Development and validation of a prognostic COVID-19 severity assessment (COSA) score and machine learning models for patient triage at a tertiary hospital

Authors: Verena Schöning[1], Evangelia Liakoni[1], Christine Baumgartner[2], Aristomenis K. Exadaktylos[3], Wolf E. Hautz[3], Andrew Atkinson[4,5], Felix Hammann[1,*]

(1) Clinical Pharmacology and Toxicology, Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland

(2) Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

(3) Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland

(4) Pediatric Pharmacology and Pharmacometrics Research Group, University Children's Hospital Basel, Switzerland

(5) Department of Infectious Diseases, Inselspital, Bern University Hospital, University of Bern, Switzerland

* Corresponding Author

Felix.Hammann@insel.ch

Counts

abstract: 290 words, main text: 3669 words, figures: 3, tables: 2, references: 36

Abstract

Background: Clinical risk scores and machine learning models based on routine laboratory values could assist in automated early identification of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) patients at risk for severe clinical outcomes. They can guide patient triage, inform allocation of health care resources, and contribute to the improvement of clinical outcomes.

Methods: In- and out-patients tested positive for SARS-CoV-2 at the Insel Hospital Group Bern, Switzerland, between February 1st and August 31st ('first wave', n=198) and September 1st through November 16th 2020 ('second wave', n=459) were used as training and prospective validation cohort, respectively. A clinical risk stratification score and machine learning (ML) models were developed using demographic data, medical history, and laboratory values taken up to three days before, or one day after, positive testing to predict severe outcomes of hospitalization (a composite endpoint of admission to intensive care, or death from any cause). Test accuracy was assessed using the area under the receiver operating characteristic curve (AUROC).

Results: Sex, C-reactive protein, sodium, hemoglobin, glomerular filtration rate, glucose, and leucocytes around the time of first positive testing (-3 to +1 days) were the most predictive parameters. AUROC of the risk stratification score on training data (AUROC = 0.94, positive predictive value (PPV) = 0.97, negative predictive value (NPV) = 0.80) were comparable to the prospective validation cohort (AUROC = 0.85, PPV = 0.91, NPV = 0.81). The most successful ML algorithm with respect to AUROC was support vector machines (median = 0.96, interquartile range = 0.85-0.99, PPV = 0.90, NPV = 0.58).

Conclusion: With a small set of easily obtainable parameters, both the clinical risk stratification score and the ML models were predictive for severe outcomes at our tertiary hospital center, and performed well in prospective validation.

Keywords

SARS-CoV-2, critical illness, risk stratification, statistical learning, artificial intelligence