SUPPLEMENTAL MATERIAL

Symptoms predicting SARS-CoV-2 Test Results in Resident Physicians and Fellows In New York City

Tania P. Chen, MBBS, MPH, Meizhen Yao, MBBS, MS, Vishal Midya, PhD, MSTAT, Betty Kolod, MD, Rabeea F. Khan, MD, Adeyemi Oduwole, MPH, Bernard Camins, MD, I. Michael Leitman, MD, Ismail Nabeel, MD, MS, MPH, Kristin Oliver, MD, MHS, Damaskini Valvi, MD, MPH, PhD

Table S1. Comparison of statistics in three prediction models.

Statistics	Model1- symptoms only[a] (n=328)	Model2- symptoms and other risk factors [b] (n=328)	Model3- symptoms only[c] (n=186)	Model4 - symptoms and other risk factors [d] (n=186)
Accuracy (95% CI)	0.878 (0.787, 0.940)	0.915 (0.832, 0.965)	0.894 (0.769, 0.965)	0.872 (0.743, 0.952)
Sensitivity	0.922	0.953	0.939	0.939
Specificity	0.722	0.778	0.786	0.714
Positive Predict Value	0.922	0.939	0.912	0.886
Negative Predict Value	0.722	0.824	0.846	0.833
AUC	0.822	0.838	0.863	0.878

Abbreviations: CI, confidence interval; AUC: area under the Receiver Operating Characteristic curve. a. Prediction model for whole data set (328 participants), only includes symptoms; b. Prediction model for whole data set (328 participants), includes symptoms and sex, age, race, change in usual patient population, medical surgical unit and training specialty; c. Prediction model for EHS data set (186 participants), includes symptoms and sex, age, race, change in usual patient population, medical surgical unit and training specialty; d. Prediction model for EHS data set (186 participants), only includes symptoms.
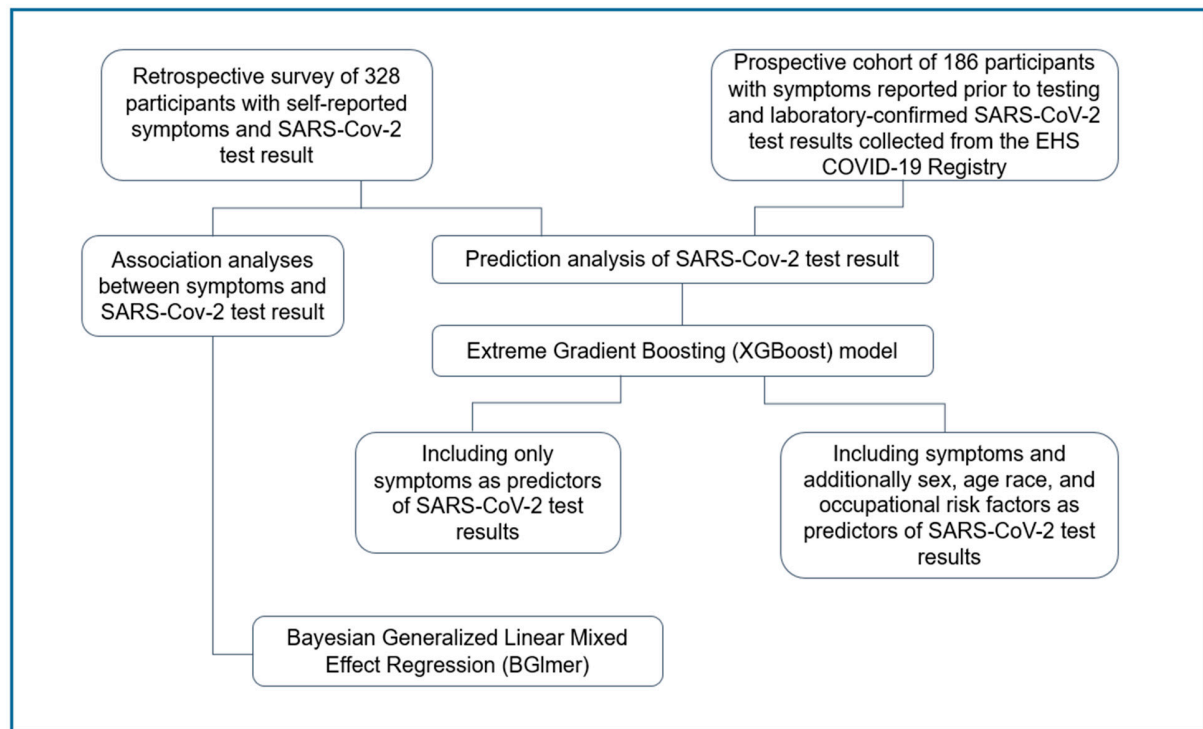
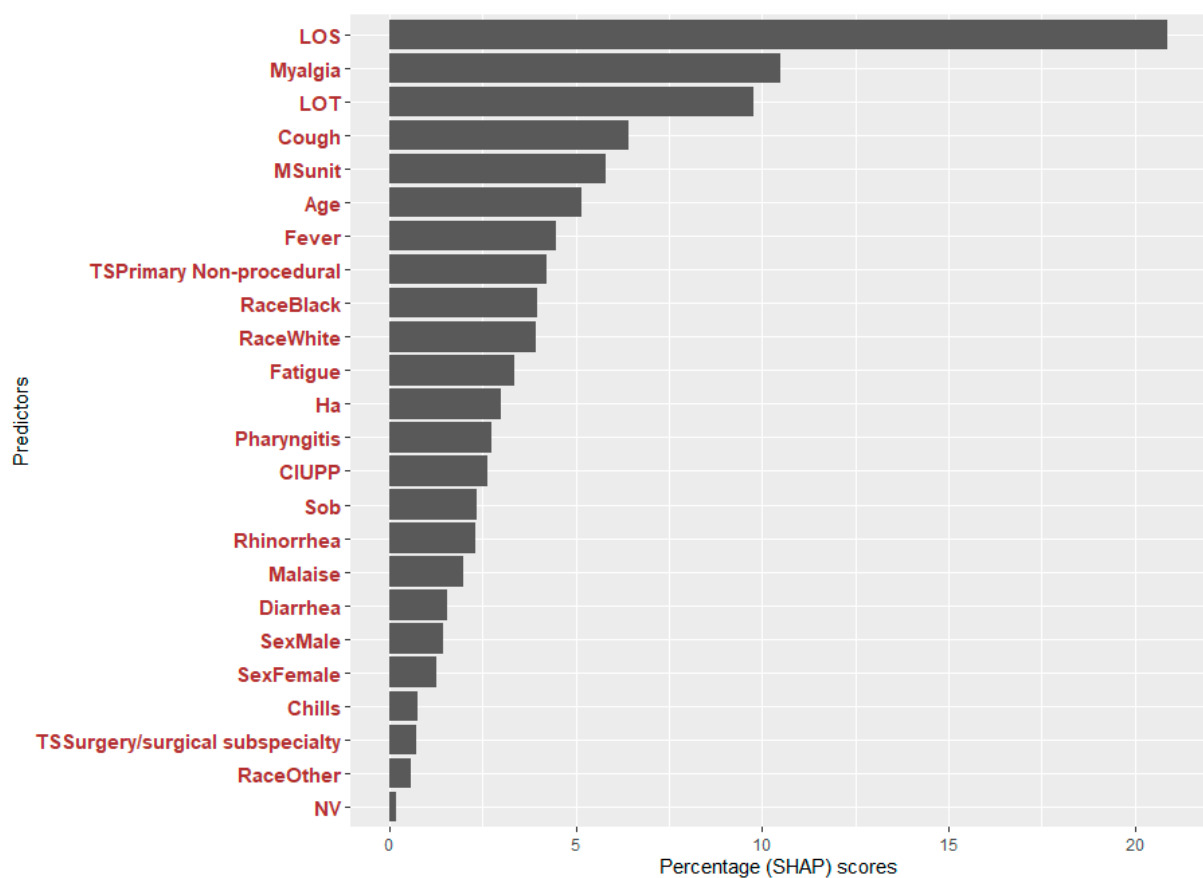**Figure S1**. Flow chart of all the main analyses performed.

Figure S2. Importance of each predictor's contribution using percentage (SHAP) scores in the prediction model among whole data set (n=328), including symptoms and other covariates.

Abbreviations: LOS, loss sense of smell; LOT, loss sense of taste; Ha, headache; Sob, shortness of breath; NV, nausea/vomiting.
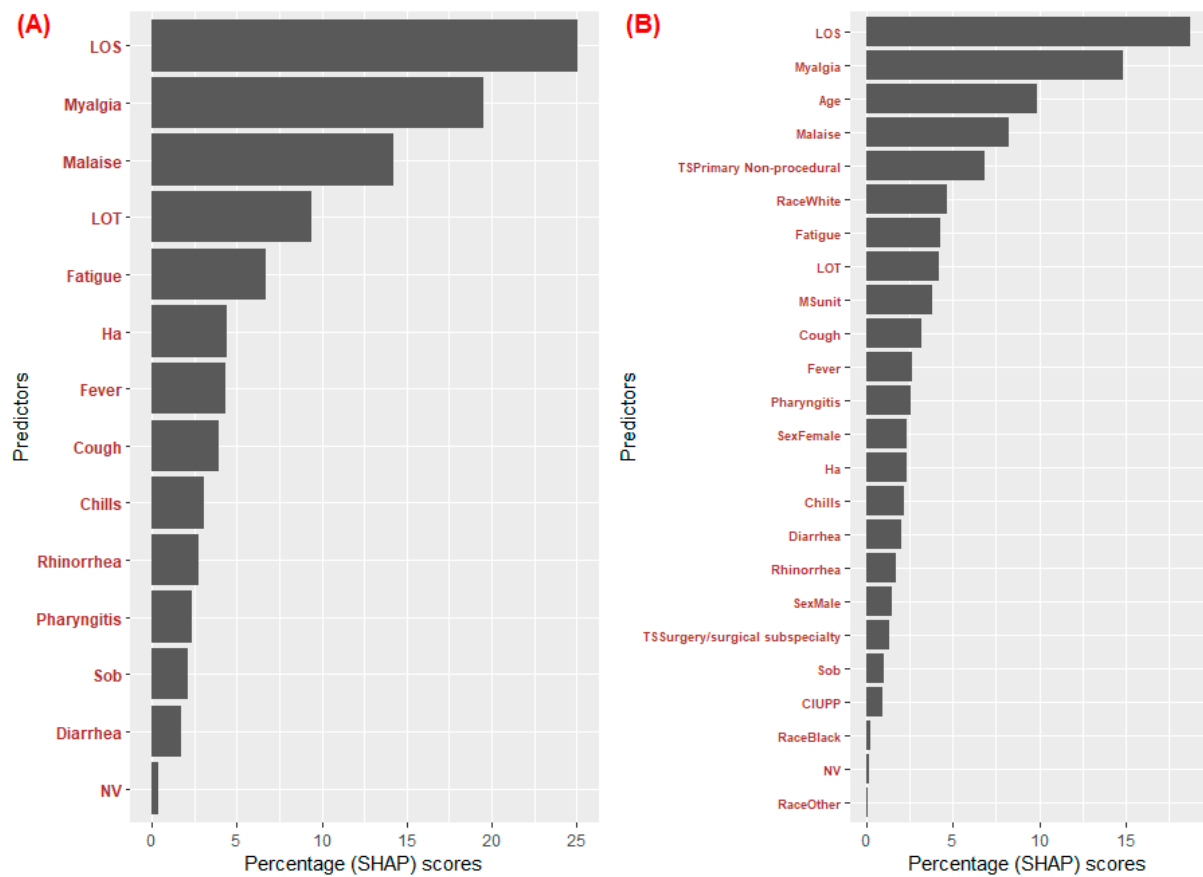
Figure S3. Comparing the importance of each predictor's contribution using percentage (SHAP) scores in the two prediction models among EHS data set.

Abbreviations: LOS, loss sense of smell; LOT, loss sense of taste; Ha, headache; Sob, shortness of breath; NV, nausea/vomiting. (A). Prediction model for EHS data set (186 participants), only includes symptoms; (B). Prediction model for EHS data set (186 participants), includes symptoms and other covariates.
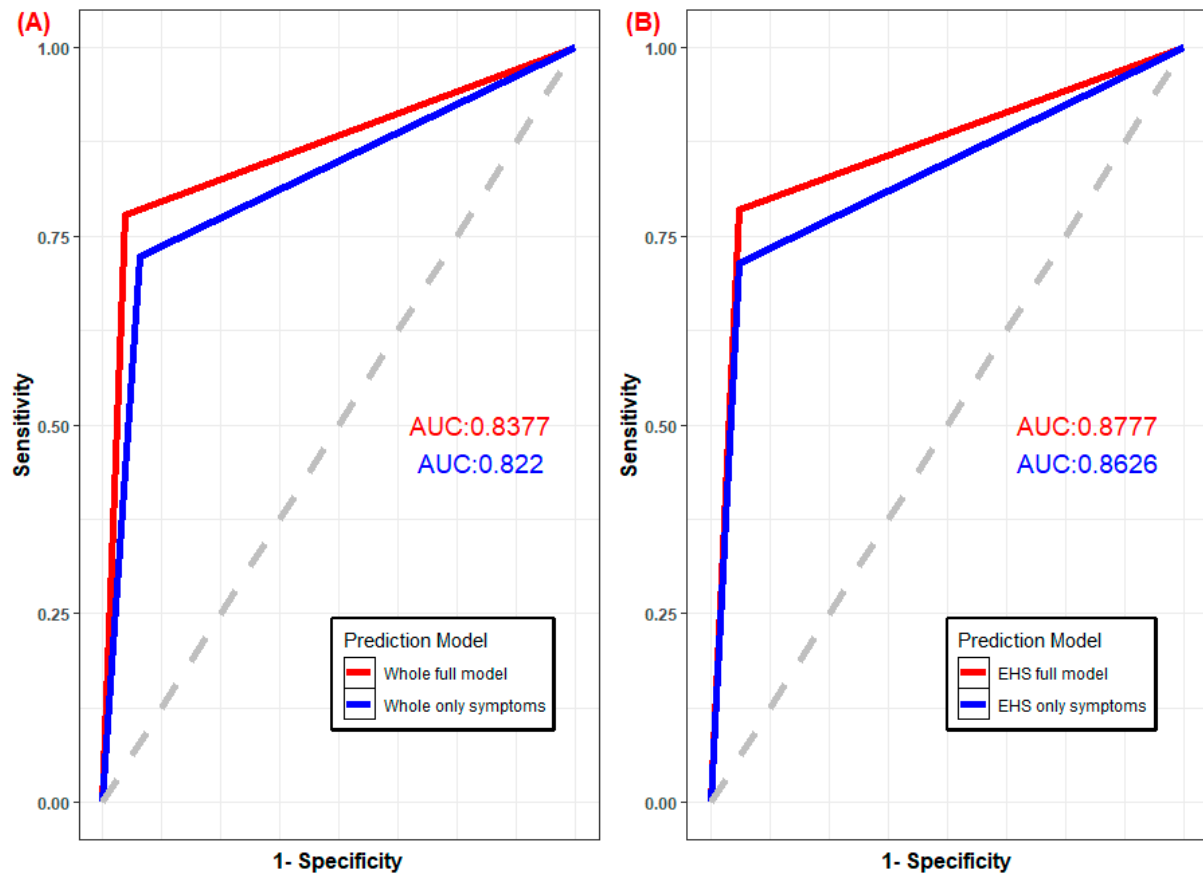
Figure S4. **ROC comparison of prediction models using different predictors in the two different data sets.**

(A). Whole data set (328 participants); (B). EHS data set (186 participants).