SUPPLEMENTAL MATERIAL FOR:

COVID-19 Outpatient Screening: a Prediction Score for Adverse Events

Contents of these Supplemental Materials include:

1. Technical Details Regarding Model Training using Nested Cross Validation

The training and validation procedures are summarized in Figure S1. Nested five-fold cross validation was carried out, with outer and inner CV loops. The outer CV loop is used to obtain the performance on the development cohort for concurrent validation. For the outer CV loop, the development cohort was split into five folds containing distinct subsets of patients. For each loop, one fold was held out for testing, while the other four were used for training. Repeating this 5 times, and concatenating predictions from each of the resulting five independent models, we get estimate performance on the development cohort. For each external CV loop, an inner CV loop is used to choose the optimal value of model hyperparameter, i.e. the strength of the LASSO penalty (lambda) that maximizes the F1 score across 5 folds of internal CV[15,16].

In both outer and inner CV loop, prior to model training, we standardized each predictor in the training set by subtracting its mean and dividing its standard deviation (SD) (z-score) across the training set. The testing set was standardized in the same way, using means and SDs compauted on the development cohort.

After CV, a final model with hyperparameters that are one-level stronger than the most common hyperparameter among the five independent models was fit on the entire development cohort. This final model was the one tested on the prospective cohort, after standardization using the means and SDs from the development cohort.

The LTR model yields probabilities for each potential adverse event occurring within the next 7 days. These are used to create acuity scores by taking the expected outcome value (weighted sum of the ordinal adverse event outcome values (0,1,2, and 3) with weights equal to the event probabilities). The acuity values are finally rescaled for ease of interpretation to the range 0 (minimum acuity) to 100 (maximum acuity). For a given acuity score, CoVAS provides a predicted probability of each adverse outcome. These predicted probabilities are obtained by ordinal Platt calibration[17], meaning that we fit an ordered logit model with the CoVAS acuity score as a single input, to provide probabilities for hospitalization, critical illness (need for ICU care or MV), or death. Platt calibrated probabilities are used instead of directly using the original model probabilities used in the acuity scores because, empirically, the original probabilities are not well calibrated.

2. Predictor encoding conventions

For predictors that are binary, we use the following encoding:

encode 1 = present or positive;

-1 = absent or negative;

0 = unavailable or no record of testing.

For smoking status, we used 0 = never or passive; and 1 = yes or quit.

3. Details of computing CoVA Scores and Risk Predictions

The following is a set of instructions for anyone wishing to write a computer program that calculates the CoVA score (e.g. in an electronic medical record).

The mean, std, and coef values are provided in Table S1 (see below).

1. Define input features
If no chest X-ray is available, set to 0; otherwise, pattern present = 1, no pattern = -1.
If no COVID status available, set to 0; otherwise, test positive = 1, negative = -1.

2. z-score standardization: substract the mean, divide by the standard deviation for each feature,
x_i ← (x_i - mean_i) / std_i

3. Compute z (a number for each patient)
z = beta0 + beta1 x1 + ... + betaK xK

4. Compute yhat (a 4-element probability distribution for each patient)
yhat = [exp(-(z-mu_none)^2) / M,
exp(-(z-mu_hosp)^2) / M,
exp(-(z-mu_icu_intub)^2) / M,
exp(-(z-mu_death)^2) / M]
where M = exp(-(z-mu_none)^2) + exp(-(z-mu_hosp)^2) + exp(-(z-mu_icu_intub)^2) + exp(-(z-mu_death)^2)

mu_none = -0.7138;
mu_hosp = 0.5514;
mu_icu_intub = 1.1932;
mu_death = 2.6278.

5. Compute acuity score, a, which the expectation of yhat

a = E[yhat] = 0*yhat_none + 1*yhat_hosp + 2*yhat_icu_intub + 3*yhat_death
where
yhat_none is the 1st element in yhat;
yhat_hosp is the 1st element in yhat;
yhat_icu_intub is the 1st element in yhat; and
yhat_death is the 1st element in yhat.

Note: For display, we rescale the acuity score to the range 0-100. This is done by scaling the acuity score, a, to obtain a new score, A = 100*a/3

6. Recalibration: compute the predicted distribution (yp) based on the acuity score

tmp0 = 0
tmp1 = sigmoid(1.7836 - 1.4130*acuity score)
tmp2 = sigmoid(4.1649 - 1.4130*acuity score)
tmp3 = sigmoid(6.4918 - 1.4130*acuity score)
tmp4 = 1
where sigmoid(x) = 1/(1+exp(-x))

P(none) = (tmp1-tmp0) *100%
P(hosp) = (tmp2-tmp1) *100%
P(icu_intub) = (tmp3-tmp2) *100%
P(death) = (tmp4-tmp3) *100%

Finally, yp = [P(none), P(hosp), P(icu_intub), P(death)]

Table S1: Model parameter values. Values used for z-normalization (mean, std) and CoVA model coefficients.

Predictor	Mean	Std	beta
Age	51.07594	19.22809	0.7353
Sex	0.51103	0.49988	0.08078
Respiratory Rate	18.80844	3.1239	0.27459
Heart Rate	81.13687	14.76491	0.12154
Temperature	98.13069	0.9537	0.1206
CCI	1.46552	2.41655	0.11423
BMI_high (>35kg/m2)	0.03475	0.18315	0.00284
BMI_low (<18.5kg/m2)	0.00917	0.09531	-0.00014
SpO2	97.26989	2.24487	-0.37759
Diastolic blood pressure	73.29219	12.25362	-0.47241
Systolic blood pressure	130.36659	20.1278	-0.11512
Ais_hx	0.05916	0.23593	0.17463
Ich_hx	0.0081	0.08964	0.10869
Sah_hx	0.00586	0.07634	0.09191
Hem-cx_hx	0.02089	0.14303	0.07649
Renal-cx_hx	0.0177	0.13184	0.0681
Pancreatitis_hx	0.01514	0.1221	0.0608
Cf_hx	0.00107	0.03263	0.04917
Arrest_hx	0.0032	0.05646	0.04914
Seizure_hx	0.02185	0.1462	0.04369
Als_hx	0.00128	0.03574	0.04054
M-acid_hx	0.01546	0.12336	0.03845
Mg_hx	0.00181	0.04253	0.03735
Pneumothorax_hx	0.00149	0.0386	0.03003
Sma_hx	0.00171	0.04126	0.02414
Pericarditis_hx	0.00842	0.09138	0.01441
ARDS_hx	0.00693	0.08295	0.00011
ever_positive_upToEvent	0	1	0.275
cxr_Multifocal	0	1	0.12929
cxr_TypicalPatternForCovid	0	1	0.00014
beta0			-0.09257

Table S2. Prior risk predictions models for COVID-19.

This table summarizes relevant studies reviewed in "Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal", BMJ 2020; 369 doi: https://doi.org/10.1136/bmj.m1328, 07 April 2020) and BMJ 2020;369:m2204, https://www.bmj.com/content/369/bmj.m2204), reviewed on June 10, 2020.

Study/location	Population	Outcome	Predictors	Sample size for model development	Type of model	Validation method	Sample size for Validation	Performance
Decaprio et al[1], United States, [Preprint]	General Population	Complications from pneumonia (proxy events, not COVID-19)	Age, sex, previous hospital admissions, diagnostic features, interaction terms	~1.5 million	LR XGB	Training test split	369,865	AUC: LR: 0.73 XGB: 0.81
Bai et al[2], China, [preprint]	Inpatients, covid+	severe disease	Clinical variables, lab values, chest CT features	133 (54)	LSTM + MLP	Unclear	N/A	AUC 0.95
Gong et al[3], China, [Preprint]	Inpatients COVID-19+	severe disease within 15 days	Clinical variables, lab values	189 (28)	LR+ LASSO, RF, DT, SVM	External validation	Two external validation sites 165 (40) and 18 (4)	AUC 0.85
Qi et al[4] China	Inpatients, COVID-19+	hospital stay >10 days	features from CT images	26 (20)	LR RF	5 fold cross validation	N/A	AUC; LR: 0.92 RF: 0.96
Shi et al[5] China	Inpatients COVID-19+	Death or severe COVID-19	Age, sex, HTN	487 (49)	LR	Validation in less severe	66(15)	Not reported
Xie et al[6] China [preprint]	Inpatients COVID-19+	mortality	Age, LDH, lymphocyte count, SpO2	299(155)	LR	External Validation	145 (69)	AUC: 0.98
Yan et al[7], China	Inpatients Suspected COVID-19+	mortality	LDH, lymphocyte, hs-CRP	375 (174)	XGB	Prospective validation	29 (17)	Sensitivity 92 PPV 95
Yuan et al[8] China	Inpatients COVID-19+	mortality	CT imaging score	27 (10)		N/a	Na/a	AUC: 0.90
Huang et al[9] China [preprint]	Inpatients COVID-19+	severe disease progression	comorbidities, RR, CRP, LDH	125(32)	LR		N/a	AUC: 0.99

Pourhomayoun[10] Multiple countries [preprint]	Inpatients COVID-19+	mortality	Not defined	117,000	SVM, ANN, RF, DT, LR, KNN	10 fold cross validation	n/a	AUC: 0.96
Sarkar[11] [preprint] Multiple countries	Inpatients with COVID symptoms	mortality	Age, days of symptoms prior to admission, travel or from Wuhan	115(37)	RF	N/a	N/a	AUC: 0.97
Wang[12] China	Inpatient COVID 19+	LOS	CT features	485	DL	External validation	290	AUC 0.81
Zeng[13] [preprint] China	Inpatients COVID-19+	Severe disease progression	CT and labs	338(76)	Multivariate survival analysis	Cross validation		AUC: 0.88
Legend: LR = logistic regression; RF = random forest; DT = decision tree; MLP = multilayer perceptron; SVM = support vector machine; LASSO = least absolute shrinkage and selection operator (aks L1 regularization); XGB = XGboost, aka gradient boosted decision trees; Covid+ = tested positive for Covid-19, LOS= length of hospital stay								

REFERENCES

1. DeCaprio D, Gartner J, Burgess T, Kothari S, Sayed S. Building a COVID-19 vulnerability index. arXiv preprint arXiv:200307347 **2020**;

2. Bai X, Fang C, Zhou Y, et al. Predicting COVID-19 malignant progression with AI techniques. **2020**;

3. Gong J, Ou J, Qiu X, et al. A Tool to Early Predict Severe 2019-Novel Coronavirus Pneumonia (COVID-19) : A Multicenter Study using the Risk Nomogram in Wuhan and Guangdong, China. medRxiv **2020**; :2020.03.17.20037515.

4. Qi X, Jiang Z, Yu Q, et al. Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: A multicenter study. medRxiv **2020**; :2020.02.29.20029603.

5. 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. Critical Care **2020**; 24:1–4.

6. Xie J, Hungerford D, Chen H, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. **2020**;

7. Yan L, Zhang H-T, Goncalves J, et al. A machine learning-based model for survival prediction in patients with severe COVID-19 infection. MedRxiv **2020**;

8. Yuan M, Yin W, Tao Z, Tan W, Hu Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. Plos one **2020**; 15:e0230548.

9. Huang H, Cai S, Li Y, et al. Prognostic factors for COVID-19 pneumonia progression to severe symptom based on the earlier clinical features: a retrospective analysis. medRxiv **2020**;

10. Pourhomayoun M, Shakibi M. Predicting mortality risk in patients with COVID-19 using artificial intelligence to help medical decision-making. medRxiv **2020**;

11. Sarkar J, Chakrabarti P. A Machine Learning Model Reveals Older Age and Delayed Hospitalization as Predictors of Mortality in Patients with COVID-19. medRxiv **2020**;

12. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. European Respiratory Journal **2020**;

13. Zeng L, Li J, Liao M, et al. Risk assessment of progression to severe conditions for patients with COVID-19 pneumonia: a single-center retrospective study. medRxiv **2020**;

Table S3. ICD codes used to define the diagnoses

Predictor category	Abbreviation	Predictor	ICD-10 codes
Demographics	Age	Age	N/A
	BMI	BMI	N/A
	covid-test	COVID-19 testing status	N/A
	Smoking	Smoking	F17.201, F17.208, F17.203, F17.209, F17.210, F17.200, F17.218, F17.291, F17.299, F17.298, Z72.0, Z87.891, O99.334, O99.335, O99.330, K03.6, T65,291
Vital signs	VS	Systolic blood pressure, Diastolic blood pressure, Temperature, Heart rate, Respiratory Rate, Oxygen saturation (SpO2)	N/A
Symptoms	Anosmia	Anosmia	R43.0, R43.8, R43.9
	Dysgeusia	Dysgeusia	R43.2
Infectious	Hiv	Human immunodeficiency virus infection (HIV)	B20, Z21, O98.71

	Tb	Tuberculosis	A15.0, A15.4, A15.5, A15.6, A15.7, A15.8, A15.9, A17.0, A17.1, A17.81, A17.82, A17.89, A17.9, A18.01, A18.02, A18.03, A18.10, A18.11, A18.12, A18.13, A18.15, A18.17, A18.18, A18.2, A18.31, A18.32, A18.39, A18.4, A18.50, A18.6, A18.7, A18.81, A18.83, A18.85, A18.89, A19.2, A19.8, A19.9
Pulmonary	Ards	Acute respiratory distress syndrome - acute	J80
	Ards-hx*	Acute respiratory distress syndrome	Z87.09
	Asthma	Asthma	J45.51, J45.40, J45.41, J45.42, J45.30, J45.31, J45.32, J45.20, J45.52, J45.22, J45.20, J45.909, J45.902, J44.9, J44.1
	Bronchitis	Bronchitis	J40, J41, J41.0, J41.1, J41.8, J42
	COPD	Chronic obstructive pulmonary disease	J44.9, J98.8
	Cf	Cystic fibrosis	E84.9
	Copd	Chronic obstructive pulmonary disease	J44.9, J44.1, J44.0, J44.9
	Ild	Interstitial lung disease	J81, J82, J84, J84.0, J84.1, J84.8, J84.9
	Pe	Pulmonary embolism	I26, I26.0, I26.9
	Pneumonia	Pneumonia	A01.03, A02.22, A37.01, A37.11, A37.81, A37.91, A48.1, A50.04, A54.84, B05.2, B06.81, B77.81, J09.X1, J10.0, J10.00, J10.01, J10.08, J11.0, J11.00, J11.08, J12.1, J12.2, J12.3, J12.8, J12.81, J12.89, J13, J14, J15.0, J15.1, J15.2, J15.20, J15.21, J15.211, J15.212, J15.29, J15.3, J15.4, J15.5, J15.7, J15.8, J15.9, J16, J16.0, J16.8, J18.0, J18.1, J18.2, J18.8, J84.111, J84.116, J12.0, J82, A22.1, J69.0, J12.9, J69.1, J69.8, J95.4, J18.9, J84.89, J68.0, J15.6, J18.89, J95.851, B01.2, A21.2, J15.29,

			A69.8, J15.8, I00
	Pneumothorax	Pneumothorax	J93.0, J95.811, J93.9, J93.83
Cardiovascular	Arrest	Cardiac arrest	I46, I46.0, I46.1, I46.9
	Arrythmia	Arrythmia	I48, I49, I49.0, I49.1, I49.2, I49.3, I49.4, I49.5, I49.8, I49.9
	CM	Cardiomyopathy	I42, I42.0, I42.1, I42.2, I42.3, I42.4, I42.5, I42.6, I42.7, I42.8, I42.9, I43, I43.0, I43.1, I43.2, I43.8
	Cad	Coronary artery disease - acute	I24.0, I24.1, I24.9, I24, I25.10, I25.9, I25.810, I25.812, I25.719, I25.729, I25.721, I25.730, I25.750, I25.751, I25.739, I25.760, I25.700, I25.709, I25.720, I25.759, I25.41, I25.42, I25.710, I25.711, I25.701, I25.761, I25.731, I25.769, I25.75, I25.701, I25.799, I25.728, I25.791, I25.798, I25.110, I25.111, I25.119, I25.118, I21.09, I21.19, I21.29, I21.4, I21.9
	Cad-hx	Coronary artery disease	Z95.1, Z98.61, Z95.5, Z86.79
	Chf	Congestive heart failure - acute	I50.32, I50.33, I50.31, I50.20, I50.40, I50.9, I50.30, I50.1, I50.23, I50.22, I50.43, I50.21, I50.41, I50.9, I11.0, I50, I50.0, I50.1, I50.9
	Chf-hx	Congestive heart failure	Z82.49

	Dm	Diabetes	E11.9, E10.9, E13.9, E11.69, E09.9, E11.69, E11.10, E11.01, E11.39, E10.10, E10.69, E10.65, E10.49, E11.49, E10.29, E11.29, E10.29, E10.51, E11.51, E11.8, E10.8, E11.40, E10.40, E11.41, E10.41, E10.641, E11.21, E11.39
	Htn	Hypertension - acute	I10, I11.9, I11.0, I12.9, I12.0, I13.10, I13.0, I13.11, I13.2, I15.2, I15.8, I15.0, G93.2, I15.0, I15.9, I10.02, I97.3, I15.1, I15.2
	Htn-hx	Hypertension	Z86.79
	Htn-obs	Hypertension - obs	O10.419, O10.42, O10.43, O10.419, O10.93, O10.919, O13.9, O13.5, O16.9, O16.5, O16.9, O10.119, O10919, O10.92, O10.319, O10.119, O10.219, O10.419, O10.23, O13.9, O10.03, O13.5, O16.9, O16.5, O10.019
	Mi	Myocardial infarction	I21, I22
	Myocarditis	Myocarditis	I40, I40.0, I40.1, I40.8, I40.9, I41, I41.0, I41.1, I41.2, I41.8
	Obesity	Obesity	E66.9, E66.01, E66.8, E66.1, E66.01
	Ow	Obesity	E66.3
	Pericarditis	Pericarditis	I30, I30.0, I30.1, I30.8, I30.9, I31, I31.0, I31.1, I31.2, I31.3, I31.8, I31.9, I32
	Pvd	Peripheral vascular disease	I73, I73.0, I73.00, I73.01, I73.1, I73.8, I73.81, I73.89, I73.9
Renal	Aki	Acute kidney injury	N17.9
	Ckd-end	Chronic kidney disease - end stage	N18.5, N18.6, Z99.2
	Ckd-iv	Chronic kidney disease - Stage IV	N18.1, N18.2, N18.3, N18.4

	Renal-cx	Renal cancer	C34.90, C34.91, C34.92, C78.01, C64.9, C64.21, C68.2, C66.1, C34.91, C64.9
Neurologic	Ais	Ischemic stroke - acute	I63.9, I63.81, I67.9, R27.0, G46.4, I63.232, I63.232, I63.231, I63.512, I63.511, I63.511, I63.522, I63.21, I63.531, I63.532, I63.29, I63.219, I63.22, I63.529, I63.539, I63.9, I63.89, I63.429, I63.12, I63.139, I63.449, I63.40, I63.422, I63.132, I63.442, I63.412, I63.432, I63.112, I63.419, I63.439, I63.10, I63.421, I63.131, I63.441, I63.411, I63.431, I63.111, I63.119, I63.329, I63.02, I63.039, I63.349, I63.30, I63.322, I63.032, I63.342, I63.312, I63.332, I63.012, I63.319, I63.339, I63.321, I63.031, I63.341, I63.311, I63.331, I63.011, I63.019
	Ais-hx	Ischemic stroke	Z86.73, Z86.69, I69.359, I69.30, I69.391, I69.398, I69.390, I69.393, I69.319, i69.321, i69.359, i69.328, i69.398, i69.349, i69.365, i69.369, i69.398, r20.9
	Als	Amyolateral Sclerosis	G12.21
	Brain-cx	Brain malignancy	C71.9, C72.30, C69.22, C69.21, C71.6, C71.1, C71.0, Z85.841
	CIM	Critical illness myopathy	G72.81
	CIN	Critical illness neuropathy	G62.81
	Ce	Cerebral edema	G93.6
	Cmt	Charcot Marie Tooth Disease	G60.0
	Coma	Coma - history of	R40.20
	Dementia	Dementia	G30.9, G30, F02.80, F03.90, G31.09, G30.0, F02.81
	Gbs	Guillain Barre Syndrome	G61.0

	Hemiplesia	Hemiplegia	G81, G81.0, G81.1, G81.9, G82, G82.0, G82.1, G82.2, G83
	Hydrocephalus	Hydrocephalus	G91, G91.0, G91.1, G91.2, G91.3, G91.8, G91.9
	Ich	Intracranial hemorrhage - acute	I62.9, i62.1, I62.00, I61.5, I61.4, S06.300A, S06.309A, S06.369A, S06.305A, S06.2X9A, S06.0XXA, S02.91XA, S02.109A, S02.0XXA, S02.109B, S02.0XXK, S02.109G, S02.0XXD, S02.109D, S02.91XK, S02.91XB, S02.91XA
	Ich-hx	Intracranial hemorrhage	I69.293, I69.262, I69.261, I69.220, I69.290, I69.223, I69.26, I69.292, I69.263, I69.362, I69.227, I69.244, I69.241, I69.249, I69.239, I69.2, I69.29, I69.232, I69.25, I69.24, I69.23, I69.21, I69.252, I69.269, I69.254, I69.228, I69.233, I69.242, I69.234, I69.231, I69.243, I69.259
	Ih	Intracranial hyertension	G93.2
	Inf-men-enc	Meningitis or encephalitis	A86, A87.2, A87.9, A85.1, A85.2, A87.1, A87.0, Z86.61, A32.11, A17.0, A39.0, A37.00, A69.21, A87.2, A85.8, A83.2, A84.1, A84.8, A83.5, A83.2, G03.8, G03.8, A32.11, A92.39, G00.1, A92.32, A52.13, G00.8, B45.1, G02, A27.81, A51.41, B38.4, G01, B00.3, G00.9, B39.9, B00.4, G03.9, B01.0, B02.1, G08, G03.1, G03.0, B01.11, B06.01, B60.2, B30.8, A07.8, A85.8, J11.81
	Mg	Myasthenia gravis	G70.00, G70.01, G70
	Movement	Movement disorder	G21, G21.0, G21.1, G21.1, G21.3, G21.8, G21.9, G22, G23, G23.0, G23.1, G23.2, G23.8, G23.9, G24, G24.0, G24.1, G24.2, G24.3, G24.4, G24.5, G24.8, G24.9, G25, G25.0, G25.1, G25.2, G25.3, G25.4, G25.5, G25.6, G25.8, G25.9, G26, G10, G11, G11.0, G11.1, G11.2, G11.3, G11.4, G11.8, G11.9, G12, G12.0, G12.1, G12.2, G12.8, G12.9, G13, G13.0, G13.1, G13.2, G13.8

	Ms	Multiple sclerosis	G35, G37.9, G37.8, G35, G36, G36.0, G37, G37.9
	Neuromuscular	Neuromuscular	G70, G70.0, G70.1, G70.2, G71, G71.0, G71.1, G71.2, G71.3, G71.8, G71.9, G72, G72.0, G72.4, G73, G73.0, G73.1, G73.3, G73.4, G73.5, G73.6, G73.7
	Neuropathy	Neuropathy	G60, G60.0, G60.1, G60.2, G60.3, G60.8, G60.9, G61, G61.1, G61.8, G61.9, G62, G62.0, G62.1, G62.2, G62.8, G62.9, G63, G63.1, G63.2, G63.3, G63.4, G63.5, G63.6, G63.8, G64, G50, G50.0, G51.0, G51.0
	PbP	Pseudobulbar palsy	G12.22
	Pls	Primary lateral sclerosis	G.12.23
	Pn	Peripheral neuropathy	G90.0, G90.09, G60.9
	Sah	Subarachnoid hemorrhage - acute	I60.9, I60.8, I60.5, I60.00, I60.2, I60.01, I60.02, I60.00, I60.51, I60.52, I60.50, I60.31, I60.32, I60.7, I60.4, I60.10, I60.11, I60.12, I60.31, I60.32, I60.3, I60.30, I60.20, I60.10, I60.6, S06.6X9A, S06.6X0A, S06.6X5A, S01.90XA
	Sah-hx	Subarachnoid hemorrhage	z86.79, I69.093, I69.021, I69.032, I69.043, I69.039, I69.044, I69.00, I69.020, I69.091, I69.033, I69.041
	Sdh	Subdural hematoma - acute	S06.5X9A, S02.0XXA, S02.0XXB, S02.109A, S02.109B, S06.5X9D, S06.5X5A, S06.5X0A, S06.6X9A, S06.4X9A, I62.00
	Sdh-hx	Subdural hematoma	Z98.890

	Seizure	Seizure / epilepsy	G40.90, G40.901, G40.909, G40.91, G40.911, G40.919, G40.811, G40.80, G40.801, G40.802, G40.811, G40.812, G40.813, G40.814, G40.821, G40.822, G40.823, G40.824, G40.89, G40.A, G40.3, G40.30, G40.31, G40.2, G40.20, G40.21, G40.10, G40.101, G40.109, G40.11, G40.111, G40.119, G40.1, G40.0, G40.00, G40.001, G40.009, G40.011, G40.019
	Sfn	Polyneuropathy	G62.9
	Sma	Spinal muscular atrophy	G12.9, G12.0, G12.1, G12.21, G12.8
	Vst	Venous sinus thrombosis	G08, I67.6, E86.0, O22.50, O22.51, O22.52, O22.53, G03.9, D68.59, H70.90, Z86.718, Q24.9, H70.93
	Parkinsons	Parkinsons	G20
Gastrointestinal	Colitis	Colitis	K50, K50.0, K50.1, K50.8, K50.9, K51, K51.0, K51.1, K51.2, K51.8, K51.9, K52, K52.1, K52.2, K52.8, K52.9
	Hepatitis	Hepatitis	B15, B15.0, B15.9, B16, B16.0, B16.1, B16.2, B16.9, B17, B17.0, B17.1, B17.2, B17.8, B18, B18.0, B18.1, B18.2, B18.8, B18.9, B19, B19.0, B19.9
	Liver	Liver disease	K70, K70.1, K70.2, K70.3, K70.4, K70.9, K71, K72, K72.0, K72.1, K72.9, K73, K73.0, K73.1, K73.2, K73.8, K73.9, K74, K74.1, K74.0, K74.2, K74.3, K74.4, K74.5, K74.6, K75, K76, K77
Hematological	Hem-cx	Hematologic malignancy	C81.90, C85.90, C85.92, C85.10, C85.93, C85.89, C85.80, C85.99, C88.4, C82.99, C82.80, C86.6, C83.19, C83.59, C91.60, C91.00, C91.01, C91.10, C92.00, C92.10, C92.01, C92.20, C90.00, C90.02, C90.30, Z94.81, D47.1, D46.9, D46.20, D63.0, D94.6
Psychiatric	Anxiety	Anxiety	F41.0, F41.1, F41.3, F41.8, F41, F40.0, F40, F40.1, F41.9

	Md	Depression	F33, F33.0, F33.1, F33.2, F33.3, F33.4, F33.8, F33.9, F32, F32.0, F32.1, F32.2, F32.3, F32.4, F32.5, F32.8, F32.9
	Ocd	Obsessive comulsive disorder	F42, F42.2, F42.3, F42.4, F42.8, F42.9
Miscellaneous	Ar	Allergic rhinitis	J30.1, J30.9
	CTD	Connective tissue disease	M30, M30.0, M30.1, M30.2, M30.3, M30.8, M31, M31.0, M31.1, M31.2, M31.3, M31.4, M31.5, M31.6, M31.8, M31.9, M32, M32.0, M32.1, M32.8, M32.9, M33, M33.0, M33.1, M33.2, M33.9, M34, M34.0, M34.1, M34.2, M34.8, M34.9, M35, M35.0, M35.1, M35.2, M35.3, M35.4, M35.5, M35.6, M35.7, M35.8, M35.9, M36, M36.0, M36.1, M36.2, M36.3, M36.4, M36.8
	Gerd	Gastroesophageal reflux disease	K21..9, Z87.81, K21.0
	Hypothyroidism	Hypothyroidism	E03.9, E03.8, E03.5, E03.4, E03.3, E03.2, E03.1, E03.0
	M-acid	Acidosis	E87.2, E87.4, E87.3
	Myositis	Myositis	M60.9, M60.10, M60.20, M60.009, M60.89, M61.00, M61.10, M61.50, M62.50, M62.59, M33.90, M33.20, M33.10, M33.00, M72.6, H05.129
	Oa	Osteoarthritis	M15, M16, M17, M18, M19, M19.0, M19.1, M19.2, M19.9, M15.1, M15.0, M15.2, M15.3, M15.4, M15.9, M15.8
	Osa	Obstructive sleep apnea	G47.33
	PU	Peptic ulcer disease	K27.0, K27.9

	Pancreatitis	Pancreatitis	K85.0, K85.00, K85.1, K85.10, K85.2, K85.20, K85.3, K85.30, K85.8, K85.80, K86.0, K86.1, K85, K85.01, K85.02, K85.11, K85.12, K85.21, K85.22, K85.31, K85.32, K85.81, K85.82, K85.9, K85.90, K85.91, K85.92
	Ra	Rheumatoid arthritis	M06.9
	Rhabdo	Rhabdomyolysis	M62.82
	Sarcoidosis	Sarcoidosis	D86.9, D86, D86.0, D86.1, D86.2, D86.3, D86.8, D86.9

* Given a particular disease, it is _hx=1 if ICD ever mentioned before the event. If it's only mentioned after the event and never before, the category without _hx =1 and the _hx one is 0. If it's never mentioned both are 0

Table S4. Key phrases and groupings used to create predictors from CXR reports.

ID	Group	Key phrases
1	Multifocal	diffuse opacities bilateral lung opacities diffuse ground glass multifocal viral pneumonia multifocal pneumonia multifocal patchy airspace opacities (multifocal)...(opacit)* viral pneumonia (compatible with|may represent|suggesting|consistent with|in keeping with) … (viral pneumonia) * patchy opacity (patchy opacities|patchy opacities) new patchy airspace opacities patchy pneumonia patchy consolidation (consolidative|patchy)...(opacities|opacity) ground glass ground glass opacity

<table>
<tr><td></td><td></td><td>ground glass opacities
(groundglass)...(opacities|opacity)*
ggo</td></tr>
<tr><td>2</td><td>Typical pattern for COVID-19</td><td>typical pattern for covid
consistent with covid pneumonia
covid 19 pneumonia
could be due to covid
(seen in) … (covid) **
(concerning for) … (covid) *
(compatible with|may represent|suggesting|consistent with|in (keeping with) ...(covid) *
consistent with pneumonia
likely representing pneumonia
likely pneumonia
concerning for infection
(infection) … (not excluded)*
can’t exclude consolidation
(compatible with|may represent|suggesting|consistent with|in keeping with) ... (pneumonia) *
(compatible with|may represent|suggesting|consistent with|in keeping with) … (bronchopneumonia) *
(compatible with|may represent|suggesting|consistent with|in keeping with) … (ards) *</td></tr>
<tr><td>3</td><td>Consolidation</td><td>focal infiltrate
infiltrate
consolidation
(consolidation) … (not excluded)*
lobe consolidation
confluent airspace opacities
confluent opacities
consolidative opacities</td></tr>
<tr><td>4</td><td>Peripheral or interstitial opacity</td><td>peripheral opacities
peripheral opacity
faint interstitial opacities
reticular opacities</td></tr>
</table>

5	Hazy or airspace opacity	hazy opacities airspace opacities (airspace) ... (opacities)

Table legend:

* = (\W+(?:\w+\W+){0,10}?)

** = (\W+(?:\w+\W+){0,5}?)

(x|y) = x or y

Table S5. Range of physiologically plausible predictor values

Predictor	Range	Unit
Age	18 - 110	year
Systolic blood pressure	50 - 225	mmHg
Diastolic blood pressure	25 - 150	mmHg
Body temperature	94 - 105	^{o}F
Heart rate	33 - 195	Per minute
Respiratory rate	8 - 55	Per minute
BMI	9 - 80	kg/m^2
SpO2	50 - 100	%

Table S6. Reasons for being included in either development or prospective cohorts

Reason	**n**	**%**
Possible COVID Related	4163	35.9
Pain|Medication Refill|Other	1638	14.1
Gastrointestinal|Urinary - Unlikely COVID Related	1043	9
Neurological|ENT|Ophthalmological	841	7.3
Surgical|Trauma|Vascular|Skin, Soft tissue, Bone, Blood Infection	777	6.7
Cardiac|Pulm - Unlikely COVID Related	546	4.7
Psychiatry|Substance Use|Overdose	537	4.6
Endocrine|Rheumatological|Dermatological	93	0.8
Abnormal Lab|Imaging	51	0.4
Gyn|Obstetrics	43	0.4
Missing	1854	16

Table S7. Univariate association of each predictor and the outcome.

Predictor	$\log_{10}$(p-value)
SpO2	-272.5
Age	-271.5
Diastolic blood pressure	-238.1
Respiratory rate	-155.6
CCI	-133.6
Chf_hx	-74.2
Systolic blood pressure	-59.2
Body temperature	-54.7
Pneumonia_hx	-50.6
Cad_hx	-49.5
Aki_hx	-47.0
Htn_hx	-47.0

Ckd-end_hx	-46.4
Ckd-iv_hx	-46.4
High BMI (>35kg/m^2)	-44.2
Dm_hx	-28.6
Ais_hx	-25.5
TobaccoUserDSC	-17.3
Mg_hx	-16.4
Copd_hx	-15.9
Renal-cx_hx	-15.2
Low BMI (<18.5kg/m^2)	-15.0
Ich_hx	-14.5
Arrythmia_hx	-13.9
sex	-13.1
M-acid_hx	-11.9
Hem-cx_hx	-11.3
Arrest_hx	-9.6
Als_hx	-9.5
Pancreatitis_hx	-9.4
Sdh_hx	-9.3
HR	-9.1
CM_hx	-8.4
Neuropathy_hx	-7.5
Sma_hx	-6.3

Sah_hx	-6.0
Anxiety_hx	-5.7
Seizure_hx	-5.2
Ards_hx	-4.8
Ild_hx	-4.7
Pneumothorax_hx	-4.6
Vst_hx	-4.4
Parkinsons_hx	-4.2
Bronchitis_hx	-3.9
Hypothyroidism_hx	-3.3
Ow_hx	-3.2
Sfn_hx	-3.0
Pericarditis_hx	-3.0
Ar_hx	-2.8
Brain-cx_hx	-2.6
CTD_hx	-2.6
Ce_hx	-2.3
Movement_hx	-2.1
Colitis_hx	-2.1
Cf_hx	-2.1
Hydrocephalus_hx	-2.0
Ra_hx	-1.9
Osa_hx	-1.6

Md_hx	-1.5
Myositis_hx	-1.2
Sarcoidosis_hx	-1.1
Pn_hx	-1.1
Anosmia_hx	-0.9
CIM_hx	-0.8
Ocd_hx	-0.8
Inf-men-enc_hx	-0.7
Oa_hx	-0.7
Gbs_hx	-0.6
Obesity_hx	-0.6
Dysgeusia_hx	-0.5
Asthma_hx	-0.5
PU_hx	-0.5
Hiv_hx	-0.5
Rhabdo_hx	-0.5
Cmt_hx	-0.4
CIN_hx	-0.3
Gerd_hx	-0.3
Ms_hx	-0.3
Hepatitis_hx	-0.3
Coma_hx	-0.2
Tb_hx	-0.2

Myocarditis_hx	-0.1
Neuromuscular_hx	0.0
Ih_hx	0.0

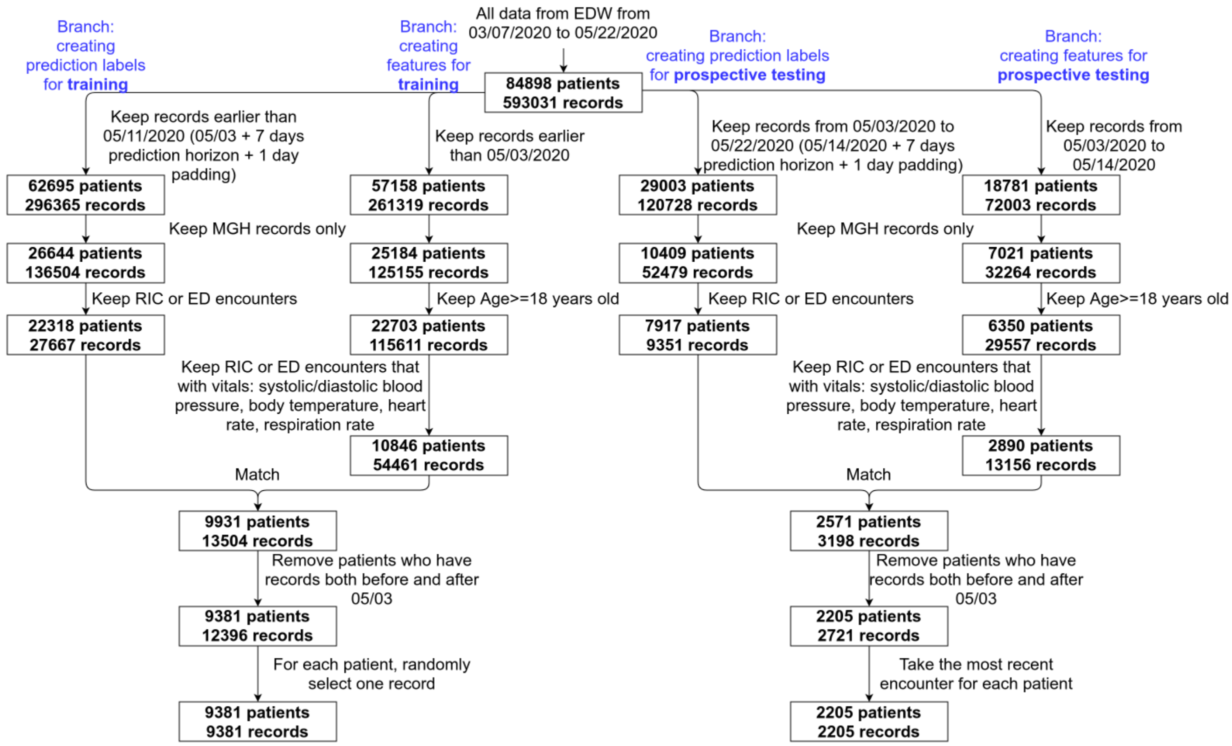

Figure S1. Data flowchart to generate the predictors and outcomes for the development and prospective cohorts.

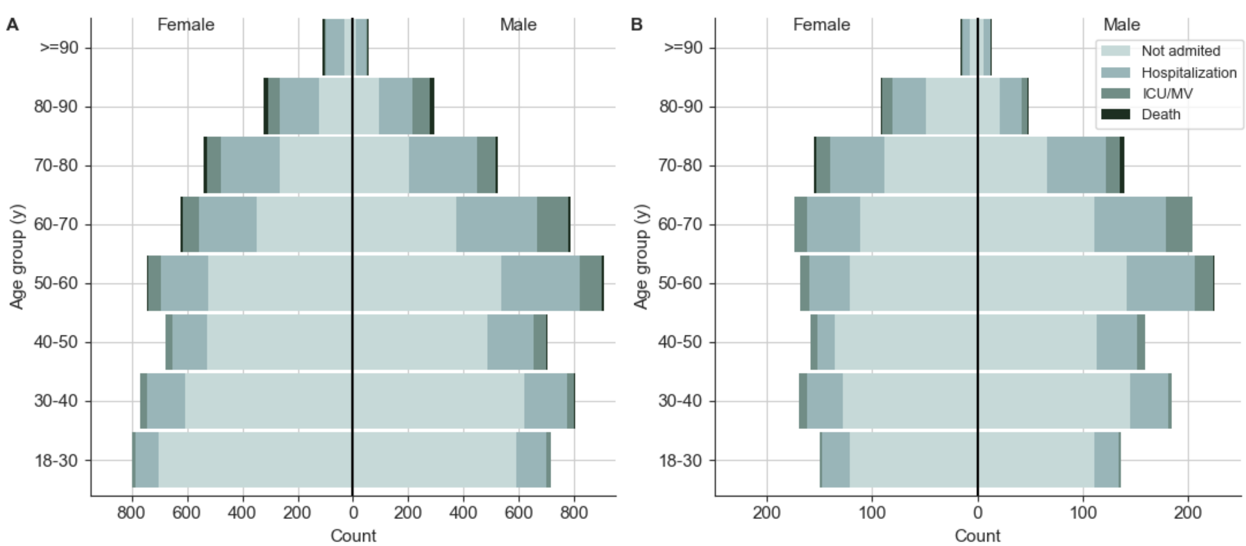

Figure S2. Distribution of adverse events at different decades and sexes. (A) Development cohort; (B) Prospective cohort.

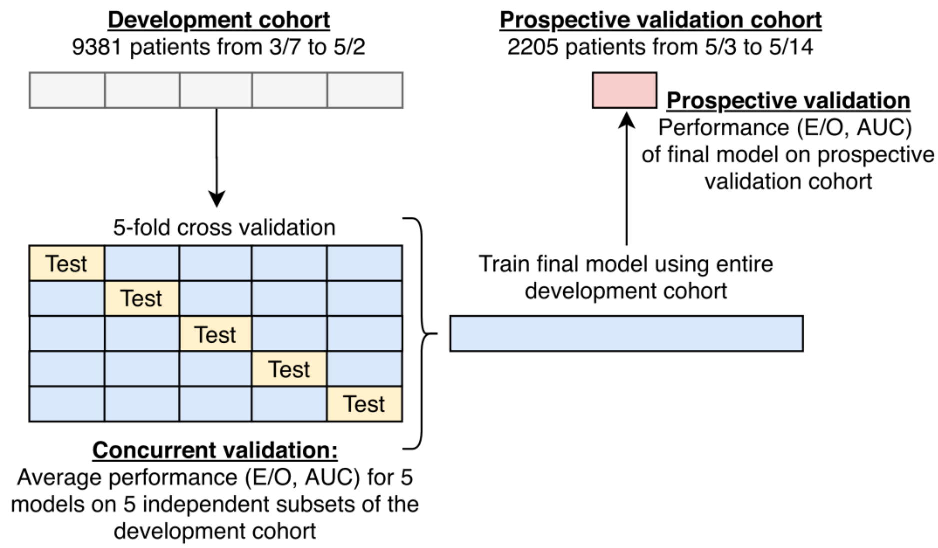

Figure S3. Model training and validation schema. Concurrent validation in the model development cohort is based on nested cross-validation. Prospective validation is done using the final model trained on the development cohort.

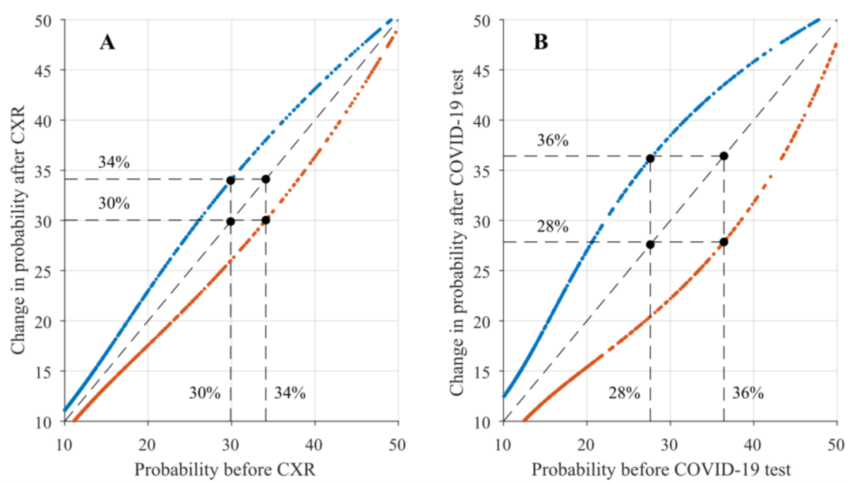

Figure S4. Relationship in the CoVA score between pre- and post-test probability (critical illness or death) following a chest X-ray (CXR) (left) or COVID-19 testing (right). The diagonal line shows the pre-test probability. The blue curve shows the post-test probability following a positive test result (e.g. both CXR findings in the CoVA model are present, or positive COVID-19 test). The red curve shows the post-test probability following a negative result. Black dots show the values for which a positive or negative test result causes the large difference between the pre- and post-test probability, i.e. the pre-test probabilities at which testing is maximally informative. For CXR, the positive results cause the largest increase (4%) when the pre-test probability is 30%, and the largest decrease (4%) when the pre-test probability is 34%. For COVID-19 testing, the maximal increase or decrease (8%) occur at pre-test probabilities of 28% or 36%, for positive vs negative test results, respectively.

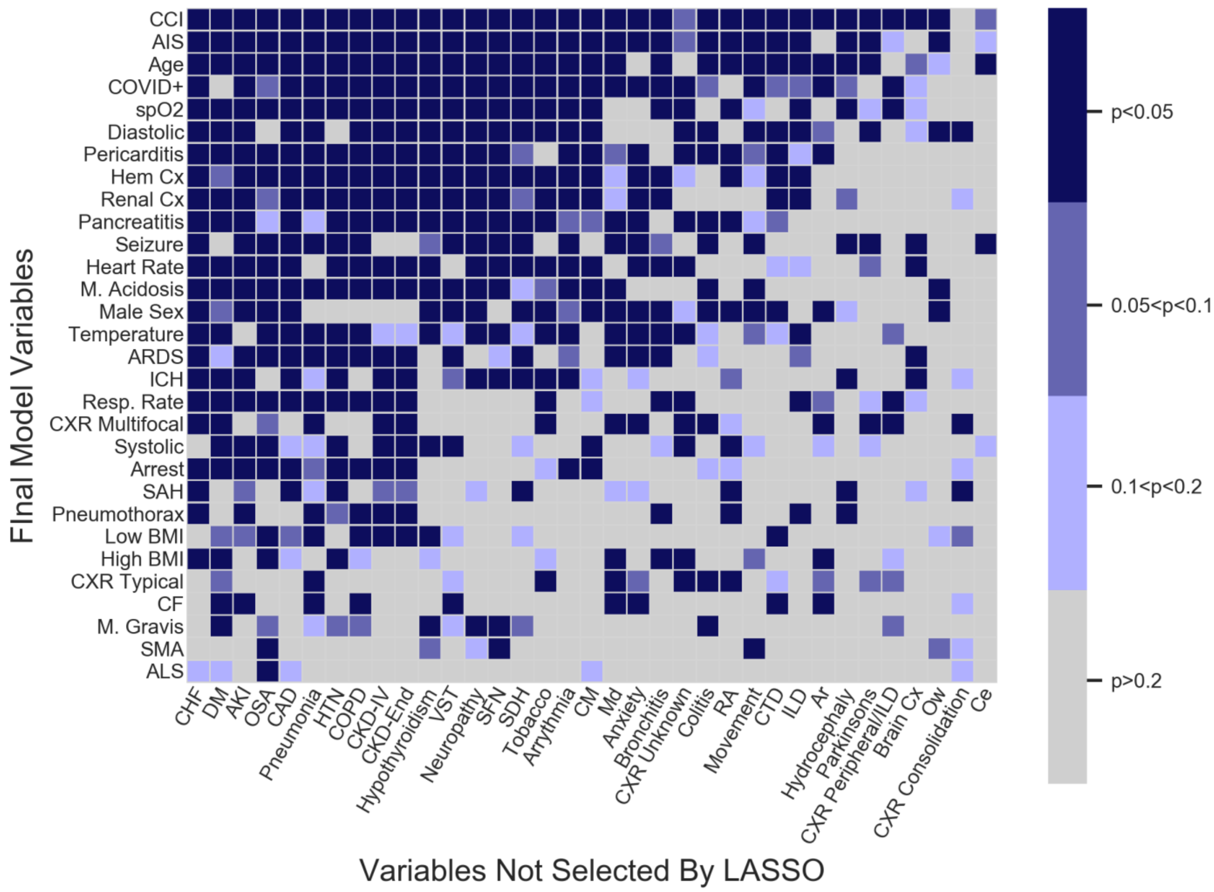

Figure S5. Heatmap of p-values representing the statistical significance of pairwise correlations between final model variables and variables not selected by LASSO. Rows represent final model variables, and columns represent variables that are univariately associated with outcomes, but were not selected by LASSO. For pairs of continuous variables we computed the p-value of the Spearman's rank correlation, for pairs of binary variables the Phi coefficient, and for continuous-binary pairs the biserial correlation. p-values are binned into $p < 0.05$, $0.05 < p < 0.01$, $0.01 < p < 0.02$, and $p > 0.02$. Shade indicates the magnitude of the p-value bin, where the darkest shade of blue indicates $p < 0.05$, and progressively lighter shades indicate bins of progressively higher p-values.

Abbreviations: CXR = chest X-ray; AIS = Acute Ischemic Stroke; AKI = Acute Kidney Injury; ALS = Amyotrophic Lateral Sclerosis; AR = Allergic Rhinitis; ARDS = Acute Respiratory Distress Syndrome; Arrest = Cardiac Arrest; Brain Cx = Brain Malignancy; CAD = Coronary Artery Disease; CE = Cerebral Edema; CCI = Charlson Comorbidity Index; CF = Cystic Fibrosis; CHF = Congestive Heart Failure; CKD-End = End Stage Renal Disease; CKD-IV = Stage Four Chronic Kidney Disease; CM = Cardiomyopathy; COPD = Chronic Obstructive Pulmonary Disease; COVID+ = COVID+ test up to point of visit; CTD = Connective Tissue Disorder; CXR Typical = CXR: Typical for COVID; Diastolic = Diastolic blood pressure; DM = Diabetes Mellitus; Hem Cx = Hematologic malignancy; HTN = Hypertension; ILD = Interstitial Lung Disease; M. Gravis = Myasthenia Gravis; MD = Major Depressive Disorder; OSA = Obstructive Sleep Apnea; OW = Overweight; RA = Rheumatoid Arthritis; Renal Cx = Renal cancer; Resp. Rate = Respiratory Rate; SAH = Subarachnoid hemorrhage; SDH = Subdural

Hematoma; Seizure = Seizure disorder; SFN = Small Fiber Neuropathy; SMA = Spinal Muscular Atrophy; Systolic = Systolic blood pressure. Tobacco = Tobacco User; VST = Venous Sinus Thrombosis. All diagnoses are pre-existing conditions based on past medical history, and are coded as present (e.g. Pneumothorax = 1) if recorded in the electronic medical record at any time before the date of presentation for COVID-19 screening.