Predicting deterioration in Covid-19 patients

Dan Coster*1, Omer Noy*1, Maya Metzger1, Itai Atar2, Shani Shenhar-Tsafraty3, Ori Rogowski3, Ron Shamir1

1Blavatnik School of Computer Science and 2Sackler Faculty of Medicine, Tel-Aviv University, 3Department of Internal Medicine “C”, Tel-Aviv Sourasky Medical Center.
*Equal contribution. omernoy@gmail.com | danoster@gmail.com

Background

The COVID-19 pandemic has been spreading worldwide since December 2019, presenting an urgent threat to global health. To date, there is no reliable predictor of disease deterioration. It is currently a clinical challenge to predict which hospitalized patients will deteriorate. Early prediction can allow physicians to take measures that prevent or lessen condition worsening.

Our goal is to develop a predictive model for early identification of patients at risk for clinical deterioration by analyzing Electronic Medical Records (EMR) data of COVID-19 inpatients and by conducting a meta-analysis of existing studies.

Dataset

The Tel-Aviv Sourasky Medical Center (TASMC) EMR data contains medical records of 760 inpatients that were diagnosed with COVID-19, including

58 ventilated patients

239 patients with severe disease

The features are derived automatically from the EMR of the entire hospitalization period. A total of ~150 features were collected per patient, including demographics, background diseases, laboratory tests, vital signs, symptoms and habits.

Model Development Overview

Our Approach

We developed models for predicting patients at risk for clinical deterioration in the next 30 hours, without using the data in 6 hours prior to the event. We used two different criteria of deterioration:

• Mechanical ventilation.
• A high COVID-19 Modified Early Warning Score (COVID-19 MEWS ≥ 7).

Results

Conclusions and Future Plans

We developed two machine learning models for predicting deterioration events within the next 7-30 hours in COVID-19 inpatients. On TASMC EMR data, the models achieved AUROC=0.76 and AUPR=0.43 in predicting severe condition, and AUROC=0.78 and AUPR=0.2 in predicting mechanical ventilation.

These results are still preliminary. We aim to achieve higher performance by improving data pre-processing, imputation and feature engineering, and by developing of survival-based methods.

Our study is limited at this point to data from a single center. We aim to validate our models on data from other hospitals. This will increase model generalizability and lower the risk of overfitting.

Study supported in part by the Israel Science Foundation (grant 1338/18 and grant 3165/19 within the Israel Precision Medicine Partnership program) and German-Israeli Project DFG RE 4193/1-1.

DC and ON were supported in part by fellowships from the Edmond J. Safra Center for Bioinformatics at Tel Aviv University.

Five top significant parameters that were observed in more than six studies