Title

Development and validation of risk prediction models for COVID-19 positivity in a hospital setting

Authors

Ng, Ming-Yen
Wan, Eric Yuk Fai
Wong, Ho Yuen Frank
et al.

Publication Date

2020-12-01

DOI

10.1016/j.ijid.2020.09.022

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Infectious Diseases 101 (2020) 74–82

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Development and validation of risk prediction models for COVID-19 positivity in a hospital setting

Ming-Yen Ng[a,b,*,1], Eric Yuk Fai Wan[c,1], Ho Yuen Frank Wong[d], Siu Ting Leung[e], Jonan Chun Yin Lee[f], Thomas Wing-Yan Chin[f], Christine Shing Yen Lo[d], Macy Mei-Sze Lui[g], Edward Hung Tat Chan[b], Ambrose Ho-Tung Fong[a], Sau Yung Fung[a], On Hang Ching[a], Keith Wan-Hang Chiu[a], Tom Wai Hin Chung[h], Varut Vardhanbhuti[a], Hiu Yin Sonia Lam[d], Kelvin Kai Wang To[h], Jeffrey Long Fung Chiu[f], Tina Poy Wing Lam[d], Pek Lan Khong[a], Raymond Wai To Liu[j], Johnny Wai Man Chan[k], Alan Ka Lun Wu[l], Kwok-Cheung Lung[m], Ivan Fan Ngai Hung[g,i], Chak Sing Lau[i], Michael D. Kuo[n], Mary Sau-Man Ip[i,o]

[a] *Department of Diagnostic Radiology, The University of Hong Kong, Hong Kong Special Administrative Region*
[b] *Department of Medical Imaging, The University of Hong Kong-Shenzhen Hospital, Shenzhen, Hong Kong Special Administrative Region*
[c] *Department of Family Medicine and Primary Care, The University of Hong Kong, Hong Kong Special Administrative Region*
[d] *Department of Radiology, Queen Mary Hospital, Hong Kong Special Administrative Region*
[e] *Department of Radiology, Pamela Youde Nethersole Eastern Hospital, Hong Kong Special Administrative Region*
[f] *Department of Radiology and Imaging, Queen Elizabeth Hospital, Hong Kong Special Administrative Region*
[g] *Department of Medicine, Queen Mary Hospital, Hong Kong Special Administrative Region*
[h] *Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region*
[i] *Department of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region*
[j] *Department of Medicine, Ruttonjee Hospital, Hong Kong Special Administrative Region*
[k] *Department of Medicine, Queen Elizabeth Hospital, Hong Kong Special Administrative Region*
[l] *Department of Clinical Pathology, Pamela Youde Nethersole Eastern Hospital, Hong Kong Special Administrative Region*
[m] *Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong Special Administrative Region*
[n] *Medical Artificial Intelligence Laboratory (MAIL) Program, Department of Diagnostic Radiology, The University of Hong Kong, Hong Kong Special Administrative Region*
[o] *Division of Respiratory & Critical Care Medicine, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China*

ARTICLE INFO

Article history:
Received 7 June 2020
Received in revised form 8 September 2020
Accepted 10 September 2020

Keywords:
COVID-19
Prediction model
Nomogram
White cell count
Chest x-ray

ABSTRACT

Objectives: To develop: (1) two validated risk prediction models for coronavirus disease-2019 (COVID-19) positivity using readily available parameters in a general hospital setting; (2) nomograms and probabilities to allow clinical utilisation.

Methods: Patients with and without COVID-19 were included from 4 Hong Kong hospitals. The database was randomly split into 2:1 for model development database (n = 895) and validation database (n = 435). Multivariable logistic regression was utilised for model creation and validated with the Hosmer–Lemeshow (H–L) test and calibration plot. Nomograms and probabilities set at 0.1, 0.2, 0.4 and 0.6 were calculated to determine sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

Results: A total of 1330 patients (mean age 58.2 ± 24.5 years; 50.7% males; 296 COVID-19 positive) were recruited. The first prediction model developed had age, total white blood cell count, chest x-ray appearances and contact history as significant predictors (AUC = 0.911 [CI = 0.880−0.941]). The second model developed has the same variables except contact history (AUC = 0.880 [CI = 0.844−0.916]). Both were externally validated on the H–L test (p = 0.781 and 0.155, respectively) and calibration plot. Models

Abbreviations: AUC, Area under the curve; COVID-19, Coronavirus Disease 2019; CT, Computed tomography; CXR, Chest x-rays; GGO, Ground glass opacity; H–L, Hosmer–Lemeshow test; NPV, Negative predictive value; OR, Odds ratio; PEff, Pleural effusion; PPV, Positive predictive value; RT-PCR, Reverse transcription polymerase chain reaction; SARS-CoV-2, Severe acute respiratory syndrome coronavirus 2; WCC, Total white blood cell count.

* Corresponding author at: Department of Diagnostic Radiology, University of Hong Kong, Room 406, Block K, Hong Kong Special Administrative Region.
E-mail address: myng2@hku.hk (M.-Y. Ng).

[1] Both authors contributed equally to the manuscript and are joint first authors.

https://doi.org/10.1016/j.ijid.2020.09.022

were converted to nomograms. Lower probabilities give higher sensitivity and NPV; higher probabilities give higher specificity and PPV.
Conclusion: Two simple-to-use validated nomograms were developed with excellent AUCs based on readily available parameters and can be considered for clinical utilisation.

Introduction

Coronavirus disease-2019 (COVID-19) has spread rapidly worldwide and as of 6th September 2020, there are now ~27 million cases worldwide and ~900,000 deaths (World Health Organization, 2020a). Respiratory and non-respiratory complications of COVID-19 are also becoming increasingly apparent (Chung et al., 2020; Ng et al., 2020a). Reverse transcription polymerase chain reaction (RT-PCR) is regarded as a vital tool in identifying the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and quarantining COVID-19 patients to prevent further spread of the disease (World Health Organization, 2020b). Furthermore, it is the definitive test in confirming the diagnosis of COVID-19. However, the availability of RT-PCR kits maybe difficult in various countries (World Health Organization, 2020b) and from specimen collection to report generation, the tests could take 48–72 h to confirm a positive or negative result (NHS England and NHS Improvement, 2020). Therefore, clinical assessment, blood tests and imaging have been recommended to help identify potential COVID-19 positive patients (Rubin et al., 2020).

Various strategies have been proposed, including widespread computed tomography (CT) scanning (Zu et al., 2020; Zhang et al., 2020; Lee et al., 2020), the greater use of chest x-rays (CXR) (American College of Radiology, 2020; British Society of Thoracic Imaging, 2020), the identification of low lymphocyte counts (Zhang et al., 2020; British Society of Thoracic Imaging, 2020) to determine patients more likely to have COVID-19 (Guan et al., 2020; Huang et al., 2020) and thus more suitable for testing. As yet, the data that support these strategies are predominantly based on the data of COVID-19 patients (Guan et al., 2020; Huang et al., 2020) but without comparisons to patients with other conditions and symptoms overlapping with COVID-19 (e.g. fever, shortness of breath and cough).

Several issues have arisen in trying to determine the likelihood of a COVID-19 diagnosis. First, in the early stages of the pandemic when the disease was limited to a few countries, travel and contact history may have been helpful to increase the suspicion of a COVID-19 diagnosis, but in some countries where there is established community transmission, this has resulted in patients being COVID-19 positive but with no knowledge of possible contact. Secondly, different countries have adopted different strategies due to socioeconomic factors and healthcare resources. Thus, a COVID-19 prediction model based on clinical, laboratory and radiological findings, which presents the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) would allow public healthcare systems to decide a suitable strategy on prioritising tests, when such RT-PCR availability is constrained.

In this study, we aimed to construct a prediction model utilising patient characteristics, commonly available haematological and biochemical blood tests and CXR findings, which can identify COVID-19 patients within a cohort of patients who presented to hospitals for various disease conditions and underwent testing for COVID-19. In addition, we aimed to create a separate model, in the event that contact history is not available to determine the presence of COVID-19.

Methods

Research ethics approval was obtained from the Hong Kong West Cluster (No. UW20-115), Hong Kong East Cluster (No. HKECREC-2020-012) and Kowloon Central Cluster (No. KC/KE-20-0052/ER-3) Institutional Review Boards for this retrospective study. Patients were recruited from 4 acute general hospitals under the publicly funded Hospital Authority: Queen Mary Hospital, Pamela Youde Nethersole Eastern Hospital, Ruttonjee Hospital and Queen Elizabeth Hospital. These four hospitals cover three out of seven areas that comprise the whole territory of the Hong Kong special administrative region. All four hospitals provided care for COVID-19 cases.

Patient cohort for model development and validation

Inclusion criteria: All patients diagnosed with COVID-19 from any of these 4 hospitals between 1 January 2020 and 1 April 2020. COVID-19 diagnosis was established through reverse transcription polymerase chain reaction tests (RT-PCR) of nasopharyngeal and throat swabs. Sixty-four of these patients have been previously reported (Wong et al., 2020). As a comparison group, patients from Queen Mary Hospital and Pamela Youde Nethersole Eastern Hospital who underwent RT-PCR tests for COVID-19 and had subsequent negative results were recruited from 1 January 2020 to 26 February 2020 and from 24 March 2020 to 1 April 2020, respectively. Exclusion criteria were the lack of documentation for contact history (as defined by the World Health Organization (2020c)), symptoms and missing data on total white blood cell count (WCC), neutrophil count, lymphocyte count, liver function tests, renal function tests or CXR examinations. See Figure 1 for CONSORT diagram. The data completion rate of the data in the cohort is shown in the Supplementary Table 1.

RT-PCR testing

Patients were tested in the accident and emergency departments or inpatient wards. COVID-19 patients were confirmed using standard Hong Kong Hospital Authority and Department of Health RT-PCR protocol for the diagnosis of SARS-CoV2, which required two swabs: a throat swab and a nasopharyngeal swab. Indications for testing were clinical suspicion with at least one or more of the following criteria: (i) contact history, (ii) travel from countries with a high prevalence of COVID-19 and (iii) fever, cough, shortness of breath and/or respiratory/gastrointestinal symptoms consistent with COVID-19. Repeat RT-PCR was performed in negative patients with a strong suspicion of COVID-19 or persistent symptoms with no alternative diagnosis. Of the 1034 negative patients, 165 (16.0%) had two or more negative RT-PCR results. Sixteen out of 296 COVID-19 patients (5.4%) had negative RT-PCR results on first testing, but had subsequent positive RT-PCR results. Previous publications have shown that RT-PCR from our units have high sensitivity of >90% (Wong et al., 2020) and low false-negative rates (Chan et al., 2020) based on laboratory-confirmed COVID-19. This cohort did not have any false-positive cases.

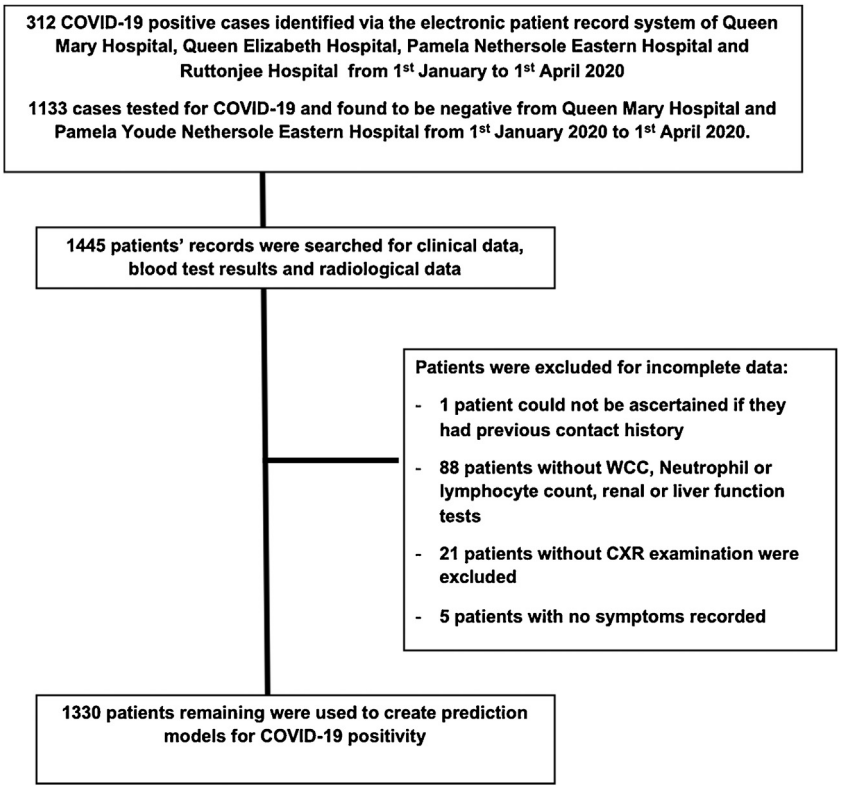

Figure 1. Consort diagram showing the process of identification of cases and exclusion of cases.

Radiological investigation

CXR images were searched through the electronic patient record system. Baseline CXR images were reviewed and interpreted by radiologists blinded to the patient's COVID-19 status. The assessment was based on identifying the common findings of COVID-19 on CXR which were (i) consolidation or ground glass opacity (GGO) and (ii) the absence of pleural effusion (Wong et al., 2020; Ng et al., 2020b). This was done in a binary format (present or absent) to make this more reproducible in the clinical environment for front-line clinicians.

Image quality was assessed in 702 randomly chosen CXRs (53% of entire cohort of CXRs) by 8 radiologists separately. We ensured that the CXRs were taken from each of the 4 hospitals. Image quality was assessed on a scale from 1 to 3. See supplementary Table 2 for examples of CXRs graded as 1, 2 and 3. Briefly, CXRs which could not be interpreted with any confidence were graded 1. CXRs with suboptimal image quality but lung changes and pleural effusion (PEff) could be interpreted with some confidence were graded 2. CXRs with good quality such that lung changes and pleural effusions can be diagnosed with high confidence were graded 3.

Data analysis

Patients positive for COVID-19 were compared to those negative for COVID-19 patients. Continuous variables were compared by using student's t-tests. Categorical variables were compared by using chi-squared tests. The database was randomly split on a 2:1 basis for the model development and validation dataset. The model development database had 893 cases (212 cases were COVID-19 positive) and the validation database had 437 cases (84 cases were COVID-19 positive). The model development database was used to develop the COVID-19 risk prediction model, and the validation cohort was used to validate the derived COVID-19 probability.

Development of COVID-19 prediction models

Prediction model development was based on multivariable logistic regression from the model development cohort. The potential predictors were considered based on their common acquisition in patients presenting to general hospitals. The variables included in the models were sex, age, symptoms (i.e. fever, cough, shortness of breath, vomiting and diarrhoea), blood test results (i.e. total white cell count [WCC], neutrophil count, lymphocyte count, neutrophil: lymphocyte cell ratio, albumin, bilirubin, alanine aminotransferase and estimated glomerular filtration rate) and CXR findings (i.e. the presence of consolidation/ground glass opacity (GGO), absence of PEff and combination of consolidation/GGO with the absence of pleural effusion). The forward stepwise selection method was conducted to select the predictors. In each step of forward stepwise selection procedure, the model included the predictor with the lowest Bayesian Information Criteria (BIC) as compared to the previous mode. The selection was finished until the difference in BIC of all remaining risk factors <10 (Raftery, 1995). To test the non-linear effect of selected clinical parameters, the quadratic term of significant continuous predictors were considered. Given that patients can present without the knowledge of contact history with an infected person, a further model was developed with one having contact history removed, to represent an event in which contact history is unknown.

To validate the model, the discrimination and calibration power of models were examined. The area under the receiver operating characteristic curve (AUC) was conducted to evaluate the discrimination power, where from 0.7 to 0.8 of AUC is considered acceptable, from 0.8 to 0.9 is considered excellent and more than

Table 1
Characteristics and laboratory findings of patients with negative and positive COVID-19 RT-PCR results in model development and validation cohort.

Characteristics	Model Development Cohort			Validation Cohort		
	COVID-19 negative (n = 681)	COVID-19 positive (n = 212)	P value	COVID-19 negative (n = 353)	COVID-19 positive (n = 84)	P value
Age, years Mean (SD)	62.6 (24.7)	40.6 (17.6)	<0.001*	64.1 (22.7)	42.3 (17.5)	<0.001*
Sex			0.172			0.430
Female	361 (53.0%)	101 (47.6%)		168 (47.6%)	44 (52.4%)	
Male	320 (47.0%)	111 (52.4%)		185 (52.4%)	40 (47.6%)	
Contact History	30 (4.4%)	82 (38.7%)	<0.001*	15 (4.2%)	36 (42.9%)	<0.001*
Signs and symptoms at presentation						
Asymptomatic	25 (3.7%)	19 (9.0%)	0.002*	20 (5.7%)	10 (11.9%)	0.042*
Fever	310 (45.5%)	107 (50.5%)	0.207	156 (44.2%)	43 (51.2%)	0.247
Cough	304 (44.6%)	119 (56.1%)	0.003*	151 (42.8%)	50 (59.5%)	0.006*
Shortness of breath	216 (31.7%)	22 (10.4%)	<0.001*	116 (32.9%)	11 (13.1%)	<0.001*
Chest Pain	38 (5.6%)	15 (7.1%)	0.421	20 (5.7%)	3 (3.6%)	0.440
Diarrhoea	31 (4.6%)	24 (11.3%)	<0.001*	22 (6.2%)	8 (9.5%)	0.284
Vomiting	44 (6.5%)	3 (1.4%)	0.004*	23 (6.5%)	2 (2.4%)	0.143
Past Medical History						
Smoking	81 (11.9%)	8 (3.8%)	<0.001*	34 (9.6%)	1 (1.2%)	0.010*
Diabetes	162 (23.8%)	8 (3.8%)	<0.001*	84 (23.8%)	8 (9.5%)	0.004*
Hypertension	269 (39.5%)	21 (9.9%)	<0.001*	155 (43.9%)	16 (19.0%)	<0.001*
COPD	28 (4.1%)	0 (0.0%)	0.003*	16 (4.5%)	0 (0.0%)	0.047*
Chronic liver disease	56 (8.2%)	7 (3.3%)	0.015*	29 (8.2%)	6 (7.1%)	0.745
Myocardial infarction	46 (6.8%)	3 (1.4%)	0.003*	37 (10.5%)	0 (0.0%)	0.002*
Active malignancy	82 (12.0%)	2 (0.9%)	<0.001*	43 (12.2%)	2 (2.4%)	0.008*
Laboratory findings						
White cell count (normal range: 4-10 $\times$ 10^9 cells/L)	9.9 (5.3)	5.5 (1.8)	<0.001*	10.4 (5.8)	5.9 (2.1)	<0.001*
Leucocytosis	266 (39.1%)	4 (1.9%)	<0.001*	147 (41.6%)	4 (4.8%)	<0.001*
Neutrophil count (normal range: 1.8-6.2 $\times$ 10^9 cells/L)	7.4 (4.5)	3.4 (1.5)	<0.001*	7.9 (5.4)	3.7 (1.9)	<0.001*
Neutrophilia	346 (50.8%)	12 (5.7%)	<0.001*	182 (51.6%)	7 (8.3%)	<0.001*
Lymphocyte count (normal range: 1.0-3.2 $\times$ 10^9 cells/L)	1.4 (0.9)	1.8 (3.4)	0.017*	1.6 (3.3)	1.5 (0.6)	0.874
ymphopenia	268 (39.4%)	52 (24.5%)	<0.001*	150 (42.5%)	20 (23.8%)	0.002*
Neutrophil: Lymphocyte ratio	8.9 (12.7)	2.9 (2.6)	<0.001*	11.3 (26.9)	3.0 (2.4)	0.005*
Platelet count (x10^9 cells/Litre)(normal range: 162–341)	250.0 (104.5)	226.0 (65.7)	0.002*	247.9 (104.6)	229.3 (75.2)	0.125
Albumin (normal range: 35·0–52·0)	37.9 (7.3)	41.8 (4.3)	<0.001*	37.7 (7.6)	41.6 (4.3)	<0.001*
Total Bilirubin (μmol/L) (normal range: 0·0–21·0)	12.0 (14.9)	7.4 (3.8)	<0.001*	12.9 (15.2)	8.5 (6.2)	0.010*
Alanine Aminotransferase (U/L) (normal range 0·0–33·0)	37.0 (126.3)	27.4 (18.8)	0.270	34.6 (67.6)	30.1 (20.5)	0.546
eGFR (ml/min/1.73m^2)	75.1 (41.8)	96.4 (26.5)	<0.001*	73.5 (61.4)	94.5 (24.7)	0.002*
Radiological Findings						
CXR appearances, lung changes consistent with consolidation/GGO	274 (40.2%)	92 (43.4%)	0.414	147 (41.6%)	35 (41.7%)	0.997
CXR Pleural effusion absent	582 (85.5%)	212 (100.0%)	<0.001*	295 (83.6%)	83 (98.8%)	<0.001*
CXR Consolidation/GGO and Pleural effusion absent	200 (29.4%)	92 (43.4%)	<0.001*	101 (28.6%)	34 (40.5%)	0.034*
Diagnosis for Presenting Illness						
URTI	106 (15.6%)	NA	NA	52 (14.7%)	NA	NA
Influenza	10 (1.5%)	NA	NA	10 (2.8%)	NA	NA
Non-Influenza Pneumonia	170 (25.0%)	NA	NA	75 (21.2%)	NA	NA

All values are presented in mean (standard deviation), median (interquartile range), or in number (percentage), as appropriate.
COPD = Chronic obstructive pulmonary disease; eGFR = estimated glomerular filtration rate; CXR = Chest X-ray; GGO = ground glass opacity; RT-PCR = Reverse transcription polymerase chain reaction and URTI = Upper respiratory tract infection.
* p-value < 0.05 for comparison between COVID-19 negative and positive using independent t-test and chi-squared test as appropriate.

0.9 is considered as outstanding discrimination power. Meanwhile, Hosmer–Lemeshow (H–L) test and calibration plot was used to test how well the percentage of observed COVID-19 positive matches the percentage of predicted COVID-19 positive over the deciles of predicted risk. A p-value >0.05 is needed to conclude that there are insignificant differences between the observed and expected outcomes, and therefore, the model has good overall calibration. Different probabilities were used to evaluate the model performance based on the sensitivity, specificity, PPV and NPV.

Sensitivity analysis was conducted to examine the robustness of the model. Multiple imputation was applied to handle missing data. The chained equation method was used to impute each missing value twenty times, adjusted for all baseline covariates and outcomes. Moreover, 10-fold cross validation was applied to evaluate the discrimination and calibration power.

Nomogram development and probability of COVID-19

To facilitate the risk prediction models used for screening in routine busy clinical practice, simple nomograms were developed. The effect of each predictor in the model was converted to a score and summation of all predictors that can be mapped to an estimated risk of COVID-19 positive. The nomograms were plotted by using the nomolog package in Stata (Zlotnik Enaliev, 2016). Sensitivity, specificity, PPV and NPV were determined for the following probabilities which were: 0.1, 0.2, 0.4 and 0.6 as well as

Table 2
Prediction models for the risk of COVID-19 by multivariable logistic regression. The model development database (n = 895) was used to create the models with area under the curve (AUC) values derived from testing the models against the validation data set (n = 435).

Characteristics	Overall Cohort Model			Unknown Contact History Model		
	Odds Ratio	CI	p-value	Odds Ratio	CI	p-value
Age	1.04	(0.99, 1.09)	0.102	1.06	(1.01, 1.11)	0.012*
Age^2	0.9991	(0.9986, 0.9996)	<0.001*	0.9989	(0.9984, 0.9994)	<0.001*
Contact History	10.0	(5.5, 18.2)	<0.001*			
Total WCC (x 10^9 cells/L)	0.58	(0.52, 0.65)	<0.001*	0.59	(0.54, 0.66)	<0.001*
CXR C/GGO & PEff Absent	5.2	(3.1, 8.7)	<0.001*	4.9	(3.0, 7.8)	<0.001*
Validation indicators						
AUC	0.911 (CI 0.880–0.941)			0.880 (CI 0.844–0.916)		
Hosmer-Lemeshow test p-value	0.781			0.155		

Note: the validation indicators were obtained based on the validation data set (n = 435).
CI = Confidence Interval; SOB = shortness of breath; WCC = white cell count; eGFR = estimated glomerular filtration rate; CXR = Chest x-ray; GGO = ground glass opacity; PEff = Pleural effusion and AUC = Area under the curve.

the optimal cut-off for each model based on the Youden method (Youden, 1950).

Statistical analysis and model building were performed by using the STATA version 14 (StataCorp LP, College Station, Texas, USA). All significance tests were two-tailed and those with a *p*-value <0.05 were considered statistically significant.

Results

A total of 1330 patients (mean age 58.2 ± 24.5 years [range: 0–106 years old]; 674 [50.7%] males) were included in the analysis. Thirty-five patients were <18 years old of whom 11 were COVID-19 positive. In all, 296 patients were COVID-19 positive. Supplementary Table 3 illustrates the patient characteristics and compares the COVID-19 positive and negative patients in the overall cohort. Of the patients positive for COVID-19, none had any synchronous infection. Patients (40.9%) who tested negative for COVID-19 had upper respiratory tract infection, influenza or non-influenza pneumonia. The characteristics in model development and validation cohort are shown in Table 1.

Of the 702 CXR assessed for image quality, the number of CXRs graded with the following image quality scores of 3, 2 and 1 were 609 (86.8%), 89 (12.7%) and 4 (0.6%), respectively.

Prediction model results for the model development cohort (n = 895) with and without contact history are shown in Table 2.

The prediction model for the overall cohort showed that the age (OR 1.04, 95% confidence interval [CI] (0.99–1.09)) and quadratic term of age (OR 0.9991 95% CI (0.9986, 0.9996)), contact history (OR 10.0, 95% CI: 5.5–18.2), total WCC (OR 0.58, 95% CI: 0.52−0.65) and CXR consolidation/GGO with absent PEff (OR 5.2, 95% CI: 3.1–8.7) were independent significant predictors of COVID-19 positivity. Other variables such as fever, cough, diarrhoea, smoking, past medical history, neutrophil count, lymphocyte count, neutrophil, lymphocyte ratio or the quadratic term of WCC were not independently significant variables. When the model was tested against the validation dataset, the AUC was excellent at 0.911 (CI:

Table 3
Test performance of different probability to identify COVID-19 for the Overall Cohort Model and Unknown Contact History Model in the nomogram. Probability of 0.1, 0.2, 0.4 and 0.6 are stated and optimum probability cut-off for the two models. As an example of this table being utilised, if we use the overall cohort model and set 0.4 as the chosen probability cut-off to determine RT-PCR testing, patients scoring ≥0.4 on the nomogram would be tested and those <0.4 would not be tested. With this cut-off, the model indicates that on the validation cohort the following results: sensitivity 66.7%, specificity 90.9%, positive predictive value 63.6% and negative predictive value 92.0%.

Probability	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
Overall Cohort Model				
Derivation cohort				
0.1	97.6 (94.6,99.2)	68.3 (64.6,71.8)	48.9 (44.1,53.8)	98.9 (97.5,99.7)
0.18†	91.5 (86.9,94.9)	78.0 (74.7,81.0)	56.4 (51.0,61.7)	96.7 (94.9,98.0)
0.2	90.1 (85.3,93.8)	79.9 (76.7,82.8)	58.2 (52.7,63.6)	96.3 (94.4,97.7)
0.4	73.1 (66.6,79.0)	89.3 (86.7,91.5)	68.0 (61.5,74.0)	91.4 (89.0,93.4)
0.6	56.1 (49.2,62.9)	95.3 (93.4,96.8)	78.8 (71.4,85.0)	87.5 (84.9,89.8)
Validation cohort				
0.1	91.7 (83.6,96.6)	69.4 (64.3,74.2)	41.6 (34.4,49.1)	97.2 (94.4,98.9)
0.18†	88.1 (79.2,94.1)	77.9 (73.2,82.1)	48.7 (40.5,56.9)	96.5 (93.6,98.3)
0.2	86.9 (77.8,93.3)	79.9 (75.3,83.9)	50.7 (42.2,59.1)	96.2 (93.4,98.1)
0.4	66.7 (55.5,76.6)	90.9 (87.4,93.7)	63.6 (52.7,73.6)	92.0 (88.6,94.6)
0.6	47.6 (36.6,58.8)	95.5 (92.7,97.4)	71.4 (57.8,82.7)	88.5 (84.8,91.5)
Unknown Contact History Model				
Derivation cohort				
0.1	97.2 (93.9,99.0)	63.4 (59.7,67.1)	45.3 (40.6,50.0)	98.6 (97.0,99.5)
0.18†	92.9 (88.6,96.0)	74.0 (70.5,77.3)	52.7 (47.5,57.8)	97.1 (95.3,98.4)
0.2	91.0 (86.4,94.5)	75.9 (72.5,79.1)	54.1 (48.7,59.3)	96.5 (94.5,97.9)
0.4	72.6 (66.1,78.5)	87.2 (84.5,89.6)	63.9 (57.5,70.0)	91.1 (88.7,93.2)
0.6	48.6 (41.7,55.5)	94.9 (92.9,96.4)	74.6 (66.5,81.7)	85.6 (82.9,88.0)
Validation cohort				
0.1	90.5 (82.1,95.8)	65.2 (59.9,70.1)	38.2 (31.4,45.3)	96.6 (93.5,98.5)
0.18†	84.5 (75.0,91.5)	73.4 (68.4,77.9)	43.0 (35.4,51.0)	95.2 (92.0,97.4)
0.2	82.1 (72.3,89.6)	74.5 (69.6,79.0)	43.4 (35.6,51.5)	94.6 (91.3,96.9)
0.4	66.7 (55.5,76.6)	88.4 (84.6,91.5)	57.7 (47.3,67.7)	91.8 (88.3,94.5)
0.6	41.7 (31.0,52.9)	96.6 (94.1,98.2)	74.5 (59.7,86.1)	87.4 (83.7,90.6)

Sensitivity, specificity, positive predictive value and negative predictive value are stated in percentage with 95% confidence intervals in brackets.
† This probability is the optimal cut-off determined by using the Youden method from the derivation cohort.

0.880–0.941). The Hosmer–Lemeshow (H–L) test p-value of 0.781 and the calibration plot (Figure 2) showed similar results between predicted and observed outcome, which indicated superb calibration power of the model.

The unknown contact history prediction model dropped contact history from the model leaving age, total WCC and CXR consolidation/GGO with absent pleural effusions in the model with minor adjustments in the odds ratios. Testing against the validation dataset revealed a lower AUC of 0.880 (CI: 0.844–0.916) as compared to the overall cohort model but can still be regarded as excellent. The H–L test p-value was 0.155 and a good fit on the calibration plot (Figure 2).

The result of sensitivity analysis using multiple imputation and 10-fold cross validation showed that the same predictors were included in the models, and similar validation indicators and calibration plot were obtained.

The sensitivity, specificity, PPV, NPV for probabilities of 0.1, 0.2, 0.4 and 0.6 for both models can be seen in Table 3. The optimal probability cut-off for the overall cohort model and unknown contact history model was 0.18. The lower probabilities (e.g. 0.1) give a higher sensitivity and NPV, whilst the higher probabilities (e.g. 0.4 and 0.6) give higher specificity and PPV in both the overall cohort model and the unknown contact history model.

The nomograms in Figures 3 and 4 were developed based on the derived risk prediction models. Using the overall cohort model nomogram (Figure 3) as an example, if a patient suspected to have COVID-19 is aged 50 years, has no contact history, WCC of 2×10^9 cells/L and a CXR with no consolidation/GGO and absent PEff, the scoring will be as follows: Age has two steps, so for age 50 at step 1, allocate 2 points; for step 2: allocate 7.5 points. For no contact history, which is step 3, allocate 0 points. For a total white cell count (WCC) of 2×10^9 cells/L at step 4, allocate 7 points. For a CXR with no consolidation/GGO and absent PEff at step 5, allocate 0 points. Therefore, they would be allocated a total score of 16.5 points, which equates to 0.6–0.7 probability (i.e. 60%–70% probability) of being COVID-19 positive.

Discussion

In our study, we have developed two risk prediction models for determining COVID-19 positive patients, which have been validated with a separate dataset. Both models have an excellent AUC with good matching with the validation dataset. The models are based on parameters (i.e. total WCC and CXR consolidation/GGO with absent pleural effusions) that are available in general hospitals and clinical data (i.e. age with or without contact history). We have also provided nomograms to determine the probability of COVID-19, with several different probabilities illustrated to show the sensitivity, specificity, PPV and NPV so that clinicians or healthcare systems can decide which probabilities would make the best cut-offs for RT-PCR testing. The development of these nomograms will hopefully improve front-line clinicians' diagnostic accuracy to help identify patients with COVID-19 where RT-PCR may not be available or rapid results cannot be provided.

Currently, recommendations have been made on determining the likelihood of patients being COVID-19 positive. However, these are based on studies that have solely focused on COVID-19 patients without a comparison group. As such lymphopaenia has been proposed as a useful marker of COVID-19 positivity (British Society of Thoracic Imaging, 2020). However, if one compares the lymphocyte counts between the positive and negative cohorts, there is no significant difference with the lymphocyte counts of both groups. The mean lymphocyte count of both cohorts are in the lower end of normal. In contrast, WCC were significantly higher in the negative group than the positive COVID-19 patients and on average above the normal range in the negative patients. WCC is likely a better predictor as compared to lymphocyte count because it incorporates the lower neutrophil count and similar lymphocyte counts between COVID positive patients as compared to negative patients. Indeed, a normal or low WCC was stated as a feature of COVID-19 in the early stages of the epidemic from China's National Health Commission. Thus, our data provide evidence that these initial observations of COVID-19 were indeed accurate.

CXR consolidation/GGO with absent pleural effusions is the typical appearance of COVID-19 radiologically (Wong et al., 2020). This model confirms that using CXR in addition to other parameters is helpful to identify COVID-19 patients. This has already been incorporated into societal recommendations, and our models provide evidence to support this approach despite the lower sensitivity of CXR as compared to CT (American College of Radiology, 2020; British Society of Thoracic Imaging, 2020). Our model did not incorporate CT as CT was not easily available for our COVID-19 positive patients and indeed the negative patients. This would likely be the scenario globally during this pandemic. CT with its higher sensitivity (Fang et al., 2020) will likely improve the diagnostic accuracy, but this is dependent on the facilities in each health service. Not all health services can dedicate CT scanners for COVID-19 diagnosis due to either a lack of scanner availability and/or the extensive cleaning required after each COVID-19 scan, which reduces the radiology department's productivity (Hope et al., 2020). In our study, we wanted to focus on parameters that would be easily accessible to all patients seen in the general hospitals, as some health systems even struggle to make CXR and WCC available (Ayebare et al., 2020).

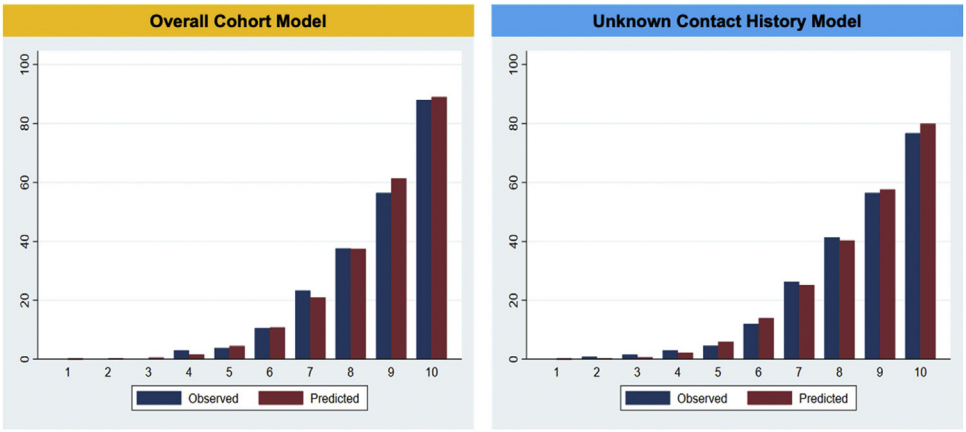

Figure 2. Calibration plot for the Overall Cohort Model and Unknown Contact History Model for the risk of COVID-19.

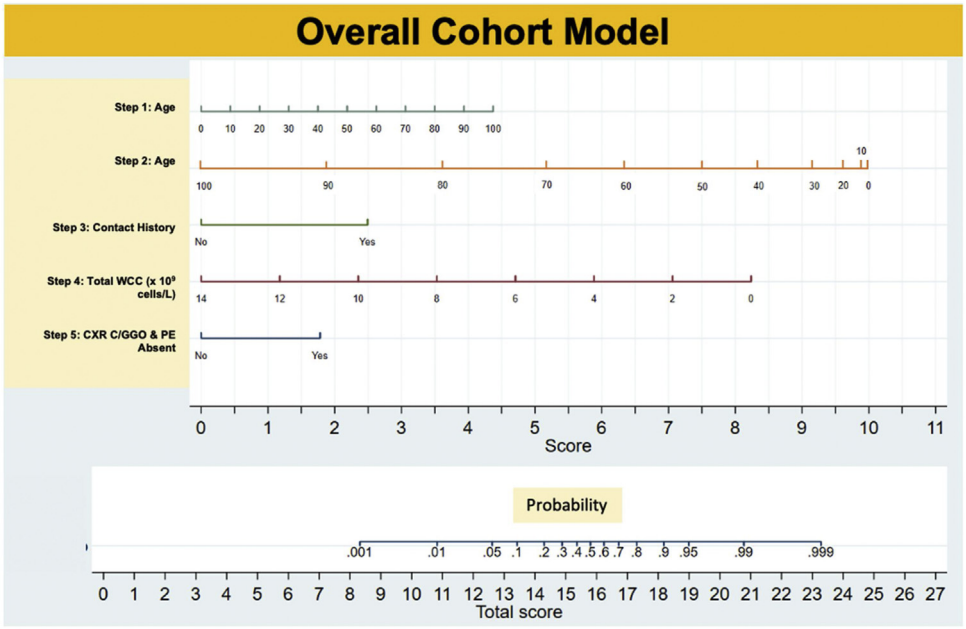

Figure 3. Nomogram of the Overall Cohort in Table 2. A total score is calculated from the addition of the scores for the variables chest x-ray (CXR) consolidation/ground glass opacity (GGO), contact history, white cell count and age. Note that age has two steps, whilst other variables have only 1 step. The total score can then be marked on the bottom row and compared with the probability scale above. For example, a patient suspected to have COVID-19 aged 50 (step 1: allocate 2 points and step 2: allocate 7.5 points), has no contact history (step 3: allocate 0 points), total white cell count (WCC) of 2×10^9 cells/L (step 4: allocate 7 points) and a CXR with no consolidation/GGO and absent pleural effusion (PE) (step 5: allocated 0 points), would receive a total score of 16.5 points, which equates to a probability of between 0.6 and 0.7. A clinician then refers to the probability table (Table 3) and decides what degree of sensitivity, specificity, positive predictive value or negative predictive value is adequate for their setting.

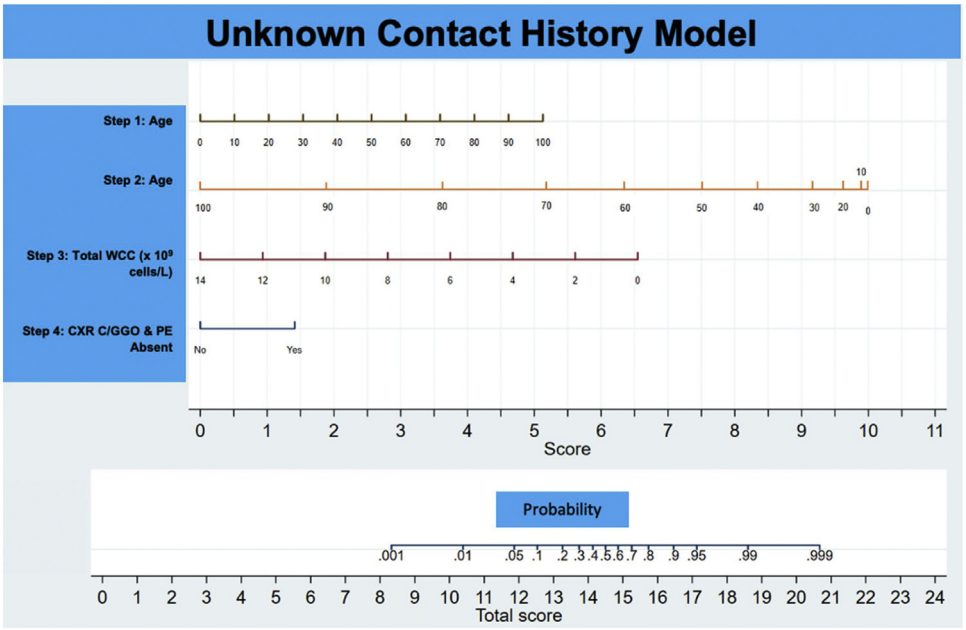

Figure 4. Nomogram of the Overall Cohort in Table 4. A total score is calculated from the addition of the scores for the variables pleural effusion, chest x-ray (CXR) consolidation/ground glass opacity (GGO), white cell count, age and vomiting symptom. Note that age has two steps, whilst other variables have only 1 step. The total score can then be marked on the bottom row and compared with the probability scale above. For example, a patient suspected to have COVID-19 aged 40 (step 1: allocate 2 points and step 2: allocate 8.5 points), total white cell count (WCC) of 8×10^9 cells/L (step 3: allocate 3 points) and a CXR with consolidation/GGO and absent pleural effusion (PE) (step 4: allocated 1.5 points), would receive a total score of 15 points which equates to a probability of between 0.6 and 0.7. A clinician then refers to the probability table (Table 3) and decides what degree of sensitivity, specificity, positive predictive value or negative predictive value is adequate for their setting.

In our cohort, age is a significant predictor for COVID-19. In this, study, the COVID-19 patients were significantly younger than the negative patients. This can be partly explained by younger patients being more mobile, and thus being more susceptible to develop COVID-19 as compared to the older population who may travel less. The review of previous publications have indicated that patients with COVID-19 are usually younger. In Korea, one paper indicated that >60% of patients (Shim et al., 2020) were <49 years old, whilst in China, 51.2%–55.1% (Guan et al., 2020; Huang et al., 2020) of patients were <49 years old. The two nomograms in this

study allocated higher scoring to the younger patients including children. This is possibly due to children having less symptoms and even less radiological changes (Qiu et al., 2020; Chen et al., 2020) making the identification of COVID-19 more difficult. Indeed, this possibly explains the noticeably less children confirmed to have COVID-19 and the statistical significance of age in the models for determining patients who are positive for SARS-CoV2. However, age as a predictor is extremely representative of this cohort. In a different healthcare system where more elderly patients present, age as a predictor will likely need to be further investigated.

The models we have established can set different probabilities to allow medical systems to self-determine the pre-test probability required for RT-PCR testing. Moreover, the nomograms have been developed to visualise the sophisticated mathematical equation, so that it can be adopted in the routine busy clinical practice. However, it should be emphasised that RT-PCR remains the gold standard for diagnosing COVID-19 and that focus should be made on making RT-PCR easily available for testing patients and increasing the time taken for results to be made available.

Our study has several limitations. First, the COVID-19 cases are reflective of practice in Hong Kong that has been active in screening for COVID-19, which has included asymptomatic patients (5.6% in this cohort) with contact history and patients with mild symptoms. This may not be representative in other health systems worldwide, so this model needs to be validated in those health systems. Secondly, the CXR were assessed by radiologists, so whether these results will be similar with front-line clinicians is uncertain. However, the assessment was simplified so that front-line clinicians can focus their search on CXR to the consolidation/GGO and absence of pleural effusions. Furthermore, some health systems have access to radiology support to review CXRs, and this model possibly justifies this practice if logistically feasible. Thirdly, inflammatory markers like C-reactive protein, creatinine kinase and lactate dehydrogenase were not included in the model as a significant proportion of patients did not have these markers measured at the time of admission. Whether these markers prove useful will require further study. Lastly, asymptomatic patients made up a very small proportion of patients and thus further validation with an asymptomatic cohort would be required to validate this model.

In conclusion, we present two models that have 3 or 4 readily available parameters to improve the accuracy of identifying COVID-19 amongst patients suspected of having COVID-19 with or without known contact history. This will help to identify patients who most likely may benefit from RT-PCR testing and thus help to better allocate RT-PCR testing where this resource is limited.

Conflict of interest

MYN has received funding from Bayer and Circle Cardiovascular Imaging. All other authors have no conflict of interest to declare.

Funding source

Sanming Project of Medicine in Shenzhen, China (SZSM201612096), and High Level-Hospital Program (Guangdong Health Commission), China.

Ethical approval

This study was approved by the Hong Kong West Cluster (No. UW20-115), Hong Kong East Cluster (No. HKECREC-2020-012) and Kowloon Central Cluster (No. KC/KE-20-0052/ER-3) Institutional Review Boards.

Author's contributions

M.Y.N and E.Y.F.W had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. M.Y.N, E.Y.F.W, M.D.K and M.S.M.I participated in study conception. H.Y.F.W, S.T.L, J.C.Y.L, T.W.Y.C, C.S.Y.L, M.M.S.L, E.H.T.C, A.H.T.F, F.S.Y, O.H.C, K.W.H.C, T.W.H.C, V.V, H.Y.S.L, K.K.W.T, L.F.J.C, T.P.W.L, P.L.K, R.W.T.L, J.W.M.C, K.L.A.W, K.C.L, F.N.I.H, C.S.L, M.D.K, and M.S.M.I recruited patients. M.Y.N, E.Y.F.W, and M.S.M.I performed data analysis. M.Y.N, E.Y.F.W, and M.S.M.I drafted and revised the paper. All authors approved the final draft of the manuscript for publication.

Acknowledgments

We dedicate this work to the Hong Kong doctors and healthcare professionals who have worked tirelessly and selflessly during this COVID-19 outbreak. In addition, we thank our families who have supported us emotionally and physically during these unprecedented times. We would also like to thank Dr Fredrik Resman from Lund University who helped review this article for us and was a source of excellent advice.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijid.2020.09.022.

References

World Health Organization. Coronavirus disease (COVID-19) Weekly Epidemiological Update. https://wwwwhoint/docs/default-source/coronaviruse/situation-reports/20200907-weekly-epi-update-4pdf?sfvrsn=f5f607ee_2. 6th Sept 2020.

Chung TW, Sridhar S, Zhang AJ, Chan KH, Li HL, Wong FK, et al. Olfactory dysfunction in coronavirus disease 2019 patients: observational cohort study and systematic review. Open Forum Infect Dis 2020;7: ofaa199.

Ng M, Ferreira VM, Leung S, Lee J, Fong A, Liu R, et al. Recovered COVID-19 patients show ongoing subclinical myocarditis as revealed by cardiac magnetic resonance imaging. JACC Cardiovasc Imaging 2020a; Published in advance on 28th August 2020.

World Health Organization. Laboratory testing strategy recommendations for COVID-19. 2020. https://appswhoint/iris/bitstream/handle/10665/331509/WHO-COVID-19-lab_testing-20201-engpdf.

NHS England and NHS Improvement. Guidance and standard operating procedure COVID-19 virus testing in NHS laboratories. 2020. . Version 1 https://wwwenglandnhsuk/coronavirus/wp-content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-laboratories-v1pdf.

Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. Radiology 2020;0:201365.

Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM, et al. Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology 2020;0:200490.

Zhang J, Zhou L, Yang Y, Peng W, Wang W, Chen X. Therapeutic and triage strategies for 2019 novel coronavirus disease in fever clinics. Lancet Respir Med 2020;8: e11–2.

Lee EYP, Ng M-Y, Khong P-L. COVID-19 pneumonia: what has CT taught us?. Lancet Infect Dis 2020;20:384–5.

American College of Radiology. ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. 2020. https://wwwacrorg/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection.

British Society of Thoracic Imaging. BSTI NHSE COVID-19 radiology decision support tool. 2020. https://wwwbstiorguk/media/resources/files/NHSE_BSTI_APPROVED_Radiology_on_CoVid19_v6_modified1__-_Read-Onlypdf.

Guan W-j, Ni Z-y, Hu Y, Liang W-h, Ou C-q, He J-x, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;.

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;.

Wong HYF, Lam HYS, Fong AH-T, Leung ST, Chin TW-Y, Lo CSY, et al. Frequency and distribution of chest radiographic finding in COVID-19 positive patients. Radiology 2020;0:201160.

World Health Organization. Global surveillance for COVID-19 caused by human infection with COVID-19 virus: Interim guidance 20th March 2020. 2020. file:///Users/mingyenng/Downloads/WHO-2019-nCoV-SurveillanceGuidance-2020.6-eng.pdf.
Chan JF-W, Yip CC-Y, To KK-W, Tang TH-C, Wong SC-Y, Leung K-H, et al. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-polymerase chain reaction assay validated *in vitro* and with clinical specimens. J Clin Microbiol 2020;00310–20.
Ng M-Y, Lee EYP, Yang M-Y, Yang M-Y, Li X, Wang M-Y, et al. Imaging profile of the COVID-19 infection: radiologic findings and literature review. Radiol Cardiothor Imaging 2020b;2:e200034.
Raftery AE. Bayesian model selection in social research. Sociol Methodol 1995;25:111–63.
Zlotnik Enaliev A. Design and evaluation of analytical tools for emergency department management based on machine learning techniques. 2016. http://oaupmes/43029/1/ALEXANDER_ZLOTNIK_ENALIEVpdf.
Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5.
Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology 2020;0:200432.
Hope MD, Raptis CA, Shah A, Hammer MM, Henry TS. A role for CT in COVID-19? What data really tell us so far. Lancet 2020;395:1189–90.
Ayebare RR, Flick R, Okware S, Bodo B, Lamorde M. Adoption of COVID-19 triage strategies for low-income settings. Lancet Respir Med 2020;8:e22.
Shim E, Tariq A, Choi W, Lee Y, Chowell G. Transmission potential and severity of COVID-19 in South Korea. Int J Infect Dis 2020;93:339–44.
Qiu H, Wu J, Hong L, Luo Y, Song Q, Chen D. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. Lancet Infect Dis 2020;.
Chen A, Huang J, Liao Y, Liu Z, Chen D, Yang C, et al. Differences in clinical and imaging presentation of pediatric patients with COVID-19 in comparison with adults. Radiol Cardiothor Imaging 2020;2:e200117.