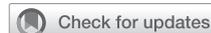≋CHEST®

Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing
Results From 11,672 Patients

Check for updates

Lara Jehi, MD, MHCDS; Xinge Ji, MS; Alex Milinovich, MS; Serpil Erzurum, MD; Brian P. Rubin, MD, PhD; Steve Gordon, MD; James B. Young, MD; and Michael W. Kattan, PhD

BACKGROUND: Coronavirus disease 2019 (COVID-19) is sweeping the globe. Despite multiple case-series, actionable knowledge to tailor decision-making proactively is missing.

RESEARCH QUESTION: Can a statistical model accurately predict infection with COVID-19?

STUDY DESIGN AND METHODS: We developed a prospective registry of all patients tested for COVID-19 in Cleveland Clinic to create individualized risk prediction models. We focus here on the likelihood of a positive nasal or oropharyngeal COVID-19 test. A least absolute shrinkage and selection operator logistic regression algorithm was constructed that removed variables that were not contributing to the model's cross-validated concordance index. After external validation in a temporally and geographically distinct cohort, the statistical prediction model was illustrated as a nomogram and deployed in an online risk calculator.

RESULTS: In the development cohort, 11,672 patients fulfilled study criteria, including 818 patients (7.0%) who tested positive for COVID-19; in the validation cohort, 2295 patients fulfilled criteria, including 290 patients who tested positive for COVID-19. Male, African American, older patients, and those with known COVID-19 exposure were at higher risk of being positive for COVID-19. Risk was reduced in those who had pneumococcal polysaccharide or influenza vaccine or who were on melatonin, paroxetine, or carvedilol. Our model had favorable discrimination (c-statistic = 0.863 in the development cohort and 0.840 in the validation cohort) and calibration. We present sensitivity, specificity, negative predictive value, and positive predictive value at different prediction cutoff points. The calculator is freely available at https://riskcalc.org/COVID19.

INTERPRETATION: Prediction of a COVID-19 positive test is possible and could help direct health-care resources. We demonstrate relevance of age, race, sex, and socioeconomic characteristics in COVID-19 susceptibility and suggest a potential modifying role of certain common vaccinations and drugs that have been identified in drug-repurposing studies.

CHEST 2020; 158(4):1364-1375

KEY WORDS: COVID-19; infectious disease; predictive modeling; testing

The first infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the novel virus responsible for coronavirus disease 2019 (COVID-19) was reported in the United States on January 21, 2020.[1] Six months later, the US health-care system and our society are struggling in an ever-changing environment of social distancing policies and projected utilization requirements, with constantly shifting treatment guidelines. A scientific approach to planning and delivering health care is sorely needed to match our limited resources with the persistently unmet demand. This supply-vs-demand gap is most obvious with diagnostic testing. Plagued with technical and regulatory challenges,[2] the production of COVID-19 test reagents and tests is lagging behind what is needed to fight a pandemic of this scale. Consequently, most hospitals are limiting testing to symptomatic patients and their own exposed health-care workers. This is occurring at a time when experts are calling for expanding testing capabilities beyond symptomatic individuals to better measure the infection's transmissibility, limit the spread by quarantine of those infected, and characterize COVID-19's epidemiologic components.[3] Recent loosening of the Food and Drug Administration testing regulations and the development of point-of-care testing will make more tests available; however, given the anticipated demand, it is unlikely that testing supply will be enough. Even if enough testing supplies become available, indications driven by scientific data are still needed. Another challenge is the suboptimal diagnostic performance of the test,[4] which raises concerns about false-negative results complicating efforts to contain the pandemic. Unless we develop intelligent targeting of our testing capabilities, we will be handicapped significantly in our ability to make progress in assessing the extent of the disease, directing clinical care, and ultimately controlling COVID-19.

We developed a prospective registry aligning data collection for research with clinical care of all patients who are tested for COVID-19 in our integrated health system. We present here the first analysis of our Cleveland Clinic COVID-19 Registry, with the aim to develop and validate a statistical prediction model to guide utilization of this scarce resource by predicting an individualized risk of a "positive test." A nomogram is a visual statistical tool that can take into account numerous variables to predict an outcome of interest for a patient.[5]

Methods

Patient Selection

We included all patients, regardless of age, who were tested for COVID-19 at all Cleveland Clinic locations in Ohio and Florida. Albeit imperfect, this provides better representation of the population than testing restricted to the Cleveland Clinic main campus. The Cleveland Clinic Institutional Review Board approval was obtained concurrently with the initiation of testing capabilities (IRB#20-283). The requirement for written informed consent was waived.

Cleveland Clinic COVID-19 Registry

Demographics, comorbidities, travel, and COVID-19 exposure history, medications, presenting symptoms, treatment, and disease outcomes are collected (e-Appendix 1). Registry variables were chosen to reflect available literature on COVID-19 disease characterization, progression, and proposed treatments, including medications proposed to have potential benefits through drug-repurposing studies.[6]

Capture of detailed research data is facilitated by the creation of standardized clinical templates that are implemented across the health-care system as patients were seeking care for COVID-19-related concerns.

Data were extracted via previously validated automated feeds[7] from our electronic health record (EPIC; EPIC Systems Corporation, Madison, WI) and manually by a study team trained on uniform sources for the study variables. Study data were collected and managed with the use of Research Electronic Data Capture (REDCap; Vanderbilt University, Nashville, TN) electronic data capture tools hosted at Cleveland Clinic.[8,9]

COVID-19 Testing Protocols

The clinical framework for our testing practice is shown in Figure 1. As testing demand increased, we adapted our organizational policies and protocols to reconcile demand with patient and caregiver safety. This occurred in three phases.

Phase I (March 12-13, 2020): We expanded primary care through telemedicine. If patients called for concerns that they had COVID-19, they were screened through a virtual visit with the use of Cleveland Clinic's Express Care Online or called their primary care provider. If they needed to travel to our locations, we asked them to call ahead before arrival. Our goal was to limit exposure to caregivers and to ensure that physicians could order testing when appropriate, while following the Center for Disease Control testing recommendations. A doctor's order was required for testing.

Phase II (March 14-17, 2020): Drive-through testing was initiated on Saturday March 14. Patients still needed to have a doctor's order for a COVID-19 test, similar to Phase I. Testing guidelines were similar to Phase I. On arrival at the drive-through location, patients stayed in their car, provided their doctor's order, and remained in their car as samples were collected. Patients were tested regardless of their ability to pay and were not charged copays.

Phase III (March 18-onwards): Given high testing demand, low initial testing yield, and backlog of tests awaiting to be processed, there was a shift to testing high-risk patients (Fig 1).
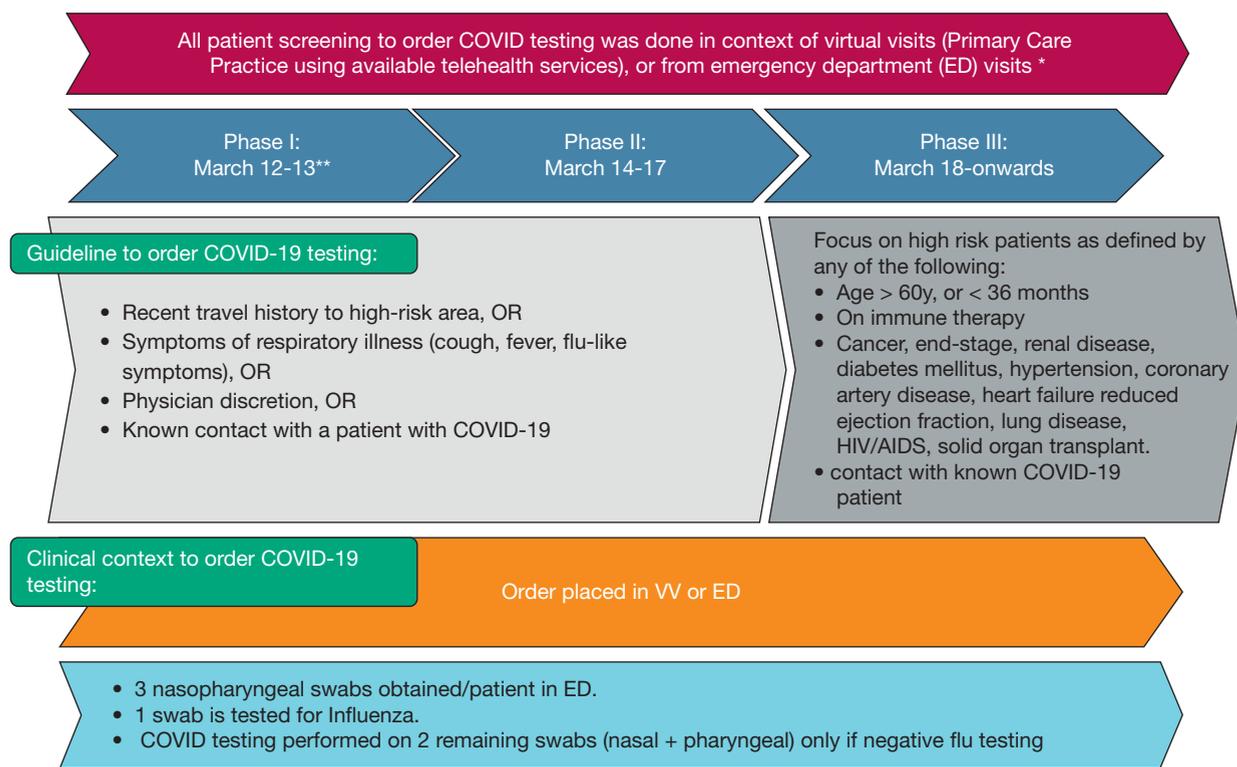
Figure 1 – *Timeline shows the evolution of clinical framework to COVID test ordering during the first 10 days of testing. The single asterisk indicates that patients were sent to the ED only if they needed evaluation of additional symptoms and not purely to obtain COVID testing. The double asterisk indicates that the guidelines to order COVID testing followed the Centers for Disease Control and Prevention recommendations. The main change in phase III was a better definition of high-risk categories, rather than reliance on "physician discretion." Of note, only 6.7% were tested in phase I + phase II because of physician discretion alone, so that number was too small to perform any modeling work in that group. COVID = coronavirus disease 2019; OR = operating room; VV = virtual visit.*

Processing of COVID Tests

Test samples were obtained through naso- and oropharyngeal swabs; both were collected and pooled for testing. Tests were run with the use of the Centers for Disease Control and Prevention assay using Roche magnapure extraction (Roche Life Science) and ABI 7500 DX PCR machines (Applied Biosystems/ThermoFisher Scientific), as per the standard laboratory testing in our organization.

Statistical Methods

Model Development: Data from 11,672 patients who were tested before April 2 were used to develop the model (development cohort). Baseline data are presented as median (interquartile range) and number (percentage). Continuous variables were compared with the use of the Mann-Whitney U test, and categoric variables were compared with the use of the chi-square test. A full multivariable logistic model was constructed initially to predict COVID-19 Nasopharyngeal Swab Test Result based on demographics, comorbidities, immunization history, symptoms, travel history, laboratory variables, and medications identified before testing. For modeling purposes, methods of missing value imputation for laboratory variables were compared with the use of median values and values from multivariate imputation by chained equations via the R package mice. Restricted cubic splines with 3 knots were applied to continuous variables to relax the linearity assumption. A least absolute shrinkage and selection operator (LASSO) logistic regression

algorithm was performed to retain the most predictive features. A 10-fold cross validation method was applied to find the regularization parameter lambda, which gave the minimum mean cross-validated concordance index. Predictors with nonzero coefficients in the LASSO regression model were chosen for calculating predicted risk.

Model Validation: The final model was first internally validated by assessment of the discrimination and calibration with 1000 bootstrap resamples. The LASSO procedure, which included 10-fold cross validation for optimizing lambda, was repeated within each resample. We then validated it in a temporally and geographically distinct cohort of 2295 patients tested at the Cleveland Clinic hospitals in Florida from April 2-16, 2020. This was done to assess the model's stability over time and its generalizability to another geographical region.

Model Performance: Discrimination was measured with the concordance index.[10] Calibration was assessed visually by plotting the nomogram predicted probabilities against the observed event proportions. The closer the calibration curve lies along the 45-degree line, the better the calibration. A scaled Brier score (index of prediction accuracy [IPA])[11] was also calculated, because this has some advantages over the more popular concordance index. The IPA ranges from -1 to 1, where a value of 0 indicates a useless model, and negative values imply a harmful model. Finally, decision curve analysis was conducted to inform clinicians about the range of
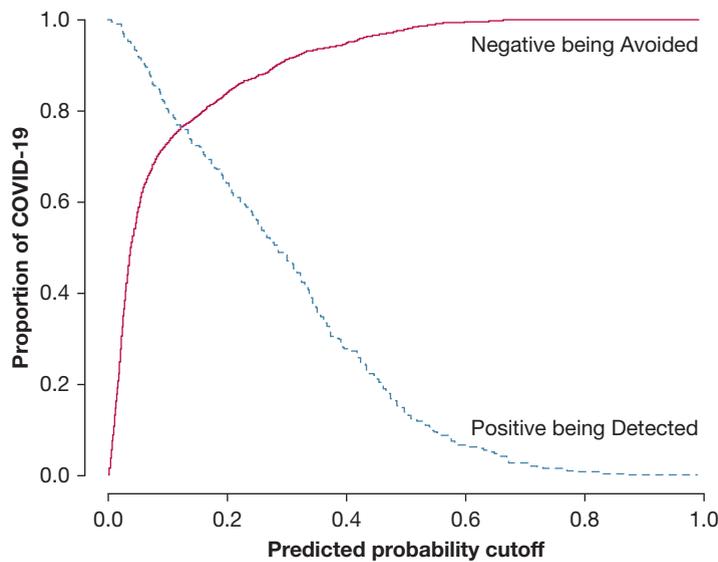
	Sensitivity	Specificity	NPV	PPV
Cutoff: 10%	0.803	0.730	0.963	0.301
Recommended cutoff: 12.3%	0.762	0.765	0.957	0.319
Cutoff: 30%	0.483	0.913	0.924	0.444

Figure 2 – *Proportion of COVID-19 negative tests being avoided (solid line, true negative rate) vs proportion of COVID-19 positive tests being identified (dashed line, true positive rate) at different nomogram predicted probability cutoffs. For example, if a predicted probability of ≥0.60 was required before testing, nearly all negative cases would have been avoided, but approximately 95% of positive cases would have been missed. At a cutoff of 12.3%, the proportion of negative tests being avoided is equal to the proportion of positive tests being detected (intersection of red and blue lines). The Table below shows the sensitivity, specificity, NPV, and PPV for this cutoff of 12.3%. For higher cutoffs, we illustrate how sensitivity decreases while specificity increases. NPV = negative predictive value; PPV = positive predictive value. See Figure 1 legend for expansion of other abbreviation.*

threshold probabilities for which the prediction model might be of clinical value.[12] We then calculated sensitivity, specificity, positive predictive value, and negative predictive value for different recommended test cutoffs (Fig 2). We adhered to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) checklist for prediction model development.

Results

Patient Characteristics

There were 11,672 patients who presented with symptoms of a respiratory tract infection or with other risk factors for COVID-19 before April 2, 2020, and who underwent testing according to the framework illustrated in Figure 1. The testing yield changed as the selection criteria became stricter (e-Fig 1). Between April 2 and 16, 2020, 2,295 patients were tested in Florida (Florida validation cohort). The clinical characteristics of the development cohort and validation cohort are found in Table 1.

Nomogram Results

Imputation methods were evaluated with 1000 repeated bootstrapped samples. We found that models based on median imputation appeared to outperform those based on data from multivariate imputation by chained

equations imputation, so median imputation was selected for the basis of the final model. Variables that we looked at that were not found to add value beyond those included in our final model for the prediction of the COVID-19 test result included being a health-care worker in Cleveland Clinic, fatigue, sputum production, shortness of breath, diarrhea, and transplantation history. The bootstrap-corrected concordance index in the development cohort was 0.863 (95% CI, 0.852-0.874), and the IPA was 20.9% (95% CI, 18.1%-23.7%). The concordance index in the Florida validation cohort was 0.839 (95% CI, 0.817-0.861), and the IPA was 18.7% (95% CI, 13.6%-23.9%). Figure 3 shows the calibration curves in the development and validation cohorts. In the development cohort, the predicted risk matches observed proportions for low predictions before the model begins to overpredict at high-risk levels.

TABLE 1] Baseline Demographic and Clinical Characteristics in 11,672 Patients Who Tested Positive vs Negative to COVID-19 in the Development Cohort in the Cleveland Clinic Health System before April 2, 2020, and a Validation Cohort of 2,295 Patients in the Florida Cleveland Clinic Health System Patients Tested Between April 2 and 16, 2020

Variable	Development Cohort			Florida Validation Cohort		
	COVID-19 Negative	COVID-19 Positive	P Value	COVID-19 Negative	COVID-19 Positive	P Value
No.	10,854	818		2005	290	
Physician discretion, No. (%)	773 (99.3)	6 (0.7)	< .001	580 (98.5)	9 (1.5)	< .001
Demographics						
Race, No. (%)			< .001			< .001
Asian	174 (98)	9 (2)		46 (85.2)	8 (14.8)	
Black	2,138 (91.1)	207 (8.9)		209 (79.8)	53 (20.2)	
Other	1,194 (92.1)	102 (7.9)		369 (84.6)	67 (15.4)	
White	7,348 (93.6)	500 (6.4)		1381 (89.5)	162 (10.5)	
Male (%)	4,192 (91.0)	415 (9.0)	< .001	831 (85.8)	138 (14.2)	.055
Ethnicity, No. (%)			< .001			< .001
Hispanic	505 (91.3)	48 (8.7)		529 (81.4)	121 (18.6)	
Non-Hispanic	9,608 (93.2)	697 (6.8)		1383 (89.6)	160 (10.4)	
Unknown	741 91.0)	73 (9.0)		93 (91.2)	9 (8.8)	
Smoking, No. (%)			< .001			< .001
Current Smoker	1,593 (97.7)	37 (2.3)		67 (91.8)	6 (8.2)	
Former Smoker	2,692 (93.0)	202 (7.0)		366 (81.3)	84 (18.7)	
No	5,141 (92.1)	440 (7.9)		626 (87.4)	90 (12.6)	
Unknown	1,428 (91.1)	139 (8.9)		946 (89.6)	110 (10.4)	
Age, median [IQR], y Missing: 0.3%	46.89 [31.57-62.85]	54.23 [38.81-65.94]	< .001	56.02 [41.95-67.52]	51.60 [36.69-63.08]	< .001
Exposure history: Yes, No. (%)						
Exposed to COVID-19?	1,510 (94.5)	88 (4.5)	.013	492 (68.5)	226 (31.5)	< .001
Family member with COVID-19?	911 (94.1)	57 (5.9)	.174	467 (68.9)	211 (31.1)	< .001
Presenting symptoms: Yes, No. (%)						
Cough?	2,782 (95.5)	130 (4.5)	< .001	609 (70.8)	251 (29.2)	< .001
Fever?	1,918 (94.6)	110 (5.4)	< .001	532 (69.9)	229 (30.1)	< .001
Fatigue?	1,472 (94.4)	87 (5.6)	< .001	406 (68.4)	188 (31.6)	< .001
Sputum production?	929 (96.0)	38 (4.0)	< .001	343 (68.2)	160 (31.8)	< .001
Flu-like symptoms?	1,813 (94.3)	108 (5.7)	.011	507 (70.7)	210 (29.3)	< .001
Shortness of breath?	1,578 (96.0)	64 (4.0)	< .001	462 (75.5)	150 (24.5)	< .001
Diarrhea?	629 (95.0)	33 (5.0)	.043	347 (69.5)	152 (30.5)	< .001
Loss of appetite?	671 (93.4)	47 (6.6)	.671	343 (67.0)	169 (33.0)	< .001
Vomiting?	536 (97.1)	16 (2.9)	< .001	309 (73.2)	113 (26.8)	< .001
Comorbidities						
BMI, median [IQR], kg/m^2 Missing: 43.3%	28.46 [23.90-33.94]	29.23 [25.86-33.78]	.001	27.60 [23.49-31.05]	28.91 [24.81-33.60]	.037
COPD/emphysema? Yes, No. (%)	304 (96.2)	12 (3.8)	.031	36 (94.7)	2 (5.3)	.257
Asthma? Yes, No. (%)	2,761 (94.9)	147 (5.1)	< .001	176 (91.7)	16 (8.3)	.078

(Continued)

TABLE 1] *(Continued)*

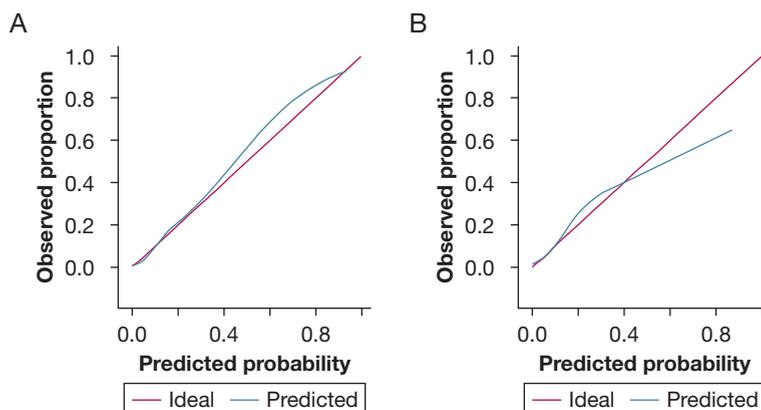| Variable	Development Cohort			Florida Validation Cohort		
	COVID-19 Negative	COVID-19 Positive	P Value	COVID-19 Negative	COVID-19 Positive	P Value
Diabetes mellitus? Yes, No. (%)	2,486 (93.0)	188 (7.0)	.993	224 (86.2)	36 (13.8)	.6
Hypertension? Yes, No. (%)	4,324 (92.7)	342 (7.3)	.283	460 (86.3)	73 (13.7)	.444
Coronary artery disease? Yes, No. (%)	1,325 (93.6)	90 (7.4)	.336	141 (97.9)	3 (2.1)	< .001
Heart failure? Yes, No. (%)	1,170 (94.7)	66 (5.3)	.018	88 (96.7)	3 (3.3)	.01
Cancer? Yes, No. (%)	1,616 (93.7)	108 (6.8)	.208	245 (92.8)	19 (7.2)	.006
Transplantation history? Yes, No. (%)	190 (96.4)	7 (3.6)	.046	43 (95.6)	2 (4.4)	.149
Multiple sclerosis? Yes, No. (%)	96 (91.4)	9 (8.6)	.661	8 (88.9)	1 (11.1)	1
Connective tissue disease? Yes, No. (%)	3,505 (94.5)	203 (5.5)	< .001	41 (89.1)	5 (10.9)	.889
Inflammatory bowel disease? Yes, No. (%)	943 (95.6)	45 (4.4)	.002	34 (81.0)	8 (19.0)	.304
Immunosuppressive disease? Yes, No. (%)	1,557 (94.5)	91 (5.5)	.012	163 (92.6)	13 (7.4)	.039
Vaccination history: Yes, No. (%)						
Influenza vaccine?	5,940 (93.9)	384 (6.1)	< .001	328 (91.6)	30 (8.4)	.011
Pneumococcal polysaccharide vaccine?	2,667 (95.2)	135 (4.8)	< .001	115 (92.0)	10 (8.0)	.143
Laboratory findings on presentation						
Pretesting platelets, median [IQR], k/uL Missing: 67.3%	245.00 [189.00-304.00]	190.00 [154.00-241.50]	< .001	236.00 [180.00-304.00]	213.50 [173.00-286.75]	.698
Pretesting AST, median [IQR], U/L Missing: 72.9%	23.00 [17.00-34.00]	32.00 [24.25-47.00]	< .001	22.00 [18.00-34.50]	31.00 [21.00-53.25]	.146
Pretesting BUN, median [IQR], mg/dL Missing: 67.2%	15.00 [11.00-23.00]	14.00 [10.00-22.00]	.099	18.00 [13.00-27.25]	12.00 [8.25-15.50]	.003
Pretesting chloride, median [IQR], mmol/L Missing: 67.2 %	101.00 [97.00-103.00]	99.00 [96.00-102.00]	< .001	100.00 [96.00-102.00]	97.50 [92.75-99.25]	.026
Pretesting creatinine, median [IQR], mg/dL Missing: 67.2%	0.90 [0.71-1.21]	1.01 [0.79-1.29]	< .001	0.94 [0.77-1.45]	0.92 [0.87-1.03]	.677
Pretesting hematocrit, median [IQR], % Missing: 67.3%	39.10 [34.20-43.00]	40.60 [37.15-43.85]	< .001	36.80 [32.20-41.00]	38.50 [36.02-43.20]	.221

(Continued)

TABLE 1] *(Continued)*

Variable	Development Cohort			Florida Validation Cohort		
	COVID-19 Negative	COVID-19 Positive	P Value	COVID-19 Negative	COVID-19 Positive	P Value
Pretesting potassium, median [IQR], mmol/L Missing: 67.3%	4.00 [3.80-4.40]	4.00 [3.70-4.20]	< .001	4.10 [3.90-4.60]	4.15 [3.90-4.35]	.808
Home medications						
Immunosuppressive treatment? Yes (%)	423 (97.2)	12 (2.8)	.001	97 (83.6)	19 (16.4)	.271
Nonsteroidal antiinflammatory drugs? Yes (%)	3,084 (95.1)	162 (5.0)	< .001	156 (94.0)	10 (6.0)	.011
Steroids? Yes (%)	2,317 (95.5)	109 (4.5)	< .001	135 (93.8)	9 (6.2)	.024
Carvedilol? Yes (%)	333 (96.2)	13 (3.8)	.022	27 (100.0)	0	.09
ACE inhibitor? Yes (%)	805 (93.3)	58 (6.7)	.784	60 (89.6)	7 (10.4)	.718
ARB? Yes (%)	585 (91.7)	53 (8.3)	.214	78 (90.7)	8 (9.3)	.434
Melatonin? Yes (%)	513 (97.0)	16 (3.0)	< .001	18 (100.0)	0	.206
Social influencers of health						
Population/km^2 median [IQR] Missing: 0.1%	3.06 [2.69-3.36]	3.08 [2.72-3.37]	.24	3.20 [3.02-3.35]	3.28 [3.12-3.42]	< .001
Median income × $1000, median [IQR], $ Missing: 0.1%	55.61 [38.73-78.56]	60.46 [42.77-84.24]	< .001	66.28 [53.41-89.11]	59.07 [47.59-75.56]	< .001
Population per housing unit, median [IQR], No. Missing: 0.1%	2.21 [1.88-2.56]	2.25 [1.89-2.59]	.038	2.47 [1.83-2.87]	2.61 [2.11-2.92]	.001

ACE = angiotensin converting enzyme; ARB = angiotensin receptor blocker; AST = aspartate aminotransferase; COVID-19 = coronavirus disease 2019; IQR = interquartile range.

Figure 3 – *Calibration curves for the model predicting likelihood of a positive test. The x-axis displays the predicted probabilities generated by the statistical model, and the y-axis shows the fraction of the patients who were COVID-19 positive at the given predicted probability. The 45-degree line therefore indicates perfect calibration, for example, a predicted probability of 0.2 is associated with an actual observed proportion of 0.2. The solid blue line indicates the model's relationship with the outcome. The closer the line is to the 45-degree line, the closer the model's predicted probability is to the actual proportion. A, The calibration curve in the development cohort of 11,672 patients tested in Cleveland Clinic Health System before April 2. B, The calibration curve in the Florida Validation Cohort (2,295 patients tested in Cleveland Clinic Florida from April 2-16, 2020). As demonstrated, there is excellent correspondence between the predicted probability of a positive test and the observed frequency of COVID-19 positive in both cohorts. See Figure 1 legend for expansion of abbreviation.*

Figure 4 – *The graphic version of the model (A) and the corresponding online risk calculator (B).*[13]

(Continued on next page)

Figure 4 – Continued. The example for both is a 60-year-old white male, former smoker, who presented with cough, fever, and a history of a known family member with COVID-19. He has coronary artery disease, did not receive vaccinations against influenza or pneumococcal pneumonia this year, and is only on melatonin to help with sleep. No laboratory tests were performed at the time of COVID-19 testing. His predicted risk of testing positive is 13.79%. If race is changed to black, with all other variables remaining constant, his relative risk almost doubles to an absolute value of 23.95%. ACE = angiotensin converting enzyme; ARB = angiotensin receptor blocker; AST = aspartate aminotransferase; NSAIDS = nonsteroidal antiinflammatory drugs. See Figure 1 legend for expansion of other abbreviation.

Calibration in the Florida validation cohort is acceptable, although predictions >40% become too high as the predicted probability increases.

CutOff Definition

Given that the tool provides a probability that an individual subject will test positive, the challenge is to use the tool in practice. This usually would require choosing a cutoff below which the risk is sufficiently low that the subject would not be tested. Figure 2 shows the tradeoff by plotting the proportion of negative tests avoided vs the proportion of positive tests retained as the cutoff is increased. A decision curve analysis showed that, if the threshold of action is ≤1.3%, the model is not better than simply assuming everyone is "high risk." However, once the threshold becomes >1.3%, using the model to determine who is high risk is preferable. The nomogram and its online version are shown in Figure 4.[13]

Discussion

The COVID-19 pandemic has impacted the world significantly, changing medical practice and our society. Some countries are now recovering from it, but many regions are just beginning to be affected. In the United States, some states are still preparing for a "surge" that may overwhelm the health-care delivery system, while others are preparing to "reopen" and lift social distancing measures. In a "presurge" situation, resources needed to address every step of a patient's trajectory through COVID-19 are limited, starting from testing through hospitalization and intensive care if needed. In a "pre-reopening" situation, tools to better identify individuals who are at risk of experiencing COVID-19 are sorely needed to inform policy.

We developed the Cleveland Clinic COVID-19 Registry to include all patients who were tested for COVID-19 (rather than just those with the disease) to better understand disease epidemiology and to develop nomograms, which are tools that go beyond cohort descriptions to individualize risk prediction for any given patient. This could empower front-line health-care providers and inform decision-making, immediately impacting clinical care. We present here our first such nomogram, one that predicts the risk of a positive COVID-19 test. We want to emphasize that our work should not be interpreted as "accepting" or rationalizing inadequate testing capacity. Our tool should not take the pressure off being able to do what is right clinically for individual patients by expanding testing capabilities.

COVID-19 Testing Challenge

Available COVID-19 clinical literature is based mostly on small case series or descriptive cohort studies of patients already documented to have COVID-19[14-23]: this provides some information on the population that may be at greatest risk of adverse outcomes if they get infected with the virus but does little to inform us on who is most at risk to get infected. The proportion of COVID negative tests fell significantly in the patient population with stricter testing guidelines (e-Fig 1), but the yield remained very low, which suggests that our ability to differentiate COVID-19 clinically from other respiratory illnesses at the early stages of the disease is limited, further supporting the need for better tools to individualize testing indications.

COVID-19 Risk Factors

Some of our predictors for developing COVID confirm previous literature. For example, we corroborate a recent World Health Organization report that suggests that men may be at higher risk of experiencing COVID-19,[24] which is thought to reflect underlying hormonal or genetic risk. Our finding of a higher COVID-19 risk with advancing age can be explained by known age-related changes in the angiotensin-renin system in mice[25] and humans[26] that may facilitate infection with the SARS-CoV-2 virus, which binds to the host cells through angiotensin receptors. A family member with COVID-19 also increased the risk of testing positive in our cohort, which is consistent with familial disease clustering observed in China and highlights the limitations of disease containment strategies that focus on home lock-down without isolation of sick individuals. In addition, our study provides several unique insights that are made possible by our large sample size and our inclusion of a control cohort of patients who tested negative for COVID. The following list includes critical findings that ultimately were relevant to our model's performance.

(1) The lower risk of being COVID positive in Asian individuals relative to white individuals in our cohort is intriguing, given the higher rates of spread and disease severity that were observed in the western hemisphere now when compared with China.

(2) The lower risk observed with pneumococcal polysaccharide vaccine and flu vaccine is also a unique finding. The mechanism could be biologic, related possibly to the documented sustained activation of Toll-Like Receptor 7 by the influenza vaccine[27]: Toll-Like Receptor 7 is critical for the binding of single-stranded RNA respiratory viruses, such as SARS-CoV-2, and may thus explain some cross protection. Alternatively, this correlation may just reflect safer health practices in general of people who seek and obtain vaccination.

(3) The higher risk observed with poor socioeconomic status. Using the zip code, our team was able to infer estimated population per square kilometer and estimated median income from the 5-year American Community Survey dataset. The end year of the 5-year dataset was 2018. The critical role played by these variables in our final model emphasize the importance of social influencers of health and their influence on disparities in health care outcomes.

(4) Most potentially impactful is the reduced risk of testing positive in patients who were on melatonin, carvedilol, and paroxetine, which are drugs identified in drug-repurposing studies to have a potential benefit against COVID-19.[6] Melatonin up-regulates angiotensin converting enzyme 2 (ACE2) expression, such that increased occupancy of ACE2 receptors competes with SARS-CoV2 viral attachment to the receptors and blocks entry.[6] Carvedilol was found recently to inhibit ACE2-induced proliferation and contraction in hepatic stellate cells through the rhoa/rho-kinase pathway.[28] It is unclear whether it has similar effects on ACE2 in lung endothelium. With ACE2 being key in the pathophysiologic findings of infection with SARS-CoV-2, our findings are intriguing.

These findings would have to be reproduced and validated in clinical trials before their full significance can be assessed. When interpreting our multivariable model, it is important to recognize that a single predictor cannot be interpreted in isolation. For example, it is artificial to claim that a drug is reducing risk because, in reality, other variables tend to be different for a patient who is on, or not on, a drug. Moving a patient on a nomogram axis, holding all other axes constant, is hypothetical, because he or she is likely moving on other axes when moved on one. This is the case for all multivariable statistical prediction models.

Nomogram Performance

Model performance, as measured by the concordance index, is very good in the development and in the validation cohort (c-statistic = 0.863 and 0.839, respectively). This level of discrimination is clearly

superior to a coin toss or assuming all patients are at equivalent risk (both c-statistics = 0.5). The internal calibration of the model is excellent at low predicted probabilities (Fig 3), but some regression to the mean is apparent at predictions >40% or so in the validation cohort. This would seem to be of little concern, that the model is overpredicting risk at that level, because this is considerably high risk clinically and likely beyond a threshold of action. Moreover, the metric that considers calibration, the IPA value, confirms that the model predicts better than chance or no model at all. The good performance of our model in a geographically distinct region (Florida), and over time (validation cohort in patients tested at a later timeframe) suggests that patterns and predictors identified in our model are likely consistent across health systems and regions, rather than specific to the unique spread of the virus within Cleveland's social structures.

Clinical Utility

As with any predictive tool, the utility of a nomogram depends on the clinical context. The decision curve analysis suggests that, if the goal is to distinguish patients with a risk of 1.3% (or a higher cutoff) vs those of higher risk, then the prediction model is useful. In other words, using the model to determine whom to test detects more true positives per test performed than does testing everyone as long as one is willing to test 1000 subjects to detect 13 cases. Any cutoff choice involves tradeoffs of avoiding negative tests vs missing positive cases (Fig 2). Using a low prediction cutoff (<1.3% from the tool) as a trigger to order testing will allow us to continue to identify a vast majority of COVID positive cases (assuming we maintain our other selection criteria for testing constant) while avoiding testing a large proportion of patients who are indeed COVID negative. This may be appropriate when testing supplies are abundant and one wants to comprehensively survey the extent of COVID-19 in the population. Conversely, in a resource-limited setting (eg, hospital facing a surge), a cutoff ≥1.3% may be more appropriate to avoid unnecessary testing.

Study Limitations

Available real-time reverse transcriptase polymerase chain reaction tests of nasopharyngeal swabs have been used typically for diagnosis, but data suggest suboptimal test performance because it detected only the SARS-CoV-2 virus in 63% of nasal swabs and 32% of pharyngeal swabs in patients with known disease.[4] In our study, we did both swabs, hoping to at least partly address this limitation. Although we performed validation of our model in a temporally and geographically distinct cohort, we acknowledge the fact that our results depend on the particular time and place that the data were collected. As the pandemic evolves, our results may not reflect updated distribution of the virus in any given region, and our model will need to be refit. To accommodate an ever-increasing COVID-19 prevalence, the model will need to be recalibrated and refit over time. Our online risk calculator is publicly available, but direct integration with the electronic health record can further improve its utility. The online calculator will reflect this updating. Our study is not designed to evaluate the very real issue of health care disparities, which would require a population-based approach for the study of health care delivery that is beyond the scope of the work presented here. Our conclusions are highly dependent on access to testing sites and doctors' orders rather than population-based predictors of positive results.

Interpretation

We provide an online risk calculator that effectively can identify individualized risk of a positive COVID-19 test. Such a tool provides immediate benefit to the patients and health care providers as we face anticipated increased demand and limited resources but does not obviate the critical need for adequate testing. The scarcity of resources must not be accepted as an unalterable fact, and we should resist the inevitability of lack of resources and inequities in health care. We also provide some mechanistic and therapeutic insights.

Acknowledgments

References

1. Callaway E, Cyranoski D, Mallapaty S, Stoye E, Tollefson J. The coronavirus pandemic in five powerful charts. Nature March 18, 2020. https://www.nature.com/articles/d41586-020-00758-2. Accessed March 21, 2020.

2. Sharfstein JM, Becker SJ, Mello MM. Diagnostic testing for the novel coronavirus. *JAMA*. 2020;323(15):1437-1438.

3. Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19 - studies needed. *N Engl J Med*. 2020;382(13):1194-1196.

4. Wang W, Xu Y, Gao R. SARS-CoV-2 in different types of clinical specimens. *JAMA*. 2020;323(18):1843-1844.

5. Kattan MW. Nomograms: introduction. *Semin Urol Oncol*. 2002;20(2):79-81.

6. Zhou Y, Hou Y, Shen J, et al. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov*. 2020;6:14.

7. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med*. 2018;6(3):42.

8. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381.

9. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software partners, *J Biomed Inform*. 2019;95:103208.

10. Harrell FE Jr, Califf RM, Pryor B, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247(18):2543-2546.

11. Kattan MW, Gerds TA. The index of prediction accuracy: an intuitive measure useful for evaluating risk prediction models. *Diagn Progn Res*. 2018;2:7.

12. Steyerberg E, Vickers A, Cook N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138.

13. Cleveland Clinic. Predict COVID-19 test result. https://riskcalc.org/COVID19/. Accessed June 30, 2020.

14. Jin X, Lian JS, Hu JH, et al. Epidemiological, clinical and virological characteristics of 74 cases of coronavirus-infected disease 2019 (COVID-19) with gastrointestinal symptoms. *Gut*. 2020;69(6):1002-1009.

15. Sun Y, Koh V, Marimuthu K, et al. Epidemiological and clinical predictors of COVID-19. *Clin Infect Dis*. 2020;ciaa322.

16. Weiss P, Murdoch DR. Clinical course and mortality risk of severe COVID-19. *Lancet*. 2020;395(10229):1014-1015.

17. Shi Y, Yu X, Zhao H, Wang H, Zhao R, Sheng J. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. *Crit Care*. 2020;24(1):108.

18. Wang Z, Yang B, Li Q, Wen L, Zhang R. Clinical features of 69 cases with coronavirus disease 2019 in Wuhan, China. *Clin Infect Dis*. 2020;ciaa272.

19. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study [published correction appears in Lancet. 2020 Mar 28;395(10229):1038]. *Lancet*. 2020;395(10229):1054-1062.

20. Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. 2020;e200994.

21. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study [published correction appears in *Lancet Respir Med*. 2020 Apr;8(4):e26]. *Lancet Respir Med*. 2020;8(5):475-481.

22. Zhang JJ, Dong X, Cao YY, et al. Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy*. In press.

23. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513.

24. World Health Organization. http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/weekly-surveillance-report. Accessed March 28, 2020.

25. Yoon HE, Kim EN, Kim MY, et al. Age-associated changes in the vascular renin-angiotensin system in mice. *Oxide Med Cell Longev*. 2016;2016:6731093.

26. Ogbadu J, Singh G, Gupta K, Mehra K, Sen P. Ageing reduces angiotensin II type 1 receptor antagonism mediated pre-conditioning effects in ischemic kidneys by inducing oxidative and inflammatory stress. *Exp Gerontol*. 2020;135:110892.

27. Goff PH, Hayashi T, Martínez-Gil L, et al. Synthetic toll-like receptor 4 (TLR4) and TLR7 ligands as influenza virus vaccine adjuvants induce rapid, sustained, and broadly protective responses. *J Virol*. 2015;89(6):3221-3235.

28. Wu Y, Li Z, Wang S, Xiu A, Zhang C. Carvedilol inhibits angiotensin ii-induced proliferation and contraction in hepatic stellate cells through the RhoA/Rho-Kinase pathway. *Biomed Res Int*. 2019;2019:7932046.