



Data-Driven COVID-19 Hospitalization Risk-Alert Clinical Decision Support Platform

Wenyu Song^{1,3}, Luwei Liu¹, Linying Zhang⁴, Michael Sainlaire¹, Mehran Karvar^{2,3}, Min-Jeoung Kang^{1,3}, Avery Pullman¹, Anthony Massaro^{1,3}, Namrata Patil^{2,3}, Ravi Jasuja^{1,3}, Patricia C. Dykes^{1,3}

¹Department of Medicine, Brigham & Women's Hospital, ²Department of Surgery, Brigham & Women's Hospital, ³Harvard Medical School, Boston, MA, ⁴Department of Biomedical Informatics, Columbia University, New York, NY

Abstract

COVID-19 outbreak has caused unprecedented burden on the health care system. With the disproportionately high morbidity and mortality among elderly patients, there is an urgent need to provide active monitoring and risk assessment for high-risk older adults. During the first phase of our study, we developed a personalized risk profile for high-risk population (defined as hospitalization) from COVID-19. 1,495 COVID-19 test-positive elderly patients (65 and older) in the MGB Health system were included in the study cohort. 27 potential predictive variables from demographics, vital signs, disease diagnosis, and lab values were selected. Models of logistic regression, support vector machine (SVM), random forest, and neural network were trained and validated with these variables as features and with whether patients are hospitalized during the defined window close to their COVID-19 test date as output. The results identified albumin (an index for nutritional status) as a strong predictor of hospitalization risk and achieved the best predictive performance of AUC=0.81 with the random forest model. For the next step, we are conducting a feasibility study in a geographically distant long-term care facility. We will recruit 60 high risk assistant living (AL) residents to wear a non-intrusive wireless wrist band with complementary ecological momentary assessment module for comprehensive monitoring. Collectively, the EHR data and wearable data stream from the residents will be integrated into our current model to develop a novel COVID-19 clinical decision support (CDS) platform.

Background and Significance

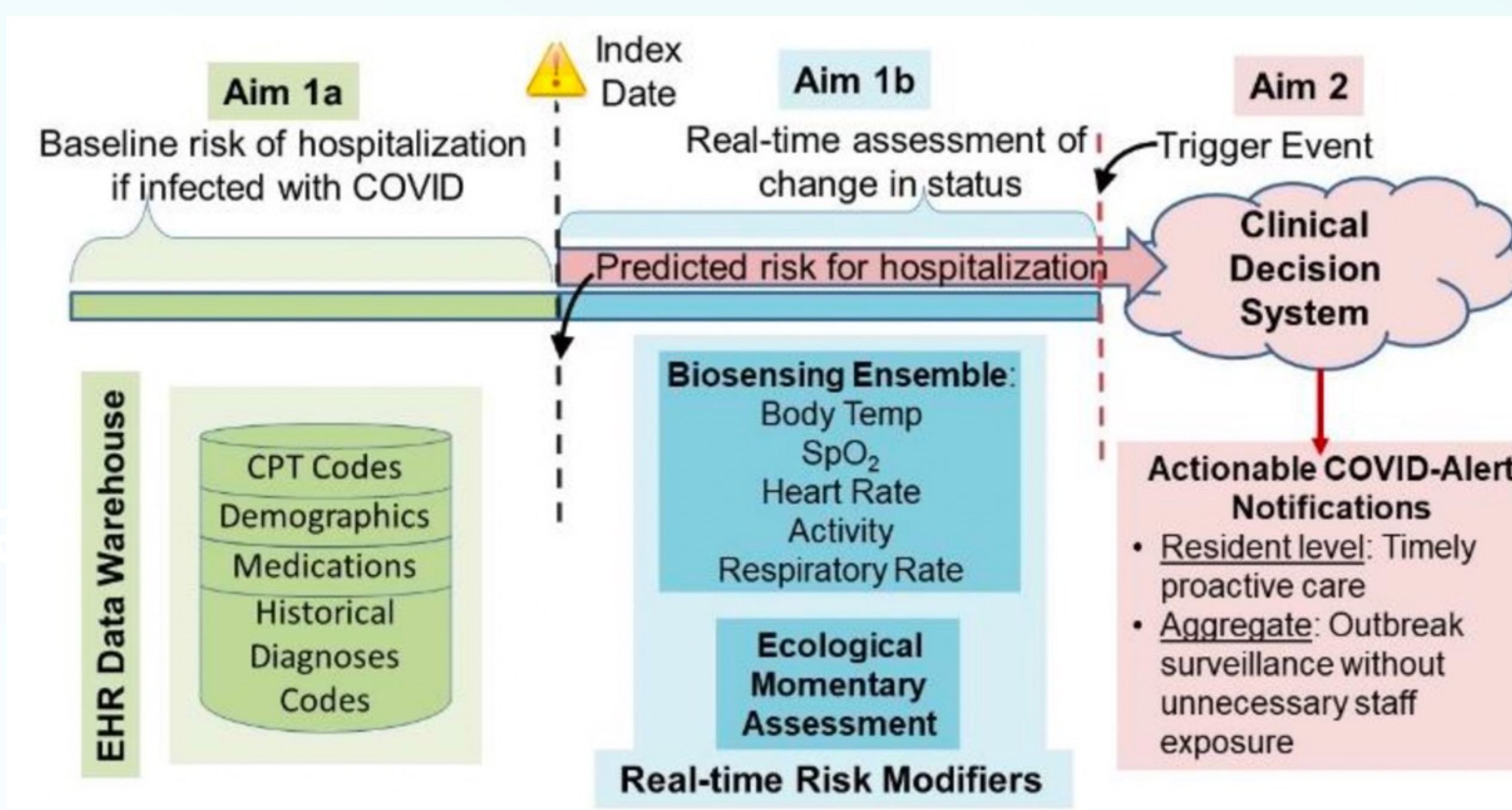
- COVID-19 outbreak is disproportionately high in long-term care (LTC) settings
- Elderly COVID-19 patients are more susceptible to severe illness and have a higher mortality rate

There is an urgent need in providing strategies for monitoring COVID-19 deterioration for residents in long term care settings with minimal person-to-person contact without adversely impacting the quality of care.

Study Cohort

- COVID-19 test-positive elderly patients (age > 65) in the MGB Health system
 - case group with hospitalization record
 - control group without hospitalization record
- LTC residents in Asbury Village in Bethesda, MD

Study Design



Overall Development of COVID-19 Alert platform

Discussion and Next Step

Our results identified albumin as a strong predictor of hospitalization risk and achieved the best predictive performance of AUC=0.81 with the random forest model. For the next step, we are conducting a feasibility study in a long-term care facility, Asbury Village in Bethesda, MD. We will use the EHR data of the residents in the facility to validate and refine the current prediction model. We will also recruit 60 residents to wear a non-intrusive wireless wrist band to capture wearable data stream in providing real-time assessment of physiological change in patients' status and develop an actionable clinical decision support system.

Results

During the first phase, we developed and validated machine learning based prediction models that estimate the personalized risk profile for hospitalization using EHR data from 1,495 COVID-19 test-positive elderly patients in the MGB Health system. The outcome of the models is whether elderly COVID-19 positive patients are hospitalized.

Model	AUC (SD)	Precision (SD)	Specificity (SD)	Sensitivity (SD)
Logistic Regression	0.80(0.03)	0.66(0.06)	0.77(0.05)	0.74(0.06)
Support Vector Machine	0.80(0.03)	0.67(0.05)	0.73(0.08)	0.77(0.08)
Random Forest	0.81(0.03)	0.69(0.06)	0.77(0.07)	0.74(0.07)
Neural Network	0.79(0.03)	0.64(0.05)	0.75(0.09)	0.74(0.09)

Top Significant Predictors (p<.05, based on Logistic Regression)

Albumin (Standardized coefficient = -1.02); Cancer (-0.20); Blood Urea Nitrogen (0.15); HIV/AIDS (0.08); Age (0.05); White Blood Count (0.06); Diabetes (0.06); BMI (0.04); Cardiomyopathy (-0.04); Heart Failure (-0.04); Liver Disease (-0.04); Metastatic Solid Tumor (-0.04); Solid Organ Transplant (0.05); Smoking Status (0.05); Coronary Artery Disease (0.03); Dementia (-0.03); Hypertensive Disease (0.02); Dyslipidemia (0.02); Immunodeficiency (0.02)