medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Development and Validation of a Survival Calculator for Hospitalized Patients with COVID-19

Running title: Survival Score for Hospitalized Patients with COVID-19

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; Stuart L. Cohen[bc], MD; Jamie S. Hirsch**[bcd], MD, MA, MSB; Theodoros P. Zanos**[ab], PhD

And the Northwell & Maimonides COVID-19 Research Consortium; Hannah Bodenstein[e], MBA, BSN, RN; Shubham Debnath[a], PhD; Andrew J. Dominello[c], BA; Louise Falzon[c], BA, PGDipInf; Michael Gitman[b], MD; Jay M. Goldstein[f], MBA; Crystal Herron[c], PhD; Eun-Ji Kim[bc], MD, MS; Lawrence Lau[b]; Zachary S. Lockerman[e], MD, MBA, FACG, CPE; Alexander Makhnevich[b], MD; Jazmin N. Mogavero[c], MA; Ernesto P. Molmenti[bc], MD, PhD, MBA, FACS; Marc d. Paradis[g], SM; Viktor Tóth[a], MSc

*Co-first authors

**Co-senior authors with equal contribution

[a]Institute of Bioelectronic Medicine, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[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]Maimonides Medical Center, Brooklyn, NY

[f]Management Information Systems, Maimonides Medical Center, Brooklyn, NY

[g]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: 2966

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background: Chinese studies reported predictors of severe disease and mortality associated with coronavirus disease 2019 (COVID-19). A generalizable and simple survival calculator based on data from US patients hospitalized with COVID-19 has not yet been introduced.

Objective: Develop and validate a clinical tool to predict 7-day survival in patients hospitalized with COVID-19.

Design: Retrospective and prospective cohort study.

Setting: Thirteen acute care hospitals in the New York City area.

Participants: Adult patients hospitalized with a confirmed diagnosis of COVID-19. The development and internal validation cohort included patients hospitalized between March 1 and May 6, 2020. The external validation cohort included patients hospitalized between March 1 and May 5, 2020.

Measurements: Demographic, laboratory, clinical, and outcome data were extracted from the electronic health record. Optimal predictors and performance were identified using least absolute shrinkage and selection operator (LASSO) regression with receiver operating characteristic curves and measurements of area under the curve (AUC).

Results: The development and internal validation cohort included 11□095 patients with a median age of 65 years [interquartile range (IQR) 54-77]. Overall 7-day survival was 89%. Serum blood urea nitrogen, age, absolute neutrophil count, red cell distribution width, oxygen saturation, and serum sodium were identified as the 6 optimal of 42 possible predictors of survival. These factors constitute the NOCOS (Northwell COVID-19 Survival) Calculator. Performance in the internal validation,

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

prospective validation, and external validation were marked by AUCs of 0.86, 0.82, and 0.82, respectively.

Limitations: All participants were hospitalized within the New York City area.

Conclusions: The NOCOS Calculator uses 6 factors routinely available at hospital admission to predict 7-day survival for patients hospitalized with COVID-19. The calculator is publicly available at https://feinstein.northwell.edu/NOCOS.

Trial registration: N/A

Funding Source: This work was supported by grants R24AG064191 from the National Institute on Aging, R01LM012836 from the National Library of Medicine, and K23HL145114 from the National Heart Lung and Blood Institute.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

INTRODUCTION

The World Health Organization designated coronavirus disease 2019 (COVID-19) as a global pandemic on March 11, 2020, with over 1 million confirmed cases worldwide (1). Estimates of severe disease range from 20% to 30%, and fatality rates range from 2% to 7% (2, 3). As healthcare facilities around the world struggle to provide care for rising numbers of critically ill patients, evidence-based tools to assist with prognosis and estimating disease severity are becoming increasingly important (4). These tools can guide conversations with patients and families, advise therapeutic decisions (e.g., admission to the intensive care unit), and align treatment plans with the likelihood of benefit (5).

Some clinical prediction tools have been established to estimate survival in patients with pneumonia or hospitalized with severe illness, including the Sequential Organ Failure Assessment (SOFA) and CURB-65 Scores. However, these tools have not been validated in patients with COVID-19. A recent study reported a clinical risk score for hospitalized patients with this disease (6). While the score relies on data routinely available at admission, it uses unstructured data, cannot predict survival alone, and is based on a cohort of patients hospitalized in China. Several other models predict outcomes in patients with COVID-19. However, many of these models are not peer-reviewed and are at a high risk of bias because of non-representative samples of control patients (7).

Our objective was to develop and externally validate a clinical prediction tool to estimate 7-day survival in patients hospitalized with COVID-19 in the United States. For this tool, we aimed to use exclusively discrete data points from the electronic health record (EHR), forgoing symptom-related records and radiology reads. By including only objective, structured data points that are routinely available

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

at hospital admission, we could reduce ambiguity, improve external performance, and ensure that the tool could be used in most acute-care settings.

METHODS

Study Design

This study includes a retrospective analysis for development, retrospective internal validation, prospective internal validation, and external validation of a model to predict survival of patients hospitalized with COVID-19. The development cohort included patients admitted to 11 of 12 acute care facilities in the Northwell Health system between March 1 and April 23, 2020. The internal retrospective validation cohort included patients admitted to the remaining acute care tertiary facility in the Northwell Health System, Long Island Jewish Medical Center, between March 1 and May 7, 2020. Long Island Jewish Medical Center has the largest number of hospitalized patients with COVID-19 in the Northwell Health system. The internal prospective validation cohort included patients admitted to all 12 acute care facilities in the Northwell Health system between April 24 and May 6, 2020. The external validation cohort included patients admitted to Maimonides Medical Center, an affiliate of the Northwell Health system, between March 1 and May 12, 2020. The final date of follow-up was May 7, 2020 for the internal validation cohorts and May 12, 2020 for the external validation cohort (Figure 1A).

Patients were included if they were adults (≥ 18 years old) admitted to the hospital with COVID-19 confirmed by a positive result from polymerase chain reaction testing of a nasopharyngeal sample. Clinical outcomes (i.e., discharges, mortality, length of stay) were monitored until the final date of follow-up. Patients were excluded if they received invasive mechanical ventilation before inpatient

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

admission, either before presentation to or during their stay in the emergency department. Patients were also excluded if their length of stay was less than 7 days and they were still hospitalized on the final date of follow-up.

This study was approved by the Institutional Review Boards at Northwell Health and Maimonides Medical Center as minimal-risk research that used data collected for routine clinical practice, and as such, waived the requirement for informed consent.

Data Acquisition

Data were collected from the enterprise electronic health record (EHR; Sunrise Clinical Manager, Allscripts, Chicago, IL). Transfers from 1 in-system hospital to another were merged and considered 1 hospital visit. Data collected for the development and internal validation of the tool included patient demographic information, comorbidities, laboratory values, and outcomes (i.e., death, length of stay, discharge). Data collected for the external validation included only the 6 predictor variables found in the development and internal validation process, length of stay, and final outcome (i.e., death, discharge).

Potential Predictive Variables

Potential predictive variables were included if they were available for more than half of study patients at the time of admission. This approach ensured that the results would contain data points routinely available at admission. Continuous variables are presented as median and interquartile range (IQR), and categorical variables are expressed as the number of patients and percentage.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Demographic variables included age, gender, race, ethnicity, and language preference as English or non-English. Vitals signs included systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, oxygen saturation, temperature, body mass index, height, and weight. Comorbidities included coronary artery disease, diabetes, hypertension, heart failure, lung disease, and kidney disease. Laboratory variables included white blood cell count, absolute neutrophil count, automated lymphocyte count, automated eosinophil count, automated monocyte count, hemoglobin, red cell distribution width, automated platelet count, serum sodium, serum potassium, serum chloride, serum carbon dioxide, serum blood urea nitrogen, serum creatinine, estimated glomerular filtration rate, serum glucose, serum albumin, serum bilirubin, serum alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, and lactate.

Outcomes

Outcomes collected included death, length of stay, and discharge. The primary outcome was 7-day survival. Patients who were discharged alive on any hospital day before or on hospital day 7 were considered to have survived. Patients who were alive and still in-hospital on hospital day 7 were considered to have survived. Patients who died before or on hospital day 7 were considered to have expired. Patients who were alive and still in-hospital at the study endpoint with a length of stay less than 7 days were excluded from the study.

Prediction Model Development

The model was developed by analyzing 42 potential predictors for the patients hospitalized in 11 of 12 hospitals within the Northwell Health system and discharged

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

on or before April 23, 2020 (Figure 1A). Patients hospitalized at Long Island Jewish Medical Center or discharged after April 23 were used for the internal retrospective and internal prospective validations, respectively. Least Absolute Shrinkage and Selection Operator (LASSO) regression was used to identify predictors that, when linearly combined, predict the survival of hospitalized patients with COVID-19 (8). Missing measurements were imputed using mean imputation. All analyses were performed in MATLAB 2019b (The Mathworks, Inc., Natick, MA).

By including an L_1-norm regularization term that promotes sparsity, LASSO regression determines a subset of measurements in which only the strongest predictors remain in the model. 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 are normalized by taking the z-score, which puts all measurements on the same scale. The mean and standard deviation of the measurements with coefficients that are not 0 are stored during training and applied to test data.

The training set was evaluated with the model using 50-fold cross-validation to prevent overfitting. The class-conditional likelihood functions of the LASSO predictions for survival past 7 days and expiration before 7 days were estimated, and the posterior probability of survival past 7 days was evaluated using Bayes Theorem. The regularization factor λ is a hyperparameter that was swept over a range while evaluating the area under the receiver operating characteristic (ROC) curve. After optimizing for λ, the number of predictors was fixed at 6 inputs. The variables identified were used to construct the Northwell COVID-19 Survival (NOCOS) Calculator, available publicly at https://feinstein.northwell.edu/NOCOS.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Prediction Model Validation

The generalizability of the NOCOS Calculator was validated with the retrospective cohort from Long Island Jewish Medical Center, the internal prospective cohort, and the external cohort. The predictive performance of the model was assessed at the time of admission and every 2 days within the hospitalization via ROC and precision recall (PR) curves and the Area Under the Curve (AUC). We also tested the predictive value of the SOFA Score and CURB-65 Score for pneumonia severity, and we compared the AUCs for each score using the nonparametric DeLong method (9, 10).

To determine the performance of survival predictions for all calculators, operating points can be established by choosing thresholds on the probability scores. We chose 3 operating points for each calculator and provided the numbers of true positives, true negatives, false positives, and false negatives, as well as the positive predictive value (PPV) and negative predictive value (NPV) for each calculator.

Calculation of SOFA and CURB-65 Scores

In 2 of the test datasets, data were assessed with the SOFA and CURB-65 Scores. The SOFA Score numerically quantifies the severity of failed organs based on PaO_2/FiO_2 (11), the Glasgow Coma Scale, mechanical ventilation (yes/no), platelets, bilirubin, mean arterial pressure or administration of vasoactive agents, and creatinine. Because PaO_2 was not captured for most patients, we used a formula with SpO_2 and FiO_2 (Appendix Table 4). For patients missing a Glasgow Coma Scale score, the patient's mental status was assessed using nursing documentation of mental status and level of consciousness. Missing data on all other variables were limited and treated similar to our imputations. No patients were missing data on ventilation status. The CURB-65 score is another mortality risk score

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

based on confusion, blood urea nitrogen, respiratory rate, blood pressure, and age. For this score, a patient's confusion level was assessed using nursing documentation.

Role of the Funding Source

This work was supported by grants R24AG064191 from the National Institute on Aging, R01LM012836 from the National Library of Medicine, and K23HL145114 from the National Heart Lung and Blood Institute. 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.

RESULTS

Patient Characteristics

A total of 11□919 adult patients were hospitalized at the 12 Northwell Health acute care facilities between March 1 and May 7, 2020. Of these patients, 360 (3.02%) were excluded because they were still in the hospital at the study end point with a length of stay less than 7 days; 460 (3.86%) were excluded because they were transferred to a hospital outside of the health system and their outcomes were unknown; and 4 (0.03%) were excluded because they expired but were not marked as discharged in the EHR. The remaining 11□095 (93.09%) patients were included in the development and internal validation cohort. These patients had a median age of 65 years [IQR 54-77], and 42% were female. Overall 7-day survival was 89%. At the study end point, 10□207 (92%) patients were discharged alive or expired and 888 (8%) were still in the hospital. Baseline characteristics of included patients are presented in Table 1.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Survival Prediction Model

Development

To determine the predictors of survival, data were collected from patients hospitalized in 11 of 12 Northwell Health hospitals (n = 6162) (Figure 1A). The optimal predictors of survival, in decreasing order of relative predictive strength, were serum blood urea nitrogen, patient age, absolute neutrophil count, red cell distribution width, oxygen saturation, and serum sodium. For each predictor, the magnitude of the coefficient indicates the relative strength of the predictor in determining the outcome, and the sign of the coefficient corresponds to the sign of the correlation between the predictor and the outcome (Figure 1B).

Validation

For the internal retrospective validation, data were collected from patients hospitalized in Long Island Jewish Medical Center (n = 2229). These data were analyzed with NOCOS, and then the ROC and PR curves and AUC values were determined. The NOCOS Calculator resulted in an AUC of 0.86, which significantly outperformed the SOFA (AUC = 0.70; $p < 0.05$) and CURB-65 (AUC = 0.81; $p < 0.05$) Scores (Figure 2A).

For the internal prospective validation, data were collected from patients discharged from all 12 Northwell hospitals (n = 3328) (Figure 1A). Based on these data, the NOCOS Calculator (AUC = 0.82) significantly outperformed the SOFA (AUC = 0.64; $p < 0.05$) and CURB-65 (AUC = 0.72; $p < 0.05$) Scores (Figure 2B).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

For the external validation, data were collected from patients hospitalized at Maimodines Medical Center (*n* = 2669) (Figure 1A). The NOCOS Calculator yielded a comparable AUC of 0.82 (Figure 2C); however, the SOFA and CURB-65 Scores were not readily available for this dataset because components of each score were not documented for all patients.

Test Characteristics of the NOCOS Calculator, SOFA Score, and CURB-65 Score

To determine the performance of survival predictions for all calculators, operating points can be established by choosing thresholds on the probability scores. We chose 3 different operating points for each calculator for the internal retrospective validation (Long Island Jewish Medical Center) data and provided the number of true positives, true negatives, false positives, and false negatives. We also provided the PPV and NPV for each case (Table 2). In all cases, the NOCOS Calculator outperformed the SOFA and CURB-65 Scores. See Appendix Table 2 for metrics at the operating points of data from the internal retrospective validation (Long Island Jewish Medical Center), and see Appendix Table 3 for metrics from the external validation (Maimonides Medical Center).

The NOCOS Calculator can also be reevaluated with updated labs and vitals as the patient's condition progresses. The performance of the NOCOS calculator remained relatively stable over 10 days (Figure 3).

DISCUSSION

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

This study is the first to develop a model, the NOCOS Calculator, that predicts survival of patients hospitalized with COVID-19 in the United States. We created and validated the NOCOS Calculator with data on almost 14 000 patients, using only 6 clinical data points typically available to clinicians within the first 60 minutes of patient presentation. All data points are available as discrete inputs in most commercial EHRs, supporting that this calculator could be readily incorporated into tools to support clinical decisions. This calculator is publicly available at https://feinstein.northwell.edu/NOCOS and can be used by clinicians to estimate the probability of survival for their patients.

Several elements of the NOCOS Calculator have been either established as prognostic markers in other populations or identified as risk factors for severe illness or death in patients with COVID-19. Older age and elevated blood urea nitrogen (a marker of kidney dysfunction) have both been associated with increased mortality risk in patients with COVID-19.(12, 13) Hypoxemia, measured by lower levels of blood oxygen, has also been linked to increased mortality in this population (14). Neutrophil count, either individually or paired in a ratio with lymphocytes, also predicts disease severity in COVID-19 patients (15). While serum sodium has not yet been linked to COVID-19, it (16) has been independently and consistently associated with negative outcomes in other populations (17) and disease states.(18, 19) Elevated values of red cell distribution width often suggest chronic disease states and inflammation.(20, 21) An increased red cell distribution width may also be an effect of COVID-19 on iron displacement of the heme molecule, leading to impaired red blood cells, free-radical formation, and a toxic effect to the lungs.(22)

The NOCOS Calculator performs well with the 6 early measurements, and it retains its predictive performance as these measurements are updated over at least

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

10 days throughout the hospitalization of the patient (Figure 3). This finding supports that the most up-to-date values of the 6 measures can accurately predict survival. Moreover, while we present the calculator output as a probability score, a specific operating point can be chosen to provide a binary outcome prediction with significant accuracy. Stakeholders can choose an operating point, and local clinical teams can adjust thresholds toward a more stringent or risk-averse solution (Table 2) based on the rapidly changing needs during this pandemic.

Due to the challenges that arise during the ongoing COVID-19 pandemic, we need robust tools to aid in making complex clinical decisions. Using well-known clinical calculators, such as the SOFA or CURB-65 Scores, can be useful. However, these scores are limited by their accuracy and the ease of collecting necessary measurements to construct the scores. They also use input variables, such as confusion (for CURB-65) and the Glasgow Coma Scale (for SOFA), both of which are ambiguous, difficult to measure, and frequently unavailable, as shown by our external validation dataset. We found that the NOCOS Calculator consistently outperformed both the CURB-65 and SOFA scores in both our validation datasets.

Limitations

The study population only included patients within the New York City metropolitan area. However, given the diverse demographic population of the region, we expect the model to generalize to patients at centers outside of this geographic area. The data were collected entirely from EHR reports, which supported robust and rapid analysis of a large cohort of patients. However, we did not include data elements that would require manual chart review. Due to the retrospective study design, not all laboratory tests were completed on all patients, and the performance

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

of these variables could not be adequately assessed. To optimize for usability and portability, the analysis was designed to be linear and to include a minimum number of predictors. Non-linear or convolutional/recurrent models may provide improved performance but might not be easily used at all centers.

Conclusion

This study is the first to develop and externally validate a simple predictive model of survival for hospitalized patients with COVID-19 based on structured, objective data that is routinely available at admission in the United States. Serum blood urea nitrogen, age, absolute neutrophil count, red cell distribution width, oxygen saturation, and serum sodium were identified as the 6 optimal predictors of survival. The NOCOS Calculator can predict survival more accurately than commonly used survival predictors.

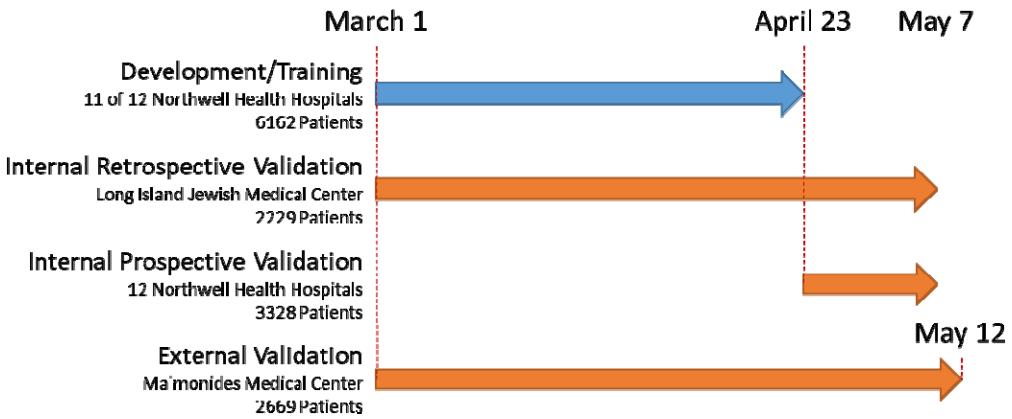
medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgments

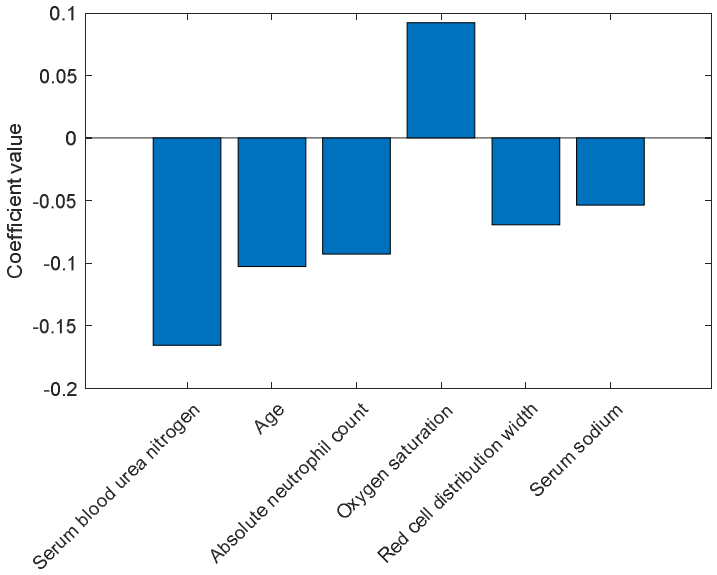
We acknowledge and honor all our Northwell team members who consistently put themselves in harm's way during the COVID-19 pandemic. We dedicate this article to them, as their vital contribution to knowledge about COVID-19 and sacrifices on the behalf of patients made it possible.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figures



A



B

Figure 1. Study design and predictive performance of the NOCOS Calculator. (A) Training and validation datasets used to create and test the predictive performance of the NOCOS Calculator. The datasets include internal retrospective, internal prospective, and external datasets comprising 13□764 patients. Note that the internal retrospective and prospective validation sets overlap by 624 patients admitted to Long Island Jewish Medical Center. (B) Importance of the 6 predictors captured by the coefficients of the NOCOS Calculator. NOCOS = Northwell COVID-19 Survival.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

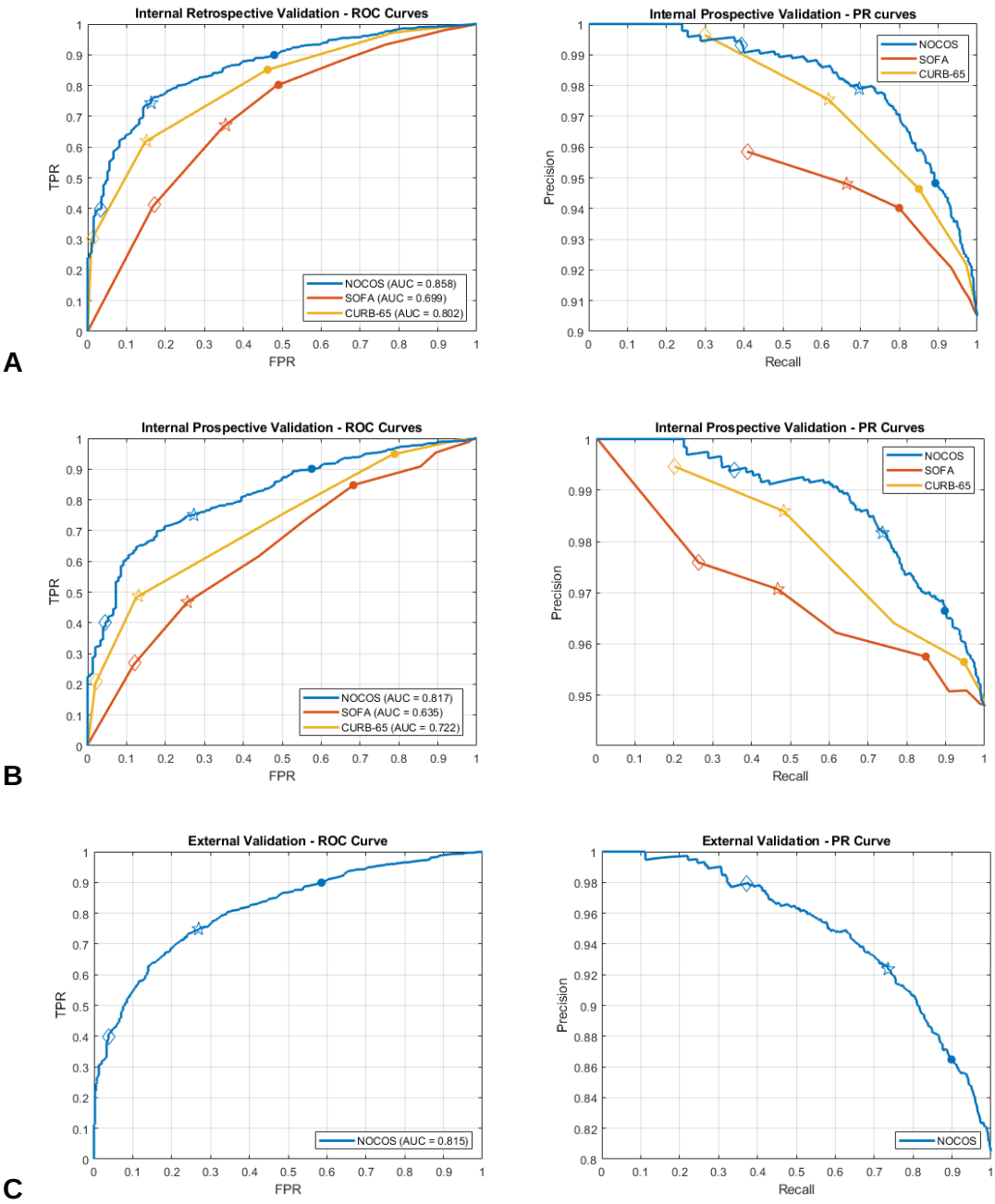


Figure 2. Predictive performance of the NOCOS Calculator on the internal retrospective, internal prospective, and external datasets. ROC and PR curves for the (A) internal retrospective validation with patients hospitalized at Long Island Jewish Hospital (*n* = 2229), (B) prospective validation with patients hospitalized across all 12 Northwell Health hospitals (*n* = 3328), and (C) external validation with patients hospitalized at Maimonides Medical Center (*n* = 2669). AUC = area under

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the curve; FPR = false positive rate; NOCOS = Northwell COVID-19 Survival; PR = precision recall; ROC = receiver operating characteristic; SOFA = Sequential Organ Failure Assessment; TPR = true positive rate.

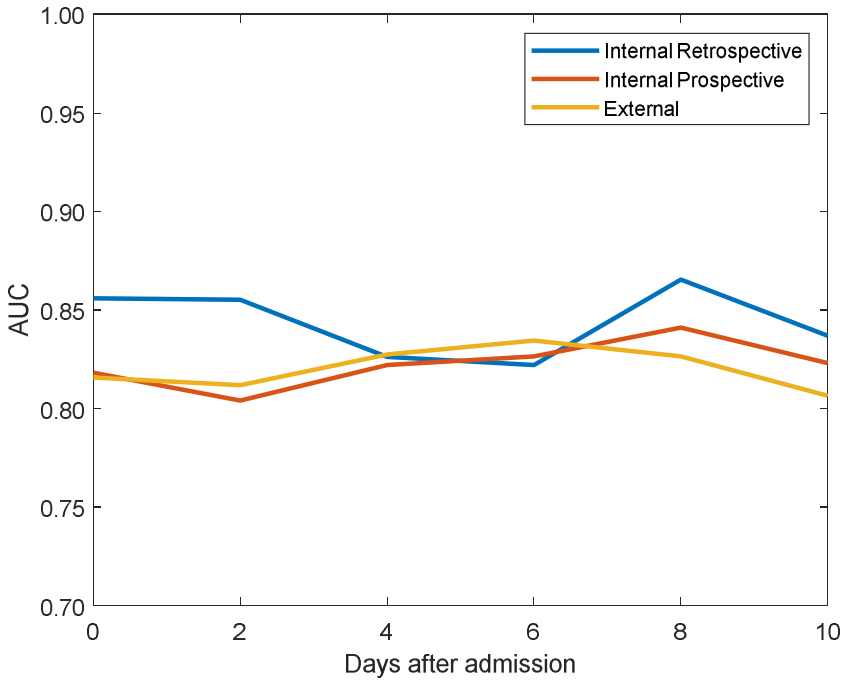


Figure 3. Stability of the predictive performance of the NOCOS Calculator across hospitalization days. AUC capturing the predictive performance of the NOCOS Calculator. The performance remained stable when the most recent values of the predictors were updated for up to 10 days during hospitalization after admission from the emergency department. AUC = Area Under the Curve; NOCOS = Northwell COVID-19 Survival.

References

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1. The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [Internet]. Baltimore, MD: Center for Systems Science and Engineering; 2020 [cited 2020 May 29]. Available from: https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.
2. Livingston E, Bucher K**.** Coronavirus Disease 2019 (COVID-19) in Italy. JAMA. Epub 2020 Mar 17.
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 72314 Cases From the Chinese Center for Disease Control and Prevention. JAMA. Epub 2020 Feb 24.
4. Myers J, Cosby R, Gzik D, et al. Provider Tools for Advance Care Planning and Goals of Care Discussions: A Systematic Review. Am J Hosp Palliat Care. 2018;35(8):1123-32.
5. Brajer N, Cozzi B, Gao M, et al. Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. JAMA Netw Open. 2020;3(2):e1920733.
6. Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Intern Med. 2020;12: e202033.
7. Wynants L, Van Calster B, Bonten MM, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ. 2020;369:m1328.
8. Tibshirani R**.** Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Series B Methodol. 1996;58(1):267-88.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

9. 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):837-45.

10. Sun X, Xu W**.** Fast Implementation of DeLong's Algorithm for Comparing the Areas Under Correlated Receiver Operating Characteristic Curves. IEEE Signal Process Lett. 2014;21(11):1389-93.

11. Jones AE, Trzeciak S, Kline JA**.** The Sequential Organ Failure Assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. Crit Care Med. 2009;37(5):1649-54.

12. 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):1054-62.

13. Cheng Y, Luo R, Wang K, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. Kidney Int. 2020;97(5):829-38.

14. Xie J, Covassin N, Fan Z, et al. Association Between Hypoxemia and Mortality in Patients With COVID-19. Mayo Clin Proc. Epub 2020 Apr 11.

15. Shang W, Dong J, Ren Y, et al. The value of clinical parameters in predicting the severity of COVID-19. J Med Virol. Epub 2020 May 22.

16. Steer J, Gibson J, Bourke SC**.** The DECAF Score: predicting hospital mortality in exacerbations of chronic obstructive pulmonary disease. Thorax. 2012;67(11):970-6.

17. Zhang X, Li XY**.** Prevalence of hyponatremia among older inpatients in a general hospital. Eur Geriatr Med. Epub 2020 May 7.

18. 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

19. Ruf AE, Kremers WK, Chavez LL, et al. Addition of serum sodium into the MELD score predicts waiting list mortality better than MELD alone. Liver Transpl. 2005;11(3):336-43.

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

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

22. Liu W, Li H**.** COVID-19: Attacks the 1-Beta Chain of Hemoglobin and Captures the Porphyrin to Inhibit Human Heme Metabolism. ChemRxiv [Preprint] 2020 [cited 2020 May 29]. Available from: https://doi.org/10.26434/chemrxiv.11938173.v5

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1. Demographic, Clinical, and Laboratory Data of All Patients Hospitalized at Northwell Health

	All Included Patients	**Alive 7 Days**	**Expired 7 Days**	**Missing No. (%)**
n	11095	8499	2596	0 (0)
Female (%)	4584 (41.3)	3641 (42.8)	943 (36.3)	
Male (%)	6511 (58.7)	4858 (57.2)	1653 (63.7)	
Age, y (%)				
21-40	977 (8.8)	919 (10.8)	58 (2.2)	0 (0)
41-60	3334 (30.0)	2952 (34.7)	382 (14.7)	
61-80	4774 (43.0)	3512 (41.3)	1262 (48.6)	
81-106	2010 (18.1)	1116 (13.1)	894 (34.4)	
Race (%)				
Asian	952 (8.6)	710 (8.4)	242 (9.3)	0 (0)
Black	2336 (21.1)	1869 (22.0)	467 (18.0)	
Declined	76 (0.7)	67 (0.8)	9 (0.3)	
Other	3048 (27.5)	2449 (28.8)	599 (23.1)	
Unknown	433 (3.9)	346 (4.1)	87 (3.4)	
White	4250 (38.3)	3058 (36.0)	1192 (45.9)	
Ethnicity (%)				
Declined	38 (0.3)	35 (0.4)	3 (0.1)	0 (0)
Hispanic or Latino	2371 (21.4)	1918 (22.6)	453 (17.4)	
Not Hispanic or Latino	8001 (72.1)	6000 (70.6)	2001 (77.1)	
Unknown	685 (6.2)	546 (6.4)	139 (5.4)	
English primary language (%)	8884 (80.1)	6781 (79.8)	2103 (81.0)	0 (0)
Length of stay, days (median [IQR])	6.32 [3.31, 11.08]	5.97 [3.16, 10.20]	7.81 [3.86, 14.18]	0 (0)
Required mechanical ventilation (%)	2250 (20.3)	925 (10.9)	1325 (51.0)	0 (0)
Last emergency department vital sign measurement (median [IQR])				
Systolic blood pressure, mmHg	127.00 [113.00, 142.00]	127.00 [115.00, 142.00]	126.00 [110.00, 142.00]	52 (0.5)
Diastolic blood pressure, mmHg	73.00 [64.00, 82.00]	74.00 [65.00, 82.00]	70.00 [61.00, 79.00]	52 (0.5)
Heart rate, beats per minute	91.00 [80.00, 103.00]	91.00 [80.00, 103.00]	92.00 [79.00, 105.00]	39 (0.4)
Respiratory rate, breaths per minute	20.00 [18.00, 24.00]	20.00 [18.00, 22.00]	22.00 [19.00, 26.00]	83 (0.7)
Oxygen saturation, %	96.00 [94.00, 98.00]	96.00 [94.00, 98.00]	96.00 [93.00, 98.00]	153 (1.4)
Temperature, Celsius	37.30 [36.90, 38.10]	37.30 [36.90, 38.00]	37.30 [36.80, 38.20]	189 (1.7)
Body mass index, kg/m^2	28.20 [24.90, 32.40]	28.30 [25.10, 32.60]	27.40 [23.90, 31.70]	5293 (47.7)
Height, cm	167.64 [160.02, 175.26]	167.64 [160.02, 175.26]	167.64 [162.56, 175.26]	5005 (45.1)
Weight, kg	79.80 [68.00,	81.60 [68.00,	77.10 [65.80,	4061 (36.6)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

	93.00]	94.08]	90.70]	
Comorbidities, %				
Coronary artery disease	1042 (9.9)	664 (8.2)	378 (15.4)	0 (0)
Diabetes	3258 (31.0)	2387 (29.6)	871 (35.5)	
Hypertension	5522 (52.5)	4005 (49.6)	1517 (61.8)	
Heart failure	535 (5.1)	318 (3.9)	217 (8.8)	
Lung disease	1287 (12.2)	954 (11.8)	333 (3.6)	
Kidney disease	820 (7.8)	525 (6.5)	295 (2.0)	
Last emergency department laboratory result (median [IQR])				
White blood cell count, K/µL	7.44 [5.53, 10.14]	7.25 [5.41, 9.72]	8.32 [5.97, 11.48]	462 (4.2)
Absolute neutrophil, No., K/µL	5.82 [4.07, 8.37]	5.59 [3.92, 7.92]	6.79 [4.67, 9.64]	778 (7.0)
Automated neutrophil, %	78.70 [71.00, 84.40]	77.70 [70.00, 83.50]	82.00 [74.60, 87.00]	768 (6.9)
Automated lymphocyte, No., K/µL	0.89 [0.62, 1.24]	0.92 [0.66, 1.27]	0.77 [0.53, 1.11]	778 (7.0)
Automated lymphocyte, %	12.10 [7.70, 18.00]	12.90 [8.50, 19.00]	9.50 [6.00, 14.60]	768 (6.9)
Automated eosinophil, No., K/µL	0.00 [0.00, 0.02]	0.00 [0.00, 0.02]	0.00 [0.00, 0.01]	787 (7.1)
Automated eosinophil, %	0.00 [0.00, 0.20]	0.00 [0.00, 0.30]	0.00 [0.00, 0.10]	776 (7.0)
Automated monocyte, No., K/µL	0.46 [0.31, 0.67]	0.46 [0.32, 0.67]	0.45 [0.30, 0.67]	778 (7.0)
Automated monocyte, %	6.30 [4.30, 8.80]	6.60 [4.50, 9.00]	5.50 [3.70, 8.00]	768 (6.9)
Hemoglobin, g/dL	13.20 [11.80, 14.40]	13.20 [11.90, 14.40]	12.90 [11.30, 14.30]	462 (4.2)
Red cell distribution width, %	13.60 [12.80, 14.60]	13.40 [12.70, 14.50]	14.10 [13.20, 15.40]	469 (4.2)
Automated platelet count, K/µL	211.00 [161.00, 276.00]	214.00 [164.00, 281.00]	197.00 [147.00, 258.00]	477 (4.3)
Serum sodium, mmol/L	136.00 [133.00, 139.00]	136.00 [133.00, 139.00]	137.00 [133.00, 141.00]	528 (4.8)
Serum potassium, mmol/L	4.10 [3.70, 4.50]	4.00 [3.70, 4.40]	4.20 [3.80, 4.80]	618 (5.6)
Serum chloride, mmol/L	99.00 [95.00, 103.00]	99.00 [95.00, 103.00]	100.00 [95.00, 105.00]	527 (4.8)
Serum carbon dioxide, mmol/L	24.00 [21.00, 26.00]	24.00 [21.00, 26.00]	23.00 [20.00, 25.00]	523 (4.7)
Serum blood urea nitrogen,	18.00 [12.00, 31.00]	16.00 [11.00, 25.00]	30.00 [18.00, 50.00]	524 (4.7)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

mg/dL				
Serum creatinine, mg/dL	1.06 [0.82, 1.51]	1.00 [0.80, 1.34]	1.40 [1.00, 2.30]	523 (4.7)
eGFR if Non-African American, mL/min/1.73 M2	67.00 [41.00, 90.00]	73.00 [48.00, 93.00]	45.00 [24.00, 70.00]	522 (4.7)
eGFR if African American, mL/min/1.73 m^2	77.00 [47.00, 104.00]	84.00 [56.00, 108.00]	52.00 [28.00, 81.00]	522 (4.7)
Serum glucose, mg/dL	126.00 [108.00, 169.00]	123.00 [107.00, 161.00]	141.00 [114.00, 197.00]	523 (4.7)
Serum albumin, g/dL	3.40 [3.00, 3.80]	3.50 [3.00, 3.80]	3.20 [2.80, 3.60]	598 (5.4)
Total serum bilirubin, mg/dL	0.50 [0.40, 0.70]	0.50 [0.40, 0.70]	0.50 [0.40, 0.80]	599 (5.4)
Serum alkaline phosphatase, U/L	75.00 [59.00, 98.00]	74.00 [59.00, 96.00]	78.00 [61.00, 104.00]	617 (5.5)
Alanine aminotransferase (ALT/SGPT), U/L	33.00 [21.00, 55.00]	34.00 [21.00, 57.00]	31.00 [20.00, 50.00]	657 (5.9)
Aspartate aminotransferase (AST/SGOT), U/L	46.00 [31.00, 72.00]	44.00 [30.00, 68.00]	54.00 [36.00, 83.00]	651 (5.9)
Lactate, mmol/L	1.80 [1.30, 2.40]	1.70 [1.30, 2.20]	2.10 [1.50, 3.20]	3987 (35.9)

eGFR = estimated glomerular filtration rate using the CKD-EPI equation; IQR = interquartile range.

Table 2. Confusion Matrices for Multiple Operating Points for the 3 Calculators Tested on Data from the Long Island Jewish Medical Center Dataset.

Operating Points A (●)			
	A1 (NOCOS)	**A2 (SOFA)**	**A3 (CURB-65)**
Operating probability of survival/score	51.6%	3	2
Predicted to survive & survived	1667	1493	1588
Predicted to expire & expired	105	101	106
Predicted to survive & expired	91	95	90
Predicted to expire & survived	201	375	280
Positive predictive value, %	94.8	94.0	94.6
Negative predictive value, %	34.3	21.2	27.5
Operating Points B (★)			
	B1 (NOCOS)	**B2 (SOFA)**	**B3 (CURB-65)**
Operating probability of survival/score	75.4%	2	1

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Predicted to survive & survived	1301	1240	1153
Predicted to expire & expired	168	128	167
Predicted to survive & expired	28	68	29
Predicted to expire & survived	567	628	715
Positive predictive value, %	97.9	94.8	97.5
Negative predictive value, %	22.9	16.9	18.9
Operating Points C (◇)			
	C1 (NOCOS)	**C2 (SOFA)**	**C3 (CURB-65)**
Operating probability of survival/score	93.1%	1	0
Predicted to survive & survived	734	762	558
Predicted to expire & expired	191	163	194
Predicted to survive & expired	5	33	2
Predicted to expire & survived	1134	1106	1310
Positive predictive value, %	99.3	95.8	99.6
Negative predictive value, %	14.4	12.8	12.9

NOCOS = Northwell COVID-19 Survival; SOFA = Sequential Organ Failure Assessment.

Appendix Table 1. Summary of the 6 Predictor Variables from External Validation Data from Patients hospitalized at Maimonides Medical Center

		All Discharged Patients	Discharged alive	Expired	Missing No. (%)
N		2649	1836	813	
Age (%)					
21-40		226 (8.5)	215 (8.1)	11 (0.4)	
41-60		483 (18.1)	372 (13.9)	111 (4.2)	
61-80		1049 (39.3)	684 (25.6)	365 (13.7)	
81-106		708 (26.5)	379 (14.2)	329 (12.3)	
Length of Stay, days (median [IQR])		5.15 [2.70, 9.59]	4.84 [2.51, 8.81]	6.51 [3.25, 11.80]	0 (0)
Oxygen saturation (%)		96 [93, 98]	96 [94, 98]	95 [92, 98]	26 (0.2)
Absolute neutrophil #		5.57 [3.81, 8.13]	5.27 [3.62, 7.51]	6.93 [4.65, 10.08]	389 (3.3)
Red cell distribution width		13.6 [12.9, 14.7]	13.5 [12.8, 14.5]	14.2 [13.3, 15.5]	158 (1.3)
Serum sodium, mmol/L		137 [134, 140]	137 [134, 140]	138 [134, 142]	262 (2.2)
Serum blood urea nitrogen, mg/dL		17 [12, 30]	15 [10, 24]	30 [18, 52]	262 (2.2)

IQR = interquartile range.

Appendix Table 2. Confusion Matrices for Multiple Operating Points for the 3 Calculators Tested on the Prospective Dataset.

Operating Points A (●)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

	A1 (NOCOS)	**A2 (SOFA)**	**A3 (CURB-65)**
Probability of survival/score	29.9%	5	3
Predicted to Survive & survived	2480	2345	2617
Predicted to expire & expired	66	48	33
Predicted to survive & expired	86	104	119
Predicted to expire & survived	283	418	146
Positive predictive value, %	96.6	95.7	95.7
Negative predictive value, %	18.9	10.3	18.4
Operating Points B (★)			
	B1 (NOCOS)	**B2 (SOFA)**	**B3 (CURB-65)**
Probability of survival/score	55.0%	2	1
Predicted to survive & survived	2037	1292	1334
Predicted to expire & expired	114	113	133
Predicted to survive & expired	38	39	19
Predicted to expire & survived	726	1471	1429
Positive predictive value, %	98.2	97.1	98.6
Negative predictive value, %	13.6	7.1	8.5
Operating Points C (◇)			
	C1 (NOCOS)	**C2 (SOFA)**	**C3 (CURB-65)**
Probability of survival/score	82.1%	1	0
Predicted to survive & survived	986	729	558
Predicted to expire & expired	146	134	149
Predicted to survive & expired	6	18	3
Predicted to expire & survived	1777	2034	2205
Positive predictive value, %	99.4	97.6	99.5
Negative predictive value, %	7.6	6.2	6.3

NOCOS = Northwell COVID-19 Survival; SOFA = Sequential Organ Failure Assessment.

Appendix Table 3. Confusion Matrices for Multiple Operating Points for the NOCOS Calculator Tested on the External Dataset.

Operating Points A (●)	
	A1 (NOCOS)
Probability of survival	36.6%
Predicted to survive & survived	1490
Predicted to expire & expired	168
Predicted to survive & expired	233

medRxiv preprint doi: https://doi.org/10.1101/2020.04.22.20075416; this version posted June 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Predicted to expire & survived	168
Positive predictive value, %	86.5
Negative predictive value, %	50.0
Operating Points B (★)	
	B1 (NOCOS)
Probability of survival	62.1%
Predicted to survive & survived	1220
Predicted to expire & expired	300
Predicted to survive & expired	101
Predicted to expire & survived	438
Positive predictive value, %	92.4
Negative predictive value, %	40.7
Operating Points C (◇)	
	C1 (NOCOS)
Probability of survival	91.3%
Predicted to survive & survived	618
Predicted to expire & expired	388
Predicted to survive & expired	13
Predicted to expire & survived	1040
Positive predictive value, %	97.9
Negative predictive value, %	27.2

NOCOS = Northwell COVID-19 Survival; SOFA = Sequential Organ Failure Assessment.

Appendix Table 4. Determination of PaO_2/FiO_2 Score Based on SpO_2 and FiO_2 Values.

$(SpO_2/FiO_2)*100$	**PaO_2/FiO_2 Score**
Null	0
≥ 302.3	0
Between 221.2 and 302.2	1
Between 142.3 and 221.1	2
Between 67 and 142.2	3
< 67	4