medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

TITLE

Predicting critical illness on initial diagnosis of COVID-19: Development and validation of the PRIORITY model for outpatient applicability.

AUTHORS AND AFFILIATIONS: Miguel Martinez-Lacalzada MD[1], Luis Adrián Viteri-Noël MD[1], Luis Manzano MD PhD[1][2]*, Martin Fabregate-Fuente MEng[1], Manuel Rubio-Rivas MD PhD[3], Sara Luis Garcia MD[4], Francisco Arnalich Fernández MD PhD[5], José Luis Beato Pérez MD[6], Juan Antonio Vargas Núñez MD PhD[7], Elpidio Calvo Manuel MD PhD[8], Alexia-Constanza Espiño MD[9], Santiago J. Freire Castro MD[10], Jose Loureiro-Amigo MD[11], Paula Maria Pesqueira Fontan MD[12], Arturo Artero MD PhD[13], Ana María Álvarez Suárez MD[14], Andrea Silva Asiain MD[15], Beatriz García López MD[16], Jairo Luque del Pino MD PhD[17], Jaime Sanz Cánovas MD[18], Paloma Chazarra Pérez MD[19], Gema María García García MD[20], Jesús Millán Núñez-Cortés MD PhD[4], José Manuel Casas Rojo MD[21], Ricardo Gómez Huelgas MD PhD[18], for the SEMI-COVID-19 Network.

[1] Internal Medicine Department, Hospital Universitario Ramón y Cajal, IRYCIS, Madrid, Spain.
[2] Faculty of Medicine. Universidad de Alcalá (UAH), Madrid, Spain.
[3] Internal Medicine Department, Bellvitge University Hospital-IDIBELL, L'Hospitalet de Llobregat (Barcelona), Spain
[4] Internal Medicine Department, Gregorio Marañon University Hospital, Madrid, Spain
[5] Internal Medicine Department, La Paz University Hospital, Madrid, Spain
[6] Internal Medicine Department, Albacete University Hospital Complex, Albacete, Spain
[7] Internal Medicine Department, Puerta de Hierro University Hospital, Majadahonda (Madrid), Spain
[8] Internal Medicine Department, San Carlos Clinical Hospital, Madrid, Spain
[9] Internal Medicine Department, La Princesa University Hospital, Madrid, Spain
[10] Internal Medicine Department, A Coruña University Hospital, A Coruña, Spain
[11] Internal Medicine Department, Moisès Broggi Hospital, Sant Joan Despí (Barcelona), Spain
[12] Internal Medicine Department, Santiago Clinical Hospital, Santiago de Compostela (A Coruña), Spain
[13] Internal Medicine Department, Dr. Peset University Hospital, Valencia. University of Valencia, Spain
[14] Internal Medicine Department, Cabueñes Hospital, Gijón (Asturias), Spain
[15] Internal Medicine Department, Nuestra Señora del Prado Hospital, Talavera de la Reina (Toledo), Spain
[16] Internal Medicine Department, Zamora Hospital Complex, Zamora, Spain
[17] Internal Medicine Department, Costa del Sol Hospital, Marbella (Málaga), Spain
[18] Internal Medicine Department, Regional University Hospital of Málaga, Biomedical Research Institute of Málaga (IBIMA), University of Málaga (UMA), Málaga, Spain
[19] General Internal Medicine Department, San Juan de Alicante University Hospital, San Juan de Alicante (Alicante), Spain
[20] Internal Medicine Department, Badajoz University Hospital Complex, Badajoz, Spain
[21] Internal Medicine Department, Infanta Cristina University Hospital, Parla (Madrid), Spain

***Correspondence to:** Luis Manzano (luis.manzano@uah.es)

Drs. Martinez-Lacalzada and Viteri-Noël contributed equally to the development of this article.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

SUMMARY

OBJECTIVE

To develop and validate a prediction model, based on clinical history and examination findings on initial diagnosis of COVID-19, to identify patients at risk of critical outcomes.

DESIGN

National multicenter cohort study.

SETTING

Data from the SEMI (Sociedad Española de Medicina Interna) COVID-19 Registry, a nationwide cohort of consecutive COVID-19 patients presenting in 132 centers between March 23 and May 21, 2020. Model development used data from hospitals with ≥300 beds, and validation used those from hospitals with <300 beds.

PARTICIPANTS

Adults (age ≥18 years) presenting with COVID-19 diagnosis.

MAIN OUTCOME MEASURE

Composite of in-hospital death, mechanical ventilation or admission to intensive care unit.

RESULTS

There were 10,433 patients, 7,850 (main outcome rate 25.1%) in the model development cohort and 2,583 (main outcome rate 27.0%) in the validation cohort. The clinical variables in the final model were: age, cardiovascular disease, moderate or severe chronic kidney disease, dyspnea, tachypnea, confusion, systolic blood pressure, and SpO_2 ≤93% or supplementary oxygen requirement at presentation. The model developed had C-statistic of 0.823 (95% confidence interval [CI] 0.813 to 0.834) and calibration slope of 0.995. The external validation had C-statistic of 0.792 (95% CI, 0.772 to 0.812) and calibration slope of 0.872. The model showed positive net benefit in terms of hospitalizations avoided for the predicted probability thresholds between 3% and 79%.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

CONCLUSIONS

Among patients presenting with COVID-19, easily-obtained basic clinical information had good discrimination for identifying patients at risk of critical outcomes, and the model showed good generalizability. A model-based online prediction calculator provided with this paper would facilitate triage of patients during the pandemic.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

INTRODUCTION

The clinical spectrum of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection ranges from an asymptomatic state to critical illness; the symptomatic profile is called coronavirus disease 2019 or COVID-19. [1,2] As of November 25, 2020, the COVID-19 pandemic has affected more than 57 million people worldwide, and has led to nearly 1 300 000 deaths. [3] Notably, Spain has been one of the countries with the highest number of patients with COVID-19. [4] To optimize the use of limited healthcare resources, it would be essential to identify, as early as possible, those patients who are at high risk of progressing to critical illness that necessitates admission to intensive care unit (ICU) or mechanical ventilation, or that may lead to mortality.

To date, studies of COVID-19 prognostic factors [5-15] have focused on laboratory test results and in-hospital data obtained following admission. They have tended not to include clinical variables that could easily be obtained from history and examination carried out on initial assessment in an outpatient setting. Where one machine learning model has addressed basic clinical features, [16] it has narrowed down the prediction to the mortality outcome only and it lacks wider generalizability. A critical appraisal of the COVID-19 models [14] has shown poor reporting and high risk of bias. Recently published well-developed models [9-13] deploy radiological examinations and laboratory measurements such as blood counts, creatinine, lactate dehydrogenase, direct bilirubin, urea and C-reactive protein levels. These tests are not available on initial assessment or in resource-limited settings.

Prediction models based on easy-to-collect data without using imaging or laboratory measures have previously been developed for other infectious diseases, e.g. during meningitis epidemics and pneumonia. [17-19] As a global health emergency, management of COVID-19 too would benefit from a readily applicable prediction model that can be applied on initial diagnosis without the need for radiological and laboratory tests. Therefore, we developed and externally validated a prediction model, based on easily obtained

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

clinical measures at presentation with confirmed COVID-19 diagnosis, to identify patients at risk of developing critical outcomes.

METHODS:

Study design and data source

This study was based on the SEMI (Sociedad Española de Medicina Interna) COVID-19 Registry, the Spanish national registry of COVID-19 patients.[20] It is an ongoing multicenter nationwide cohort of consecutive patients hospitalized for COVID-19 across Spain. At the inception of the cohort, patients were confirmed to be COVID-19 cases defined as a positive result on real-time reverse-transcription-polymerase-chain-reaction (RT-PCR) for the presence of SARS-CoV-2 in nasopharyngeal swab specimens or sputum samples. Exclusion criteria were age under 18 years, subsequent admissions of the same patient and refusal or withdrawal of informed consent. Clinical baseline data, history of previous medication, known comorbidities, laboratory and imaging variables were collected on admission. In addition, treatments administered, complications during hospitalization, status on day of discharge and/or 30 days after diagnosis were obtained. Registry's characteristics have been previously described in detail.[20] The SEMI-COVID-19 Registry was approved by the Provincial Research Ethics Committee of Malaga (Spain) and by Institutional Research Ethics Committees of each participating hospital.

For the study, we used data from patients admitted in 132 hospitals between March 23 and May 21, 2020 provided by the Registry. Development and validation cohorts were defined according to the size of hospitals. Model development was performed on a cohort of patients from hospitals with at least 300 beds, and validated on a separate cohort from hospitals with less than 300 beds. This approach was taken to examine the external validity of the prognostic model [21] in a lower complexity level setting compared to the development setting [22]. The study was reported following the TRIPOD (Transparent Reporting of a

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

multivariable prediction model for Individual Prognosis or Diagnosis) guidelines (*eTable 1*).[23]

Outcome description

Primary outcome, critical illness during hospitalization, was defined as the composite of in-hospital death, mechanical ventilation or admission to ICU, according to previously published studies that assessed the severity of COVID-19 and other infectious diseases.[10,24] Moreover, this composite outcome captures the prioritized outcomes included in the COVID-19 core outcomes set. [25]

Potential predictors

To develop a predictive model useful at out-of-hospital settings, only easily measurable variables registered at admission were selected. For this purpose, from the available variables at SEMI-COVID-19 Registry, we considered clinical signs and symptoms, demographic variables, and medical history, and excluded imaging and laboratory findings. Our approach for selecting predictors was developed to meet the recommendation that new prediction models, rather than using purely data-driven selection, should build on previous literature and expert opinion. [14]

An initial list of 29 candidate variables was selected based on review of the existing evidence, [5-16] clinical plausibility and relevance to clinical care. Demographic variables included age, sex, ethnicity (defined as Caucasian, Latino or others), history of smoking and previous medication as angiotensin converting enzyme inhibitors (ACEi) and angiotensin receptor blockers (ARBs). Medical history included hypertension, cardiovascular disease (history of cerebrovascular disease, peripheral arterial disease, myocardial infarction, angina pectoris, heart failure or atrial fibrillation), moderate or severe dependency for activities of daily living (Barthel index score < 60), diabetes mellitus, obesity and chronic respiratory diseases (asthma, chronic obstructive pulmonary disease, obstructive sleep apnea or hypopnea syndrome), severe chronic kidney disease (previously registered serum creatinine level > 3 mg/dl or

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

history of dialysis), malignancy (solid tumor, leukemia or lymphoma), chronic liver disease, immunocompromised status (autoimmune diseases, solid-organ transplant recipients, HIV infection or previous immunosuppressive treatment including systemic steroids). Clinical signs and symptoms were cough, arthromyalgia, ageusia/anosmia, asthenia/anorexia, headache, gastrointestinal symptoms, fever (defined as temperature ≥ 38 °C or history of fever), systolic blood pressure, heart rate, tachypnea (respiratory rate > 20 breaths per minute), pulmonary rales, confusion, dyspnea and peripheral oxygen saturation by pulse oximetry (SpO_2) ≤ 93% at room air or supplementary oxygen requirement at admission.[26]

To improve consensus on model applicability, a 1-round online questionnaire was conducted among a multidisciplinary panel of 24 physicians involved in COVID-19 clinical management at nursing homes, emergency departments, primary care centers and hospitalization wards (6 per each setting). The panelists were asked to rate (on a 1 to 9 Likert scale) the availability/reliability of each predictor, as well as its ability to predict the outcome, the best way to merge predictors of rare occurrence and the maximum number of variables this model should contain. Agreement was considered when ≤ 7 panelists rated outside the 3-point region containing the median.[27]

Patient and public involvement

Due to the global public health emergency status of the COVID-19 pandemic this research study was conducted without the opportunity for patient and public involvement.

Statistical analysis

The predictive model, called PRIORITY, was presented as the formula for estimating the probability of COVID-19 critical illness outcome, as well as an associated web-based calculator. To develop and validate the model patients' characteristics were summarized in terms of frequencies and percentages for

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

categorical variables and by the mean and standard deviation (SD) for continuous variables. Statistical analysis was performed with R software version 4.0.0 (The R Foundation for Statistical Computing), with the mice, mfp, glmnet, pROC, and rmda packages.

Model development: Missing values in the potential predictors were imputed using single imputation. A stochastic single imputation dataset was created for both cohorts (development and validation) as the first of a series of datasets created by multiple imputation by chained reactions. Single imputation was selected as a reasonable alternative to dealing with multiple completed datasets with relatively few missings.[28] Quantitative variables were kept as continuous to avoid loss of prognostic information, and non-linear relationships were modelled using the multivariate fractional polynomials with a maximum of 2 degrees of freedom.[29] The least absolute shrinkage and selection operator (LASSO) method [30] was used to identify a parsimonious set of potential predictors of critical illness. We selected the regulation penalty parameter (λ) that minimized the 10-fold cross-validation mean squared error (MSE) for a maximum number of predictive features in the model settled by the expert panel agreement. Then, this subset of predictors was entered into a logistic regression model, and those that were statistically significant ($p<0.05$) were retained for the final model. The model coefficients were represented as odds ratios (OR), and 95% confidence intervals (95% CI) were obtained using 1000 bootstrap samples.

Model performance: We used Nagelkerke's R^2 to evaluate the overall predictive accuracy of the model. The overall discriminatory ability was assessed using the C-statistic, as the area under the receiver operating characteristic curve (AUC ROC), with 95% CI by stratified bootstrap resampling. Calibration of the model was assessed graphically in a plot with predictions by deciles of risk on the x-axis and observed proportions of outcomes on the y-axis, as well as the locally estimated scatterplot smoothing (LOESS). We got an overfitting-corrected estimate of the calibration slope from the calibration plot by bootstrapping 1000 resamples, with well-calibrated models having a slope of 1. [31]

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Model validation: To estimate the reproducibility of the model's predictions for the underlying population from which the data were originated (internal validation), the potential overfitting and optimism in the model was assessed by 10-fold cross-validation. [30] Moreover, to assess the model´s stability and generalizability to different settings, we externally validated the final model in a separate cohort including patients admitted at smaller hospitals (< 300 beds). The use of less complex hospital setting also helped to assess model generalizability. [21] We reported the same measures of performance as used in the model development cohort.

Sensitivity analysis: To assess the impact of imputation of missing values, we carried out a complete-case analysis, using for model development only those patients with complete data in the potential predictors. We also developed a full model with no restriction in the maximum number of predictors (selecting λ at which the MSE was within one standard error of the minimal MSE). Then, we developed an alternative model using linear continuous predictors instead of fractional polynomial terms.

Decision curve analysis: We undertook decision curve analysis (DCA) [32] to assess the clinical usefulness of the predictive model in terms of net benefit (NB) if used to prioritize hospital referrals that are most likely to require critical care. For the whole range of decision threshold probabilities (p_t), the net benefit of the model was compared to default strategies of treating all or no patients. The NB was calculated as the percentage of true positives minus the percentage of false positives and weighted by the "harm-to-benefit" ratio ($p_t/(1-p_t)$). We represented the NB vs. p_t in a decision curve plot. The benefit of the prediction model was also quantified in terms of reduction in avoidable hospitalization referrals per 100 patients, calculated as: (NB of the model – NB of treat all)/(harm-to-benefit ratio) × 100.[32]

The choice of threshold probability p_t will vary across different regions, according to changing epidemiological situations and availability of health resources, taking into consideration that the intervention would consist of referring the patient to a hospital. At a low threshold, false negatives are

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

minimized at the expense of unnecessary referrals. At a high threshold, patients would be referred less frequently, but some high-risk patients may not be derived to the hospital.

RESULTS

From a total of 11,523 patients of the SEMI-COVID-19 Registry, 10,433 were considered in this study. The development cohort included 7,850 (75.2%) patients, of which 1,967 (25.1%) presented critical illness (650 [8.3%] admitted to the ICU and 1,598 [20.4%] died). The mean age was 65.8 ± 16.4 years (57.2% male), and 66.5% presented comorbidities. Demographics and clinical characteristics for the development cohort are shown on *Table 1.*

Model development and performance

From an initial list of 29 candidate variables, the expert panel forged an agreement on 21 potential predictors for further evaluation in the predictive model. So, chronic liver disease, previous medication (ACEi and ARBs), cough, arthromyalgia, ageusia/anosmia, asthenia/anorexia, headache, gastrointestinal symptoms were excluded. Moreover, consensus was achieved for including a range between 5 and 9 variables on the final model. These 21 potential predictors of critical illness were included in the LASSO predictor selection process. A subset of 9 variables were retained as the best predictors of critical illness (*eFigure 1*), including age squared, moderate or severe dependency, cardiovascular disease, moderate or severe chronic kidney disease, dyspnea, tachypnea, confusion, reciprocal of systolic blood pressure squared, and SpO2 ≤93% or supplementary oxygen requirement. A multivariable logistic regression model was then fitted with these 9 variables. All of them, except for moderate or severe dependency, were statistically significant independent predictors of critical illness and were therefore included in the final prediction model *(Table 2).*

Based on the logistic regression model, the probability of critical COVID-19 illness could be calculated as:

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Probability (%) = 100 / (1 + Exp(-z)), where z = -4.710 + 2.756·[(Age/100)^2] + 0.331·[Cardiovascular disease] + 0.595·[CKD] + 0.508·[Dyspnea] + 0.860·[1/(SBP/100)^2] + 0.905·[Tachypnea] + 1.204·[SpO_2 ≤ 93% or oxygen requirement] + 0.740·[Confusion]. All predictors were coded as binary variables (1 when present and 0 when absent) except for age (years) and systolic blood pressure (mmHg). We also developed an online calculator based on this model (*eFigure 2*), that is accessible at https://www.evidencio.com/models/show/2344.

The final model had an R^2 of 0.346 and an apparent C-statistic of 0.823 (95% CI 0.812, 0.833) to discriminate between patients with and without critical illness in the development cohort *(Figure 1a*). By internal 10-fold cross-validation, estimated performance of the model was 0.822 (95% CI 0.789, 0.848). After bootstrap resampling, the agreement between the observed outcomes and predicted probabilities in the development cohort showed good calibration with a slope of 0.995 *(Figure 2a)*.

External validation

The validation cohort included 2583 (24.8%) patients, of which 698 (27.0%) presented critical illness (200 [7.7%] admitted to the ICU and 594 [23.0%] died). The mean age was 69.5 ± 16.0 years, (54.8% male), and 69.8% presented comorbidities (Table 1). The final model showed good discrimination when externally validated, with a C-statistic of 0.795 (95% CI 0.776, 0.814) *(Figure 1b)*, and a calibration slope of 0.883 (*Figure 2b).*

Sensitivity analysis

We carried out a complete-case analysis selecting as development cohort the 5513 patients with complete data on the 21 potential predictors and the outcome. The resulting model had the same predictors as the final model with imputed data. R^2 was 0.324, with an apparent C-statistic of 0.813 (95% CI 0.800, 0.823) and a slope of 0.992. Next, using the original development cohort, we fitted a model with no restriction

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in maximum number of predictors in the model, resulting in a model with 15 variables, adding sex, moderate or severe dependency, diabetes mellitus, malignancy, immunocompromised status, pulmonary rales and heart rate cubed to the 8 predictors in the PRIORITY model. R^2 was 0.348, with a C-statistic of 0.832 (95% CI 0.821, 0.842) and a calibration slope of 0.995. Likewise, we fitted an alternative model using linear continuous predictors instead of fractional polynomial terms. The linear term of systolic blood pressure was not found to be a significant predictor of critical illness, while moderate or severe dependency was included in the model. R^2 was 0.339, C-statistic of 0.819 (95% CI 0.809, 0.830) and a slope of 0.996.

Net benefit of model use

The decision curve analysis (*Figure 3*) showed a positive net benefit for threshold probabilities (p_t) between 3% and 79%, compared to default strategies (treat-all or treat-none). For low thresholds, below 3%, the net benefit of the model was comparable to managing all COVID-19 patients as if they will progress to critical illness (treat-all strategy). *Table 3* presents estimates of the net benefit of using the model and the reduction in avoidable hospitalization referrals for different probability thresholds.

DISCUSSION

Statement of principal findings

We developed and validated a new clinical risk model to predict COVID-19 critical illness based on eight simple clinical features easily available on initial assessment in out-of-hospital settings. The model was well calibrated, had good discrimination, and performed robustly in an external validation cohort. Moreover, it showed a potential clinical benefit in a variety of scenarios covering different healthcare situations over a range of threshold probabilities, highlighting its practical usefulness. Its web-based calculator can facilitate its immediate application for frontline clinicians facing the current COVID-19 peak.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Strengths and weakness of the study

This study has several methodological strengths maximizing internal and external validity. [23] To the best of our knowledge, this is the first generalizable COVID-19 predictive model built with simple clinical information for use in the outpatient settings, excluding imaging and laboratory data. We developed and validated the model in a large multicenter, national cohort. Ours was a cohort twice as large as the previous model using simple information. [16] It was also one of the largest cohorts of all previous models published to date.[9-16] Our model excluded readmissions, a feature that focusses the analysis on the question of interest, i.e. the need of triage in patients at their first COVID-19 presentation. Moreover, methodology was rigorous, avoiding biases that affected previous studies. [14] We complied with the recommendation made regarding avoidance of data-driven predictor selection. [14] The practical application of the model was maximized by forging an agreement among an expert panel on key issues, such as the predictors most suitable for an outpatient setting and the size (number of variables) of the final model. The performance of the model was strong which allowed discrimination between progression or not to critical illness in COVID-19. Moreover, the model was validated in a separate cohort of patients admitted in smaller hospitals, showing good reproducibility in a healthcare setting of a different complexity level. [21,22] Our decision curve analysis (DCA) showed that the model could be useful in triage of patients under pandemic pressure, providing underpinning evidence to guide policymakers' decisions.

The strength of our findings should be interpreted in light of some limitations. We carefully selected easily available clinical and demographic variables that could be applied at outpatient setting, the data were collected at the time of hospital admission. In this regard it should be kept in mind that during the first COVID-19 peak many patients were hospitalized despite low symptom severity as part of prudent management since not much was known about clinical disease course. We used registry data collected in a situation of healthcare pressure due to the pandemic peak, so the data quality may be variable across

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

centers. In this regard, it is notable that missing data per predictor variable were relatively low. To reduce the impact of data loss we used imputation. The sensitivity analysis found that our model with imputation was robust compared to the performance of the model with the complete cases. The complete-case dataset was 27% smaller than the imputed dataset, a feature that was favourable compared to a previous model using radiology and laboratory tests [12] where the complete dataset was 35% smaller. So, our rate of patients with missing data is even lower. The impact of other assumptions adopted in the model development were also evaluated. For example, restricting the maximum number of predictors to 8 (as recommended by the expert panel to enhance usability in clinical practice) was found not to limit model performance compared to a 15-predictor model developed without restrictions (R^2 from 0.346 to 0.348; C-statistic from 0.823 to 0.832, respectively). Considering the balance between strengths and limitations, our model is ready for application as a triage tool within the context of an evaluative study to allow solidification of evidence about model effectiveness in practice.

Comparison with previous studies

An external validation study of 22 previously published prognostic models including laboratory and radiological had shown that oxygen saturation and age were the most discriminating univariate predictors for in-hospital mortality, and that none of the multivariate models had superior performance than these individual predictors. [15] It is important to point out that the PRIORITY model, despite its simplicity, showed a performance similar to the previously published models that included imaging and laboratory data for prediction. For example, our model (apparent C-statistic 0.823, 95% CI 0.813 to 0.834; external validation C-statistic 0.792, 95% CI 0.772 to 0.812) would be expected to dominate in health economic terms the model of Knight et al. [12] (apparent C-statistic of 0.79, 95% CI 0.78 to 0.79; external validation C-statistic 0.77, 95% CI 0.76 to 0.77) on the basis that it would not incur costs involved in imaging and laboratory tests.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Implication for clinical practice, research and policymakers

Our model could be applied in triage, using easily measurable variables available in outpatient settings, identifying high-risk patients for referral to hospital. The DCA (Figure 3) provides information to underpin clinical management and policy-making under COVID-19 pandemic pressure. The PRIORITY model has potential value, resulting in higher net benefit than the default strategies of treat-all or treat-none (hospitalize all or hospitalize none), over a range of risk thresholds which could be considered as relevant in clinical practice. For example, in situations under pandemic peak pressure or low-resource healthcare systems, policy-makers may consider a cut-off point up to 20%, a threshold that will be associated with higher reduction in unnecessary critical care admissions. However, in situations with low numbers of COVID-19 cases and little risk of overwhelming the critical care capacity, a lower threshold may be considered. For example, a 5% cut-off could be appropriate to make decisions on early referral to hospital attention, minimizing the risk of critical illness without ward level in-patient monitoring. We recommend objectively defining specific cut-off points considering the circumstances and the availability of health resources. This approach would allow for patients under the risk threshold to be as safely managed within the community as possible.

Conclusion

We developed and validated a new prediction model, called PRIORITY, to estimate the risk of critical illness in patients with COVID-19, based on eight clinical variables easily measurable in out-of-hospital settings. This model could help in triage of outpatients at risk for critical COVID-19 illness. The study provides underpinning evidence to inform decision-making in health systems under pandemic pressure.

ACKNOWLEDGEMENTS

We gratefully acknowledge all the investigators who participate in the SEMI-COVID-19 Registry, especially

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

those from Hospital Universitario Ramón y Cajal (Luis F. Abrego-Vaca, Ana Andreu-Arnanz, Octavio A. Arce-García, Marta Bajo-González, Pablo Borque-Sanz, Alberto Cózar-Llistó, Beatriz Del Hoyo-Cuenda, Alejandra Gamboa-Osorio, Isabel García-Sánchez, Óscar A. López-Cisneros, Borja Merino-Ortiz, Elisa Riera-González, Jimena Rey-García, Cristina Sánchez-Díaz, Grisell Starita-Fajardo, Cecilia Suárez-Carantoña, Svetlana Zhilina Zhilina). We especially thank to our colleagues at Hospital Universitario Ramon y Cajal, IRYCIS: Nuria Bara Ledesma, Andrés González García and José Luis Calleja López (Dept. of Internal Medicine), and Javier Zamora and Borja M. Fernandez-Felix (Clinical Biostatistics Unit), for their valuable contributions to the review of the manuscript. The authors would also like to gratefully acknowledge Professor Khalid S. Khan, Distinguished Investigator at the University of Granada, Spain, for his support and advice on the manuscript. Finally, we also thank the SEMI-COVID-19 Registry Coordinating Center, S&H Medical Science Service, for their quality control data, logistic and administrative support. The authors declare that there are no conflicts of interest.

AUTHORS CONTRIBUTIONS

MML, LAVN and MMF planned, conceived the study, analysed and interpreted the data. MML, LAVN, MMF and LM wrote the original draft of the manuscript. MRR, SLG, FAF, JLBP, JAVC, JLA, PMPF, AA, AMAS, ASA, BGL, JLP, JSC, PCP, GMGG, JMNC, JMCR, RGH contributed to read and approved the final version of the manuscript. MML and LAVN are joint first authors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. LM is the guarantor.

FUNDING

This research did not receive external funding.

DISCLOSURES

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The authors declare no conflict of interest.

APPENDIX. LIST OF THE SEMI-COVID-19 NETWORK MEMBERS

Coordinator of the SEMI-COVID-19 Registry: José Manuel Casas Rojo.

SEMI-COVID-19 Scientific Committee Members: José Manuel Casas Rojo, José Manuel Ramos Rincón, Carlos Lumbreras Bermejo, Jesús Millán Núñez-Cortés, Juan Miguel Antón Santos, Ricardo Gómez Huelgas.

SEMI-COVID-19 Registry Coordinating Center: S & H Medical Science Service.

Members of the SEMI-COVID-19 Group

Hospital Universitari de Bellvitge. L'Hospitalet de Llobregat. Xavier Corbella, Narcís Homs, Abelardo Montero, Jose María Mora-Luján, Manuel Rubio-Rivas. **H. U. Gregorio Marañón. Madrid**. Laura Abarca Casas, Álvaro Alejandre de Oña, Rubén Alonso Beato, Leyre Alonso Gonzalo, Jaime Alonso Muñoz, Crhistian Mario Amodeo Oblitas, Cristina Ausín García, Marta Bacete Cebrián, Jesús Baltasar Corral, Maria Barrientos Guerrero, Alejandro Bendala Estrada, María Calderón Moreno, Paula Carrascosa Fernández, Raquel Carrillo, Sabela Castañeda Pérez, Eva Cervilla Muñoz, Agustín Diego Chacón Moreno, Maria Carmen Cuenca Carvajal, Sergio de Santos, Andrés Enríquez Gómez, Eduardo Fernández Carracedo, María Mercedes Ferreiro-Mazón Jenaro, Francisco Galeano Valle, Alejandra Garcia, Irene Garcia Fernandez-Bravo, María Eugenia García Leoni, Maria Gomez Antunez, Candela González San Narciso, Anthony Alexander Gurjian, Lorena Jiménez Ibáñez, Cristina Lavilla Olleros, Cristina Llamazares Mendo, Sara Luis García, Víctor Mato Jimeno, Clara Millán Nohales, Jesús Millán Núñez-Cortés, Sergio Moragón Ledesma, Antonio Muiño Miguez, Cecilia Muñoz Delgado, Lucía Ordieres Ortega, Susana Pardo Sánchez, Alejandro Parra Virto, María Teresa Pérez Sanz, Blanca Pinilla Llorente, Sandra Piqueras Ruiz, Guillermo Soria Fernández-Llamazares, María Toledano Macías, Neera Toledo Samaniego, Ana Torres do Rego, Maria

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Victoria Villalba Garcia, Gracia Villarreal, María Zurita Etayo. **H. U. La Paz-Cantoblanco-Carlos III. Madrid.** Jorge Álvarez Troncoso, Francisco Arnalich Fernández, Francisco Blanco Quintana, Carmen Busca Arenzana, Sergio Carrasco Molina, Aranzazu Castellano Candalija, Germán Daroca Bengoa, Alejandro de Gea Grela, Alicia de Lorenzo Hernández, Alejandro Díez Vidal, Carmen Fernández Capitán, Maria Francisca García Iglesias, Borja González Muñoz, Carmen Rosario Herrero Gil, Juan María Herrero Martínez, Víctor Hontañón, Maria Jesús Jaras Hernández, Carlos Lahoz, Cristina Marcelo Calvo, Juan Carlos Martín Gutiérrez, Monica Martinez Prieto, Elena Martínez Robles, Araceli Menéndez Saldaña, Alberto Moreno Fernández, Jose Maria Mostaza Prieto, Ana Noblejas Mozo, Carlos Manuel Oñoro López, Esmeralda Palmier Peláez, Marina Palomar Pampyn, Maria Angustias Quesada Simón, Juan Carlos Ramos Ramos, Luis Ramos Ruperto, Aquilino Sánchez Purificación, Teresa Sancho Bueso, Raquel Sorriguieta Torre, Clara Itziar Soto Abanedes, Yeray Untoria Tabares, Marta Varas Mayoral, Julia Vásquez Manau. **C. H. U. de Albacete. Albacete.** Jose Luis Beato Pérez, Maria Lourdes Sáez Méndez. **H. Clínico San Carlos. Madrid.** Inés Armenteros Yeguas, Javier Azaña Gómez, Julia Barrado Cuchillo, Irene Burruezo López, Noemí Cabello Clotet, Alberto E. Calvo Elías, Elpidio Calvo Manuel, Verónica Cano, Carmen María Cano de Luque, Cynthia Chocron Benbunan, Laura Dans Vilan, Claudia Dorta Hernández, Ester Emilia Dubon Peralta, Vicente Estrada Pérez, Santiago Fernandez-Castelao, Marcos Oliver Fragiel Saavedra, José Luis García Klepzig, Maria del Rosario Iguarán Bermúdez, Esther Jaén Ferrer, Alejandro Maceín Rodríguez, Alejandro Marcelles de Pedro, Rubén Ángel Martín Sánchez, Manuel Méndez Bailón, Sara Miguel Álvarez, Maria José Nuñez Orantos, Carolina Olmos Mata, Eva Orviz García, David Oteo Mata, Cristina Outon González, Juncal Perez-Somarriba, Pablo Pérez Mateos, Maria Esther Ramos Muñoz, Xabier Rivas Regaira, Laura Mª Rodríguez Gallardo, Iñigo Sagastagoitia Fornie, Alejandro Salinas Botrán, Miguel Suárez Robles, Maddalena Elena Urbano, Andrea María Vellisca González, Miguel Villar Martínez. **Hospital Royo Villanova. Zaragoza.** Nicolás Alcalá Rivera, Anxela Crestelo Vieitez, Esther del Corral Beamonte, Jesús Díez Manglano, Isabel Fiteni Mera, Maria del Mar Garcia Andreu, Martin Gerico Aseguinolaza, Claudia Josa Laorden, Raul

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Martínez Murgui, Marta Teresa Matía Sanz. **H. U. Puerta de Hierro. Majadahonda.** María Álvarez Bello, Ane Andrés Eisenhofer, Ana Arias Milla, Isolina Baños Pérez, Laura Benítez Gutiérrez, Javier Bilbao Garay, Silvia Blanco Alonso, Jorge Calderón Parra, Alejandro Callejas Díaz, José María Camino Salvador, Mª Cruz Carreño Hernández, Valentín Cuervas-Mons Martínez, Sara de la Fuente Moral, Miguel del Pino Jimenez, Alberto Díaz de Santiago, Itziar Diego Yagüe, Ignacio Donate Velasco, Ana María Duca, Pedro Durán del Campo, Gabriela Escudero López, Esther Expósito Palomo, Ana Fernández Cruz, Esther Fiz Benito, Andrea Fraile López, Amy Galán Gómez, Sonia García Prieto, Claudia García Rodríguez-Maimón, Miguel Ángel García Viejo, Javier Gómez Irusta, Edith Vanessa Gutiérrez Abreu, Isabel Gutiérrez Martín, Ángela Gutiérrez Rojas, Andrea Gutiérrez Villanueva, Jesús Herráiz Jiménez, Pedro Laguna del Estal, Mª Carmen Máinez Sáiz, Cristina Martín Martín, María Martínez Urbistondo, Fernando Martínez Vera, Susana Mellor Pita, Patricia Mills Sánchez, Esther Montero Hernández, Alberto Mora Vargas, Cristina Moreno López, Alfonso Ángel-Moreno Maroto, Victor Moreno-Torres Concha, Ignacio Morrás De La Torre, Elena Múñez Rubio, Ana Muñoz Gómez, Rosa Muñoz de Benito, Alejandro Muñoz Serrano, Jose María Palau Fayós, Lina Marcela Parra Ramírez, Ilduara Pintos Pascual, Arturo José Ramos Martín-Vegue, Antonio Ramos Martínez, Isabel Redondo Cánovas del Castillo, Alberto Roldán Montaud, Lucía Romero Imaz, Yolanda Romero Pizarro, Mónica Sánchez Santiuste, David Sánchez Órtiz, Enrique Sánchez Chica, Patricia Serrano de la Fuente, Pablo Tutor de Ureta, Ángela Valencia Alijo, Mercedes Valentín-Pastrana Aguilar, Juan Antonio Vargas Núñez, Jose Manuel Vázquez Comendador, Gema Vázquez Contreras, Carmen Vizoso Gálvez. **H. Miguel Servet. Zaragoza.** Gonzalo Acebes Repiso, Uxua Asín Samper, María Aranzazu Caudevilla Martínez, José Miguel García Bruñén, Rosa García Fenoll, Jesús Javier González Igual, Laura Letona Giménez, Mónica Llorente Barrio. **H. U. La Princesa. Madrid.** María Aguilera García, Ester Alonso Monge, Jesús Álvarez Rodríguez, Claudia Alvarez Varela, Miquel Berniz Gòdia, Marta Briega Molina, Marta Bustamante Vega, Jose Curbelo, Alicia de las Heras Moreno, Ignacio Descalzo Godoy, Alexia Constanza Espiño Alvarez, Ignacio Fernández Martín-Caro, Alejandra Franquet López-Mosteiro, Gonzalo Galvez

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Marquez, María J. García Blanco, Yaiza García del Álamo Hernández, Clara García-Rayo Encina, Noemí Gilabert González, Carolina Guillamo Rodríguez, Nicolás Labrador San Martín, Manuel Molina Báez, Carmen Muñoz Delgado, Pedro Parra Caballero, Javier Pérez Serrano, Laura Rabes Rodríguez, Pablo Rodríguez Cortés, Carlos Rodriguez Franco, Emilia Roy-Vallejo, Monica Rueda Vega, Aresio Sancha Lloret, Beatriz Sánchez Moreno, Marta Sanz Alba, Jorge Serrano Ballester, Alba Somovilla, Carmen Suarez Fernández, Macarena Vargas Tirado, Almudena Villa Marti. **H. U. de A Coruña. A Coruña.** Alicia Alonso Álvarez, Olaya Alonso Juarros, Ariadna Arévalo López, Carmen Casariego Castiñeira, Ana Cerezales Calviño, Marta Contreras Sánchez, Ramón Fernández Varela, Santiago J. Freire Castro, Ana Padín Trigo, Rafael Prieto Jarel, Fátima Raad Varea, Ignacio Ramil Freán, Laura Ramos Alonso, Francisco Javier Sanmartín Pensado, David Vieito Porto. **Hospital Clínico de Santiago. Santiago de Compostela.** Maria del Carmen Beceiro Abad, Maria Aurora Freire Romero, Sonia Molinos Castro, Emilio Manuel Paez Guillan, María Pazo Nuñez, Paula Maria Pesqueira Fontan. **H. de Cabueñes. Gijón.** Ana María Álvarez Suárez, Carlos Delgado Vergés, Rosa Fernandez-Madera Martínez, Eva Fonseca Aizpuru, Alejandro Gómez Carrasco, Cristina Helguera Amezua, Juan Francisco López Caleya, María del Mar Martínez López, Aleida Martínez Zapico, Carmen Olabuenaga Iscar, María Luisa Taboada Martínez, Lara María Tamargo Chamorro. **H. Moisès Broggi. Sant Joan Despí.** Judit Aranda Lobo, Jose Loureiro Amigo, Isabel Oriol Bermúdez, Melani Pestaña Fernández, Nicolas Rhyman, Nuria Vázquez Piqueras. **Hospital Universitario Dr. Peset. Valencia**. Juan Alberto Aguilera Ayllón, Arturo Artero, María del Mar Carmona Martín, María José Fabiá Valls, Maria de Mar Fernández Garcés, Ana Belén Gómez Belda, Ian López Cruz, Manuel Madrazo López, Elisabeth Mateo Sanchis, Jaume Micó Gandia, Laura Piles Roger, Adela Maria Pina Belmonte, Alba Viana García. **H. U. Ramón y Cajal. Madrid.** Luis Fernando Abrego Vaca, Ana Andréu Arnanz, Octavio Arce García, Marta Bajo González, Pablo Borque Sanz, Alberto Cozar Llisto, Sonia de Pedro Baena, Beatriz Del Hoyo Cuenda, María Alejandra Gamboa Osorio, Isabel García Sánchez, Andrés González García, Oscar Alberto López Cisneros, Luis Manzano, Miguel Martínez Lacalzada, Borja Merino Ortiz, Jimena Rey-García, Elisa Riera González,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Cristina Sánchez Díaz, Grisell Starita Fajardo, Cecilia Suárez Carantoña, Adrian Viteri Noel, Svetlana Zhilina Zhilina. **H. Nuestra Señora del Prado. Talavera de la Reina**. Sonia Casallo Blanco, Jeffrey Oskar Magallanes Gamboa, Cristina Salazar Mosteiro, Andrea Silva Asiain. **C. Asistencial de Zamora. Zamora.** Carlos Aldasoro Frias, Luis Arribas Perez, María Esther Fraile Villarejo, Beatriz García López, Victor Madrid Romero, Emilia Martínez Velado, Victoria Palomar Calvo, Sara Pintos Otero, Carlota Tuñón de Almeida. **H. Virgen de la Salud. Toledo.** Ana Maria Alguacil Muñoz, Marta Blanco Fernández, Veronica Cano, Ricardo Crespo Moreno, Fernando Cuadra Garcia-Tenorio, Blanca Díaz-Tendero Nájera, Raquel Estévez González, María Paz García Butenegro, Alberto Gato Díez, Verónica Gómez Caverzaschi, Piedad María Gómez Pedraza, Julio González Moraleja, Raúl Hidalgo Carvajal, Patricia Jiménez Aranda, Raquel Labra González, Áxel Legua Caparachini, Pilar Lopez Castañeyra, Agustín Lozano Ancin, Jose Domingo Martin Garcia, Cristina Morata Romero, María Jesús Moya Saiz, Helena Moza Moríñigo, Gemma Muñiz Nicolás, Enriqueta Muñoz Platon, Filomena Oliveri, Elena Ortiz Ortiz, Raúl Perea Rafael, Pilar Redondo Galán, María Antonia Sepulveda Berrocal, Vicente Serrano Romero de Ávila, Pilar Toledano Sierra, Yamilex Urbano Aranda, Jesús Vázquez Clemente, Carmen Yera Bergua. **Hospital Costa del Sol. Marbella.** Javier García Alegría, Nicolás Jiménez-García, Jairo Luque del Pino, María Dolores Martín Escalante. **H. U. Infanta Cristina. Parla.** Juan Miguel Antón Santos, Ana Belén Barbero Barrera, Coralia Bueno Muiño, Ruth Calderón Hernaiz, Irene Casado Lopez, José Manuel Casas Rojo, Andrés Cortés Troncoso, Mayte de Guzmán García-Monge, Francesco Deodati, Gonzalo García Casasola Sánchez, Elena Garcia Guijarro, Davide Luordo, María Mateos González, Jose A Melero Bermejo, Lorea Roteta García, Elena Sierra Gonzalo, Javier Villanueva Martínez. **Hospital Regional Universitario de Málaga. Málaga.** Mª Mar Ayala Gutiérrez, Rosa Bernal López, José Bueno Fonseca, Verónica Andrea Buonaiuto, Luis Francisco Caballero Martínez, Lidia Cobos Palacios, Clara Costo Muriel, Francis de Windt, Ana Teresa Fernandez-Truchaud Christophel, Paula García Ocaña, Ricardo Gómez Huelgas, Javier Gorospe García, Maria Dolores López Carmona, Pablo López Quirantes, Almudena López Sampalo, Elizabeth Lorenzo Hernández, Juan José Mancebo Sevilla, Jesica Martin Carmona, Luis

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Miguel Pérez-Belmonte, Araceli Pineda Cantero, Carlos Romero Gómez, Michele Ricci, Jaime Sanz Cánovas. **H. U. San Juan de Alicante. San Juan de Alicante**. Marisa Asensio Tomás, David Balaz, David Bonet Tur, Ruth Cañizares Navarro, Paloma Chazarra Pérez, Jesús Corbacho Redondo, Eliana Damonte White, Leticia Espinosa Del Barrio, Pedro Jesús Esteve Atiénzar, Carles García Cervera, David Francisco García Núñez, Vicente Giner Galvañ, Angie Gómez Uranga, Javier Guzmán Martínez, Isidro Hernández Isasi, Lourdes Lajara Villar, Verónica Martínez Sempere, Juan Manuel Núñez Cruz, Sergio Palacios Fernández, Juan Jorge Peris García, Andrea Riaño Pérez, José Miguel Seguí Ripoll, Azucena Sempere Mira, Philip Wikman-Jorgensen. **H. del Henares. Coslada.** Jesús Ballano Rodríguez-Solís, Luis Cabeza Osorio, María del Pilar Fidalgo Montero, Mª Isabel Fuentes Soriano, Erika Esperanza Lozano Rincon, Ana Martín Hermida, Jesus Martinez Carrilero, Jose Angel Pestaña Santiago, Manuel Sánchez Robledo, Patricia Sanz Rojas, Nahum Jacobo Torres Yebes, Vanessa Vento. **H. U. La Fe. Valencia.** Dafne Cabañero, María Calabuig Ballester, Pascual Císcar Fernández, Ricardo Gil Sánchez, Marta Jiménez Escrig, Cristina Marín Amela, Laura Parra Gómez, Carlos Puig Navarro, José Antonio Todolí Parra. **H. de Mataró. Mataró.** Raquel Aranega González, Ramon Boixeda, Javier Fernández Fernández, Carlos Lopera Mármol, Marta Parra Navarro, Ainhoa Rex Guzmán, Aleix Serrallonga Fustier. **H. San Pedro. Logroño.** Diana Alegre González, Irene Ariño Pérez de Zabalza, Sergio Arnedo Hernández, Jorge Collado Sáenz, Beatriz Dendariena, Marta Gómez del Mazo, Iratxe Martínez de Narvajas Urra, Sara Martínez Hernández, Estela Menendez Fernández, Jose Luís Peña Somovilla, Elisa Rabadán Pejenaute. **Complejo Hospitalario Universitario Ourense. Ourense.** Raquel Fernández González, Amara Gonzalez Noya, Carlos Hernández Ceron, Isabel Izuzquiza Avanzini, Ana Latorre Diez, Pablo López Mato, Ana María Lorenzo Vizcaya, Daniel Peña Benítez, Milagros María Peña Zemsch, Lucía Pérez Expósito, Marta Pose Bar, Lara Rey González, Laura Rodrigo Lara. **H. Juan Ramón Jiménez. Huelva.** Francisco Javier Bejarano Luque, Francisco Javier Carrasco-Sánchez, Mercedes de Sousa Baena, Jaime Díaz Leal, Aurora Espinar Rubio, Maria Franco Huertas, Juan Antonio García Bravo, Andrés Gonzalez Macías, Encarnación Gutiérrez Jiménez, Alicia Hidalgo Jiménez,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Constantino Lozano Quintero, Carmen Mancilla Reguera, Francisco Javier Martínez Marcos, Francisco Muñoz Beamud, Maria Perez Aguilera, Alícia Perez Jiménez, Virginia Rodríguez Castaño, Alvaro Sánchez de Alcazar del Río, Leire Toscano Ruiz. **H. U. Reina Sofía. Córdoba.** Antonio Pablo Arenas de Larriva, Pilar Calero Espinal, Javier Delgado Lista, Francisco Fuentes-Jiménez, María Jesús Gómez Vázquez, Jose Jiménez Torres, José López-Miranda, Laura Martín Piedra, Javier Pascual Vinagre, Pablo Pérez-Martinez, María Elena Revelles Vílchez, Juan Luis Romero Cabrera, José David Torres Peña. **Hospital Infanta Margarita. Cabra.** María Esther Guisado Espartero, Lorena Montero Rivas, Maria de la Sierra Navas Alcántara, Raimundo Tirado-Miranda. **Hospital Alto Guadalquivir.** Andújar. Begoña Cortés Rodríguez. **C. H. U. de Ferrol. Ferrol.** Hortensia Alvarez Diaz, Tamara Dalama Lopez, Estefania Martul Pego, Carmen Mella Pérez, Ana Pazos Ferro, Sabela Sánchez Trigo, Dolores Suarez Sambade, Maria Trigas Ferrin, Maria del Carmen Vázquez Friol, Laura Vilariño Maneiro**. C. H. U. de Badajoz. Badajoz.** Rafael Aragon Lara, Inmaculada Cimadevilla Fernandez, Juan Carlos Cira García, Gema Maria García García, Julia Gonzalez Granados, Beatriz Guerrero Sánchez, Francisco Javier Monreal Periáñez, Maria Josefa Pascual Perez. **Complejo Asistencial Universitario de León. León.** Rosario Maria García Diez, Manuel Martin Regidor, Angel Luis Martínez Gonzalez, Alberto Muela Molinero, Raquel Rodríguez Díez, BeatrizVicente Montes. **Hospital Marina Baixa. Villajoyosa.** Javier Ena, Jose Enrique Gómez Segado. **Hospital del Tajo. Aranjuez**

Ruth Gonzalez Ferrer, Raquel Monsalvo Arroyo. **H. de Pozoblanco. Pozoblanco.** José Nicolás Alcalá Pedrajas, Antonia Márquez García, Inés Vargas. **H. U. Marqués de Valdecilla. Santander.** Marta Fernández-Ayala Novo, José Javier Napal Lecumberri, Nuria Puente Ruiz, Jose Riancho, Isabel Sampedro Garcia. **H. U. Severo Ochoa. Leganés.** Yolanda Casillas Viera, Lucía Cayuela Rodríguez, Carmen de Juan Alvarez, Gema Flox Benitez, Laura García Escudero, Juan Martin Torres, Patricia Moreira Escriche, Susana Plaza Canteli, M Carmen Romero Pérez. **Hospital Platón. Barcelona**. Ana Suarez Lombraña. **Hospital Valle del Nalón. Riaño. Langreo**. Sara Fuente Cosío, César Manuel Gallo Álvaro, Julia Lobo García, Antía Pérez Piñeiro. **H. General Defensa.** Anyuli Gracia Gutiérrez, Leticia Esther Royo Trallero. **H. U. del Vinalopó.**

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Elche. Francisco Amorós Martínez, Erika Ascuña Vásquez, José Carlos Escribano Stablé, Adriana Hernández Belmonte, Ana Maestre Peiró, Raquel Martínez Goñi, M.Carmen Pacheco Castellanos, Bernardino Soldan Belda, David Vicente Navarro. **H. Francesc de Borja. Gandia.** Alba Camarena Molina, Simona Cioaia, Anna Ferrer Santolalia, José María Frutos Pérez, Eva Gil Tomás, Leyre Jorquer Vidal, Marina Llopis Sanchis, M Ángeles Martínez Pascual, Alvaro Navarro Batet, Mari Amparo Perea Ribis, Ricardo Peris Sanchez, José Manuel Querol Ribelles, Silvia Rodriguez Mercadal, Ana Ventura Esteve. **H. G. U. de Castellón. Castellón de la Plana.** Jorge Andrés Soler, Marián Bennasar Remolar, Alejandro Cardenal Álvarez, Daniela Díaz Carlotti, María José Esteve Gimeno, Sergio Fabra Juana, Paula García López, María Teresa Guinot Soler, Daniela Palomo de la Sota, Guillem Pascual Castellanos, Ignacio Pérez Catalán, Celia Roig Martí, Paula Rubert Monzó, Javier Ruiz Padilla, Nuria Tornador Gaya, Jorge Usó Blasco. **C. A. U. de Salamanca. Salamanca.** Gloria María Alonso Claudio, Víctor Barreales Rodríguez, Cristina Carbonell Muñoz, Adela Carpio Pérez, María Victoria Coral Orbes, Daniel Encinas Sánchez, Sandra Inés Revuelta, Miguel Marcos Martín, José Ignacio Martín González, José Ángel Martín Oterino, Leticia Moralejo Alonso, Sonia Peña Balbuena, María Luisa Pérez García, Ana Ramon Prados, Beatriz Rodríguez-Alonso, Ángela Romero Alegría, Maria Sanchez Ledesma, Rosa Juana Tejera Pérez. **Hospital Doctor José Molina Orosa. Arrecife Lanzarote. Virginia Herrero García, Berta Román Bernal. H. U. del Sureste. Arganda del Rey**. Jon Cabrejas Ugartondo, Ana Belén Mancebo Plaza, Arturo Noguerado Asensio, Bethania Pérez Alves, Natalia Vicente López. **Hospital de Palamós. Palamós.** Ana Alberich Conesa, Mari Cruz Almendros Rivas, Miquel Hortos Alsina, José Marchena Romero, Anabel Martin-Urda Diez-Canseco. **H. U. Quironsalud Madrid. Pozuelo de Alarcón. Madrid**. Pablo Guisado Vasco, Ana Roda Santacruz, Ana Valverde Muñoz. **H. Parc Tauli. Sabadell.** Francisco Epelde, Isabel Torrente. **H. Virgen de los Lirios. Alcoy. Alicante**. Mª José Esteban Giner. **Hospital Clínico Universitario de Valladolid. Valladolid.** Xjoylin Teresita Egües Torres, Sara Gutiérrez González, Cristina Novoa Fernández, Pablo Tellería Gómez. **H. G. U. de Elda. Elda.** Carmen Cortés Saavedra, Jennifer Fernández Gómez, Borja González López, María Soledad Hernández Garrido Ana Isabel López Amorós,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Santiago López Gil, Maria de los Reyes Pascual Pérez, Nuria Ramírez Perea, Andrea Torregrosa García. **Hospital Público de Monforte de Lemos. Monforte de Lemos.** José López Castro, Manuel Lorenzo López Reboiro. **H. Virgen del Mar. Madrid.** Thamar Capel Astrua, Paola Tatiana Garcia Giraldo, Maria Jesús González Juárez, Victoria Marquez Fernandez, Ada Viviana Romero Echevarry**. Hospital do Salnes. Vilagarcía de Arousa.** Vanesa Alende Castro, Ana María Baz Lomba, Ruth Brea Aparicio, Marta Fernandez Morales, Jesús Manuel Fernández Villar, María Teresa López Monteagudo, Cristina Pérez García, Lorena Rodríguez Ferreira, Diana Sande Llovo, Maria Begoña Valle Feijoo. **Hospital Quironsalud A Coruña. A Coruña.** Hector Meijide Miguez. **H. U. 12 de Octubre. Madrid.** Paloma Agudo de Blas, Coral Arévalo Cañas, Blanca Ayuso, José Bascuñana Morejón, Samara Campos Escudero, María Carnevali Frías, Santiago Cossio Tejido, Borja de Miguel Campo, Carmen Díaz Pedroche, Raquel Diaz Simon, Ana García Reyne, Lucia Jorge Huerta, Antonio Lalueza Blanco, Jaime Laureiro Gonzalo, Jaime Lora-Tamayo, Carlos Lumbreras Bermejo, Guillermo Maestro de la Calle, Barbara Otero Perpiña, Diana Paredes Ruiz, Marcos Sánchez Fernández, Javier Tejada Montes.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

REFERENCES:

1 Docherty AB, Harrison EM, Green CA, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985. doi:10.1136/ bmj.m1985

2 Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;382:1708-20. doi:10.1056/NEJMoa2002032

3 WHO. COVID-19 situation reports. 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed 25 Nov 2020).

4 Ministerio De Sanidad, Consumo y Bienestar Social. Enfermedad por el coronavirus (COVID-19). 2020 (in Spanish). https://www.mscbs.gob.es/en/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Actualizacion_257_COVID-19.pdf (accessed 25 Nov 2020).

5 Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020;323:1061-69. doi:10.1001/jama.2020.1585

6 Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054-62. doi:10.1016/S0140-6736(20)30566-3

7 Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* 2020;323:2052-2059. doi:10.1001/jama.2020.6775

8 Berenguer J, Ryan P, Rodríguez-Baño J, et al. Characteristics and predictors of death among 4,035 consecutively hospitalized patients with COVID-19 in Spain. *Clin Microbiol Infect* 2020;26:1525-1536. doi:10.1016/j.cmi.2020.07.024 26(11):1525-1536

9 Fumagalli C, Rozzini R, Vannini M, et al. Clinical risk score to predict in-hospital mortality in COVID-19 patients: a retrospective cohort study. *BMJ Open* 2020;10:e040729. doi:10.1136/bmjopen-2020-040729

10 Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Intern Med* 2020.180(8):1081-1089. doi:10.1001/jamainternmed.2020.2033

11 Ji D, Zhang D, Xu J, et al. Prediction for Progression Risk in Patients with COVID-19 Pneumonia: the CALL Score. *Clin* Infect Dis 2020;71:1393-1399. doi: 10.1093/cid/ciaa414

12 Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020; 370:m3339. doi:10.1136/bmj.m3339

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

13 Bartoletti M, Giannella M, Scudeller L, et al. Development and validation of a prediction model for severe respiratory failure in hospitalized patients with SARS-Cov-2 infection: a multicenter cohort study (PREDI-CO study). *Clin Microbiol Infect* 2020;26:1545-1553. doi:10.1016/j.cmi.2020.08

14 Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of COVID-19 infection: systematic review and critical appraisal. *BMJ* 2020;369:m1328. doi:10.1136/bmj.m1328

15 Gupta, Rishi K, Marks M, Samuels T, et al. Systematic Evaluation and External Validation of 22 Prognostic Models among Hospitalised Adults with COVID-19: An Observational Cohort Study. *Eur Respir Journal* 2020;2003498. doi:10.1183/13993003.03498-2020

16 Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health* 2020;2:e516-e525.doi: 10.1016/S2589-7500(20)30217-X

17 Ajayi-Obe EK, Lodi E, Alkali AS, et al. Prognostic scores for use in African meningococcal epidemics. *Bull World Health Organ* 1998;76:149-52.

18 Bassat Q, Lanaspa M, Machevo S, et al. Hypoxaemia in Mozambican children <5 years of age admitted to hospital with clinical severe pneumonia: clinical features and performance of predictor models. *Trop Med Int Health* 2016;21:1147-56. doi:10.1111/tmi.12738

19 Lanaspa M, Valim C, Acacio S, et al. High reliability in respiratory rate assessment in children with respiratory symptomatology in a rural area in Mozambique. *J Trop Pediatr* 2014;60:93-98. doi:10.1093/tropej/fmt081

20 Casas Rojo JM, Antón Santos JM, Nuñez-Cortés J, et al. Clinical characteristics of patients hospitalized with COVID-19 in Spain: Results from the SEMI-COVID-19 Network. *Rev Clin Esp* 2020;220:480-494. doi:10.1016/j.rceng.2020.07.003

21 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515-24. doi: 10.7326/0003-4819-130-6-199903160-00016

22 Ministerio de Sanidad, Consumo y Bienestar Social. Registro de Altas de los Hospitales Generales del Sistema Nacional de Salud. CMBD. Norma Estatal. 2018 (in Spanish). www.mscbs.gob.es/estadEstudios/estadisticas/cmbd.htm.

23 Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162:W1-73. doi:10.7326/M14-0698

24 Metlay JP, Waterer GW, Long AC, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med* 2019;200:e45-e67. doi:10.1164/rccm.201908-1581ST

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

25 Evangelidis, Nicole Tong A, Howell M, et al. International Survey to Establish Prioritized Outcomes for Trials in People With Coronavirus Disease 2019. *Critical care medicine* 2020;48:1612-1621. doi:10.1097/CCM.0000000000004584

26 UK National Health Service. Clinical guide for the management of emergency department patients during the coronavirus pandemic. *NHS England* 2020. https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/Specialty-guide_ED-and-coronavirus_V1_17-March.pdf

27 Fitch K, Bernstein S, Aguilar MD, et al. The RAND/UCLA Appropriateness Method User's Manual. 1st Edition. Santa Monica, CA: RAND Corporation (2001).

28 Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. 2nd ed. New York: Springer (2019).

29 Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Appl Stat* 1994;43:429–67. doi:10.2307/2986270

30 Hastie T, Tibshirani R, Friedman JH, et al. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer (2009).

31 Steyerberg E.W, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014; 35:1925-31. doi:10.1093/eurheartj/ehu207

32 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565-74. doi:10.1177/0272989X06295361

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

TABLES AND FIGURES

Table 1. Demographics and clinical characteristics among patients included in the development cohort and validation cohort.

		Development cohort		Validation cohort	
		No of patients (%) or mean ± SD	**Total No (%)**	**No of patients (%) or mean ± SD**	**Total No (%)**
Characteristics of the population					
Critical Illness		1967 (25.1%)	7850 (100%)	698 (27.0%)	2583 (100%)
Age (years)		65.8 ± 16.4	7816 (99.6%)	69.5 ± 16.0	2575 (97.3%)
Male		4483 (57.2%)	7834 (99.8%)	1415 (54.8%)	2580 (99.9%)
Ethnicity	Caucasian	6836 (89.1%)	7677 (98.8%)	2340 (91.0%)	2572 (99.6%)
	Latino	693 (9.0%)		193 (7.5%)	
	Other	148 (1.9%)		39 (1.5%)	
Smoking history	Never	5270 (70.9%)	7433 (94.7%)	1625 (65.7%)	2475 (95.8%)
	Former smoker	1764 (23.7%)		718 (29.0%)	
	Active Smoker	399 (5.4%)		139 (5.3%)	
Medical history					
Obesity		1665 (23.7%)	7012 (89.3%)	584 (24.3%)	2401 (93.0%)
Hypertension		3803 (48.6%)	7833 (99.8%)	1444 (56.1%)	2576 (99.7%)
Diabetes mellitus		1440 (18.4%)	7820 (99.6%)	509 (19.8%)	2570 (99.5%)
Cardiovascular disease		1974 (25.3%)	7800 (99.4%)	806 (31.7%)	2545 (98.5%)
Pulmonary diseases		1625 (20.9%)	7776 (99.1%)	576 (22.6%)	2583 (98.9%)
Severe chronic kidney disease		488 (6.2%)	7825 (99.7%)	163 (6.3%)	2583 (99.7%)
Malignancy		793 (10.2%)	7803 (99.4%)	259 (10.1%)	2571 (99.5%)
Immunocompromised status		650 (8.6%)	7549 (96.2%)	187 (7.6%)	2473 (95.7%)
Dependency (moderate/severe)		1129 (14.7%)	7701 (98.1%)	605 (23.7%)	2555 (98.9%)
Symptoms at admission					
Fever		5138 (67.0%)	7663 (97.6%)	1670 (65.6%)	2544 (98.5%)
Dyspnea		4427 (56.7%)	7805 (99.4%)	1523 (59.4%)	2562 (99.2%)
Clinical signs and physical exploration at admission					
SBP (mmHg)		129.0 ± 21.5	7430 (94.6%)	127.6 ± 21.0	2451 (94.9%)
HR (beats/minute)		88.6 ± 17.4	7500 (95.5%)	87.5 ± 17.5	2504 (96.9%)
Tachypnea (> 20 breaths/min)		2271 (29.9%)	7604 (96.9%)	879 (35.1%)	2504 (96.9%)
$SpO_2 \leq 93\%$ or O_2 requirement		4152 (52.9%)	7842(99.9%)	1605 (62.1%)	2583 (100%)
Pulmonary rales		4630 (60.7%)	7626 (97.1%)	1588 (63.6%)	2495 (96.6%)
Confusion		849 (11.0%)	7736 (98.5%)	384 (15.1%)	2546 (98.6%)

SD: standard deviation.
Obesity is defined as Medical history or BMI ≥30kg/m^2.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Pulmonary disease is defined as chronic obstructive pulmonary disease, obstructive sleep apnea / hypopnea syndrome and asthma.
CKD: Chronic kidney disease (serum creatinine level > 3 mg/dl or history of dialysis).
Fever: Temperature ≥ 38 ºC or history of fever.
HR: Heart rate
SBP: Systolic blood pressure.
SpO_2: Peripheral oxygen saturation

Table 2. Multivariable LASSO logistic regression of critical illness prediction in COVID-19.

Predictors	Odds ratio	Bootstrap 95% CI
$(Age/100)^2$ [years]	15.738	10.957, 22.714
Cardiovascular disease	1.392	1.214, 1.590
Severe CKD severe	1.813	1.391, 2.319
Dyspnea	1.662	1.456, 1.908
$1/(SBP/100)^2$ [mmHg]	2.363	1.885, 3.010
Tachypnea (>20 breaths/min)	2.473	2.206, 2.841
$SpO_2 \leq 93\%$ or O_2 requirement	3.335	2.889, 3.838
Confusion	2.097	1.761, 2.542
Dependency (Moderate or severe)	--	--

Continuous predictors were presented as fractional polynomial algorithms, including a squared term for age and systolic blood pressure as the inverse of a quadratic term.
CKD: Chronic kidney disease.
SBP: Systolic blood pressure.
SpO_2: Peripheral oxygen saturation.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3. Net benefit of using the model compared to a policy managing COVID-19 patients with the assumption all of cases could progress to critical illness.

Probability threshold (p_t, %)	Net benefit		Advantage of using the model	
	Treat all	Prediction Model	Difference in the net benefit	Reduction in avoidable hospitalization referrals per 100 patients
1	0.243	0.243	0	0
2	0.235	0.235	0	0
3	0.227	0227	0	0
4	0.219	0.220	0.001	2
5	0.211	0.214	0.003	6
10	0.167	0.188	0.021	19
15	0.118	0.164	0.046	26
20	0.063	0.144	0.081	32
25	0.001	0.122	0.121	36
50	-0.499	0.005	0.504	50
75	-1.998	0.004	2.002	67
79	-2.569	0	2.569	68
99	-73.943	0	73.943	75

Net benefit: percentage of true positives minus the percentage of false positives weighted by the ratio (pt /(1-pt)).
Reduction in avoidable hospitalization referrals per 100 patients: (net benefit of the model – net benefit of treat all)/(pt/(1− pt)) × 100.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

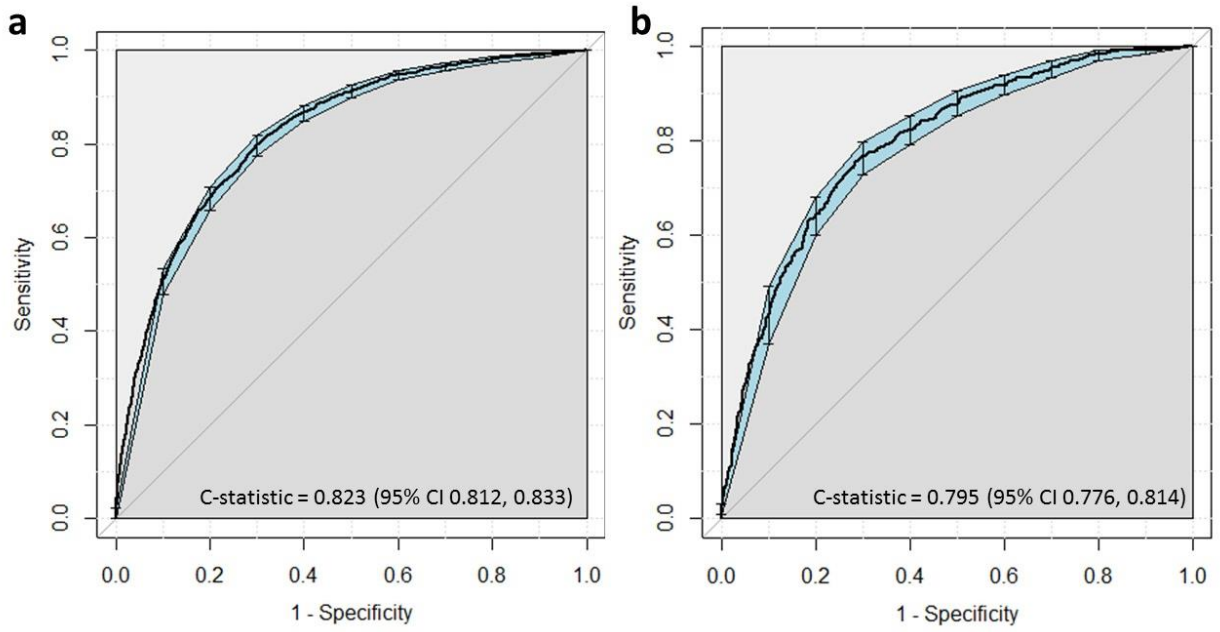

Figure 1. Area under the receiver-operator characteristic curve (AUC ROC) of the predictive model for critical illness among patients hospitalized with COVID-19. (a). AUC ROC in the development cohort, n=7850 patients from hospitals with equal or more than 300 beds. (b). AUC ROC in the validation cohort, n=2583 patients from hospitals with less than 300 beds. 95% coefficient intervals (CI) computed with 1000 bootstrap replicates.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

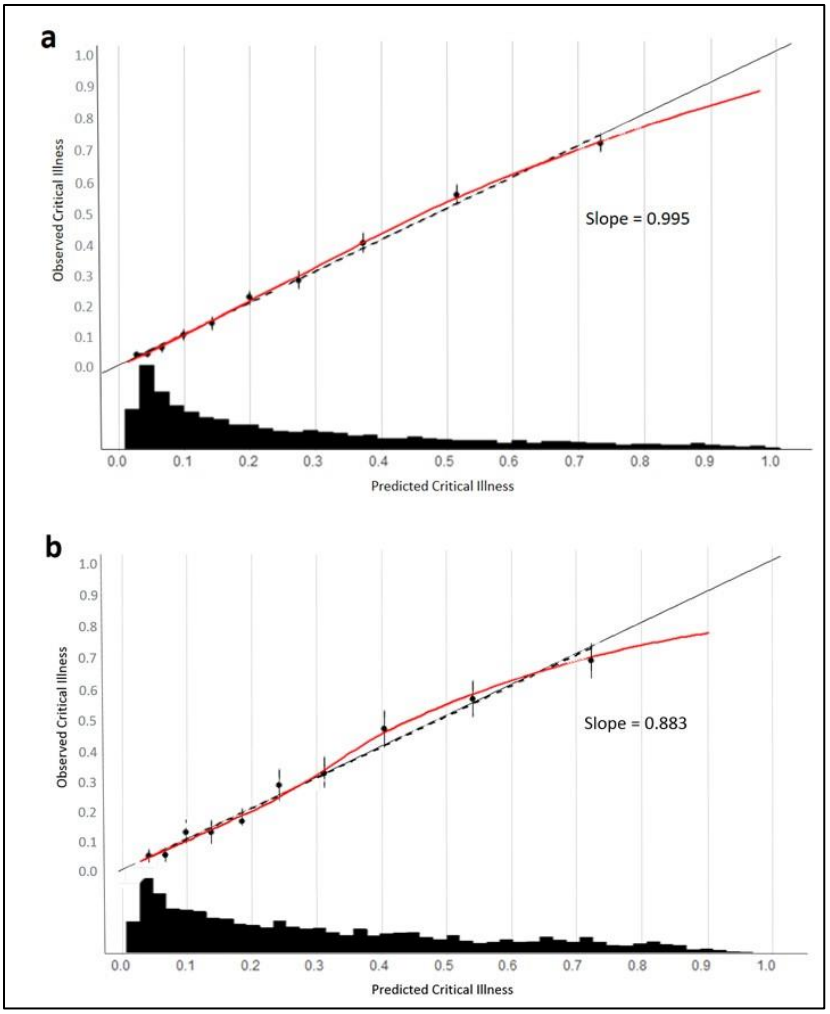

Figure 2. Calibration curves of the model predicting COVID-19 critical illness. (a). AUC ROC in the development cohort, n=7850 patients from hospitals with equal or more than 300 beds. (b). AUC ROC in the validation cohort, n=2583 patients from hospitals with less than 300 beds. Upper, X-axis representing model predictions and y-axis observed critical illness rates. Circles representing deciles of risk according to the model predictions, plotted against the observed critical illness rate. Linear and local (LOESS) regression were represented to visualize the agreement between observed and predicted values. Below, histogram of predicted critical illness across the range of risk predictions.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20237966; this version posted November 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

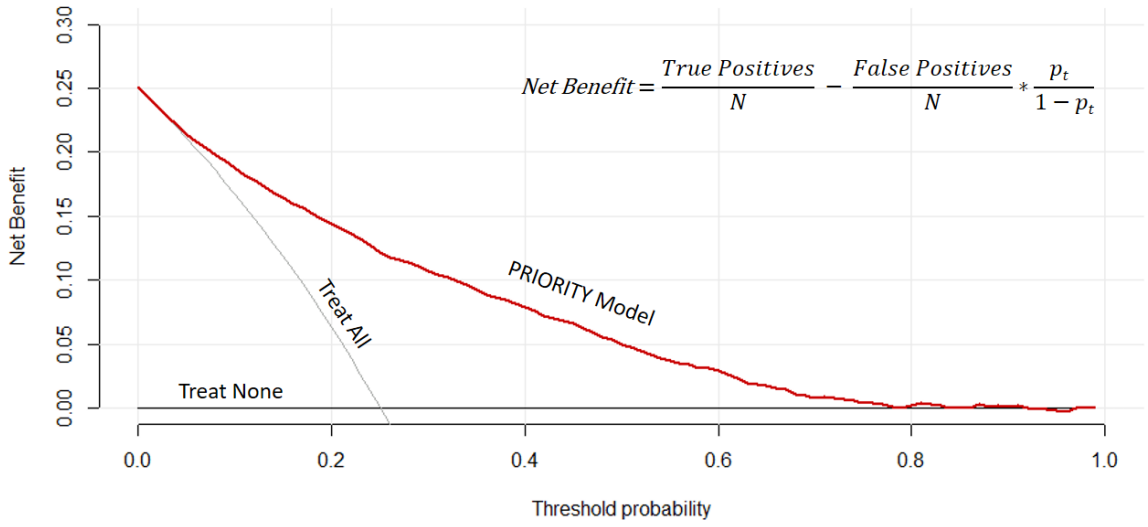

Figure 3. Decision curves of the predictive model for severe COVID-19. The x-axis represents threshold probabilities and the y-axis the net benefit.