Development and internal validation of a diagnostic prediction model for COVID-19 at time of admission to hospital

Running title: diagnostic prediction model for COVID-19

Fink D L[1*], Khan P Y[2*], Goldman N[3], Cai J[1], Hone L[3], Mooney C[3], El-Shakankery K H[3], Sismey G[3], Whitford V[3], Marks M[2], Thomas S[1 4]

[1] Department of Infectious Diseases, Whipps Cross Hospital, Bart's Health NHS Trust, London

[2] Department of Clinical Research, London School of Hygiene and Tropical Medicine, London

[3] Department of Respiratory Medicine, Whipps Cross Hospital, Bart's Health NHS Trust, London

[4] Department of Acute Medicine, Whipps Cross Hospital, Bart's Health NHS Trust, London

* Co-first authors

Corresponding author: Dr Douglas Fink, Department of Infectious Diseases, Whipps Cross University Hospital, Whipps Cross road, Leytonstone, London, E11 1NR; douglas.fink@nhs.net; +447815142926

Author contributions: DLF, PYK and MM conceived the idea. DLF and PYK designed the model. PYK performed the analyses and computations. DLF took the lead in writing the manuscript. PYK, MM and ST made key contributions to early drafts of the manuscript. NG, JC, LH, CM, KHE, GS and VW collected data, suggested analyses and edited drafts of the final manuscript. ST supervised the project.

Abstract

Background: Early COVID-19 diagnosis prior to laboratory testing results is crucial for infection control in hospitals. Models exist predicting COVID-19 diagnosis, but significant concerns exist regarding methodology and generalisability.

Aim: To generate the first COVID-19 diagnosis risk score for use at the time of hospital admission using the TRIPOD (transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) checklist.

Design: A multivariable diagnostic prediction model for COVID-19 using the TRIPOD checklist applied to a large single-centre retrospective observational study of patients with suspected COVID-19.

Methods: 581 individuals were admitted with suspected COVID-19; the majority had laboratory-confirmed COVID-19 (420/581, 72.2%). Retrospective collection was performed of electronic clinical records and pathology data.

Results: The final multivariable model demonstrated AUC 0.8535 (95% confidence interval (0.8121 – 0.8950). The final model used 6 clinical variables that are routinely available in most low and high resource settings. Using a cut-off of 2, the derived risk score has a sensitivity of 78.1% and specificity of 86.8%. At COVID-19 prevalence of 10% the model has a negative predictive value (NPV) of 96.5%.

Conclusions: Our risk score is intended for diagnosis of COVID-19 in individuals admitted to hospital with suspected COVID-19. The score is the first developed for COVID-19 diagnosis using the TRIPOD checklist. It may be effective as a tool to rule out COVID-19 and function at different pandemic phases of variable COVID-19 prevalence. The simple score could be used by any healthcare worker to support hospital infection control prior to laboratory testing results.

Introduction

Coronavirus disease 2019 (COVID-19) is a potentially life-threatening acute respiratory infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)(1). The virus, originally identified in Wuhan, China, is responsible for pandemic disease and unprecedented global pressures on acute hospital services(2). Pneumonia appears to be the most common presentation of COVID-19 although a range of non-respiratory symptoms are common(3,4). Hospitalisation rates for individuals with COVID-19 increase with age, up to approximately 18%(5).

Prompt diagnosis of COVID-19 at admission is fundamental to acute management, infection control and prevention of nosocomial transmission (6). Worldwide, secondary care isolation facilities have been saturated meaning that patients admitted to hospital must frequently be cohorted in shared ward spaces(7). Inappropriate cohorting of patients with and without COVID-19 risks nosocomial transmission. The current gold standard for diagnosis of acute COVID-19 remains laboratory-based PCR testing of respiratory samples, most commonly swabs from the upper respiratory tract (URT)(8,9). Even in high-resource settings, the turnaround time from sampling to result is often more than 24 hours, and testing of URT samples has a recognised false-negative rate(10,11). COVID-19 management and infection control decisions at admission are therefore founded on routinely available investigations and clinical judgement.

A number of diagnostic prediction models are available for COVID-19 however systematic review suggested that none are suitable for clinical use(12). In particular, no model has been derived from a real-world population of individuals with suspected COVID-19 who require admission to hospital. Our diagnostic prediction model is intended for use by acute hospital

staff to support clinical diagnosis of individuals with COVID-19 to guide infection control decisions within the first 24 hours of admission in the absence of laboratory testing for SARS-CoV-2. We used TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) to develop and validate our model using routinely available data from a cohort of patients with suspected COVID-19 requiring hospital admission.

Methods

Participants

Participants included in this study, which was conducted in an east London hospital, were identified from a pathology database of all respiratory specimens sent for laboratory SARS-CoV-2 PCR testing between 16th March and 12th April 2020 (inclusive). The electronic records of all adults (aged >18 years) identified from this database were then interrogated to identify those individuals who were admitted to secondary care. Adults tested in maternity services were excluded as the department did not adhere to national COVID-19 testing guidelines at that time(13). Patients who received COVID-19 testing more than 72 hours after admission were also excluded given the small possibility positive tests results represented nosocomial infection. Admission to secondary care was defined as hospital inpatient stay exceeding 24 hours. Therefore by definition all participants included in the analysis are patients with suspected COVID-19 consistent with national guidelines who received laboratory SARS-CoV-2 testing at least once within 72 hours of admission (13).

Source of data

Retrospective collection was undertaken of all routinely available clinical observations and blood test results at time of admission from Clinical Records Service Millennium electronic patient records. Nadir observations were selected as the least physiologically favourable

measurement at any time prior to time of patient arrival on an inpatient ward. Admission chest radiography reports by consultant radiologists were classified according to the British Institute of Radiology and British Society of Thoracic Imaging templates: normal, suggestive of COVID-19, indeterminate for COVID-19, suggestive of alternative diagnosis(14). Local guidelines actively discouraged use of CT imaging at this time and we did not collect these data. All results of laboratory SARS-CoV-2 PCR testing of respiratory samples for each individual were recorded. At the time of this study there was no local guidance regarding the number of samples that should be sent from individual patients after an initial negative result. Retrospective electronic notes audit was undertaken to establish history of fever, duration of symptoms and final clinical diagnosis at 30 days post admission.

Outcome

Our outcome was defined as either respiratory sample positive for SARS-Cov2 (on either the admission sample or on a subsequent swab taken within 72 hours of admission) OR a clinical diagnosis of COVID-19. A clinical diagnosis was defined by a senior clinician (a consultant) formally documenting a diagnosis of COVID-19 in electronic records in the absence of a positive SARS-CoV-2 laboratory test. A clinical diagnosis of COVID-19 was ascertained from a retrospective review of the electronic notes undertaken at discharge or death, or during the admission if still an inpatient at 30 days.

Predictors

As a group, authors prioritised a minimum number of predictors that the model should contain to enhance implementation in the clinical setting weighed against the effective sample size (see below). Decisions about which predictors to retain in the final model were therefore based

primarily on clinical reasoning and availability of predictor measurement at the time the model would be used (within 24 hours of admission).

We included age as it is independently associated with severity of symptoms resulting from COVID-19 infection, and we reasoned that the likelihood of hospital admission with suspected COVID-19 increases with age. We created a composite variable for fever to encompass history of fever (subjective reporting at any time during acute illness) and any documented fever ≥37.8 degrees Celsius. We reasoned that this composite would be more sensitive to capture fever as a feature of COVID-19, in particular in the presence of anti-pyretic therapy. Hypoxemic respiratory failure, in the absence of other clinical features, is increasingly recognised as a common presentation of COVID-19 pneumonia(15,16). We decided *a priori* to use either maximal FiO_2 requirements within the first 24 hours of admission or lowest recorded 0_2 saturations rather than any respiratory symptoms e.g. shortness of breath or cough. On examination of the data, $Fi0_2$ was used in preference to 0_2 saturations as there was a wider range of values with $Fi0_2$ (see supplementary data section). We also decided *a priori* that heart rate and respiratory rate will have limited discriminatory power with regards to diagnostic prediction model for acutely unwell adults. We also included chest radiograph findings which are commonly used by clinicians to investigate any respiratory or febrile illness. We reasoned that most clinicians are able to distinguish a normal chest x-ray from one that is non-specifically abnormal or indeterminate so created a binary category.

CRP rises in both viral and bacterial causes of acute respiratory illness with some literature suggesting that on average CRP level is higher in the bacterial causes of community-acquired pneumonia than viral causes. Conversely, the lack of a rise in neutrophil count has been shown to correlate well with viral causes(15,17). We therefore decided *a priori* that CRP and

neutrophil count should both be included in the model. On examination of the data distribution of the predictors, neutrophil count was also found to have a wider range (as compared to lymphocytes or neutrophil to lymphocyte ratio). We did not include absolute lymphocyte counts as a variable owing to the narrow distribution of data which would be expected to have poor discriminatory function (see supplementary data). Predictors which had many missing values and were therefore excluded were Troponin T (49.5% missing); LDH (67.6% missing); ferritin (43.3%); D-dimer (50.7%).

Our choice of predictors was therefore not based on potentially biased univariable selection of predictors. Predictors examined during modelling included: age, fever (history of fever OR any documented fever ≥37.8 degrees Celsius within 24h), CXR finding (normal or abnormal/indeterminate), maximal oxygen requirements during first 24 hours of admission, lowest recorded oxygen saturations, CRP, neutrophil count, neutrophils to lymphocyte ratio, and highest temperature recorded during first 24 hour. For the continuous predictors, e.g. age, CRP, neutrophil count, a linear relationship with the outcome was modelled after assessment of non-linearity.

Sample size:

All available data were used to maximise the power and generalisability of the results. The sample size was determined by the availability of existing data at the time and not statistical considerations. The well-known rule of thumb for sample size for prediction models is to have at least 10 events per variable (EPV) although this has been called into question recently(18). For binary outcomes, the number of events is the number of cases in the smallest of the two outcome levels so for this analysis, that equates to 90 events. We therefore *a priori* decided to restrict the number of variables in the final prediction model to ≤ 6 (15 EPV).

Model development, calibration and internal validation

Multivariable logistic regression sequentially removing the variable with the largest Wald p-value >0.05 (stepwise backward elimination) was undertaken to generate our final model. A complete-case analysis was conducted, excluding participants with missing information relating to any of the predictors.

Calibration and discrimination

We assessed model calibration, the agreement between probability of COVID-19 predicted by the model and observed probability of COVID-19 within quantiles of predicted risk, graphically in a calibration plot and statistically using the Hosmer-Lemeshow goodness-of-fit test(19). We then assessed discrimination, the ability of our model to differentiate patients with COVID-19 from those without using the area under the receiver-operating characteristic curve (AUC).

Internal Model Validation

The bootstrap procedure was employed for internal validation(20). The predictor selection was applied to each bootstrap sample to obtain a final model, and the optimism was estimated by comparing the final model performance to the original data for each bootstrap sample (N=200). The bootstrap corrected area under the curve was computed by subtracting the optimism from the original area under the curve.

Transformation from regression model to risk score

We used the beta coefficients and intercept from the final regression model, generated from the complete dataset, to calculate the risk score for each participant. We then used the risk score

to create a cut-off threshold to identify ‘high risk COVID-19’ versus ‘low risk COVID-19’ patients and calculated sensitivity, specificity, negative and positive predictive values for each cut-off. The negative predictive value (NPV) and positive predictive value (PPV) were then calculated for each cut-off threshold for varying prevalence of COVID-19 in the study population.

All statistical analyses were performed using STATA/SE 16 (Stata Corporation, College Station, USA).

Study approval

The study was reviewed by the Joint Research Management Office for Barts Health NHS Trust and Queen Mary University of London. The UK Government Coronavirus Act provides for hospitals to utilise anonymised routinely collected health data as part of the response to the COVID-19 pandemic and ethical approval was therefore not required.

Results

Patient characteristics

Between 16th March and 12th April, 985 individuals were tested for SARS CoV-2 at our hospital. 581 patients, who were admitted to hospital with suspected COVID-19, were included for analysis (Figure 1). Most participants were diagnosed with confirmed or probable COVID-19 (491/581, 84.5%; Table 1); the majority had COVID-19 confirmed by laboratory testing (420/581, 72.2%) and most positive results were confirmed on the first respiratory specimen tested (381/420, 90.7%). The median age of the study population was 67 years (range 19,101) and 58.7% were male (Table 1). For individuals with data available, the median duration of symptoms prior to admission was 7 days (N=499; range 0,32). Observation variables and a full blood count were available for all patients at the time of admission. Chest radiographs were not performed at admission in 9 (1.6%) and CRP was missing for 12 patients (2.1%).

Diagnostic model and performance measures

The development of this model was primarily guided *a priori* by clinical reasoning as outlined in methods. Our final model comprised of age (continuous, linear), neutrophil count (continuous, linear), CRP (continuous, linear), maximal FiO2 requirements (continuous, linear), documented or reported fever (yes= 1, no=0) and chest x-ray findings (normal=1, not normal (including indeterminate)=0). There was very little missing data for the predictors included in the final model. Only 21 (3.6%) patients were dropped from the complete case analysis. The final multivariate model is show in Table 2 (regression coefficients).

Internal validation

The final multivariable model demonstrated adequate calibration and discrimination with Hosmer-Lemeshow statistic p=0.41 and AUC 0.8535 (95% confidence interval (0.8121 –

0.8950). The AUC and calibration plots are shown in Figure 2 and E1 respectively. The optimism-corrected AUC was 0.8465 (95% CI 0.7814 – 0.9038). The model performed comparably well for patients aged less than 80 years (AUC 0.8736, 95% confidence interval 0.8291 – 0.9181) and greater than 80 years (AUC 0.8364, 95% confidence interval 0.7492 – 0.9236).

Risk scores

The risk scores were rounded to the nearest integer and Table 3 shows the proportion of patients diagnosed with COVID-19 at each value of the rounded risk score. A histogram of the risk scores is shown in Figure E2. Using a cut-off threshold of 2 for the risk score, the diagnostic prediction model has a sensitivity of 78.1% and specificity of 86.8%. At COVID-19 prevalence of 85%, the diagnostic prediction model has a positive predictive value (PPV) of 95.1% and negative predictive value (NPV) of 36.0% (Table E1). At COVID-19 prevalence of 10% the PPV falls to 28.1% and NPV rises to 96.5% (Table E1).

Discussion:

Our retrospective cross-sectional study is the first to derive a COVID-19 diagnostic prediction model intended to support infection control decision-making at the time of hospital admission in the absence of laboratory testing. To our knowledge, we have also derived and internally validated the first COVID-19 diagnostic prediction model that applies the TRIPOD checklist for prediction model development(21). Our model performs well in predicting COVID-19 diagnosis at times of high COVID-19 prevalence, such as during the period of data collection for this study (AUC=0.8535). Crucially, extrapolated from our data, the model demonstrates a high NPV when applied to populations with lower COVID-19 prevalence which may speak to its utility as a rule-out tool for COVID-19 at other phases of the pandemic (NPV=93.7%). Our prediction tool may also help guide clinician decision making about whether repeat SARS-CoV-2 testing is warranted following an initial negative test. We would have liked to attempt external validation of other published COVID-19 diagnostic prediction models purposed for COVID-19 infection control but found only one related article that has not been peer-reviewed and did not generate a prediction model(22).

Strengths and limitations:

Multiple other diagnostic prediction models, derived from different populations and clinical scenarios, are published or in pre-print, however there is significant concern regarding the rigour of model development and their real-world utility(12). Our diagnostic prediction model is a cross-sectional study based on participants selected on the basis of symptoms or signs suggestive of the condition of interest (COVID-19) which should intrinsically minimise risk of bias. Selection of predictors is a point of controversy consistently raised in commentaries relating to prediction modelling(23). There is no formal consensus regarding the best method for selecting predictors for such models but suggested approaches include literature review,

clinical experience and statistical selection of variables, all of which informed our model(12,20,24,25). Given the risk of bias associated with univariate analyses to select predictors, we analysed data distribution to select predictors with high likelihood of discriminatory function (Figures E3-4; Table E2). We also prioritised clinical availability of measurements in our predictor selection. All of our model predictors should be routinely available in all healthcare settings, including those of low-resource. Further, the use of investigation results in the model has also been intentionally designed for simple use by a non-specialist healthcare worker, including those unfamiliar with COVID-19. In particular, the predictor of chest radiography findings has been dichotomised to normal or abnormal for the final risk score which requires very little familiarity with COVID-19 imaging or indeed chest radiography generally. Specialist and non-routine investigations, such as lung ultrasound, used in other COVID-19 diagnosis prediction studies, were also purposefully not assessed in this study(26). An alternative diagnosis to COVID-19, as determined by the admitting team, was not assessed as a predictor given inherent complexity and subjectivity as a parameter, and given the retrospective design of the study.

One important limitation of our study is that data collection was undertaken at approximately the time of peak prevalence of COVID-19 in the UK. The model is therefore likely to be over-optimistic in predicting COVID-19 disease. However, our simple model can be easily tested in different settings. The small proportion of participants diagnosed with probable COVID-19, within the composite COVID-19 diagnosis outcome, might be expected to introduce bias given that the clinical diagnosis of probable COVID-19 would not have been blinded to the same predictors used in our model. However, the model performed comparably whether outcomes encompassed confirmed COVID-19 cases alone or a composite of confirmed and probable COVID-19 cases (data not shown). It is also important to note that prevalence of other

respiratory viruses (such as influenza) was low in this population at this time (3 respiratory specimens tested positive for non-SARS-CoV-2 pathogens out of 115 requested). Our predictors may not prove to be highly specific for COVID-19 as compared with infections caused by other respiratory viruses. However, the same principles of early identification and infection control apply to any viral respiratory illness. It will be instructive to validate the model during winter seasons when we might traditionally expect higher prevalence of other seasonal respiratory viruses compared to SARS-CoV-2(27). However, COVID-19-related changes to global travel and public health may significantly disrupt future global patterns of respiratory virus seasonality and prevalence.

The study was conducted at a single site but is expected to be representative of UK and western European acute hospital settings, and urban UK and western European patient populations. The study was designed to predict COVID-19 diagnosis in individuals presenting with syndromes consistent with suspected COVID-19. Our definition of suspected COVID-19 was derived from national guidelines which emphasise testing in patients presenting with acute respiratory or influenza-like illnesses(13). Our study was conducted prior to formal addition of anosmia or dysgeusia as symptoms that indicate COVID-19 laboratory testing. We would expect most patients requiring hospitalisation would have other symptoms besides these relatively minor complications. However, our study may have missed individuals with atypical presentations of COVID-19, such as delirium in elderly patients, which are not included in national COVID-19 testing guidelines as clinical syndromes indicating SARS-CoV-2 laboratory testing(28). Our data suggest that diagnosis of COVID-19 in elderly populations, particularly those aged over 80 years, with respiratory or influenza-like symptoms is comparable to younger populations. Diagnostic prediction models aimed at elderly populations and other sub-populations who

might be expected to present with atypical or pauci-respiratory syndromes of COVID-19, such as immunocompromised patients, warrant dedicated study.

Implications:

Necessarily the wider utility of our model requires validation to define its role in clinical practice. External validation of the model is needed for similar populations requiring hospitalisation with suspected COVID-19 both inside and outside the UK. Temporal validation of our model will be crucial to explore its utility at different phases of the COVID-19 pandemic. While clinical acumen and appropriate laboratory testing should dictate management of COVID-19, if validation in a dataset where COVID-19 prevalence is low does confirm the high NPV of our model then it promises to be a useful tool to support decision-making for patients with suspected COVID-19 at the time of admission to hospital. The model may serve to rule out COVID-19 in these patients and extricate them from respiratory isolation, thus preserving infection control resources and patient safety. At times of high prevalence of COVID-19, such as might be expected during a second epidemic wave, our model will add little to the diagnosis of individuals presenting with suspected COVID-19 when the pre-test probability is high. But during dynamic phases of increasing and reducing incidence, the model, which is dependent on cheap and widely available investigations, may prove very useful in both high and low resource countries.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References:

1. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Vol. 5, Nature Microbiology. Nature Research; 2020. p. 536–44.
2. COVID-19 situation reports [Internet]. [cited 2020 May 6]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
3. Vetter P, Vu DL, L'Huillier AG, Schibler M, Kaiser L, Jacquerioz F. Clinical features of covid-19. Vol. 369, The BMJ. BMJ Publishing Group; 2020.
4. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 16,749 hospitalised UK patients with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. medRxiv [Internet]. 2020 Apr 28 [cited 2020 May 11];2020.04.23.20076042. Available from: https://www.medrxiv.org/content/10.1101/2020.04.23.20076042v1
5. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis. 2020 Mar 30;0(0).
6. World Health Organization. Clinical management of severe acute respiratory infection when COVID-19 is suspected (v1.2) [Internet]. Who. 2020 [cited 2020 May 10]. p. 1–21. Available from: https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected
7. Verelst F, Kuylen E, Beutels P. Indications for healthcare surge capacity in European countries facing an exponential increase in coronavirus disease (COVID-19) cases, March 2020. Eurosurveillance [Internet]. 2020 Apr 2 [cited 2020 May

11];25(13):2000323. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.13.2000323

8. CDC. Interim Guidance: Healthcare Professionals 2019-nCoV | CDC [Internet]. Center for Disease Control and Prevention. 2020 [cited 2020 May 11]. Available from: https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html

9. World Health Organization. National laboratories [Internet]. Who. 2020 [cited 2020 May 11]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance

10. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 2020 Feb 26;200642.

11. Wang X, Tan L, Wang X, Liu W, Lu Y, Cheng L, et al. Comparison of nasopharyngeal and oropharyngeal swabs for SARS-CoV-2 detection in 353 patients received tests with both specimens simultaneously. Int J Infect Dis. 2020 May 1;94:107–9.

12. Wynants L, Van Calster B, Bonten MMJ, Collins GS, Debray TPA, De Vos M, et al. Prediction models for diagnosis and prognosis of covid-19 infection: Systematic review and critical appraisal. BMJ. 2020 Apr 7;369.

13. Public Health England. COVID-19: investigation and initial clinical management of possible cases [Internet]. gov.uk. 2020 [cited 2020 May 11]. p. 1–5. Available from: https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases/investigation-and-initial-clinical-management-of-possible-cases-of-wuhan-novel-coronavirus-wn-cov-infection#interim-definition-possible-cases

14. Hare SS, Rodrigues JCL, Nair A, Jacob J, Upile S, Johnstone A, et al. The continuing

evolution of COVID-19 imaging pathways in the UK: a British Society of Thoracic Imaging expert reference group update. Clinical Radiology. W.B. Saunders Ltd; 2020.

15. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020 Feb 28;

16. Bertran Recasens B, Martinez-Llorens JM, Rodriguez-Sevilla JJ, Rubio MA. Lack of dyspnea in patients with Covid-19: another neurological conundrum? Eur J Neurol [Internet]. 2020 May 12 [cited 2020 May 27];ene.14265. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ene.14265

17. Ruiz-González A, Falguera M, Vives M, Nogués A, Porcel JM, Rubio-Caballero M. Community-acquired pneumonia: Development of a bedside predictive model and scoring system to identify the aetiology. Respir Med. 2000;94(5):505–10.

18. van Smeden M, Moons KGM, de Groot JAH, Collins GS, Altman DG, Eijkemans MJC, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Stat Methods Med Res. 2019 Aug 1;28(8):2455–74.

19. Ensor J, Snell KI, Martin EC. PMCALPLOT: Stata module to produce calibration plot of prediction model performance. Stat Softw Components. 2020 Jan 4;

20. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001;54(8):774–81.

21. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med. 2015 Jan 6;162(1):55–63.

22. Patterson B, Marks M, Martinez-Garcia G, Bidwell G, Luintel A, Ludwig D, et al. A Novel Cohorting and Isolation Strategy for Suspected COVID-19 Cases during a Pandemic. medRxiv [Internet]. 2020 [cited 2020 May 27];2020.05.14.20091843.

Available from: https://doi.org/10.1101/2020.05.14.20091843

23. Sperrin M, Grant SW, Peek N. Prediction models for diagnosis and prognosis in Covid-19. Vol. 369, The BMJ. BMJ Publishing Group; 2020.

24. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med. 2015 Jan 6;162(1):W1–73.

25. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an ABCD for validation [Internet]. Vol. 35, European Heart Journal. 2014 [cited 2020 May 27]. p. 1925–31. Available from: https://academic.oup.com/eurheartj/article-abstract/35/29/1925/2293109

26. Peyrony O, Marbeuf-Gueye C, Truong V, Giroud M, Rivière C, Khenissi K, et al. Accuracy of Emergency Department Clinical Findings for Diagnosis of Coronavirus Disease 2019. Ann Emerg Med [Internet]. 2020 Oct 1 [cited 2020 Oct 13];76(4):405–12. Available from: /pmc/articles/PMC7241345/?report=abstract

27. Public Health England. Annual flu reports - GOV.UK [Internet]. Annual flu reports. 2020 [cited 2020 Jun 17]. Available from: https://www.gov.uk/government/statistics/annual-flu-reports

28. Ouslander JG. Coronavirus Disease19 in Geriatrics and Long-Term Care: An Update. Journal of the American Geriatrics Society. Blackwell Publishing Inc.; 2020.

Tables:

Characteristics	N (col %)
Age at admission (years): median (IQR)	67 (19 – 101)
Male	341 (58.7)
Time between symptom onset and admission (days): median (IQR) (N=499)	7 (3 – 10)
History of fever or documented fever in the first 24h from admission	420 (72.3)
Observations in the first 24 hours from admission: median (IQR)	
Highest temperature	38.0 (37.3 – 38.6)
Lowest 0_2 saturations	93 (90 – 95)
Highest $Fi0_2$	0.36 (0.24 – 0.60)
Highest respiratory rate	27 (22 – 32)
Highest heart rate	104 (93 – 115)
Lowest systolic blood pressure	107 (99-117)
Chest x-ray at admission	
Normal	122 (21.0)
Suggestive of COVID-19	303 (52.1)
Indeterminate	103 (17.7)
Alternative diagnosis	44 (7.6)
Not done	9 (1.6)
Blood tests on admission: median (IQR)	
CRP, mg/L (N=569)	92 (46 - 167)
Neutrophils, $x10^9$ cells/L	6.4 (4.6 – 9.2)
Lymphocytes, $x10^9$ cells/L	1.0 (0.7 – 1.4)

ALT, U/L (N=494)	30 (20 – 52)
Ferritin, ng/ml (N=366)	891 (430 – 1692)
LDH, U/L (N=222)	400 (280 – 502)
Troponin T, ng/L (N=335)	16 (7 – 42)
D-dimer, mcg/mL (N=319)	1.0 (0.59 – 2.3)
SARS-CoV-2 detected by RT-PCR	420 (72.3)
SARS-CoV-2 detected by RT-PCR on first sample (N=381)	381 (90.7)
Clinical diagnosis of COVID-19 only	71 (12.2)
Outcome: either laboratory-confirmed or clinical diagnosis	491 (84.5)

Table 1. Patient characteristics.

N=number; IQR=interquartile range; FiO2=fraction of inspired oxygen; CRP=C-reactive protein; LDH=lactate dehydrogenase

	Coefficient	Standard error	z	P>z	95% Confidence Interval	
Age	0.0176018	0.0080696	2.18	0.029	0.0017857	0.0334179
Fever	1.547593	0.2931327	5.28	0.000	0.9730635	2.122122
Maximal FiO2	0.0223121	0.0075972	2.94	0.003	0.0074219	0.0372023
CRP	0.0045913	0.0016767	2.74	0.006	0.001305	0.0078776
Normal CXr	-1.113497	0.2879424	-3.87	0.000	-1.677854	-0.5491403
Neutrophils	-0.1611107	0.0321778	-5.01	0.000	-0.2241779	-0.0980434
Intercept	-0.0367507	0.6817362	-0.05	0.957	-1.372929	1.299428

Table 2. Regression co-efficients from the final prediction model.

FiO2= maximal fraction of inspired oxygen in the first 24h of admission; CXr = chest x-ray

	COVID-19 diagnosis		
Score	**NO**	**YES**	**Total**
	N (%)		
-3	1 (50)	1 (50)	2
-2	2 (66.7)	1 (33.3)	3
-1	13 (36.6)	10 (43.5)	23
0	24 (38.1)	39 (61.9)	63
1	31 (29.3)	75 (70.7)	106
2	10 (8.3)	111 (91.7)	121
3	4 (2.7)	145 (97.3)	149
4	1 (1.5)	64 (98.5)	65
5	0 (0)	26 (100)	26
6	0 (0)	1 (100)	1
17	0(0)	1 (100)	1
	86	474	560

Table 3. Risk scores organised by COVID-19 diagnosis.

Supplementary tables:

Prevalence (%)	PPV (%)	NPV (%)
85	95.1	36.0
60	84.1	67.2
50	77.9	75.5
40	70.1	82.2
30	60.1	87.8
20	46.8	92.5
10	28.1	96.5
5	15.6	98.3

Table E1. Performance of the risk score using a cut-off score of 2 at different prevalence of COVID-19.

PPV=positive predictive value; NPV=negative predictive value

Variable	Odds ratio	p value	95% confidence interval
Gender	1.365	0.176	0.869-2.144
History of fever	6.054	<0.001	3.756-9.762
Duration of symptoms	1.045	0.128	0.987-1.106

Observations at admission			
First temperature	1.103	0.361	0.894-1.361
First RR	1.154	<0.001	1.095-1.216
First FiO_2	1.028	<0.001	1.012-1.043
First SpO_2	0.860	<0.001	0.793-0.932
First HR	0.989	0.078	0.978-1.001
First sBP	0.989	0.024	0.978-0.998
Observations within the first 24h of admission			
Highest temperature	2.162	<0.001	1.616-2.893
Highest RR	1.149	<0.001	1.095-1.204
Highest FiO2	1.034	<0.001	1.020-1.049
Lowest SpO2	0.850	<0.001	0.777-0.930
Highest HR	0.988	0.050	0.976-1.000
Lowest sBP	0.996	0.027	0.982-1.011
Blood tests at admission			
Lymphocytes	0.771	0.041	0.601-0.990
Low lymphocytes <$1.0x10^9$/L	1.115	0.637	0.709-1.755

Very low lymphocytes <0.5 x10^9/L	0.712	0.386	0.330-1.534
Neutrophils	0.892	<0.001	0.851-0.935
Neutrophil:lymphocyte ratio	0.982	0.094	0.962-1.001
CRP	1.005	0.001	1.002-1.008
Ferritin	1.001	0.001	1.001-1.002
High ferritin >1000 ng/ml	4.732	0.002	1.779-12.586
Very high ferritin >2000 ng/ml	2.381	0.163	0.704-8.053
High LDH >250 U/L	8.873	<0.001	3.469-22.694
High D dimer >0.5 mcg/mL	0.302	0.110	0.069-1.311
High Troponin T F >14ng/L M >22 ng/L	0.388	0.012	0.186-0.809
High ALT >35 U/L	2.599	0.002	1.430-4.724
Chest x-ray at admission			
Normal	0.225	<0.001	0.139-0.365

Suggestive of COVID-19	7.017	<0.001	3.909-12.600
Indeterminate	1.297	0.420	0.689-2.441

Table E2. Univariate analyses for study variables for COVID-19 diagnosis.

RR=respiratory rate; FiO2= fraction of inspired oxygen; SpO2 = saturation of oxygen by pulse oximetry; HR=heart rate; sBP=systolic blood pressure; LDH = lactate dehydrogenase; F=female; M=male; ALT = alanine aminotransferase

Acronyms:

COVID-19 (coronavirus infectious disease 2019)

SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2)

TRIPOD (transparent reporting of a multivariable prediction model for individual prognosis or diagnosis)

AUC (area under the receiver-operating characteristic curve)

PCR (polymerase chain reaction)

CRP (C-reactive protein)

LDH (lactate dehydrogenase)

PPV (positive predictive value)

NPV (negative predictive value)

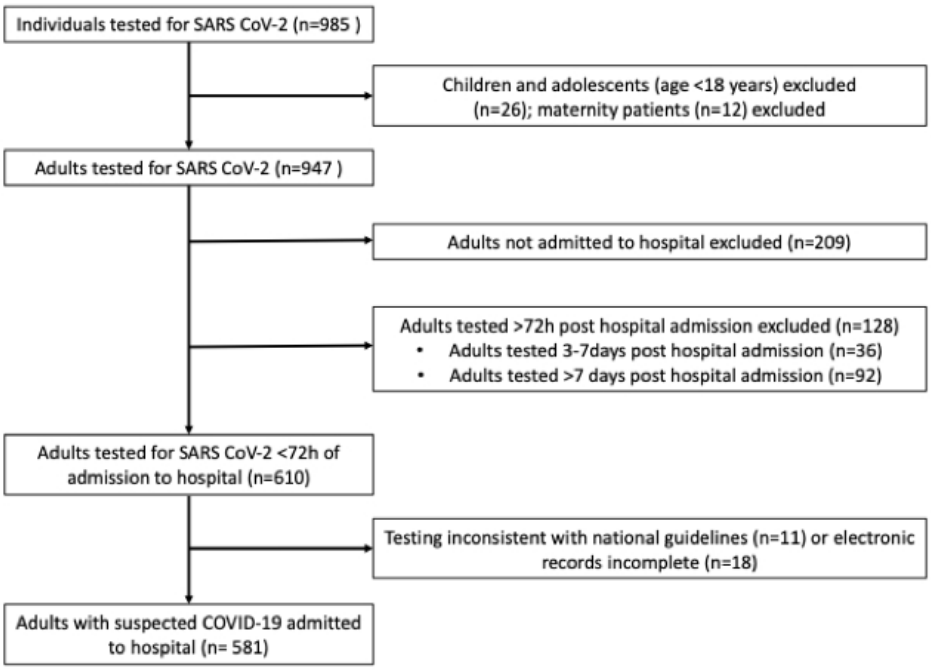

Figure 1. Participant flow.

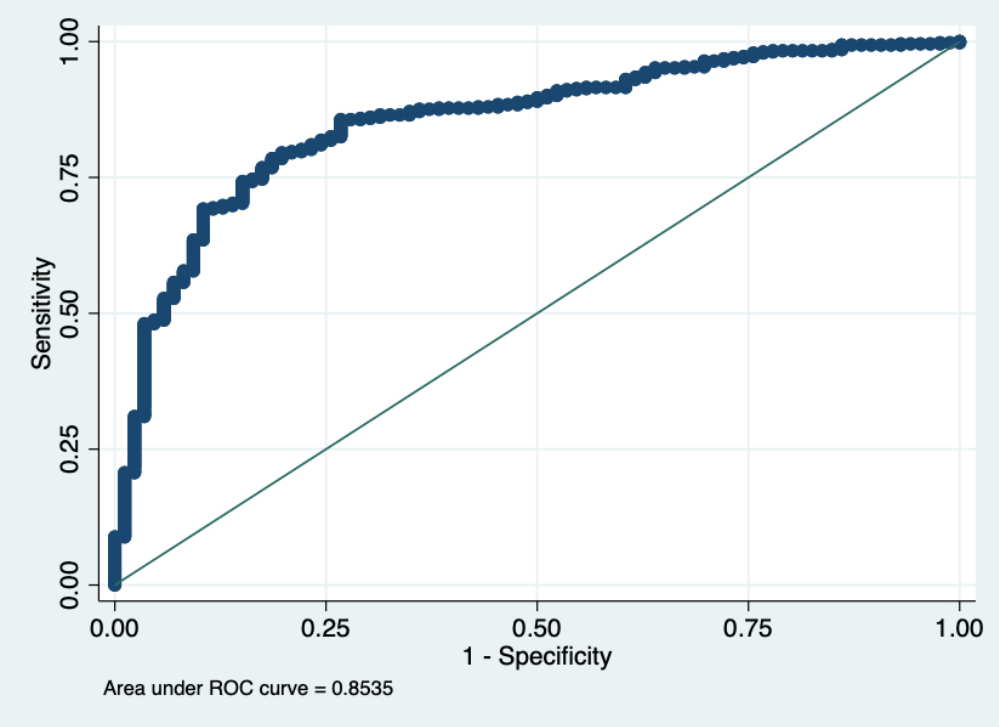

Figure 2. AUC plot for final multivariate model.

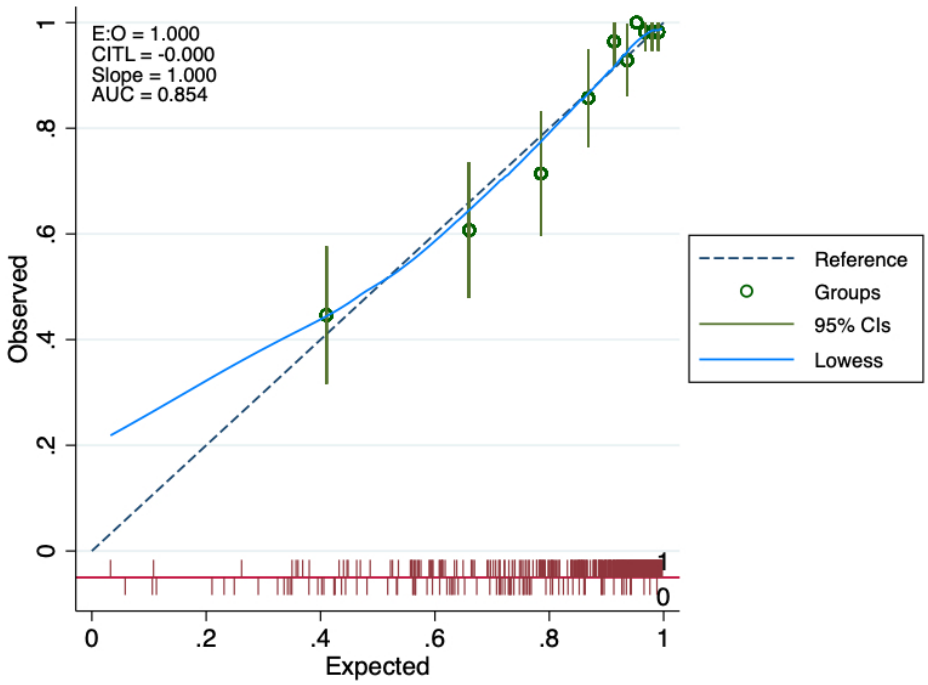

Figure E1. Calibration plot of prediction model performance.

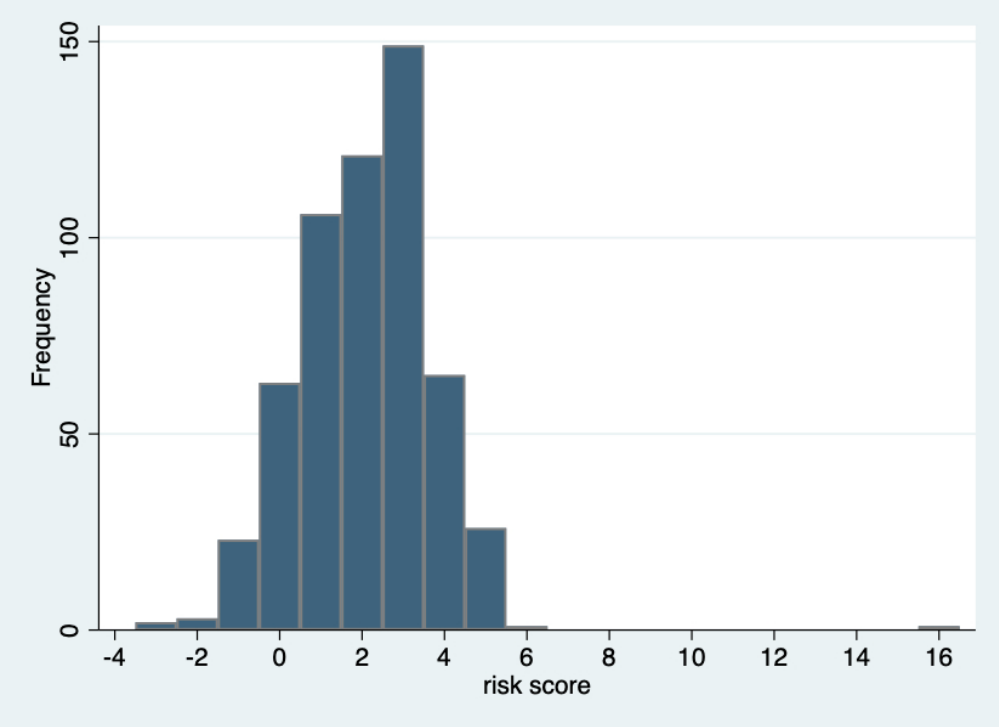

Figure E2. Histogram of the risk score generated by prediction model.

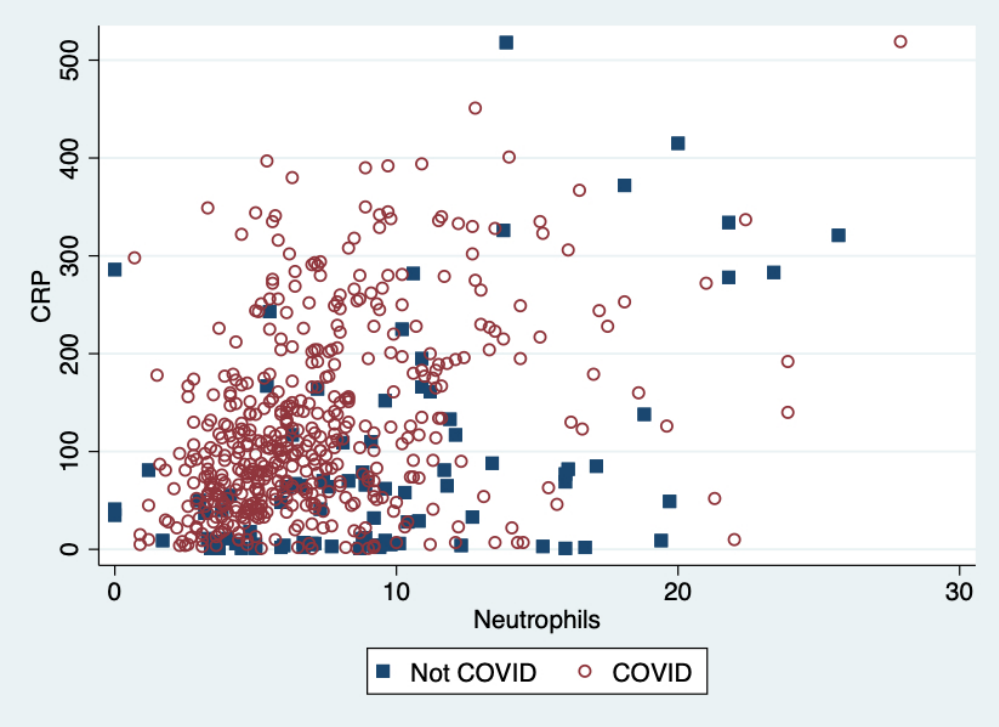

Figure E3. Scatterplot of CRP against neutrophil count organized by COVID and non-COVID diagnosis. CRP=C-reactive protein

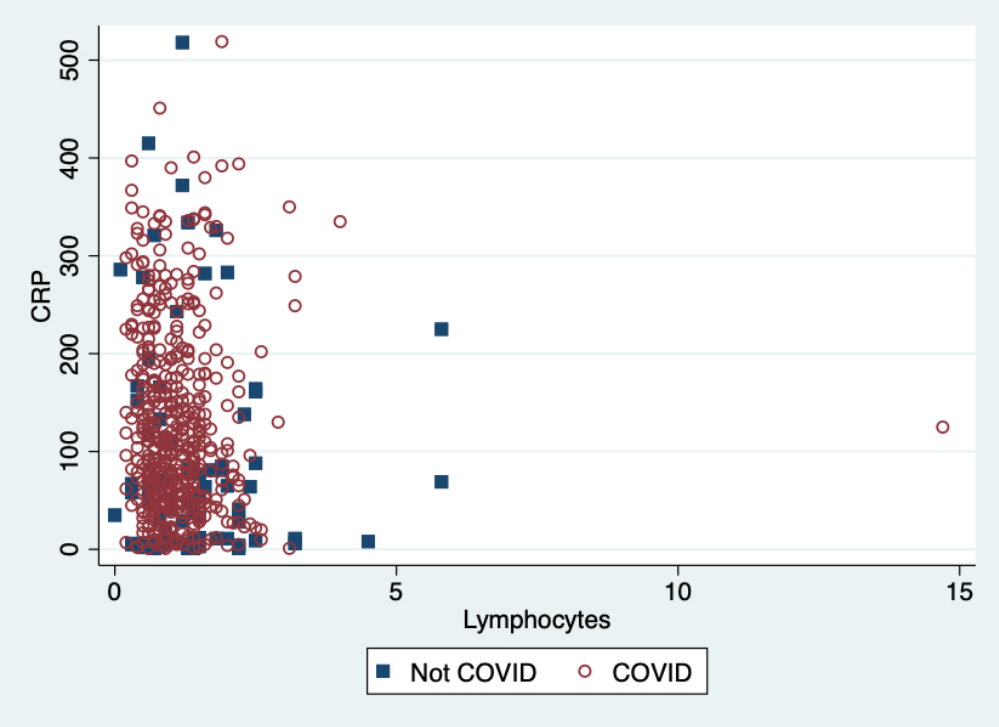

Figure E4. Scatterplot of CRP against lymphocyte count organized by COVID and non-COVID diagnosis. CRP=C-reactive protein