Original Investigation

Estimating Survival of Hospitalized COVID-19 Patients from Admission Information

Todd J. Levy[a], MS; Safiya Richardson[bc], MD, MPH; Kevin Coppa[d], BS; Douglas P. Barnaby[bc], MD, MSc; Thomas McGinn[bc], MD, MPH; Lance B. Becker[ab], MD; Karina W. Davidson[bc], PhD, MASc; Jamie S. Hirsch[bcd], MD, MA, MSB; Theodoros P. Zanos[ab], PhD

And the Northwell COVID-19 Research Consortium: Stuart L. Cohen[bc], MD; Shubham Debnath[a], PhD; Andrew J. Dominello[c], BA; Louise Falzon[c], BA; Michael Gitman[b], MD; PGDipInf; Eun-Ji Kim[bc], MD, MS; Alexander Makhnevich[b], MD; Jazmin N. Mogavero[c], MA; Ernesto P. Molmenti[b], MD, PhD, MBA, FACS; Marc d. Paradis[e], SM; Viktor Tóth[a], MSc

[a]Institute of Bioelectronic Medicine, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY
[b]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Northwell Health, Hempstead, NY
[c]Institute of Health Innovations and Outcomes Research, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY
[d]Department of Information Services, Northwell Health, New Hyde Park, NY
[e]Department of Data Strategy & Ventures, Northwell Health, Manhasset, NY

Corresponding author:
Theodoros P. Zanos, PhD
Assistant Professor
The Feinstein Institutes of Medical Research
Northwell Health
350 Community Drive, Room 1257
Manhasset, NY 11030
T: 516-562-0484
E: tzanos@northwell.edu

Word count: 2964

Abstract

Background
While clinical characteristics and a range of mortality risk factors of COVID-19 patients have been reported, a practical early clinical survival calculator specialized for the unique cohort of patients has not yet been introduced. Such a tool would provide timely and valuable guidance in clinical care decision-making during this global pandemic.

Methods
Demographic, laboratory, clinical, and treatment data (from 13 acute care facilities at Northwell Health) were extracted from electronic medical records and used to build and test the predictive accuracy of a survival probability calculator—the Northwell COVID-19 Survival (NOCOS) calculator—for hospitalized COVID-19 patients. The NOCOS calculator was constructed using multivariate regression with L1 regularization (LASSO). Model predictive performance was measured using Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) of the calculators tested.

Results
A total of 5,233 inpatients were included in the study. Patient age, serum blood urea nitrogen (BUN), Emergency Severity Index (ESI), red cell distribution width (RCDW), absolute neutrophil count, serum bicarbonate, and glucose were identified as the optimal early predictors of survival by multivariate LASSO regression. The predictive performance of the Northwell COVID-19 Survival (NOCOS) calculator was assessed for 14 consecutive days.

Conclusions
We present a rapidly developed and deployed estimate of survival probability that outperforms other general risk models. The 7 early predictors of in-hospital survival can help clinicians identify patients with increased probabilities of survival and provide critical decision support.

Introduction

The World Health Organization designated coronavirus disease 2019 (COVID-19) a global pandemic on March 11th, 2020, with over 1 million confirmed worldwide cases.[1] Estimates of severe disease range from 20-30% and case fatality rates from 2-7%.[2,3] As healthcare facilities across the world struggle to provide care for increasing numbers of critically ill patients, many countries are reporting or anticipating significant ventilator and equipment shortages.[4-6] The development of evidence-based resource allocation tools and processes will be necessary to ensure that we meet our ethical duty to provide the most benefit for the largest number of people.

In cities across the globe, physicians faced with resource limitations are independently deciding which patients to aggressively resuscitate and ventilate and for whom to withhold artificial respiratory support.[6,7] Aiding healthcare workers with robust predictive survival models ensures more informed decision-making and efficient, just resource allocation while reducing physician stress and burnout. An early, simple, and clinically relevant model to predict survival in hospitalized COVID-19 patients brings objectivity to emotionally fraught decisions and conversations with patients and families. There have been no published multivariate models predicting survival in larger cohorts (>100) of patients with COVID-19 for at the time of this study, although reports from China have identified age, Sequential Organ Failure Assessment (SOFA) score, and d-dimer level as potential predictors.[8]

Our objectives were to use parameters available early to clinicians to characterize and predict survival for hospitalized COVID-19 patients within the largest health system in New York State, the current epicenter of the global COVID-19 pandemic. We consider significant variables reported from previous work and describe the demographics, baseline

comorbidities, presenting clinical studies, and outcomes of hospitalized patients with COVID-19. We then present a simple, powerful, and clinically relevant predictive model of patient survival—the Northwell COVID-19 Survival (NOCOS) calculator—for all non-mechanically ventilated patients at the time of hospital admission with parameters available early in the care of all patients. The model utilizes routinely collected data typically available within 60 minutes of patient arrival in the emergency department and predicts hospital survival at a time that permits planning and proper decision-making around goals of care and resource allocation. This actionable model can be easily implemented and used to support providers during the current worldwide crisis.

Methods

This analysis of a COVID-19 survival calculator uses data from a retrospective cohort study that was approved by the Northwell Health Institutional Review Board. It includes all adult hospitalized patients (i.e., those aged 18 and up) with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection by positive result by polymerase chain reaction testing of a nasopharyngeal sample. Patients were excluded if they were placed on mechanical ventilation before presentation to or in the emergency department. These patients' clinical characteristics and outcomes are described more completely in a prior publication on this cohort study.[9,10] Patients were admitted to 1 of 13 Northwell Health acute care hospitals on or after March 1st, 2020, and were discharged or died before April 12th, 2020. Clinical outcomes (i.e., discharges, mortality, and length of stay) were monitored until April 12th, 2020, the final date of follow-up. With approximately 4,844 hospital beds and 672 intensive-care-unit (ICU) beds and serving approximately 11 million persons in Long Island, Westchester, and New York City, Northwell Health is the largest academic health system in

New York. Notably, during the current pandemic, the number of general hospital beds and ICU beds has increased substantially and fluctuates daily.

Data

Data were collected from the enterprise electronic health record (EHR; Sunrise Clinical Manager, Allscripts, Chicago, IL) reporting database. Transfers from one in-system hospital to another were merged and considered as one visit. Data collected included patient demographic information, comorbidities, home medications, Emergency Severity Index (ESI; an objective marker of emergency department presenting acuity), initial laboratory values and studies, prescribed medications, treatments (including oxygen therapy and mechanical ventilation), and outcomes (including length of stay, discharge, and mortality). Initial laboratory testing was defined as having been obtained while the patient was in the emergency department. Continuous variables are presented as median and interquartile range (IQR), and categorical variables are expressed as number of patients (percentage). Acute kidney injury was identified according to the Kidney Disease: Improving Global Outcomes (KDIGO) definition.[11] Acute hepatic injury was defined an elevation in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) of greater than 15 times the upper limit of normal. Oxygen requirements were collected for the highest requirement level during the emergency department stay. We used the chi-squared test for categorical variables and Kruskal-Wallis for continuous variables across all groups to test for differences by survival status.

Predictive Modeling

LASSO regression was used to identify a small subset out of 85 EHR measurements that, when linearly combined, predict the survival of hospitalized COVID-19 patients (Table 1).[12]

By including an L_1-norm regularization term that promotes sparsity, LASSO regression is well suited for determining the optimal subset of measurements. The magnitudes of the coefficients relate to the predictive values of the normalized measurements while coefficients of non-predictive measurements converge exactly to 0.

The data is normalized by taking the z-score so that all measurements are sampled from a distribution with 0 mean and a standard deviation of 1. The mean and standard deviation of the measurements with non-0 coefficients are stored as model hyperparameters during training and applied to test data. Missing measurements were imputed to the mean.

The regularization factor λ is another hyperparameter that is determined by sweeping λ over a range, evaluating the performance, and choosing the value that corresponds to the optimal tradeoff between maximizing performance and minimizing the number of predictors. After optimizing for λ, the number of predictors was fixed at 7 inputs. The performance is measured as the area under the Receiver Operating Characteristic (ROC) curve.

The training set is evaluated with the model using leave-one-out cross-validation to prevent overfitting in order to estimate the class conditional distributions (survived and expired) of LASSO predictions as Gaussian likelihood functions. The posterior probability that the patient will survive is

$$p\left(\mu_{survived}, \sigma_{survived}, \mu_{expired}, \sigma_{expired} \middle| x\right) = \frac{p_{survived}(x|\mu_{survived}, \sigma_{survived})}{\sum_{c \in C} p_c(x|\mu_c, \sigma_c) P(c = C_c)} P(c = C_{survived})$$

where $p_c(x|\mu_c, \sigma_c)$ is the Gaussian likelihood function estimated from the LASSO predictions that have outcomes for class c that is an element of the set containing survived

and expired, $P(c = C_c)$ is the prior probability of class c derived from the training set, and x is the LASSO prediction for a patient.

Two instances of the calculator were tested: one fixed, trained on data acquired until March 29th, 2020, and tested daily on new patients; and one retrained daily to incorporate new data. We also tested the predictive value of the SOFA and CURB-65 Score for Pneumonia Severity as well as a linear regression model termed SOFA+ that uses the SOFA score, age, and D-Dimer>1 μg/mL based on a recently published study.[8] All models are tested across all days, from March 30th, 2020, to April 12th, 2020, using ROC curves and the Area Under the Curve (AUC) metrics with statistical differences in predictive performance tested using the nonparametric DeLong method.[13,14] All analyses were performed in Matlab 2019b (Mathworks Inc.).

Results

Between March 1st, 2020, and April 12th, 2020, of the 5,233 patients admitted with COVID-19, 1,185 died while in the hospital (Table 1). As reported previously, [9] patients who died were more frequently older, white, and non-Hispanic males with a higher comorbidity burden, including coronary artery disease, diabetes mellitus, hypertension, heart failure, and kidney disease. With lower diastolic blood pressure, faster respiratory rate, and lower oxygen saturation, they were generally more acutely ill on emergency department arrival (based upon ESI score). The initial labs were almost all significantly different between survivors and non-survivors (Table 1), although many non-routine labs were not available for all patients. While the length of stay was not different between the groups, expired patients had been far more likely to require mechanical ventilation.

The proposed NOCOS calculator was built after optimizing for L1 regularization parameter lamda, based on out-of-sample AUCs, with multivariate logistic regression choosing 7 out of the 85 possible inputs available in the emergency department as the best predictors of survival upon hospitalization: patient age, serum blood urea nitrogen (BUN), ESI, red cell distribution width (RCDW), absolute neutrophil count, serum bicarbonate, and glucose. The fixed NOCOS calculator was trained using all cases hospitalized until March 30th, 2020. The NOCOS calculator was trained every day and tested using data only from the following day. Both fixed and daily retrained versions of NOCOS were compared to clinical benchmarks SOFA and CURB-65 as well as a variation of the SOFA score.[8] Based on the ROC and the AUC values, the daily retrained NOCOS calculator—with an AUC of 0.832 while the fixed NOCOS and SOFA+ variation followed very closely (AUC of 0.825 and 0.830 respectively)—outperformed all other calculators (Figure 1). CURB-65 and SOFA score had significantly lower predictive performance than the three aforementioned calculators (AUC of 0.739 and 0.732 respectively, DeLong's, $p<0.05$ when compared to the daily retrained NOCOS); they couldn't always be calculated due to some missing values for the patients.

Operating points to determine performance of survival predictions for all calculators can be established by choosing thresholds on the probability scores. We chose three different operating points for each calculator and provide the numbers of true positives, true negatives, false positives and false negatives, as well as Positive Predictive Value (PPV) and Negative Predictive Value (NPV) for each case (Table 2). In all cases, daily retrained NOCOS outperformed all other calculators.

The NOCOS calculator also demonstrated stability both in its predictive ability and the selection of the predictors across multiple days. As shown in Figure 2, panel A, the NOCOS

calculator maintains an AUC value roughly between 0.8 and 0.9 from March 30th, 2020, through April 12th, 2020, regardless of whether it was trained once or retrained daily. The daily trained NOCOS calculator was significantly more predictive than CURB-65 on 10 out of the 14 days, significantly more predictive than SOFA on 7 out of the 14 days, and significantly more predictive than the fixed NOCOS calculator on 5 out of the 14 days (DeLong's method, $p<0.05$). It was not significantly more predictive than SOFA+ on any of the days.

The coefficients of the daily retrained NOCOS calculator, chosen by the LASSO regularization across 7 days, are shown with the counts of the times selected in "Figure 2": "B." The final 7 parameters were patient age, ESI, BUN, serum bicarbonate, absolute neutrophil count, RCDW, and serum glucose. Five of these 7 predictors were chosen on at least 13 of the 14 days with the exception of serum bicarbonate and serum glucose, which were both chosen on 6 out of 14 days. Other measurements such as platelet count, body temperature, serum albumin, oxygen saturation, and epidermal growth factor inhibitor (eGFRi) were also chosen on fewer days but were not included in the final build of the model. In the latest iteration of the daily retrained NOCOS calculator (trained with data up to April 11th, 2020), the negative predictors of survival in order of their contribution to the probability estimate are: patient age, BUN, RCDW, absolute neutrophil count and serum bicarbonate (Figure 2 panel C). The positive predictors of survival are ESI (lower scores are more acute) and serum glucose (Figure 2, panel C).

The performance of the NOCOS calculator was also tested when not limited only to the ED values of the 7 parameters of a patient, but also when the latest measurements are used as

inputs. Figure 3 shows the performance of fixed NOCOS when tested using the up-to-date values of the seven measurements, with the AUC increasing steadily to values close to 0.91.

Discussion

In this study, we successfully developed a simple and practical survival calculator for hospitalized COVID-19 patients using only discrete and objective data values acquired during the patient's initial time in the emergency department. Our Northwell COVID-19 Survival (NOCOS) calculator, modeled on over 5,200 COVID-19–positive patients, had an AUC of 0.83 and outperformed other well-established risk calculators, including CURB-65, SOFA, while it performed similarly to COVID-19–specific enhancements to SOFA.[8] Developed to be parsimonious and easy to use, the predicted survival probability can be used to assist clinical decision-making and ease physician burden in this unprecedented situation. The output of this calculator (which is freely available at https://feinstein.northwell.edu/nocos) provides an easily comprehensible probability, which can be communicated to physicians and nurses, families, and other administrative teams.

The choice of variables included in our model, which were ascertained from the LASSO regularization, all have clinical face validity. It is well established with many diseases, and particularly with COVID-19, that older age confers an increased mortality risk.[8] ESI, a well-established ED triage tool, is an early indicator of presenting severity of illness. Abnormal laboratory values included in our model have all been independently associated with negative outcomes in other populations,[15,16] and an elevated BUN (as a maker of kidney dysfunction, in particular) was recently shown to increase mortality risk in COVID-19 patients. [17] Elevated values of RCDW, often suggesting chronic disease states and inflammation,[18,19] can also be due to recently reported effects of COVID-19 on iron displacement of the heme

molecule, leading to impaired red blood cells as well as free radical formation and toxic effect to the lungs.[20] These findings suggest potential therapeutic approaches to reduce sudden decompensation, organ failure, and death of these patients.

A major strength of this work is the development of a powerful predictive model typically usable for clinicians within 60 minutes of a patient's initial presentation. Although the calculator performs well with these very early measurements, it improves its predictive performance when these measurements are updated throughout the hospitalization of the patient (Figure 3), showing that, as expected, the most accurate prediction is given with the most up-to-date values of the seven measures. We also restricted inputs to commonly collected, discrete, and objective data. Its sheer simplicity and reliance on quantitative measurements makes it generalizable and easy to deploy to all interested stakeholders, including front-line providers and hospital administrators organizing distribution of scarce and limited resources. While we present the calculator output as a probability score, a specific operating point can also be chosen to provide a binary outcome prediction with significant accuracy. Choosing an operating point is left up to stakeholders; local clinical teams have flexibility to adjust thresholds toward a more stringent or risk-averse solution (Table 2), based on the rapidly changing needs during this pandemic.

Calculating estimates of survival or mortality using clinical measurements can extend from simple algorithmic rules and thresholds to linear regression models and more complex machine learning (ML) algorithms. Attempting to augment medical decision-making, studies ranging from modulating single parameters to advanced predictive modeling have been applied to forecast decompensation, mortality, and survival among other clinical outcomes.[21-23] Early work with small patient cohorts of COVID-19 has led to models that identify some

clinical characteristics that can be applied to predict severe cases (Yan et al., 2020, Jiang et al., 2020).[24,25] However, these studies are limited to small numbers of patients as well as the inclusion of qualitative and subjective variables, are prone to mislabeling, and are not always readily available. Our approach benefits from a simple, straightforward formula of typical measurements acquired from ED patients; a patient base at least 20-fold larger than previous studies; and an approach of data-true feature selection based on their predictive value through the LASSO regularization.

Due to the challenging situation during the ongoing COVID-19 global health crisis, there is a need for robust tools to aid in complex clinical decision-making. Using well-known clinical calculators such as SOFA or CURB-65 shows ostensible promise; however, these calculators have limitations in both their accuracy and the ease of collecting necessary measurements to construct these scores. Input variables such as confusion (for the CURB-65 score) and Glasgow Coma Scale (for the SOFA score) are ambiguous, hard to measure, and frequently unavailable. Similar difficulties are encountered when trying a novel combination of SOFA score with age and D-dimer values.[8] In our study, 78.3% of patients were missing the D-dimer measurement in the emergency department. In contrast, the NOCOS calculator is based on commonly collected laboratory results and a guideline based ESI triage acuity score. Moreover, the calculator is trained and tested on the patient cohort of interest and can account for the evolving nature of this pandemic by daily or more frequent updates and model retraining.[26]

The proposed calculator has some limitations. It was designed to be linear with only essential predictors included, and non-linear or convolutional/recurrent models may provide improved performance. Moreover, the model is not integrating additional, more complex information

such as radiology X-ray or CT-scan reads. Due to the retrospective study design, not all laboratory tests—including lactate dehydrogenase, interleukin-6, and serum ferritin—were done on all patients, and the performance of these variables could not be adequately assessed. These data were automatically extracted from the EHR database, and some patient-level details could not be extracted. However, our NOCOS calculator aimed to leverage easily obtainable data, obviating the need for sifting through charts to obtain a predictive result.

Given the complexity of data acquisition and model development in the midst of a pandemic, we prioritized the creation and rapid dissemination of a more straightforward, clinically relevant implementation. While the model validation contained patients admitted to hospitals within the New York metropolitan area, we believe it will generalize well given the diverse demographic composition of the region and the Northwell Health patient population.

In an unprecedented way, the severity of the SARS-CoV-2 pandemic has strained hospitals' resources, including space, materials, and front-line healthcare workers. Providers are often forced to take important clinical decisions under immense time pressure and limited information. Tools that could aid them in these circumstances are timely and important. The Northwell COVID-19 Survival calculator answers a clinical need and provides early information to physicians making a range of difficult-but-critical decisions every day.

Acknowledgments

Financial Disclosures: The authors report no real or apparent conflicts of interest.

Funding Sources: This work was supported by grants R24AG064191 from the National Institute on Aging and R01LM012836 from the National Library of Medicine of the National Institutes of Health.

Role of the Funding Sources: The views expressed in this paper are those of the authors and do not represent the views of the National Institutes of Health, the United States Department of Health and Human Services, or any other government entity.

Other declarations: The investigators were independent from the funders; Todd J. Levy and Theodoros P. Zanos had full access to the data and can take responsibility for the integrity of the data and the accuracy of the data analysis; Theodoros P. Zanos affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Data Availability Statement: The data that support the findings of this study are available on request from COVID19@northwell.edu. The data are not publicly available due to restrictions as it could compromise the privacy of research participants.

References

1. The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Coronavirus COVID-19 Global Cases Center for Systems Science and Engineering, 2020. Baltimore, MD: Center for Systems Science and Engineering; 2020. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e
2. Livingston E, Bucher K. Coronavirus Disease 2019 (COVID-19) in Italy. *JAMA* 2020 March 17 (Epub ahead of print).
3. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72□314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* 2020 February 24 (Epub ahead of print).
4. Kliff S, Satariano A, Silver-Greenberg J, Kulish N. There Aren't Enough Ventilators to Cope With the Coronavirus. *The New York Times* 2020 March 26.
5. Legido-Quigley H, Mateos-Garcia JT, Campos VR, Gea-Sanchez M, Muntaner C, McKee M. The resilience of the Spanish health system against the COVID-19 pandemic. *Lancet Public Health* 2020 March 18 (Epub ahead of print).
6. Rosenbaum L. Facing Covid-19 in Italy - Ethics, Logistics, and Therapeutics on the Epidemic's Front Line. *N Engl J Med* 2020 March 18 (Epub ahead of print).
7. Ramachandran S, Palazzolo J. NYU Langone Tells ER Doctors to 'Think More Critically' About Who Gets Ventilators. *The Wall Street Journal* 2020 March 31.
8. 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.

9. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities and Outcomes Among 5,700 Patients Hospitalized with COVID-19 in the New York City Area. *JAMA.* In Press.
10. Bauchner H, Golub RM, Zylke J. Editorial Concern-Possible Reporting of the Same Patients With COVID-19 in Different Reports. *JAMA.* 2020 March 16 (Epub ahead of print).
11. Kellum JA, Lameire N, Aspelin P, et al. Improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl* 2012;2(1):1-138.
12. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J *R StatSoc Series B (Methodological)* 1996;58(1):267-288.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(3):837-845.
14. Sun X, Xu W. Fast Implementation of DeLong's Algorithm for Comparing the Areas Under Correlated Receiver Operating Characteristic Curves. *IEEE Signal Processing Letters* 2014;21(11):1389-1393.
15. Kim EJ, Jeong MH, Kim JH, et al. Clinical impact of admission hyperglycemia on in-hospital mortality in acute myocardial infarction patients. *Int J Cardiol* 2017;236:9-15.
16. Steer J, Gibson J, Bourke SC. The DECAF Score: predicting hospital mortality in exacerbations of chronic obstructive pulmonary disease. *Thorax* 2012;67(11):970-976.
17. Cheng Y, Luo R, Wang K, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int* 2020 March 20 (Epub ahead of print).

18. Salvagno GL, Sanchis-Gomar F, Picanza A, Lippi G. Red blood cell distribution width: A simple parameter with multiple clinical applications. *Crit Rev Clin Lab Sci* 2015;52(2):86-105.

19. Wang B, Gong Y, Ying B, Cheng B. Relation between Red Cell Distribution Width and Mortality in Critically Ill Patients with Acute Respiratory Distress Syndrome. *Biomed Res Int* 2019;2019:1942078.

20. Liu W, Li H. COVID-19: Attacks the 1-Beta Chain of Hemoglobin and Captures the Porphyrin to Inhibit Human Heme Metabolism *ChemRxi*v 2020. (https://chemrxiv.org/articles/COVID-19_Disease_ORF8_and_Surface_Glycoprotein_Inhibit_Heme_Metabolism_by_Binding_to_Porphyrin/11938173)

21. Brekke IJ, Puntervoll LH, Pedersen PB, Kellett J, Brabrand M. The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review. *PLoS One* 2019;14:e0210875.

22. Churpek MM, Edelson DP. Moving Beyond Single-Parameter Early Warning Scores for Rapid Response System Activation. *Crit Care Med* 2016;44(12):2283-2285.

23. Kipnis P, Turk BJ, Wulf DA, et al. Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the ICU. *J Biomed Inform* 2016;64:10-19.

24. Jiang X, Coffee M, Bari A, et al. Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. Computers Materials *Continua* 2020;63(1):537--551.

25. Yan L, Zhang H-T, Goncalves J, et al. A machine learning-based model for survival prediction in patients with severe COVID-19 infection. *medRxiv*

2020:2020.02.27.20028027.
https://www.medrxiv.org/content/10.1101/2020.02.27.20028027v2

26. Santosh KC. AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data. *J Med Syst* 2020;44(5):93.

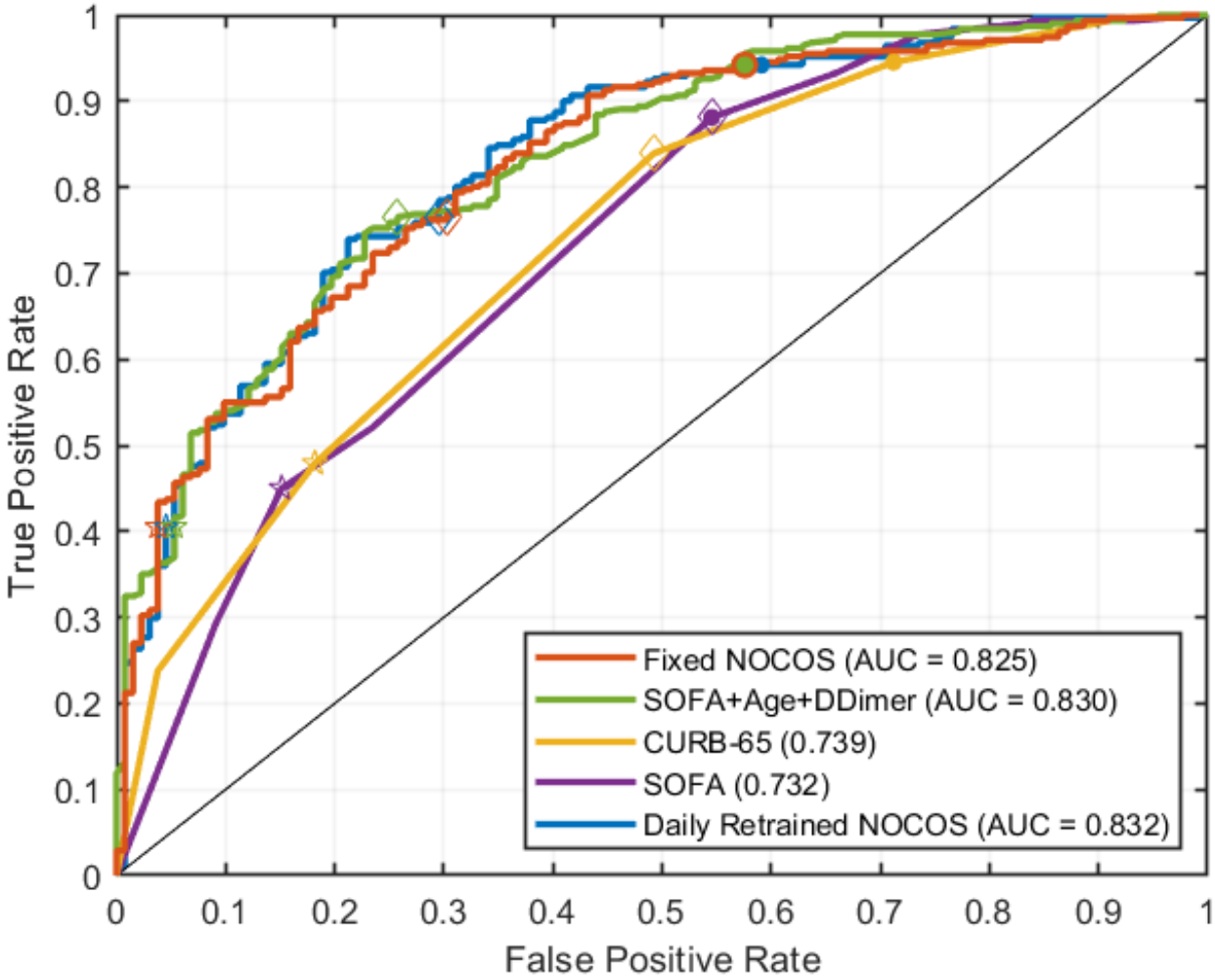

Figure 1. Predictive performance of the calculator. The ROC curve for the daily retrained NOCOS model (blue) and the Fixed NOCOS (red) is compared against the ROC curves for SOFA (purple), CURB-65 (yellow). The model was trained on data from patients admitted on or before April 11th, and evaluated on 443 patients admitted on April 12th. The AUC for the daily retrained NOCOS calculator is 0.832, for the fixed from March 29th NOCOS calculator is 0.825, for SOFA+Age+DDimer is 0.830, for CURB-65 is 0.739 and for SOFA is 0.732.

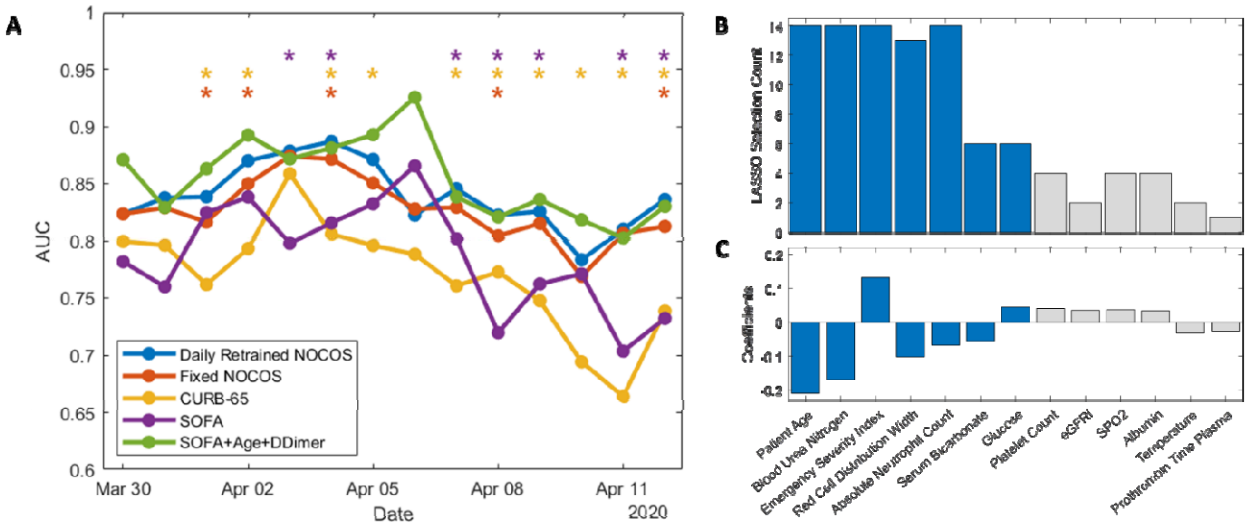

Figure 2. Stability of the calculator. (A) Stability of the calculator's performance between March 30th and April 12th, for the Daily Retrained NOCOS calculator (blue curve) and the fixed NOCOS, trained with data up to March 29th (red curve). The model is significantly superior in its predictive performance from both the CURB-65 score (yellow curve) for 10 out of the 14 days and SOFA score (purple curve) for 7 out of 14 days. The model outperforms the modified SOFA that also includes Age and D-Dimer values half the time (green curve). (B) Stability of the predictors of the calculator, when retrained daily, with 5 out of the 7 predictors being consistently chosen across at least 13 out of the 14 days and 2 out of the 7 predictors being chosen 6 out of the 14 days. (C) The coefficients of the calculator, trained on all cumulative data until April 12th.

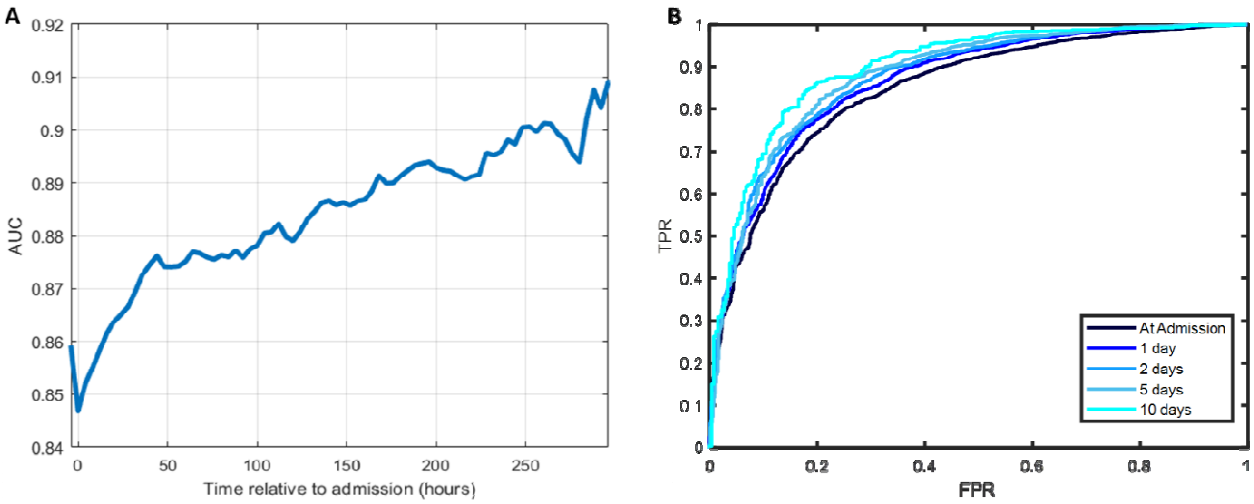

Figure 3. Predictive performance of the calculator with latest measurement. (A) Calculator's predictive performance, as captured by the AUC, increases when it uses the latest measurement values instead of only ED values. (B) ROCs of the calculator predictive performance for various days post admission.

Table 1. Demographic, clinical and laboratory data from hospitalized patients.

	All discharged patients	Discharged Alive	Expired	p-Value	Missing No. (%)
N	5233	4048	1185		
Female (%)	2176 (41.6)	1727 (42.7)	449 (37.9)	0.004	0 (0)
Age At Admission (%)					0 (0)
21-40	554 (10.6)	531 (13.1)	23 (1.9)	<0.001	
41-60	1713 (32.7)	1556 (38.4)	157 (13.2)		
61-80	2087 (39.9)	1538 (38.0)	549 (46.3)		
81-106	879 (16.8)	423 (10.4)	456 (38.5)		
Race (%)					0 (0)
White	2035 (38.9)	1435 (35.4)	600 (50.6)	< 0.001	
Black	1163 (22.2)	945 (23.3)	218 (18.4)		
Other	1410 (26.9)	1185 (29.3)	225 (19.0)		
Asian	399 (7.6)	300 (7.4)	99 (8.4)		
Unknown	195 (3.7)	154 (3.8)	41 (3.5)		
Declined	31 (0.6)	29 (0.7)	2 (0.2)		
Ethnicity (%)					0 (0)
Not Hispanic or Latino	3833 (73.2)	2865 (70.8)	968 (81.7)	<0.001	
Hispanic or Latino	1087 (20.8)	924 (22.8)	163 (13.8)		
Unknown	287 (5.5)	235 (5.8)	52 (4.4)		
Declined	26 (0.5)	24 (0.6)	2 (0.2)		
English Primary Language (%)	4310 (82.4)	3315 (81.9)	995 (84.0)	0.109	0 (0)
Length of Stay	4.88 [2.67, 7.99]	4.77 [2.69, 7.79]	5.59 [2.60, 9.14]	0.001	0 (0)
Emergency Severity Index (ESI)					

1	169 (3.3)	57 (1.4)	112 (9.5)	<0.001	71 (1.4)
2	2615 (50.7)	1849 (46.4)	766 (64.9)		
3	2283 (44.2)	1984 (49.8)	299 (25.3)		
4	93 (1.8)	89 (2.2)	4 (0.3)		
5	2 (0.0)	2 (0.1)	0 (0.0)		
Required mechanical ventilation (%)	755 (14.4)	207 (5.1)	548 (46.2)	<0.001	0 (0)
Last emergency department vital sign measurement (median [IQR])					
Systolic blood pressure (mmHg)	127.00 [113.00, 142.00]	127.00 [115.00, 141.00]	125.00 [110.00, 143.00]	0.011	19 (0.4)
Diastolic blood pressure (mmHg)	73.00 [64.00, 82.00]	74.00 [66.00, 82.00]	69.00 [60.00, 79.00]	<0.001	20 (0.4)
Heart rate (beats per minute)	90.00 [80.00, 103.00]	90.00 [80.00, 102.00]	92.00 [79.00, 105.00]	0.011	10 (0.2)
Respiratory rate (breaths per minute)	20.00 [18.00, 22.00]	20.00 [18.00, 22.00]	21.00 [18.00, 26.00]	<0.001	28 (0.5)
Oxygen saturation (%)	96.00 [94.00, 98.00]	96.00 [95.00, 98.00]	96.00 [93.00, 98.00]	0.001	58 (1.1)
Temperature (Celsius)	37.40 [36.90, 38.10]	37.40 [36.90, 38.10]	37.50 [36.90, 38.30]	0.108	88 (1.7)
Body mass index (kg/m^2)	28.30 [25.10, 32.90]	28.50 [25.10, 32.90]	27.50 [24.10, 32.30]	0.001	2443 (46.7)
Height (cm)	167.64 [162.56, 175.26]	167.64 [162.56, 175.26]	167.64 [162.56, 175.26]	0.893	2310 (44.1)

Weight (kG)	81.60 [68.90, 95.30]	81.60 [70.30, 95.30]	77.10 [65.58, 93.00]	<0.001	1892 (36.2)
Comorbidities (%)					300 (5.7)
Coronary artery disease	454 (9.2)	270 (7.1)	184 (16.5)	<0.001	
Diabetes	1414 (28.7)	1006 (26.3)	408 (36.7)	<0.001	
Hypertension	2474 (50.2)	1776 (46.5)	698 (62.8)	<0.001	
Heart failure	219 (4.4)	105 (2.7)	114 (10.3)	<0.001	
Lung disease	630 (12.8)	470 (12.3)	160 (14.4)	0.074	
Kidney disease	326 (6.6)	177 (4.6)	149 (13.4)	<0.001	
Last emergency department laboratory result (median [IQR])					
WBC Count	6.95 [5.21, 9.36]	6.67 [5.11, 8.92]	7.93 [5.81, 11.10]	<0.001	194 (3.7)
Auto Neutrophil #	5.33 [3.74, 7.57]	5.05 [3.59, 7.09]	6.36 [4.37, 9.13]	<0.001	365 (7.0)
Auto Neutrophil %	77.20 [69.47, 83.30]	76.25 [68.38, 82.30]	81.00 [73.47, 86.50]	<0.001	365 (7.0)
Auto Lymphocyte #	0.91 [0.63, 1.24]	0.94 [0.67, 1.27]	0.74 [0.51, 1.12]	<0.001	365 (7.0)
Auto Lymphocyte %	13.10 [8.40, 19.60]	14.20 [9.50, 20.80]	9.80 [6.00, 15.00]	<0.001	365 (7.0)
Auto Eosinophil #	0.00 [0.00, 0.02]	0.00 [0.00, 0.02]	0.00 [0.00, 0.01]	<0.001	365 (7.0)
Auto Eosinophil %	0.00 [0.00, 0.20]	0.00 [0.00, 0.30]	0.00 [0.00, 0.10]	<0.001	365 (7.0)
Auto Monocyte #	0.46 [0.32, 0.66]	0.46 [0.32, 0.65]	0.47 [0.30, 0.68]	0.668	365 (7.0)
Auto	6.70 [4.60, 9.40]	7.00 [4.90, 9.60]	5.85 [3.90, 8.30]	<0.001	365

Monocyte %					(7.0)
Hemoglobin	13.30 [12.00, 14.50]	13.40 [12.20, 14.50]	12.70 [11.10, 14.30]	<0.001	194 (3.7)
Red Cell Distribution Width	13.50 [12.70, 14.60]	13.30 [12.60, 14.30]	14.30 [13.30, 15.70]	<0.001	196 (3.8)
Platelet Count - Automated	201.00 [156.00, 258.00]	206.00 [161.00, 263.00]	189.00 [142.00, 243.00]	<0.001	203 (3.9)
Sodium, Serum	136.00 [133.00, 139.00]	136.00 [133.00, 138.00]	137.00 [133.00, 141.00]	<0.001	226 (4.3)
Potassium, Serum	4.10 [3.70, 4.50]	4.00 [3.70, 4.40]	4.30 [3.80, 4.80]	<0.001	276 (5.3)
Chloride, Serum	99.00 [96.00, 103.00]	99.00 [96.00, 102.00]	100.00 [96.00, 106.00]	<0.001	227 (4.3)
Carbon Dioxide, Serum	24.00 [22.00, 26.00]	24.00 [22.00, 26.00]	22.00 [19.00, 25.00]	<0.001	227 (4.3)
Blood Urea Nitrogen, Serum	16.00 [11.00, 28.00]	15.00 [10.00, 21.00]	31.00 [20.00, 49.50]	<0.001	227 (4.3)
Creatinine, Serum	1.03 [0.82, 1.44]	0.97 [0.79, 1.25]	1.48 [1.06, 2.40]	<0.001	227 (4.3)
eGFR if Non-African American	70.00 [43.50, 91.00]	77.00 [54.00, 95.00]	41.00 [22.00, 63.00]	<0.001	227 (4.3)
eGFR if African American	81.00 [50.00, 106.00]	89.00 [63.00, 110.00]	48.00 [25.00, 73.50]	<0.001	227 (4.3)
Glucose, Serum	121.00 [105.00, 159.00]	118.00 [104.00, 147.25]	141.00 [115.00, 198.00]	<0.001	227 (4.3)
Albumin, Serum	3.50 [3.00, 3.80]	3.50 [3.10, 3.90]	3.20 [2.80, 3.70]	<0.001	272 (5.2)
Bilirubin Total, Serum	0.50 [0.40, 0.70]	0.50 [0.40, 0.70]	0.50 [0.40, 0.80]	<0.001	272 (5.2)

Alkaline Phosphatase, Serum	73.00 [58.00, 93.00]	71.00 [57.00, 91.00]	78.00 [62.00, 103.00]	<0.001	284 (5.4)
Alanine Aminotransferase (ALT/SGPT)	45.00 [30.00, 69.00]	43.00 [30.00, 65.00]	53.00 [34.00, 83.00]	<0.001	302 (5.8)
Aspartate Aminotransferase (AST/SGOT)	33.00 [21.00, 55.00]	33.00 [22.00, 57.00]	30.00 [20.00, 48.00]	<0.001	302 (5.8)
pH, Arterial	7.43 [7.34, 7.47]	7.45 [7.41, 7.48]	7.39 [7.31, 7.44]	<0.001	5004 (95.6)
pCO2, Arterial	34.00 [30.00, 40.00]	33.00 [29.00, 38.00]	34.00 [30.00, 43.00]	0.074	4882 (93.3)
pO2, Arterial	74.00 [63.00, 112.00]	71.00 [63.00, 107.00]	76.00 [61.00, 119.00]	0.503	4882 (93.3)
Ferritin, Serum	769.00 [412.00, 1419.00]	734.40 [396.75, 1307.00]	930.00 [468.00, 1960.00]	<0.001	4188 (80.0)
Lactate	1.70 [1.30, 2.20]	1.60 [1.20, 2.00]	2.10 [1.50, 3.10]	<0.001	1821 (34.8)
Lactate Dehydrogenase, Serum	406.00 [317.00, 549.00]	387.00 [305.00, 501.00]	520.00 [385.00, 690.50]	<0.001	3920 (74.9)
D-Dimer Assay, Quantitative	407.00 [246.00, 767.00]	352.00 [226.00, 608.00]	700.50 [382.00, 1928.75]	<0.001	4100 (78.3)
Procalcitonin, Serum	0.18 [0.09, 0.43]	0.14 [0.08, 0.31]	0.46 [0.20, 1.41]	<0.001	3904 (74.6)
Prothrombin Time, Plasma	13.40 [12.50, 14.70]	13.30 [12.40, 14.40]	13.80 [12.67, 15.80]	<0.001	2756 (52.7)
Troponin I, Serum	0.02 [0.01, 0.06]	0.01 [0.01, 0.03]	0.07 [0.01, 0.24]	<0.001	4342 (83.0)
Troponin T, High	13.00 [7.00, 36.00]	10.50 [6.00, 18.00]	42.00 [20.00, 129.00]	<0.001	4868 (93.0)

Sensitivity					
Troponin T, High Sensitivity Result	12.00 [6.00, 30.00]	10.00 [6.00, 18.00]	60.00 [24.00, 113.00]	<0.001	4980 (95.2)
Troponin T, Serum	0.01 [0.01, 0.02]	0.01 [0.01, 0.01]	0.02 [0.01, 0.07]	<0.001	4470 (85.4)

Table 2. Confusion matrices for multiple operating points for the five calculators tested on data from the final day of the testing week, ending April 12th.

Operating Points A (●)					
	A1 (DR NOCOS)	**A2 (FIX NOCOS)**	**A3 (SOFA+)**	**A4 (SOFA)**	**A5 (CURB-65)**
Probability of Survival / Score	48.3 %	37.4 %	N/A	5	4
Predicted to Survive & Survived	293	293	293	274	294
Predicted to Expire & Expired	54	56	56	60	38
Predicted to Survive & Expired	78	76	76	72	94
Predicted to Expire & Survived	18	18	18	37	17
Positive Predictive Value (%)	79.0	79,4	79.4	79.2	75.8
Negative Predictive Value (%)	75.0	75.7	75.7	61.9	69.1
Operating Points B (★)					
	B1 (DR NOCOS)	**B2 (FIX NOCOS)**	**B3 (SOFA+)**	**B4 (SOFA)**	**B5 (CURB-65)**
Probability of Survival / Score	91.4%	92.7%	N/A	3	1
Predicted to Survive & Survived	126	126	126	140	149
Predicted to Expire & Expired	126	127	125	112	108
Predicted to Survive & Expired	6	5	7	20	24
Predicted to Expire & Survived	185	185	185	171	162

Positive Predictive Value (%)	95.5	96.2	94.7	87.5	86.1
Negative Predictive Value (%)	40.5	40.7	40.3	39.6	40.0
Operating Points C (◇)					
	C1 (DR NOCOS)	**C2 (FIX NOCOS)**	**C3 (SOFA+)**	**C4 (SOFA)**	**C5 (CURB-65)**
Probability of Survival / Score	79.3%	79.1%	N/A	4	2
Predicted to Survive & Survived	238	238	238	274	261
Predicted to Expire & Expired	93	92	98	60	67
Predicted to Survive & Expired	39	40	34	72	65
Predicted to Expire & Survived	73	73	73	37	50
Positive Predictive Value (%)	85.9	85.6	87.5	79.2	80.1
Negative Predictive Value (%)	56.0	55.8	57.3	61.9	57.3