A Pandemic's Lifecycle: A Longitudinal Validation of Predictors in 26,249 Severe COVID-19 NYC Admissions

G. Crowley[1], S. Kwon[1], J. Kim[1], M. Liu[2], A. Nolan[1]; [1]Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, NYU Grossman School of Medicine, New York, NY, United States, [2]Division of Biostatistics, Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States.

Corresponding author's email: george.crowley@nyumc.org

RATIONALE. Coronavirus Disease-2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has yielded a pandemic that has threatened our health, and depleted resources. The resulting mass casualties made obvious the need for early decision making in distribution of critical care resources. Given the strong association of hospital over-occupancy and all-cause mortality, clinicians need a dedicated tool to streamline triage/resource allocation and reduce clinical deliberation time.METHODS. A retrospective chart review of non-hospice admissions (n=26,249) for COVID-19 positive adult patients admitted between March 2020-April 2021 was conducted. Biomarker profiles and clinical phenotype collected as part of their standard of care (n=63 variables) was analyzed using random forests for variable selection via a novel modified hamming distance between variable importance rankings of models with identical hyperparameters. We applied a novel analysis pipeline to COVID-19 patients (n=6,606) admitted to NYU Langone Health March 2020-April 2020 with respect to ICU admission and—independently—mortality to discover the most important variables (n=15 per outcome) and used gradient-boosted decision trees to build predictive models of ICU admission and mortality based on these variables. Models were validated prospectively on patients (n=19,643) admitted after the 1st wave (May 2020-April 2021). All data was collected in compliance with the Code of Federal Regulations, Title 21, Part 11 and approved by the NYU IRB#20-00473.RESULTS. The classifier of ICU admission had 5-fold cross-validated AUC_{ROC} = 0.85 on training data, and the most important variables in predicting this endpoint included procalcitonin, respiratory rate, and lactate. The classifier of mortality had 5-fold cross-validated AUC_{ROC} = 0.87 on training data, and the most important variables included age, blood urea nitrogen, and troponin. We then validated these models prospectively on the next year of data and visualized model performance over time. Additionally, we analyzed how risk of individual predictors varied longitudinally as disease presentation and standard of care evolved.CONCLUSION. Our study leverages a NYC cohort of well-phenotyped COVID-19 patients from the primary surge and subsequent year. We have presented evidence that novel biomarkers and their combinations may be important in assessment and triage of COVID-19. Finally, we have analyzed how our predictors and overall model have evolved over the course of the pandemic and found that the performance of some plateaued while others rose or declined.

This abstract is funded by: NHLBI R01HL119326, CDC/NIOSH U01-OH011300, U01-OH011855, U01-OH012069