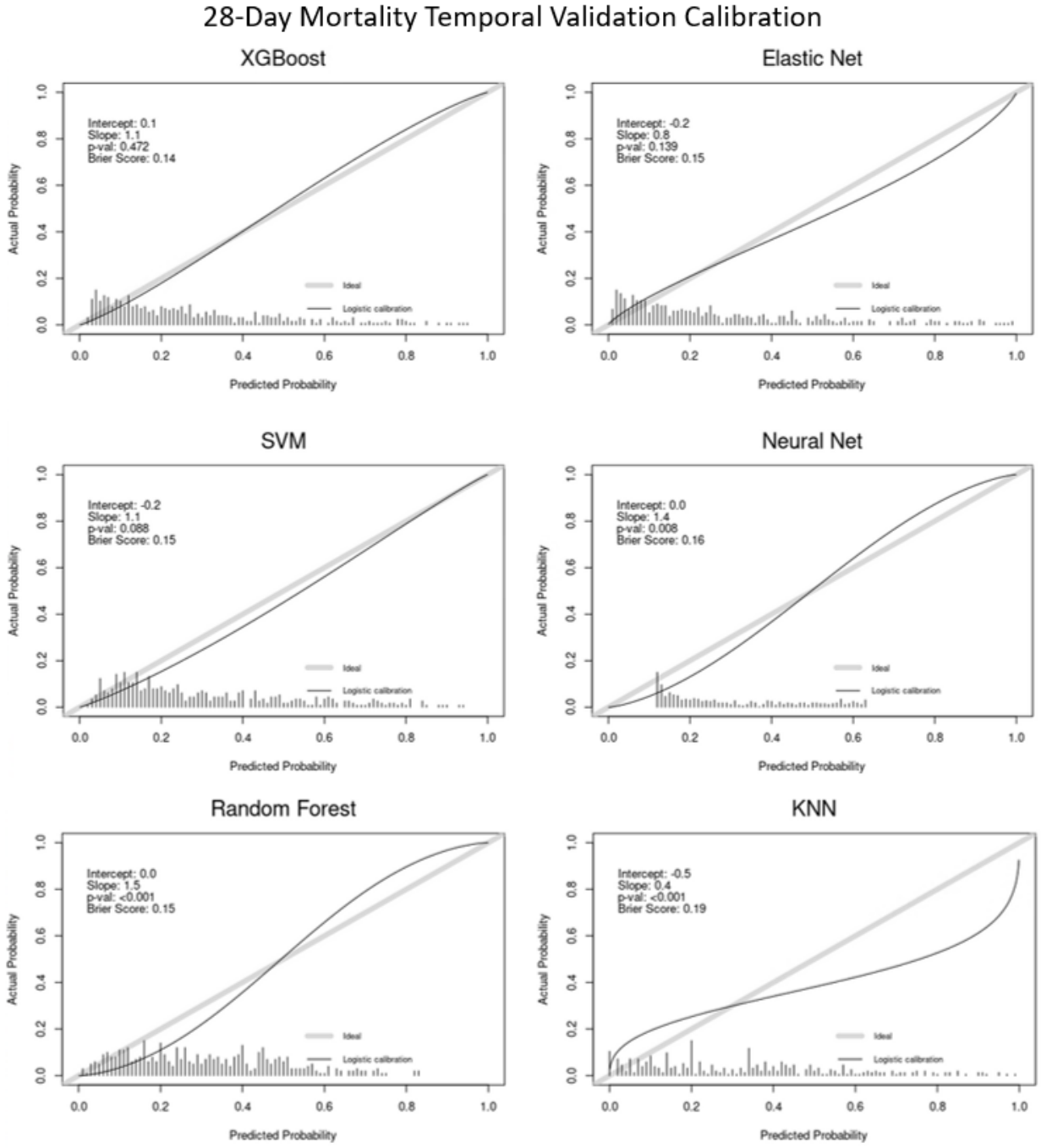
eFigure 2. Comparison of model discrimination between the different models for in-hospital mortality.



Abbreviations: AUC: area under the receiver operating characteristic curve; XGBoost = eXtreme Gradient Boosting; SVM = support vector machine; SCMI = STOP-COVID Mortality Index; SOFA = sequential organ failure assessment; KNN = K-nearest neighbors; NEWS = national early warning score

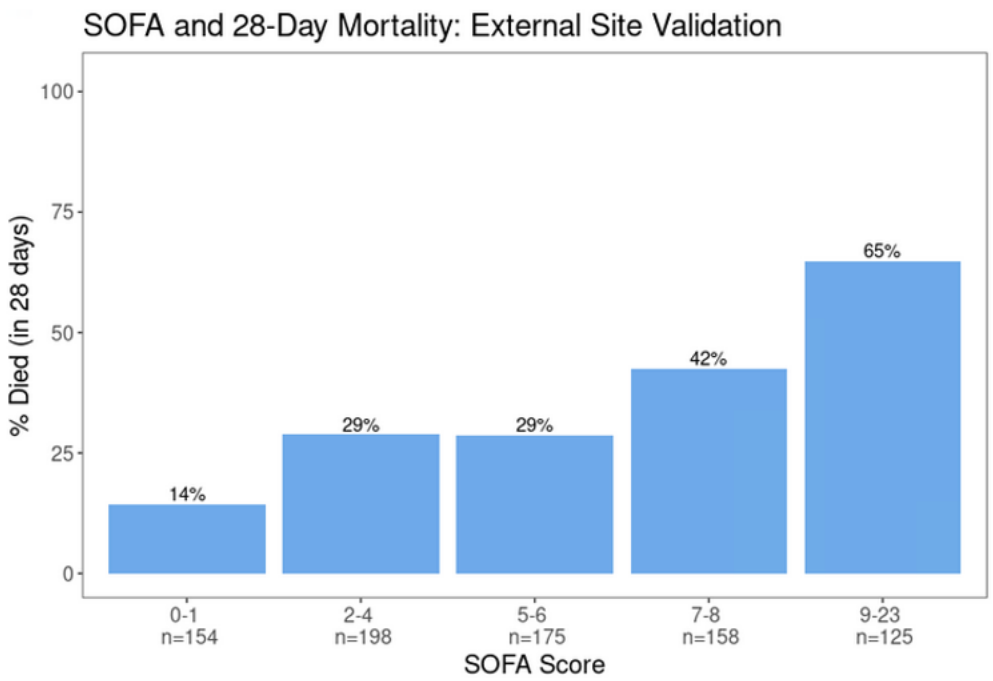
eFigure 3. Calibration plots and statistics for the different machine learning models for the external (top) and temporal (bottom) validations. Perfect calibration is shown by the shaded gray line, where actual (y-axis) and predicted (x-axis) probabilities are the same, p-values are from the unreliability index, and the vertical bars are histograms of predicted probabilities. XGBoost, Elastic Net, and SVM were well-calibrated (unreliability p>0.05), with a calibration intercept of 0 and slope of 1 and the plot showing that the actual probability of mortality is similar to the predicted probability of death from the model (as illustrated by the dark line falling within the shaded gray line). All other models demonstrated poor calibration (unreliability p<0.01).





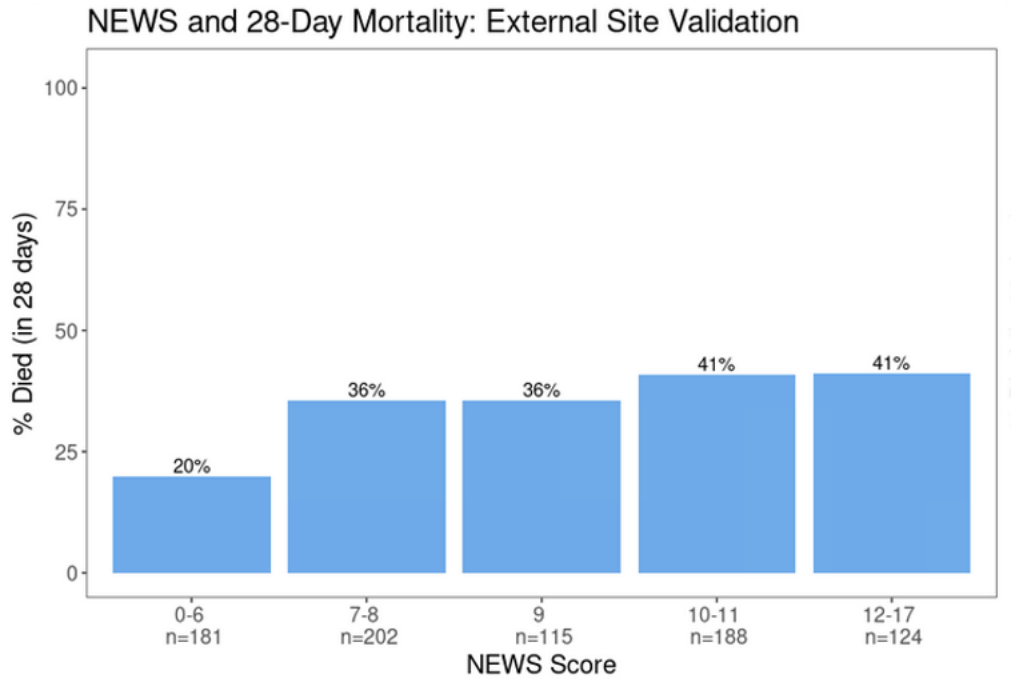
Abbreviations: XGBoost = eXtreme Gradient Boosting; SVM = support vector machine; KNN = K-nearest neighbors

eFigure 4. Relationship between the modified SOFA Score and 28-day mortality in the external validation.



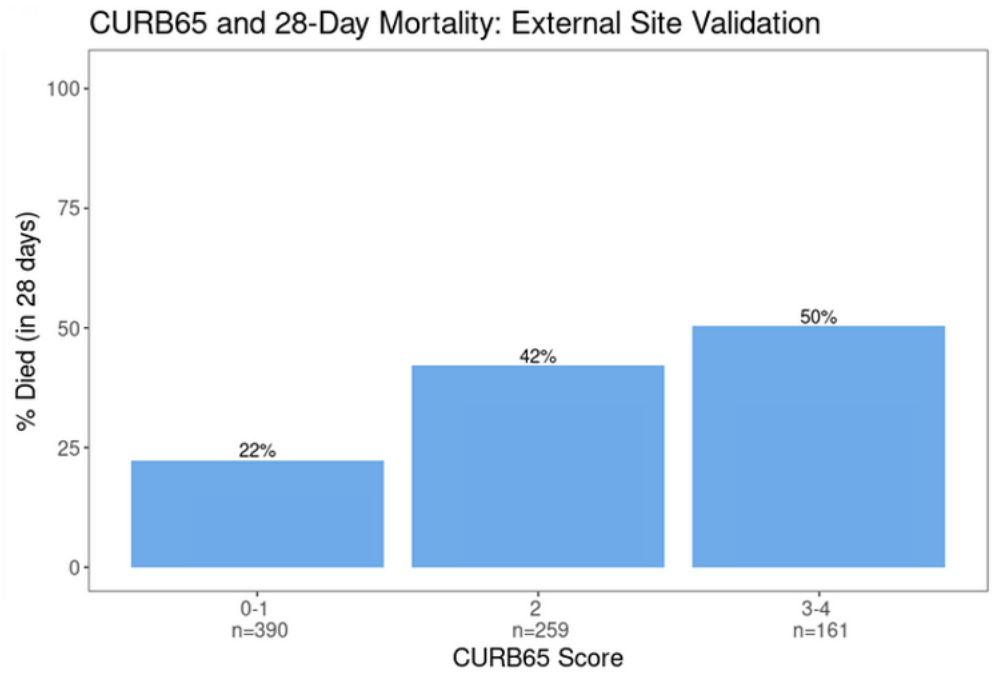
Abbreviations: SOFA = sequential organ failure assessment

eFigure 5. Relationship between the modified NEWS score and 28-day mortality in the external validation.

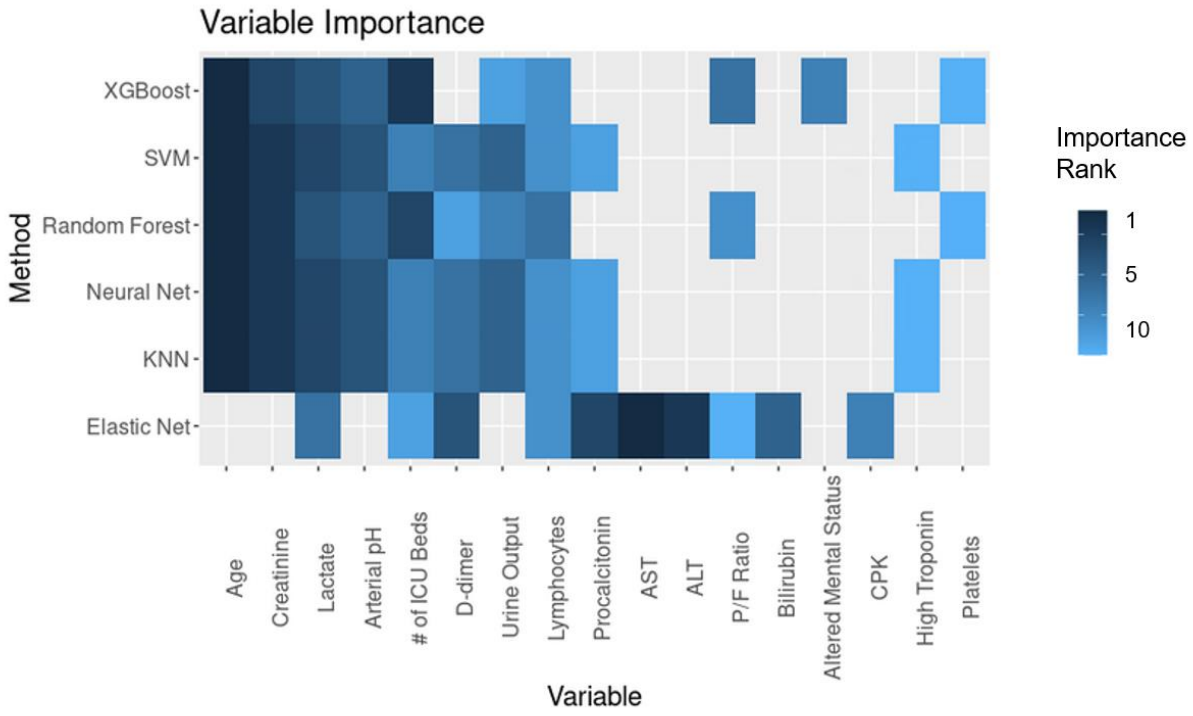


Abbreviations: NEWS = National Early Warning Score

eFigure 6. Relationship between the modified CURB-65 score and 28-day mortality in the external validation.



eFigure 7. Variable importance heat map illustrating the ten most important variables across the different models. Variables are shown from most important (1 = dark blue) to less important (10 = light blue) to unimportant (white).



Abbreviations: SVM = support vector machine; KNN = K-nearest neighbors; ICU = intensive care unit; P/F ratio = PaO2/FiO2 ratio; PEEP = positive end-expiratory pressure; ALT = alanine aminotransferase

eTable 1. Participating sites

Northeast
Beth Israel Deaconess Medical Center
Brigham and Women's Faulkner Hospital
Brigham and Women's Hospital
Cooper University Health Care
Hackensack Meridian Health Hackensack University Medical Center
Hackensack Mountainside Hospital
Johns Hopkins Hospital
Kings County Hospital Center
Lowell General Hospital
Massachusetts General Hospital
MedStar Georgetown University Hospital
Montefiore Medical Center
Mount Sinai
Newton Wellesley Hospital
New York-Presbyterian Queens Hospital
New York-Presbyterian/Weill Cornell Medical Center
New York University Langone Hospital
Rutgers/New Jersey Medical School
Rutgers/Robert Wood Johnson Medical School
Temple University Hospital
Thomas Jefferson University Hospital
Tufts Medical Center
United Health Services Hospitals
University of Pennsylvania Health System
University of Pittsburgh Medical Center
Westchester Medical Center
Yale University Medical Center
South
Baylor College of Medicine, Houston
Baylor University Medical Center/Baylor Scott White and Health
Duke University Medical Center
Mayo Clinic, Florida
Memphis VA Medical Center
Methodist University Hospital
Ochsner Medical Center
Tulane Medical Center
University of Alabama-Birmingham Hospital
University of Florida Health-Gainesville
University of Florida Health-Jacksonville
University of Miami Health System
University of North Carolina Hospitals
University of Texas Southwestern Medical Center
University of Virginia Health System
Midwest
Barnes-Jewish Hospital
Cook County Health
Froedtert Hospital
Indiana University Health Methodist Hospital
Mayo Clinic, Rochester Minnesota
Northwestern Memorial Hospital
Promedica Health System
Rush University Medical Center
University Hospitals Cleveland Medical Center
University of Chicago Medical Center
University of Illinois Hospital and Health Sciences System
University of Kentucky Hospital
University of Michigan Hospital
University of Oklahoma Health Sciences Center
West
Loma Linda University Medical Center
Mayo Clinic, Arizona
Oregon Health and Science University Hospital
Renown Health
Stanford Healthcare
University of California-Davis Medical Center
University of California-Los Angeles Medical Center
University of California-San Diego Medical Center
University of California-San Francisco Medical Center
UCHealth University of Colorado
University Medical Center of Southern Nevada
University of Washington Medical Center

eTable 2. Definitions of key variables and outcomes

Baseline Characteristics	
Baseline serum creatinine	Lowest value (mg/dl) within 365 to 7 days prior to hospital admission. If not available, serum creatinine on hospital admission
Healthcare worker	Physician, nurse, technician, or other medical professional who provides direct care to patients (does not include ancillary staff such as clerks, pharmacists, or kitchen/cleaning staff)
Home medications	Medications that the patient was taking at home within 1 week prior to admission. Does not include those started at an outside hospital if the patient was transferred.
Anticoagulation	Therapeutic anticoagulants, not including anti-platelet agents such as aspirin or clopidogrel
Immunosuppressant drugs	Chemotherapy (in the 30 days prior to admission), corticosteroids >10 mg prednisone/day (or equivalent), calcineurin inhibitors (systemic, not topical), mycophenolate mofetil, azathioprine, rituximab, other
Coexisting Conditions	
Asthma	Per chart review
Atrial fibrillation/flutter	Per chart review
Bone marrow transplant	Per chart review
Cancer	Per chart review; active malignancy (other than non-melanoma skin cancer) treated in the past year. Defined as cancer of the lung, breast, colorectal, prostate, gastric, pancreatic, melanoma, ovarian, brain, or other
Chronic kidney disease	Baseline eGFR< 60 on at least two consecutive values at least 12 weeks apart prior to hospital admission. If not available, defined as per chart review
Chronic liver disease	Cirrhosis, alcohol-related liver disease, nonalcoholic fatty liver disease, autoimmune hepatitis, hepatitis B or hepatitis C, primary biliary cirrhosis, or other
Chronic obstructive pulmonary disease	Per chart review
Congestive heart failure	Per chart review; heart failure with preserved versus reduced ejection fraction
Coronary artery disease	Per chart review; any history of angina, myocardial infarction, or coronary artery bypass graft surgery
Diabetes mellitus	Per chart review; insulin versus non-insulin dependent
End stage renal disease	Per chart review; on hemodialysis or peritoneal dialysis
History of alcohol abuse	Per chart review
HIV/AIDS	Per chart review
Homelessness	Per chart review
Hypertension	Per chart review
Solid organ transplant	Per chart review (kidney, liver, heart, lung, other)
Smoking	Per chart review; does not include vaping or smoking of non-tobacco products. Non-smoker, former smoker, current smoker
Longitudinal Parameters and Treatments[a]	
Extracorporeal membrane oxygenation	Veno-venous, veno-arterial, or veno-arterial-venous
Mechanical cardiac support	Impella, intra-aortic balloon pump, LVAD, RVAD, other
Mechanical ventilation	Invasive mechanical ventilation
Renal replacement therapy	CRRT, intermittent hemodialysis, peritoneal dialysis, other
PaO2[b]	Lowest PaO2 available during each 24 hour day (midnight to midnight)
FiO2[b]	FiO2 corresponding to the lowest PaO2
PEEP[b]	Highest PEEP available during each 24 hour day (midnight to midnight)
Vasopressors	Maximum number of vasopressors required each day
Outcomes[a]	
Acute kidney injury[c]	Doubling of serum creatinine from baseline or need for renal replacement therapy (RRT), corresponding with stages 2 and 3 of the Kidney Disease: Improving Global Outcomes Criteria.[1] Baseline serum creatinine was defined as the lowest value from within 365 to 7 days prior to hospital admission. If unavailable, the hospital admission value was used as the baseline.
Acute liver injury	Modified version of the CTCAE criteria[2]: bilirubin >3.0 mg/dl *and* either AST>100 units per liter or ALT>100 units per liter
Acute cardiac injury	Troponin T or I > the 99th percentile upper reference limit of normal for that lab
Acute respiratory distress syndrome	Modified Berlin criteria[3] (all three of the following were required): PaO_2:FiO_2 ratio<300 mm Hg and mechanically ventilated and a diagnosis of ARDS per chart review
Arrhythmia (new onset)	Per chart review; includes atrial fibrillation/flutter, ventricular tachycardia (sustained versus non-sustained), and ventricular fibrillation
Cardiac arrest	Per chart review
Coagulopathy	INR>2 or PTT>40 seconds in the absence of therapeutic anticoagulation

Congestive heart failure (new onset)	Per chart review; includes both heart failure with preserved and reduced ejection fraction
Disseminated intravascular Coagulation	Per chart review
Major bleed	Per chart review; bleeding in a critical area or organ (e.g., intracranial, retroperitoneal, pericardial, or intramuscular bleeding with compartment syndrome) or bleeding requiring a procedural intervention (e.g., EGD or IR embolization)
Myocarditis	Per chart review
Pericarditis	Per chart review
Respiratory failure	Requirement for invasive mechanical ventilation
Secondary Infection	Per chart review; suspected or confirmed new infection other than COVID-19 that developed after admission to the ICU. Pneumonia (including ventilator-associated), urosepsis, biliary sepsis, bacteremia, other
Shock	Requirement for 2 or more vasopressors
Thromboembolic event	Per chart review; deep venous thrombosis, pulmonary embolism, stroke, heparin-induced thrombocytopenia, other
Cause of Death	Per chart review; ARDS/respiratory failure, congestive heart failure, septic shock, kidney failure, liver failure, other

Abbreviations: ALT = alanine transaminase; AST = aspartate transaminase; COVID-19 = coronavirus-19; CRRT = continuous renal replacement therapy, CTCAE = Common Terminology Criteria for Adverse Events; DVT = deep vein thrombosis; EGD = esophagogastroduodenoscopy; eGFR = estimated glomerular filtration rate; FiO2 = fraction of inspired oxygen; HIV/AIDs = human immunodeficiency syndrome/acquired immunodeficiency virus; IR = interventional radiology; LVAD = left ventricular assist device; PaO2 = partial pressure of oxygen; PE = pulmonary embolism; PEEP = positive end-expiratory pressure; RVAD = right ventricular assist device.

[a]Longitudinal treatments and outcomes were recorded daily for the first 14 days following admission to the ICU. If multiple values were present, the lowest PaO_2 available, along with the corresponding FiO_2 at the time, was recorded, while the highest PEEP on each day was recorded. If the patient had an outcome, the date of the outcome was recorded.
[b]Only applies to patients on mechanical ventilation with an arterial blood gas available.
[c]Excludes patients with end stage renal disease.

References:
[1]Kidney Disease; Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl 2*, 2012: 1-138.
[2]US Department of Health and Human Services. Common Terminology Criteria for Adverse Events (CTCAE). Bethesda, MD: National Institute of Health, National Cancer Institute, 2017.
[3]The ARDS Definition Task Force. Acute Respiratory Distress Syndrome. The Berlin Definition. *JAMA*, 2012; 307(23): 2526-2533.

eTable 3. Final variables used in the machine learning models. The direction of high risk clarifies which value would be taken if more than one value is collected on day 1 and day 2 of the study and only applies to variables collected on both days (e.g., if two P/F ratios are available, we used the lower of the two values in the models).

Category	Variable	Direction of high risk*	% Missing in Training Data	% Missing in Internal Test Data	% Missing in External Test Data
Demographics					
	Age		0.0	0.0	0.0
	Body Mass Index		3.9	5.3	1.1
	Sex		0.0	0.0	0.0
Vital signs (collected on day 1 only)					
	Highest Heart Rate		0.0	0.0	0.0
	Highest Respiratory Rate		0.0	0.0	0.2
	Lowest Systolic Blood Pressure		0.1	0.0	0.0
	Max Temperature		0.1	0.0	0.2
	P/F Ratio	lower	35.1	31.6	65.2
Coexisting Conditions					
	Asthma		0.0	0.0	0.0
	Atrial fibrillation		0.0	0.0	0.0
	Bacterial Pneumonia		0.0	0.0	0.0
	CKD		0.0	0.0	0.0
	Congestive Heart Failure		0.0	0.0	0.0
	COPD		0.0	0.0	0.0
	Coronary Artery Disease		0.0	0.0	0.0
	Hypertension		0.0	0.0	0.0
	Insulin-dependent Diabetes		0.0	0.0	0.0
	No Cardiovascular Comorbidities		0.0	0.0	0.0
	No Infection		0.0	0.0	0.0
	No Other Comorbidities		0.0	0.0	0.0
	Non-Insulin-dependent Diabetes		0.0	0.0	0.0
	Other Lung Diseases		0.0	0.0	0.0
	Smoking		14.5	13.4	9.1
Symptoms	Altered Mental Status		8.1	5.1	5.7
	Chills		0.0	0.0	0.0
	Confusion		0.0	0.0	0.0
	Cough		0.0	0.0	0.0
	Diarrhea		0.0	0.0	0.0
	Fatigue		0.0	0.0	0.0
	Fever		0.0	0.0	0.0
	Headache		0.0	0.0	0.0
	Myalgia Arthralgia		0.0	0.0	0.0
	Nasal Congestion		0.0	0.0	0.0
	Nausea/Vomiting		0.0	0.0	0.0

	Prior Symptom Days		0.6	0.3	0.9
	Short of breath		0.0	0.0	0.0
	Sore Throat		0.0	0.0	0.0
	Sputum Production		0.0	0.0	0.0
Labs					
	Albumin	lower	9.6	9.8	8.4
	ALT	higher	7.9	10.6	10.0
	Arterial pH	lower	20.6	19.1	34.1
	AST	higher	8.1	10.6	10.0
	Bilirubin	higher	8.6	10.6	10.0
	CPK (recorded day 1 only)		48.5	59.5	48.1
	Creatinine	higher	2.2	0.6	4.1
	CRP	higher	23.3	27.6	18.8
	D-dimer	higher	30.5	37.7	16.8
	Ferritin	higher	25.8	30.8	25.2
	Hemoglobin	lower	2.3	0.7	3.9
	Indicator for High Troponin	higher	36.4	41.2	36.6
	Lactate	higher	27.4	34.8	30.5
	Lymphocyte Count	lower	11.3	9.5	15.4
	Platelets	lower	2.6	0.7	4.3
	Procalcitonin		38.2	34.6	30.5
	Sodium		0.5	0.3	0.9
	Urine Output	lower	27.7	26.8	49.5
	White Blood Cell Count	higher	2.3	0.7	3.9
Treatments					
	Mechanically Ventilated		0.1	0.0	0.2
	Number of Vasopressors	higher	0.0	0.0	0.0
	PEEP if on a Ventilator (day1 and day2)		Day 1: 34.6 Day 2: 4.7	Day 1: 29.2 Day 2: 4.5	Day 1: 52.5 Day2: 2.3
	Renal Replacement Therapy		0.0	0.0	0.0
Other					
	Hospital Type		0.0	0.0	0.0
	Number of ICU beds		0.0	0.0	0.0
	Source Admit		0.0	0.0	0.0

*If two or more values were available, then the worst value was chosen based on the direction of high risk.

Abbreviations: ICU = intensive care unit; CKD = chronic kidney disease; COPD = chronic obstructive pulmonary disease; P/F ratio = PaO2/FiO2 ratio; PEEP = positive end-expiratory pressure; ALT = alanine aminotransferase; AST = aspartate aminotransferase

eTable 4. Modified SOFA score. The SOFA score was modified from the original published version because of missing Glasgow coma scale, mean arterial pressure, and vasopressor doses in the database.

	0	**1**	**2**	**3**	**4**
SOFA Renal (Cr [mg/dl], UOP [ml/day], and acute RRT)	Cr<1.2 and UOP≥500	Cr 1.2-1.9 and UOP≥500	Cr 2-3.4 and UOP≥500	Cr 3.5-4.9 or UOP 200-499	Cr ≥5 or UOP<200 or acute RRT or ESRD
SOFA Liver (Bilirubin, mg/dl)	<1.2	1.2-1.9	2.0-5.9	6.0-11.9	≥12
SOFA Coagulation (Platelets, K/mm^3)	≥150	100-149	50-99	20-49	<20
SOFA Respiratory (PaO_2:FiO_2)	≥400 or not intubated	300-399	200-299	100-199	<100
SOFA Cardiovascular (#vasopressors/inotropes)	0	1	2	3	≥4
SOFA CNS	No AMS		AMS		

Abbreviations: CNS = central nervous system; Cr = creatinine; ESRD = end stage renal disease; RRT = renal replacement therapy; SBP = systolic blood pressure; SOFA = sequential organ failure assessment; UOP = urine output; AMS = altered mental status

eTable 5. Modified NEWS. The NEWS was modified from its original version due to missing oxygen saturation and because nearly all patients were on some form of supplemental oxygen.

	0	1	2	3
Respiratory Rate (per minute)	12-20	9-11	21-24	≤8 or ≥25
Systolic Blood Pressure (mmHg)	111-219	101-110	91-100	≤90 or ≥ 220
Pulse (per minute)	51-90	41-50 or 91-110	111-130	≤40 or ≥ 131
Temperature (°C)	36.1-38.0	35.1-36.0 or 38.1-39.0	≥39.1	≤35
Air or Oxygen	Not Ventilated Day 1		Ventilated Day 1	
Consciousness	No AMS			AMS

Abbreviations: NEWS = National Early Warning Score; AMS = altered mental status

eTable 6. Comparison of patient characteristics for those who survived vs. died at 28 days after ICU admission.

Variable	Summary Measure	All Patients (n = 5075)	Survivors (n = 3229)	Non-Survivors (n = 1846)
Demographics				
Age (years)	Median (P25-P75)	62 (52-71)*	59 (49-68)	67 (58-76)
Body Mass Index (kg/m^2)	Median (P25-P75)	30.3 (26.4-35.9)*	30.7 (26.6-36.0)	29.8 (25.8-35.3)
Male	n (%)	3198 (63.0)*	1981 (61.4)	1217 (65.9)
Symptoms				
Cough	n (%)	3661 (72.1)*	2430 (75.3)	1231 (66.7)
Diarrhea	n (%)	1039 (20.5)*	720 (22.3)	319 (17.3)
Fatigue or malaise	n (%)	1606 (31.6)	1034 (32.0)	572 (31.0)
Fever	n (%)	3353 (66.1)*	2218 (68.7)	1135 (61.5)
Nausea/Vomiting	n (%)	793 (15.6)*	583 (18.1)	210 (11.4)
Prior Symptom Days	Median (P25-P75)	7 (4-10)*	7 (4-11)	7(3-10)
Dyspnea	n (%)	3793 (74.7)	2417 (74.9)	1376 (74.5)
Sputum Production	n (%)	532 (10.5)*	368 (11.4)	164 (8.9)
Coexisting Conditions				
Asthma	n (%)	544 (10.7)*	378 (11.7)	166 (9.0)
Chronic Kidney Disease	n (%)	665 (13.1)*	343 (10.6)	322 (17.4)
Congestive Heart Failure	n (%)	516 (10.2)*	284 (8.8)	232 (12.6)
Chronic Obstructive Pulmonary Disease	n (%)	439 (8.7)*	230 (7.1)	209 (11.3)
Coronary Artery Disease	n (%)	685 (13.5)*	347 (10.7)	338 (18.3)
Hypertension	n (%)	3113 (61.3)*	1809 (56.0)	1304 (70.6)
Insulin Dependent Diabetes	n (%)	843 (16.6)*	471 (14.6)	372 (20.2)
Non-Insulin Dependent Diabetes	n (%)	1295 (25.5)*	780 (24.2)	515 (27.9)
Other Pulmonary Disease	n (%)	334 (6.6)	203 (6.3)	131 (7.1)
Smoking				
Former	n (%)	1229 (28.1)*	723 (25.5)	506 (32.9)
Current	n (%)	259 (5.9)	164 (5.8)	95 (6.2)
Vital signs on ICU Day 1				
Highest Heart Rate (beats/min)	Median (P25-P75)	104 (90-120)*	103 (90-117)	108 (93-124)
Lowest Systolic Blood Pressure (mm Hg)	Median (P25-P75)	97 (85-111)*	99 (87-111)	94 (82-109)
Max Temperature (°C)	Median (P25-P75)	37.9 (37.2-38.8)*	37.9 (37.2-38.8)	37.8 (37.1-38.7)
Labs on ICU Days 1-2				
Creatinine (mg/dl)	Median (P25-P75)	1.2 (0.9-2.1)*	1.1 (0.8-1.6)	1.6 (1.0-2.9)

C-reactive protein (mg/L)	Median (P25-P75)	167 (90-256)*	162 (87-246)	181 (100-273)
D-dimer (ng/mL)	Median (P25-P75)	1590 (751-4340)*	1260 (660-3180)	2390 (1078-7955)
Lactate (mmol/L)	Median (P25-P75)	1.6 (1.2-2.4)*	1.5 (1.1-2.0)	2.0 (1.3-3.0)
Lymphocytes (%)	Median (P25-P75)	8.5 (5.0-13.4)*	9.6 (6.0-14.4)	7.0 (4.0-11.0)
Platelets (K/mm^3)	Median (P25-P75)	203 (154-263)*	208 (162-267.0)	192.0 (139.0-254.0)
Total Bilirubin (mg/dl)	Median (P25-P75)	0.6 (0.4-0.9)*	0.6 (0.4-0.9)	0.6 (0.4-1.0)
White Blood Cell Count (per mm^3)	Median (P25-P75)	9.6 (6.8-13.6)*	9.1 (6.6-12.5)	10.8 (7.6-15.4)
Severity of Illness on ICU Days 1-2				
PEEP Day 1	Median (P25-P75)	10 (5-15)*	10 (0-14)	12 (8-15)
P/F Ratio (mm Hg)	Median (P25-P75)	116 (80-171)*	122 (84-178)	106 (76-159)
Ventilation				
Invasive Mechanical Ventilation	n (%)	3029 (59.7)*	1712 (53.1)	1317 (71.3)
BiPaP/CPAP/High flow Nasal Cannula	n (%)	1372 (27.1)*	962 (29.8)	410 (22.2)
Renal Replacement Therapy	n (%)	381 (7.5)	200 (6.2)*	181 (9.8)
Vasopressors				
1	n (%)	1788 (35.2)*	1102 (34.1)	686 (37.2)
2 or more	n (%)	829 (16.3)*	374 (11.6)	455 (24.6)
Other				
Number of pre-COVID ICU beds	Median (P25-P75)	87 (48-115)*	98 (55-120)	58 (47-100)

*P-value <0.05 for difference between survivors and non-survivors (Wilcoxon rank-sum test for continuous variables and chi-squared for categorical variables). Abbreviations: PEEP = positive end-expiratory pressure; P/F = PaO2/FiO2; ICU = intensive care unit.

eTable 7. Other baseline and first 48-hour characteristics among patients who died vs. survived by day 28

Variable	Summary Measure	All Patients (n = 5075)	Survivors (n = 3229)	Non-Survivors (n = 1846)
Symptoms				
Altered Mental Status	n (%)	1183 (25.2)*	524 (17.4)	659 (39.2)
Chills	n (%)	984 (19.4)*	679 (21.0)	305 (16.5)
Confusion	n (%)	649 (12.8)*	317 (9.8)	332 (18.0)
Headache	n (%)	447 (8.8)*	352 (10.9)	95 (5.1)
Myalgia Arthralgia	n (%)	1104 (21.8)*	831 (25.7)	273 (14.8)
Nasal Congestion	n (%)	299 (5.9)*	219 (6.8)	80 (4.3)
Sore Throat	n (%)	391 (7.7)*	272 (8.4)	119 (6.4)
Coexisting Conditions				
No Cardiovascular/Pulmonary Comorbidities[2]	n (%)	1139 (22.4)*	872 (27.0)	267 (14.5)
No Secondary Infection[3]	n (%)	3942 (77.7)*	2565 (79.4)	1377 (74.6)
No Other Comorbidities[4]	n (%)	3740 (73.7)*	2532 (78.4)	1208 (65.4)
Vital signs[1]				
Highest Respiratory Rate (beats/min)	Median (P25-P75)	31 (26-38)	32 (26-38)	31 (26-38)
Laboratory values				
Albumin (g/dl)	Median (P25-P75)	2.9 (2.5-3.3)*	3.0 (2.6-3.3)	2.8 (2.4-3.1)
Alanine aminotransferase - ALT (U/L)	Median (P25-P75)	38 (23-65)*	37 (23-63)	38 (24-67)
Arterial pH	Median (P25-P75)	7.3 (7.3-7.4)*	7.4 (7.3-7.4)	7.3 (7.2-7.4)
Aspartate aminotransferase - AST (U/L)	Median (P25-P75)	56 (37-89)*	52 (36-82)	64 (41-107)
Creatinine Phosphokinase[1] - CPK (U/L)	Median (P25-P75)	192 (87-518)*	179 (83-467)	230 (100-663)
Ferritin (ng/ml)	Median (P25-P75)	1052 (522-2000)*	951(483-1795)	1268 (597-2667)
Hemoglobin (g/dl)	Median (P25-P75)	11.8 (10.3-13.1)*	11.9 (10.5-13.2)	11.5 (9.9-13.0)
High Troponin Indicator	n (%)	1578 (49.5)*	836 (41.1)	742 (64.3)
Procalcitonin[1] (ng/ml)	Median (P25-P75)	0.4 (0.2-1.4)*	0.3 (0.1-0.9)	0.7 (0.2-2.6)
Sodium[1]	Median (P25-P75)	137 (134-140)*	137 (134-140)	137 (134-141)
Urine Output (mL)	Median (P25-P75)	700 (315-1135)*	800 (410-1250)	525 (210-950)
Severity of Illness				
PEEP Day 2	Median (P25-P75)	10 (0-14)*	8 (0-14)	10 (0-14)
Mechanical Ventilator Day 2	n (%)	3317 (65.4)*	1938 (60.0)	1379 (74.7)
Other				
Hospital Type: main vs affiliate	n (%)	4389 (86.5)	2813 (87.1)	1576 (85.4)
Source of Admit				
Emergency Department	n (%)	2787 (54.9)	1755 (54.4)	1032 (55.9)
Hospital Ward	n (%)	1522 (30.0)*	926 (28.7)	596 (32.3)

*P-value <0.05 for difference between survivors and non-survivors (Wilcoxon rank sum test for continuous variables and chi-squared for categorical)

[1]Available day 1 only; peep and ventilation value included for both day 1 and day 2 due to differences in granularity of ventilator type

[2]Indicator for lack of any of the following cardiovascular and pulmonary comorbidities: diabetes mellitus, hypertension, coronary artery disease, congestive heart failure, atrial fibrillation/flutter, COPD, Asthma, other lung disease

[3]Indicator for lack of any of the following infections: bacterial pneumonia, viral respiratory infection, urosepsis, biliary sepsis, cellulitis, bacteremia/endocarditis, other

[4]Indicator for lack of any of the following other comorbidities: chronic kidney disease, ESRD, chronic liver disease, HIV/AIDS, active malignancy, solid organ transplant, bone marrow transplant, other immunodeficiency

Abbreviations: PEEP = positive end-expiratory pressure

eTable 8. Model details. Details of R packages and tuning grids utilized. Bold indicates chosen hyperparameter in final model.

Model	Model within caret package	Hyperparameters Tested	Method to obtain probabilities*
Elastic Net	glmnet with rcs and 3 knots for continuous variables	alpha = c(0,0.1,0.2,0.4,0.6,0.8,**1**), lambda = sequence from **0** to .1 20 digits long	elastic net logistic regression
XGBoost	xgbtree	nrounds=c(500, 1000, **1500**), max_depth=c(**2**, 5, 10) , eta=c(0.001, **0.01**,0.1), gamma = c(**0**),, colsample_bytree = c(**1**), min_child_weight = c(**1**,2, 5), subsample = c(**0.5**)	sum of leaf's weights
Random Forest	rf	mtry = c(2,4,5,**20**)	proportion of trees voting for a given class
KNN	kknn	kmax = c(5,7,**9**,12,15), kernel = c("**optimal**"), distance = c(1,**2**)	proportion of neighbors voting for a given class
Neural Net	mlpKerasDropout	random search length 100 chose size = 8, dropout = 0.6242863, batch_size = 869, lr = 0.07409052, rho = 0.6956519, decay = 0.2133154 and activation = sigmoid	sigmoid
SVM	svmRadial	sigma = c(**.01**), C = seq(0.25, 2, length = 20) **0.639**	rescaled version of the original classifiers scored through a logistic transformation

*The default method for calculating probabilities for each model's package was used, with a brief description in this column.
Abbreviations: KNN = K-nearest neightbors, SVM = support vector machine