A Scoring Tool and Predictive Model to Detect Risk of Hospitalization from COVID-19

Nicholas Russo BS
USF MCOM- LVHN Campus, nicholas.russo@lvhn.org

Zhe Chen MD
Lehigh Valley Health Network, zhe.chen@lvhn.org

Matthew M. Miller DO
Lehigh Valley Health Network, Matthew_M.Miller@lvhn.org

Robert X. Murphy JR, MD, MS
Lehigh Valley Health Network, Robert.Murphy@lvhn.org

David B. Burmeister DO, FACEP, CPE
Lehigh Valley Health Network, david_b.burmeister@lvhn.org

Published In/Presented At

Russo, N., Chen, Z., Miller, M., Murphy, R., & Burmeister, D. (2021, February). *A Scoring Tool and Predictive Model to Detect Risk of Hospitalization from COVID-19.* [Poster presentation]. Health Datapalooza and National Health Policy Conference, Virtual.

A Scoring Tool and Predictive Model to Detect Risk of Hospitalization from COVID-19

Nicholas W. Russo, BS;[a] Zhe Chen, MD;[a] Matthew M. Miller, DO, MBA, MS;[a] Robert X. Murphy Jr., MD;[b] David B. Burmeister, DO, MBA[a]

[a]Lehigh Valley Health Network Department of Emergency and Hospital Medicine/University of South Florida Morsani College of Medicine, Lehigh Valley Campus, Allentown, Pa.

[b]Lehigh Valley Health Network, Department of Surgery, Division of Plastic Surgery, Allentown, Pa.

Health or Health Care Problem/Challenge Being Tackled

COVID-19 has caused an unprecedented global health emergency. While the typical intensive care unit occupancy is 60-80%, the strains of a pandemic such as that of COVID-19 can overwhelm hospital capacity. Efficient clinical decision-making is crucial for proper health care resource utilization in this crisis.

Innovation/Advance Achieved

We created a machine learning (ML) predictive model using retrospective data to try to predict which COVID-19 patients will likely be admitted. This resulted in the subsequent creation of a feasible scoring tool which can be used in the clinical setting.

6,485

COVID-19 PATIENTS ANALYZED

707

COVID-19 PATIENTS HOSPITALIZED

Types of Data Used

We used retrospective data of COVID-19 patients. Analysis was limited to age, gender, and historical variables. We created a Variable Importance Plot and chose a selection of the best predictors to train a logistic regression and random forest model. We then created a scoring tool and validated the score on the test set data.

Findings/Lessons Learned

A total of 6,485 COVID-19 patients were included in our analysis, of which 707 were hospitalized. The biggest predictors of patient hospitalization included age, a history of hypertension, diabetes, chronic heart disease, gender, tobacco use, and chronic kidney disease. The logistic regression and random forest models demonstrated an AUC of 0.81 and 0.76, respectively. A second logistic regression model was created which was used to develop a scoring tool. Patients with low-, intermediate-, and high-risk were deemed to have a 3.5%, 26%, and 38% chance of hospitalization, respectively.

Next Steps/Market Strategy

Our scoring needs validation at other sites in order to be generalizable. Next steps should involve validating our model in other institutions. Future direction can be aimed at the impact of closer monitoring may have in the different risk groups. Further data will inform the ML algorithm which may result in a more effective model. Additionally, age-adjusted risk stratification may provide more efficient predictive capabilities in the non-elderly and elderly populations.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Erin Shigo, BA for her scholarly assistance with this abstract.

USF HEALTH MORSANI COLLEGE OF MEDICINE UNIVERSITY OF SOUTH FLORIDA

LVHN.org

Lehigh Valley Health Network