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 García 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], Alexa-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 Asiani 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 Marañón 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, Moisés Broggi Hospital, Sant Joan Despi (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.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
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 \( \geq 300 \) beds, and validation used those from hospitals with <300 beds.

PARTICIPANTS

Adults (age \( \geq 18 \) 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 \( \text{SpO}_2 \leq 93\% \) 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%.
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.
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. 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. Notably, Spain has been one of the countries with the highest number of patients with COVID-19. 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 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, it has narrowed down the prediction to the mortality outcome only and it lacks wider generalizability. A critical appraisal of the COVID-19 models has shown poor reporting and high risk of bias. Recently published well-developed models 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. 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...
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. 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. 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 in a lower complexity level setting compared to the development setting. The study was reported following the TRIPOD (Transparent Reporting of a
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
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 (SpO₂) ≤ 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
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. 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. The least absolute shrinkage and selection operator (LASSO) method was used to identify a parsimonious set of potential predictors of critical illness. We selected the regulation penalty parameter ($\lambda$) 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.
**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. 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. 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) 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 (pₜ), 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 (pₜ/(1-pₜ)). We represented the NB vs. pₜ 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: \((\text{NB of the model} – \text{NB of treat all})/(\text{harm-to-benefit ratio}) \times 100\).

The choice of threshold probability pₜ 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
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:
Probability (%) = 100 \ / \ (1 + \text{Exp}(-z)), \text{where } z = -4.710 + 2.756 \cdot ((\text{Age}/100)^2) + 0.331 \cdot \text{[Cardiovascular disease]} + 0.595 \cdot \text{[CKD]} + 0.508 \cdot \text{[Dyspnea]} + 0.860 \cdot (1/\text{(SBP/100)}^2) + 0.905 \cdot \text{[Tachypnea]} + 1.204 \cdot (\text{SpO}_2 \leq 93\% \text{ or oxygen requirement}) + 0.740 \cdot \text{[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
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.
Strengths and weakness of the study

This study has several methodological strengths maximizing internal and external validity. 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. It was also one of the largest cohorts of all previous models published to date. Our model excluded readmissions, a feature that focuses 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. We complied with the recommendation made regarding avoidance of data-driven predictor selection. 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. 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
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\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.
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
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
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**


Nicolás Alcalá Rivera, Anxela Crestelo Vieitez, Esther del Corral Beamonte, Jesús Díez Manglano, Isabel Fiteni Mera, María del Mar García Andreu, Martín Gerico Aseguinolaza, Claudia Josa Laorden, Raul
Santiago López Gil, María 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.  
**Hospital Quironsalud A Coruña. A Coruña.** Hector Meijide Miguez.  
REFERENCES:


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

<table>
<thead>
<tr>
<th>Characteristics of the population</th>
<th>Development cohort</th>
<th>Validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of patients (%) or mean ± SD</td>
<td>Total No (%)</td>
</tr>
<tr>
<td>Critical Illness</td>
<td>1967 (25.1%)</td>
<td>7850 (100%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>65.8 ± 16.4</td>
<td>7816 (99.6%)</td>
</tr>
<tr>
<td>Male</td>
<td>4483 (57.2%)</td>
<td>7834 (99.8%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>6836 (89.1%)</td>
<td>7677 (98.8%)</td>
</tr>
<tr>
<td>Latino</td>
<td>693 (9.0%)</td>
<td>193 (7.5%)</td>
</tr>
<tr>
<td>Other</td>
<td>148 (1.9%)</td>
<td>39 (1.5%)</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>5270 (70.9%)</td>
<td>7433 (94.7%)</td>
</tr>
<tr>
<td>Former smoker</td>
<td>1764 (23.7%)</td>
<td>718 (29.0%)</td>
</tr>
<tr>
<td>Active smoker</td>
<td>399 (5.4%)</td>
<td>139 (5.3%)</td>
</tr>
<tr>
<td>Medical history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity</td>
<td>1665 (23.7%)</td>
<td>7012 (89.3%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>3803 (48.6%)</td>
<td>7833 (99.8%)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>1440 (18.4%)</td>
<td>7820 (99.6%)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>1974 (25.3%)</td>
<td>7800 (99.4%)</td>
</tr>
<tr>
<td>Pulmonary diseases</td>
<td>1625 (20.9%)</td>
<td>7776 (99.1%)</td>
</tr>
<tr>
<td>Severe chronic kidney disease</td>
<td>488 (6.2%)</td>
<td>7825 (99.7%)</td>
</tr>
<tr>
<td>Malignancy</td>
<td>793 (10.2%)</td>
<td>7803 (99.4%)</td>
</tr>
<tr>
<td>Immunocompromised status</td>
<td>650 (8.6%)</td>
<td>7549 (96.2%)</td>
</tr>
<tr>
<td>Dependency (moderate/severe)</td>
<td>1129 (14.7%)</td>
<td>7701 (98.1%)</td>
</tr>
<tr>
<td>Symptoms at admission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fever</td>
<td>5138 (67.0%)</td>
<td>7663 (97.6%)</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>4427 (56.7%)</td>
<td>7805 (99.4%)</td>
</tr>
<tr>
<td>Clinical signs and physical exploration at admission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>129.0 ± 21.5</td>
<td>7430 (94.6%)</td>
</tr>
<tr>
<td>HR (beats/minute)</td>
<td>88.6 ± 17.4</td>
<td>7500 (95.5%)</td>
</tr>
<tr>
<td>Tachypnea (&gt; 20 breaths/min)</td>
<td>2271 (29.9%)</td>
<td>7604 (96.9%)</td>
</tr>
<tr>
<td>SpO2 ≤ 93% or O2 requirement</td>
<td>4152 (52.9%)</td>
<td>7842 (99.9%)</td>
</tr>
<tr>
<td>Pulmonary rales</td>
<td>4630 (60.7%)</td>
<td>7626 (97.1%)</td>
</tr>
<tr>
<td>Confusion</td>
<td>849 (11.0%)</td>
<td>7736 (98.5%)</td>
</tr>
</tbody>
</table>

SD: standard deviation.
Obesity is defined as Medical history or BMI ≥30kg/m².
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.

SpO₂: Peripheral oxygen saturation

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Odds ratio</th>
<th>Bootstrap 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Age/100)² [years]</td>
<td>15.738</td>
<td>10.957, 22.714</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>1.392</td>
<td>1.214, 1.590</td>
</tr>
<tr>
<td>Severe CKD severe</td>
<td>1.813</td>
<td>1.391, 2.319</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>1.662</td>
<td>1.456, 1.908</td>
</tr>
<tr>
<td>1/ (SBP/100)² [mmHg]</td>
<td>2.363</td>
<td>1.885, 3.010</td>
</tr>
<tr>
<td>Tachypnea (&gt;20 breaths/min)</td>
<td>2.473</td>
<td>2.206, 2.841</td>
</tr>
<tr>
<td>SpO₂ ≤ 93% or O₂ requirement</td>
<td>3.335</td>
<td>2.889, 3.838</td>
</tr>
<tr>
<td>Confusion</td>
<td>2.097</td>
<td>1.761, 2.542</td>
</tr>
<tr>
<td>Dependency (Moderate or severe)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

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.

SpO₂: Peripheral oxygen saturation.
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.

<table>
<thead>
<tr>
<th>Probability threshold (p, %)</th>
<th>Net benefit</th>
<th>Advantage of using the model</th>
<th>Reduction in avoidable hospitalization referrals per 100 patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treat all</td>
<td>Prediction Model</td>
<td>Difference in the net benefit</td>
</tr>
<tr>
<td>1</td>
<td>0.243</td>
<td>0.243</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.235</td>
<td>0.235</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.227</td>
<td>0.227</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0.219</td>
<td>0.220</td>
<td>0.001</td>
</tr>
<tr>
<td>5</td>
<td>0.211</td>
<td>0.214</td>
<td>0.003</td>
</tr>
<tr>
<td>10</td>
<td>0.167</td>
<td>0.188</td>
<td>0.021</td>
</tr>
<tr>
<td>15</td>
<td>0.118</td>
<td>0.164</td>
<td>0.046</td>
</tr>
<tr>
<td>20</td>
<td>0.063</td>
<td>0.144</td>
<td>0.081</td>
</tr>
<tr>
<td>25</td>
<td>0.001</td>
<td>0.122</td>
<td>0.121</td>
</tr>
<tr>
<td>50</td>
<td>-0.499</td>
<td>0.005</td>
<td>0.504</td>
</tr>
<tr>
<td>75</td>
<td>-1.998</td>
<td>0.004</td>
<td>2.002</td>
</tr>
<tr>
<td>79</td>
<td>-2.569</td>
<td>0</td>
<td>2.569</td>
</tr>
<tr>
<td>99</td>
<td>-73.943</td>
<td>0</td>
<td>73.943</td>
</tr>
</tbody>
</table>

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.
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.
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.
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.