

# An algorithm to predict the need for invasive mechanical ventilation in hospitalized COVID-19 patients: the experience in Sao Paulo

Eduardo Atsushi Osawa<sup>1,2</sup>, Alexandre Toledo Maciel<sup>1,2</sup>

<sup>1</sup>Research Department, Imed Group, Sao Paulo; <sup>2</sup>Adult Intensive Care Unit, Sao Camilo Hospital-Unidade Pompeia, Sao Paulo, Brazil

**Background:** We aimed to characterize patients hospitalized for coronavirus disease 2019 (COVID-19) and identify predictors of invasive mechanical ventilation (IMV).

**Methods:** We performed a retrospective cohort study in patients with COVID-19 admitted to a private network in Sao Paulo, Brazil from March to October 2020. Patients were compared in three subgroups: non-intensive care unit (ICU) admission (group A), ICU admission without receiving IMV (group B) and IMV requirement (group C). We developed logistic regression algorithm to identify predictors of IMV.

**Results:** We analyzed 1,650 patients, the median age was 53 years (42–65) and 986 patients (59.8%) were male. The median duration from symptom onset to hospital admission was 7 days (5–9) and the main comorbidities were hypertension (42.4%), diabetes (24.2%) and obesity (15.8%). We found differences among subgroups in laboratory values obtained at hospital admission. The predictors of IMV (odds ratio and 95% confidence interval [CI]) were male (1.81 [1.11–2.94],  $P=0.018$ ), age (1.03 [1.02–1.05],  $P<0.001$ ), obesity (2.56 [1.57–4.15],  $P<0.001$ ), duration from symptom onset to admission (0.91 [0.85–0.98],  $P=0.011$ ), arterial oxygen saturation (0.95 [0.92–0.99],  $P=0.012$ ), C-reactive protein (1.005 [1.002–1.008],  $P<0.001$ ), neutrophil-to-lymphocyte ratio (1.046 [1.005–1.089],  $P=0.029$ ) and lactate dehydrogenase (1.005 [1.003–1.007],  $P<0.001$ ). The area under the curve values were 0.860 (95% CI, 0.829–0.892) in the development cohort and 0.801 (95% CI, 0.733–0.870) in the validation cohort.

**Conclusions:** Patients had distinct clinical and laboratory parameters early in hospital admission. Our prediction model may enable focused care in patients at high risk of IMV.

**Key Words:** algorithm; coronavirus; COVID-19; intensive care units; mechanical ventilation; SARS-CoV-2

## Original Article

Received: March 1, 2022

Revised: June 22, 2022

Accepted: July 21, 2022

### Corresponding author

Eduardo Atsushi Osawa  
Research Department, Imed Group,  
Av. Angélica, 2.530-9º andar, Sao  
Paulo 01228-200, Brazil  
Tel: +55-119-9737-6100  
E-mail: [eduardo.osawa@imedgroup.com.br](mailto:eduardo.osawa@imedgroup.com.br)

## INTRODUCTION

The Brazilian health system has been severely affected by the coronavirus disease 2019 (COVID-19) pandemic. As of February 2022, more than 26 million confirmed cases of COVID-19 infection and more than 630,000 deaths have been reported. In the city of Sao Paulo, over 1.0 million cases and 40,000 deaths have been reported.

Copyright © 2022 The Korean Society of Critical Care Medicine

This is an Open Access article distributed under the terms of Creative Attributions Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

The ongoing pandemic has overwhelmed health systems across the country and urged a better allocation of human and technical resources. In this context, simple approaches to understand the disease pattern and identify predictors of worse outcomes are needed. Early investigations reported that laboratory abnormalities have been associated with a greater mortality, including elevated ferritin, C-reactive protein (CRP), D-dimer levels and low lymphocyte count [1-4]. A recent meta-analysis demonstrated the importance of alterations in several biomarkers as a means of predicting mortality [5].

Despite the relevance of such derangements upon hospital admission as predictors of poor prognosis, there is scarce data addressing the impact of the trends of laboratory biomarkers on patient outcomes. Moreover, some of the predictive models available in the literature lack robustness, thereby limiting their clinical applicability [6]. Reliable information generated by such models would assist clinicians to stratify patient's risk upon hospital admission and identify those most likely to require invasive therapies. Thus, we conducted a retrospective cohort to investigate patients hospitalized for COVID-19 in a private system in Sao Paulo, Brazil. Specifically, we described the population demographics and comorbidities, trends of laboratory tests and developed a model to predict the risk of mechanical ventilation.

## MATERIALS AND METHODS

### Study Design and Participants

This retrospective cohort study was conducted in a private network in Sao Paulo, Brazil (Hospital Sao Camilo) comprising three tertiary care hospitals (Pompeia, Santana, and Ipiranga Units). The study was performed according to the Helsinki Declaration, was approved by the local Ethics Committee of São Camilo Hospital and written informed consent was waived due to the use of retrospective deidentified data. We identified consecutive adult patients (aged  $\geq 18$  years) admitted to one of the three hospitals from March 5, 2020, to October 26, 2020 who were diagnosed with laboratory-confirmed COVID-19. Laboratory testing for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection was performed using reverse transcription-polymerase chain reaction (RT-PCR) of nasopharyngeal or oropharyngeal swab sample. We excluded patients who did not have at least two sets of blood tests obtained during hospital stay, those whose symptom onset commenced more than 14 days prior to hospital admission and patients admitted for

### KEY MESSAGES

- In our cohort hospitalized for coronavirus disease 2019 (COVID-19), we observed a good accuracy for mechanical ventilation prediction with some laboratory tests assessed at hospital admission including C-reactive protein and lactate dehydrogenase.
- Our algorithm identified in the multivariate analysis eight variables as significant predictors for mechanical ventilation: male sex, age, obesity, duration from symptom onset to hospital admission, oxygen saturation, C-reactive protein, neutrophil-to-lymphocyte ratio and lactate dehydrogenase all assessed at admission.
- Despite the existence of many predictive models in COVID-19, our algorithm adds novelty to the literature by including a simple but often neglected variable: duration from symptom onset to hospital admission.

reasons other than respiratory infection.

Our institutional protocol underwent temporal modifications to incorporate evidence emerging from clinical trials. After the publication of Randomised Evaluation of COVID-19 Therapy (RECOVERY) study [7], corticosteroid was prescribed for all hospitalized patients requiring oxygen therapy. Hydroxychloroquine was administered as a compassionate therapy early in the pandemic for hospitalized patients and stopped being used after the publication of the Coalition Covid-19 Brazil I study [8]. Tocilizumab was administered for hypoxemic patients fulfilling the criteria for cytokine storm (elevated levels of CRP, ferritin, lactate dehydrogenase [LDH] and interleukin 6) in the absence of bacterial infection. Furthermore, oxygen was titrated to achieve an oxygen saturation between 88% and 92%. During the first month of the pandemic, neither non-invasive ventilation (NIV) nor high-flow nasal cannula (HFNC) was used due to guidelines reporting the risk of aerosol dispersion. Except for this period, the modalities of oxygen delivery comprised nasal cannula, non-rebreather mask, Venturi mask, NIV, HFNC, invasive mechanical ventilation (IMV) and venovenous extracorporeal membrane oxygenation.

### Data Collection

For all patients, we collected data on baseline characteristics including demographics, comorbidities, peripheral oxygen saturation upon hospital admission, date of symptom onset, outcome (discharge, death or transfer) and dates of admission and discharge. In patients who required intensive care unit (ICU) admission, we also examined Simplified Acute Physiolog-

ogy Score (SAPS) III, intubation requirement, and dates of ICU admission and discharge.

### Sequential Measurements of Laboratory Data

Laboratory data was extracted from electronic medical records. To minimize the influence of other issues complicating the initial COVID-19 infection in patients with a prolonged hospital stay, we limited the laboratory tests data collection timeframe to the first 14 days within hospital admission. Sequential serum laboratory values of the following parameters were obtained: CRP, LDH, D-dimer and creatinine levels; lymphocyte, neutrophil and platelet counts. We calculated two parameters: the neutrophil-to-lymphocyte ratio (NLR) was obtained by dividing the absolute neutrophil count by the absolute lymphocyte count, and lymphocyte-to-C-reactive protein ratio (LCR) was calculated by dividing lymphocyte count by the CRP level. Blood samples were usually obtained once daily. In cases of more than one test on a single day, the worst value was considered. In patients who required renal replacement therapy, we collected serum creatinine levels until the last measurement before the therapy commenced. Also, in patients who received tocilizumab, we obtained CRP levels up to the day when this therapy was administered.

To analyze patients across different severity groups, we split our study population into three subgroups: those who were not admitted to ICU (group A), those who were admitted to ICU but did not require IMV (group B), and those who required IMV (group C).

### Statistical Analysis

Continuous variables were presented as means and standard deviations or median (interquartile range), as appropriate. Normal distribution of continuous variables was checked with the Kolmogorov-Smirnov test. Categorical variables were summarized as counts and percentages. No imputation was made for missing data. Chi-square test or Fisher exact test was used to compare categorical variables as appropriate. Kruskal-Wallis or Mann-Whitney rank sum test was used to compare non-parametric continuous variables and t-student test was used to compare parametric continuous variables.

To evaluate the efficacy of each serum biomarker in the prediction of mechanical ventilation, we calculated the area under the receiver operator characteristics (ROC) curve (AUC) for measurements obtained at hospital admission. For D-dimer and LDH levels, we observed a high frequency of missing values on day 1 attributed to logistic issues upon hospital ad-

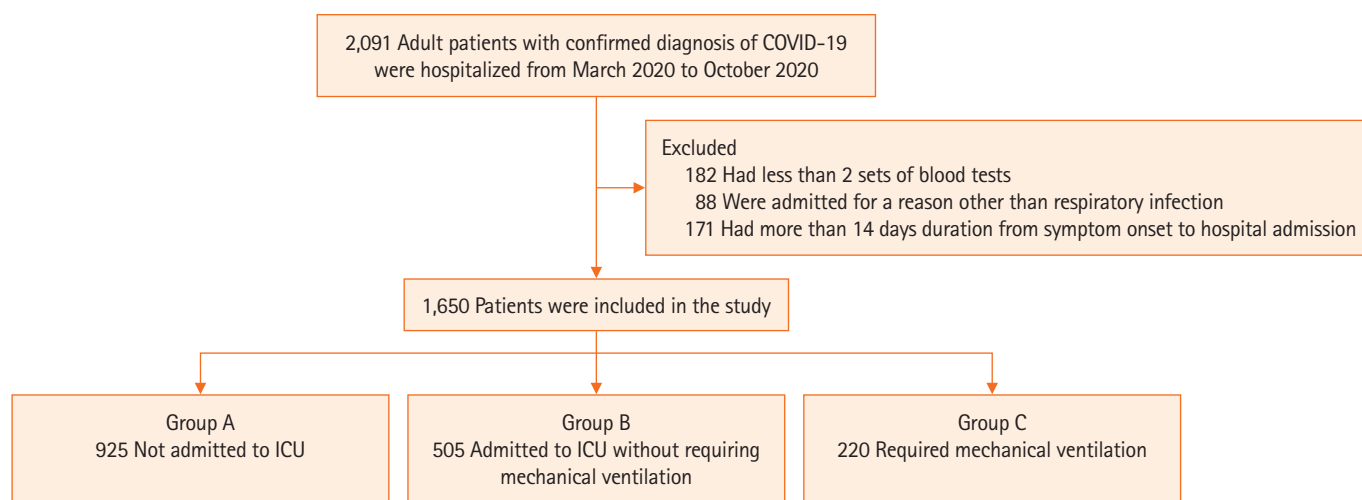
mission. Therefore, we considered the worst value obtained from either day 1 or day 2.

A bidirectional stepwise logistic regression analysis was undertaken to determine the predictive factors for mechanical ventilation. We used 80% of our study population to develop the model and applied it to the remaining 20% to validate its performance. Variables that were significantly associated with the outcome in the univariate analysis were included in the multivariate model according to the following criteria: clinical relevance, lack of colinearity and missing data <15% of cases. Overall goodness of fit was verified by Akaike information criterion and discrimination of the model was evaluated by ROC curve of predicted probability. All statistical tests were two-tailed, and a P-value <0.05 was considered statistically significant. Statistical tests were performed using R version 4.0.2 (R Foundation, Vienna, Austria).

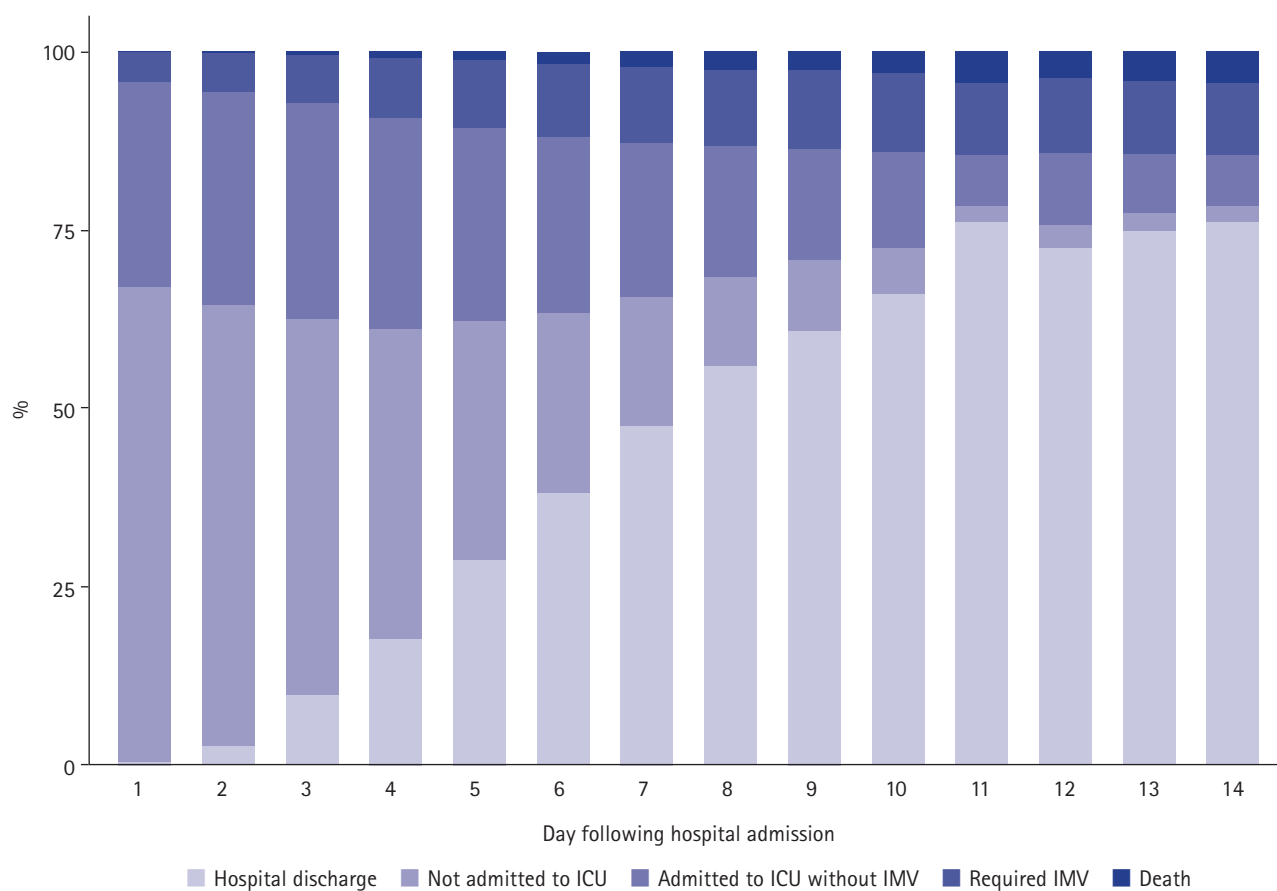
## RESULTS

A total of 1,650 patients with a SARS-CoV-2 RT-PCR assay confirmed COVID-19 diagnosis were included in our study ([Figure 1](#)). The distribution of the outcome status along the first 14 days of hospitalization is illustrated in [Figure 2](#). [Table 1](#) demonstrates the demographic and clinical characteristics of the study patients. Overall, 986 patients (59.8%) were male and the median age was 53 years (range, 42–65). The median duration from symptom onset to hospital admission was 7 days (5–9). The most predominant comorbidities were hypertension (42.4%), diabetes (24.2%), hyperlipidemia (16.2%), and obesity (15.8%). A greater proportion of patients undergoing ICU admission were male in comparison to the population not admitted to ICU. Also, patients who became critically ill were older and had a lower oxygen saturation at hospital admission ([Table 1](#)). Acute respiratory failure was the reason for ICU admission in the majority of patients.

Measurements of laboratory biomarkers obtained at hospital admission are presented in [Table 2](#). We found that such values differed among all subgroups. Furthermore, a two-by-two comparison between groups has also yielded statistically significant difference. Sequential measurements of the main laboratory biomarkers are represented in [Figure 3](#). The median (interquartile range) hospital length of stay was 7 days (5–9) in group A, 10 days (8–15) in group B and 22 days (16–31) in group C. The number of patients who died during hospitalization was the following: 2 patients (<1%) in group A, 25 patients (4.9%) in group B and 102 patients (46.4%) in group C.



**Figure 1.** Flowchart of the study enrollment. COVID-19: coronavirus disease 2019; ICU: intensive care unit.



**Figure 2.** Daily distribution of study patients according to the outcome status. ICU: intensive care unit; IMV: invasive mechanical ventilation.

**Table 1.** Baseline clinical characteristics of study patients

Variable	All patients (n=1,650)	Group A (n=925)	Group B (n=505)	Group C (n=220)	P-value			
					General	A vs. B	A vs. C	B vs. C
Sex					0.002	0.029	0.001	0.152
Female	664 (40.2)	404 (43.7)	190 (37.6)	70 (31.8)				
Male	986 (59.8)	521 (56.3)	315 (62.4)	150 (68.2)				
Age (yr)	53 (42–65)	48 (39–60)	56 (44–69)	63 (51–74)	<0.001	<0.001	<0.001	<0.001
Comorbidity								
Hypertension	699 (42.4)	304 (32.9)	256 (50.7)	139 (63.2)	<0.001	<0.001	<0.001	0.002
Diabetes	399 (24.2)	172 (18.6)	140 (27.7)	87 (39.5)	<0.001	<0.001	<0.001	0.002
COPD	46 (2.8)	8 (0.9)	23 (4.6)	15 (6.8)	<0.001	<0.001	<0.001	0.209
Smoking	27 (1.6)	14 (1.5)	10 (2.0)	3 (1.4)	0.788			
Heart failure	55 (3.3)	10 (1.1)	30 (5.9)	15 (6.8)	<0.001	<0.001	<0.001	0.620
Cirrhosis	5 (0.3)	2 (0.2)	2 (0.4)	1 (0.5)	0.850			
Previous stroke	35 (2.1)	9 (1.0)	20 (4.0)	6 (2.7)	<0.001	<0.001	0.050	0.517
Obesity	261 (15.8)	83 (9.0)	114 (22.6)	64 (29.1)	<0.001	<0.001	<0.001	0.074
Previous bariatric surgery	19 (1.2)	13 (1.4)	4 (0.8)	2 (0.9)	0.517			
Hyperlipidemia	267 (16.2)	115 (12.4)	105 (20.8)	47 (21.4)	<0.001	<0.001	0.001	0.921
Dementia	59 (3.6)	16 (1.7)	38 (7.5)	5 (2.3)	<0.001	<0.001	0.578	0.006
Autoimmune disorder	3 (0.2)	2 (0.2)	0 (0.0)	1 (0.5)	0.571			
HIV	13 (0.8)	8 (0.9)	1 (0.2)	4 (1.8)	0.066			
Solid neoplasm	15 (0.9)	4 (0.4)	6 (1.2)	5 (2.3)	0.028	0.180	0.016	0.323
Hematologic neoplasm	8 (0.5)	4 (0.4)	1 (0.2)	3 (1.4)	0.085			
Asthma	75 (4.5)	39 (4.2)	24 (4.8)	12 (5.5)	0.705			
Non-dialysis CKD	51 (3.1)	13 (1.4)	24 (4.8)	14 (6.4)	<0.001	<0.001	<0.001	0.369
CKD on dialysis	10 (0.6)	1 (0.1)	3 (0.6)	6 (2.7)	<0.001	0.129	<0.001	0.026
Previous solid organ transplant	6 (0.4)	1 (0.1)	4 (0.8)	1 (0.5)	0.101			
Coronary artery disease	89 (5.4)	27 (2.9)	38 (7.5)	24 (10.9)	<0.001	<0.001	<0.001	0.149
Duration from symptom onset to hospital admission	7 (5–9)	7 (5–9)	6 (4–8)	6 (4–8)	0.001	<0.001	<0.001	0.784
SpO <sub>2</sub> at hospital presentation	96 (93–97)	96 (95–98)	95 (92–96)	91 (86–95)	0.001	<0.001	<0.001	<0.001
SAPS III			43 (39–50)	49 (44–55)				<0.001

Values are presented as number (%) or median (interquartile range). Group A: patients who were not admitted to intensive care unit (ICU); Group B: patients who were admitted to ICU but did not require mechanical ventilation; Group C: patients who required mechanical ventilation.

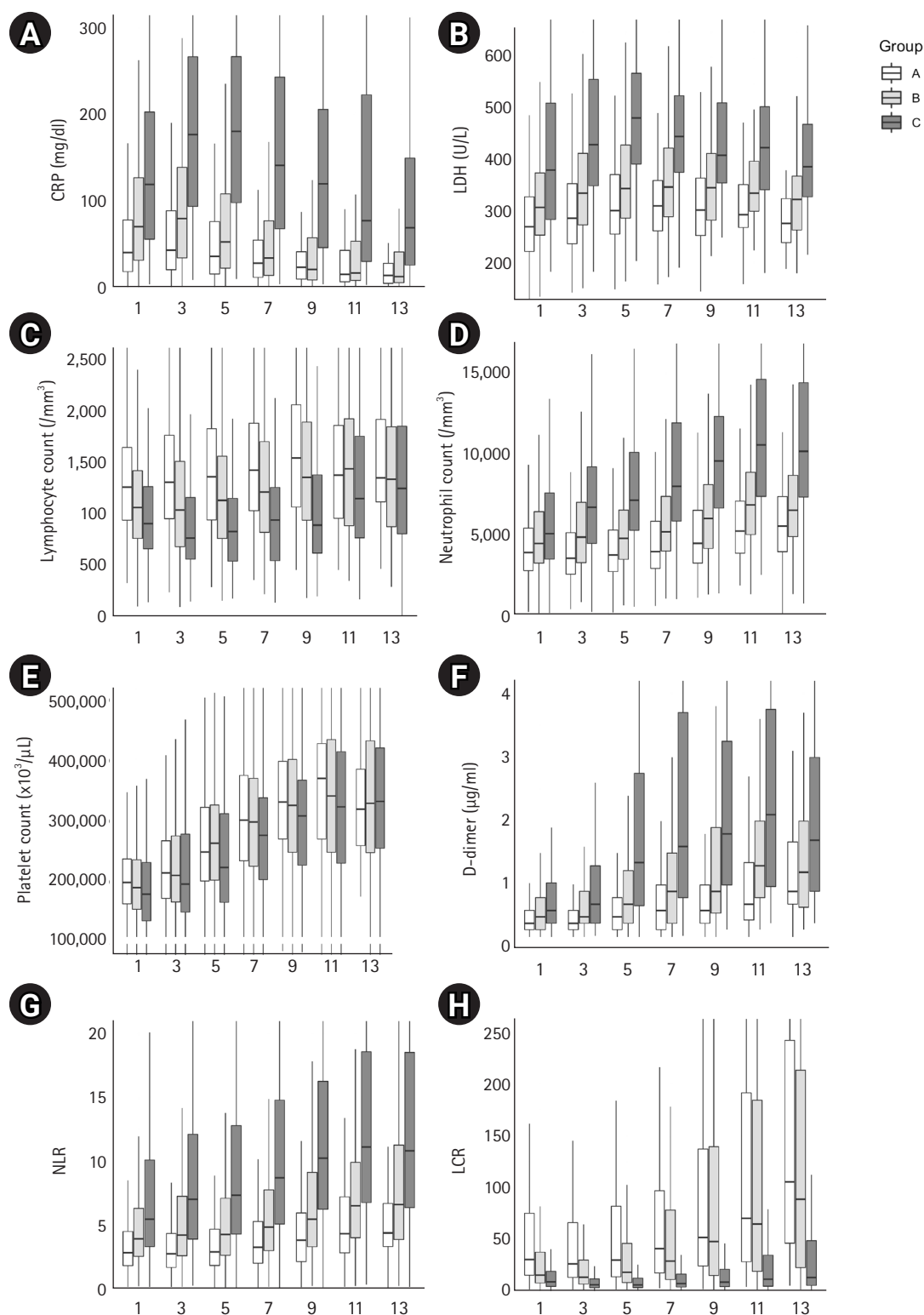
COPD: chronic obstructive pulmonary disease; HIV: human immunodeficiency virus; CKD: chronic kidney disease; SpO<sub>2</sub>: arterial oxygen saturation; SAPS: Simplified Acute Physiology Score.

**Table 2.** Measurements of laboratory biomarkers at hospital admission

Variable	All patients (n=1,650)	Group A (n=925)	Group B (n=505)	Group C (n=220)	P-value			
					General	A vs. B	A vs. C	B vs. C
CRP (mg/dl)	55 (24–107)	40 (18–78)	70 (31–127)	119 (55–203)	<0.001	<0.001	<0.001	<0.001
LDH (U/L)	288 (234–360)	269 (221–328)	307 (253–375)	380 (284–512)	<0.001	<0.001	<0.001	<0.001
Serum creatinine (mg/dl)	1.0 (0.8–1.1)	0.9 (0.8–1.1)	1.0 (0.8–1.2)	1.0 (0.9–1.4)	<0.001	0.027	<0.001	0.001
Lymphocyte count (/mm <sup>3</sup> )	1,150 (819–1,541)	1,260 (936–1,649)	1,062 (758–1,421)	903 (657–1,265)	<0.001	<0.001	<0.001	<0.001
Neutrophil count (/mm <sup>3</sup> )	4,152 (2,936–5,923)	3,866 (2,757–5,383)	4,438 (3,224–6,402)	5,046 (3,467–7,555)	<0.001	<0.001	<0.001	0.007
Platelet count (x10 <sup>3</sup> /μL)	194 (158–237)	199 (164–238)	191 (155–237)	180 (136–232)	0.001	0.038	<0.001	0.047
D-dimer (μg/ml)	0.4 (0.3–0.7)	0.4 (0.3–0.6)	0.5 (0.3–0.8)	0.6 (0.4–1.0)	<0.001	<0.001	<0.001	0.003
LCR	21 (9–56)	30 (14–79)	15 (7–39)	8 (4–18)	<0.001	<0.001	<0.001	<0.001
NLR	4 (2–6)	3 (2–5)	4 (3–7)	6 (3–11)	<0.001	<0.001	<0.001	<0.001

Values are presented as median (interquartile range).

CRP: C-reactive protein; LDH: lactate dehydrogenase; LCR: lymphocyte-to-C-reactive protein ratio; NLR: neutrophil-to-lymphocyte ratio.



**Figure 3.** Sequential measurements of laboratory biomarkers over the course of hospital stay. (A-H) Boxplots of values of laboratory biomarkers obtained during hospital stay. (A) C-reactive protein (CRP), (B) lactate dehydrogenase (LDH), (C) lymphocyte count, (D) neutrophil count, (E) platelet count, (F) D-dimer, (G) neutrophil-to-lymphocyte ratio (NLR), (H) lymphocyte-to-C-reactive protein ratio (LCR). Group A: patients who were not admitted to intensive care unit (ICU); Group B: patients who were admitted to ICU but did not require mechanical ventilation; Group C: patients who required mechanical ventilation.

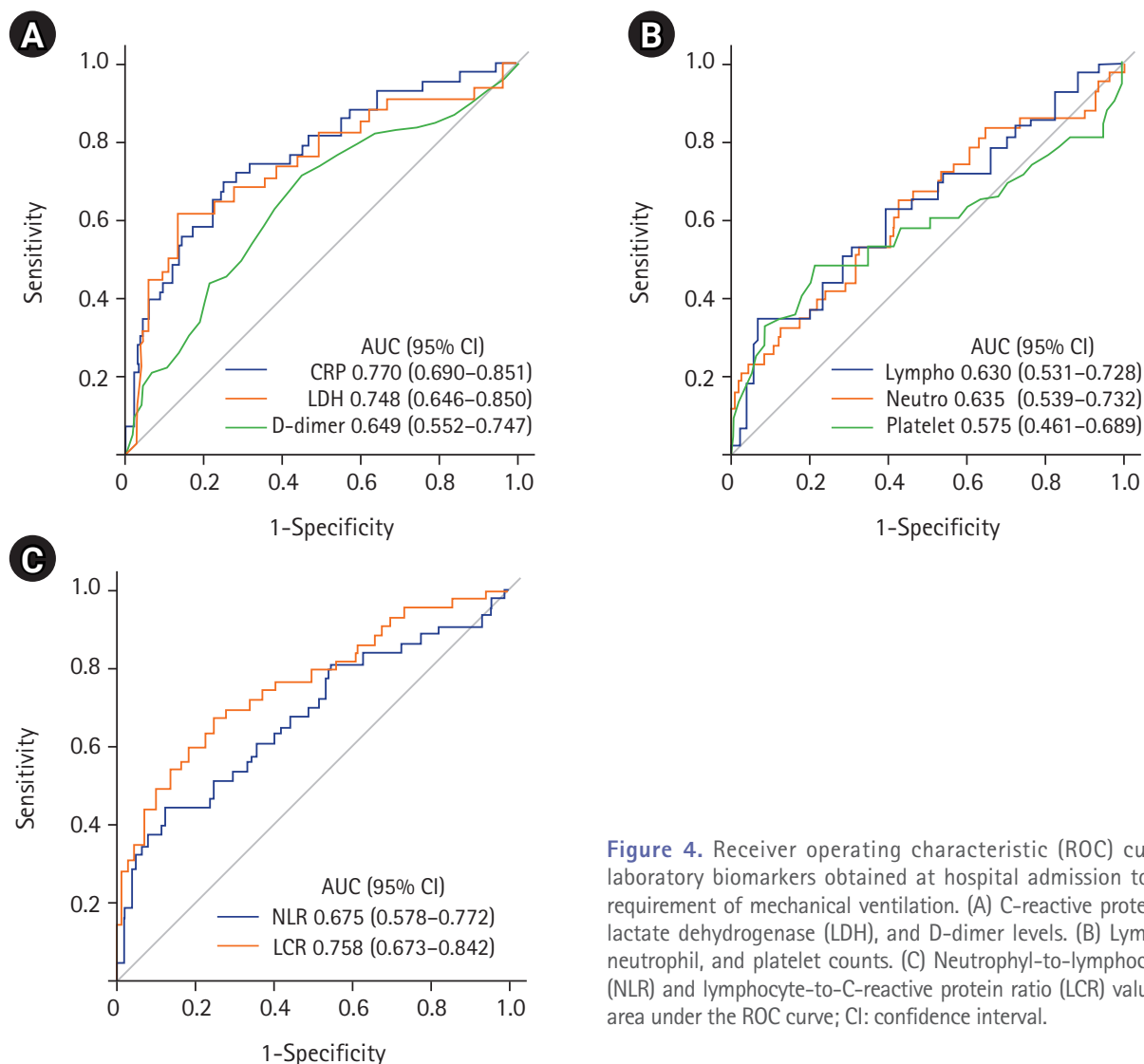


To assess the performance of each laboratory biomarker in the prediction of mechanical ventilation, we plotted ROC curves of laboratory biomarkers obtained at hospital admission (Figure 4). We observed a better ability to predict the outcome with CRP, LDH and LCR levels. The following variables fulfilled the criteria in the univariate analysis in order to be included in the multivariate analysis (Table 3): age, sex, hypertension, diabetes, chronic obstructive pulmonary disease, heart failure, obesity, hyperlipidemia, solid neoplasm, duration from symptom onset to hospital admission, coronary artery disease, oxygen saturation, and laboratory biomarkers (CRP, NLR, platelet count, creatinine, D-dimer and LDH). Of these, a total of eight variables were selected as significant predictors for the requirement of mechanical ventilation: male sex, age, obesity, duration from symptom onset to hospital

admission, oxygen saturation, CRP, NLR and LDH at hospital admission (Table 4). In the development cohort, the receiving operating characteristic curve was drawn with an AUC of 0.860 (95% confidence interval, 0.829–0.892) (Figure 5A). The performance of the model was internally validated and the AUC of the validation cohort was 0.801 (95% confidence interval, 0.733–0.870) (Figure 5B).

## DISCUSSION

We described the clinical characteristics of patients hospitalized for COVID-19 according to three subgroups (not admitted to ICU, admitted to ICU but not mechanically ventilated and admitted to ICU with requirement of mechanical ventilation). Also, we found that the main serum biomarkers differed



**Figure 4.** Receiver operating characteristic (ROC) curves for laboratory biomarkers obtained at hospital admission to predict requirement of mechanical ventilation. (A) C-reactive protein (CRP), lactate dehydrogenase (LDH), and D-dimer levels. (B) Lymphocyte, neutrophil, and platelet counts. (C) Neutrophyl-to-lymphocyte ratio (NLR) and lymphocyte-to-C-reactive protein ratio (LCR) values. AUC: area under the ROC curve; CI: confidence interval.

**Table 3.** Univariate logistic regression analysis for predicting the requirement of mechanical ventilation in 1,650 patients hospitalized for COVID-19 in Sao Paulo, Brazil

Variable	Total (n=1,650)	Did not receive IMV (n=1,430)	Received IMV (n=220)	P-value
Female	664 (40.2)	594 (41.5)	70 (31.8)	0.006
Age (yr)	52.5 (41.5–64.5)	51.1 (40.7–63.0)	62.9 (51.3–73.7)	<0.001
Hypertension	699 (42.4)	560 (39.2)	139 (63.2)	<0.001
Diabetes mellitus	399 (24.2)	312 (21.8)	87 (39.5)	<0.001
COPD	46 (2.8)	31 (2.2)	15 (6.8)	<0.001
Smoking	27 (1.6)	24 (1.7)	3 (1.4)	1.000
Heart failure	55 (3.3)	40 (2.8)	15 (6.8)	0.004
Cirrhosis	5 (0.3)	4 (0.3)	1 (0.5)	0.512
Stroke	35 (2.1)	29 (2.0)	6 (2.7)	0.454
Obesity	261 (15.8)	197 (13.8)	64 (29.1)	<0.001
Bariatric surgery	19 (1.2)	17 (1.2)	2 (0.9)	>0.990
Hyperlipidemia	267 (16.2)	220 (15.4)	47 (21.4)	0.030
Dementia	59 (3.6)	54 (3.8)	5 (2.3)	0.331
Autoimmune disease	3 (0.2)	2 (0.1)	1 (0.5)	0.349
HIV	13 (0.8)	9 (0.6)	4 (1.8)	0.083
Solid neoplasm	15 (0.9)	10 (0.7)	5 (2.3)	0.039
Hematological neoplasm	8 (0.5)	5 (0.3)	3 (1.4)	0.079
Asthma	75 (4.5)	63 (4.4)	12 (5.5)	0.486
ESKD not on dialysis	51 (3.1)	37 (2.6)	14 (6.4)	0.006
ESKD on dialysis	10 (0.6)	4 (0.3)	6 (2.7)	<0.001
Solid organ transplant	6 (0.4)	5 (0.3)	1 (0.5)	0.577
Coronary artery disease	89 (5.4)	65 (4.5)	24 (10.9)	<0.001
Duration from symptom onset to hospital admission	7.0 (5.0–9.0)	7.0 (5.0–9.0)	6.0 (4.0–8.0)	0.001
SpO <sub>2</sub> (%)	96.0 (93.0–97.0)	96.0 (94.0–97.0)	91.0 (86.0–95.0)	<0.001
C-reactive protein (mg/dl)	55 (24–107)	48 (21–94)	119 (55–203)	<0.001
Serum creatinine (mg/dl)	1.0 (0.8–1.1)	0.9 (0.8–1.1)	1.0 (0.9–1.4)	<0.001
Lymphocyte count (/mm <sup>3</sup> )	1,150 (819–1,541)	1,196 (853–1,590)	903 (657–1,265)	<0.001
Neutrophil count (/mm <sup>3</sup> )	4,152 (2,936–5,923)	4,054 (2,879–5,712)	5,046 (3,467–7,555)	<0.001
LCR	21 (9–56)	23 (11–62)	8 (4–18)	<0.001
NLR	4 (2–6)	4 (2–6)	6 (3–11)	<0.001
LDH (U/L)	300 (241–373)	292 (236–352)	404 (304–531)	<0.001
D-dimer (µg/ml)	0.5 (0.3–0.7)	0.4 (0.3–0.7)	0.6 (0.4–1.3)	0.001

All laboratory data were assessed at hospital admission.

COVID-19: coronavirus disease 2019; IMV: invasive mechanical ventilation; COPD: chronic obstructive pulmonary disease; HIV: human immunodeficiency virus; ESKD: end-stage kidney disease; SpO<sub>2</sub>: arterial oxygen saturation; LCR: lymphocyte-to-C-reactive protein ratio; NLR: neutrophil-to-lymphocyte ratio; LDH: lactate dehydrogenase.

among study subgroups. Furthermore, a good predictive performance was observed with CRP, LDH and LCR levels obtained at hospital admission. Finally, we built a logistic regression algorithm to identify predictors for the requirement of mechanical ventilation.

Our study supports the findings reported by other investigators with regards to the pattern of laboratory biomarkers derangements. Elevated CRP levels were associated with greater disease severity in observational cohorts [9–12]. Lymphopenia

was commonly present in an early pandemic report [13] and was found to be linked to worse outcomes [3,14–16]. Studies have also documented the prognostic importance of elevated LDH on disease severity [17–19]. Furthermore, other studies reported critically ill patients had higher neutrophil count and D-dimer concentrations in comparison with mild cases [2,13,15,19,20].

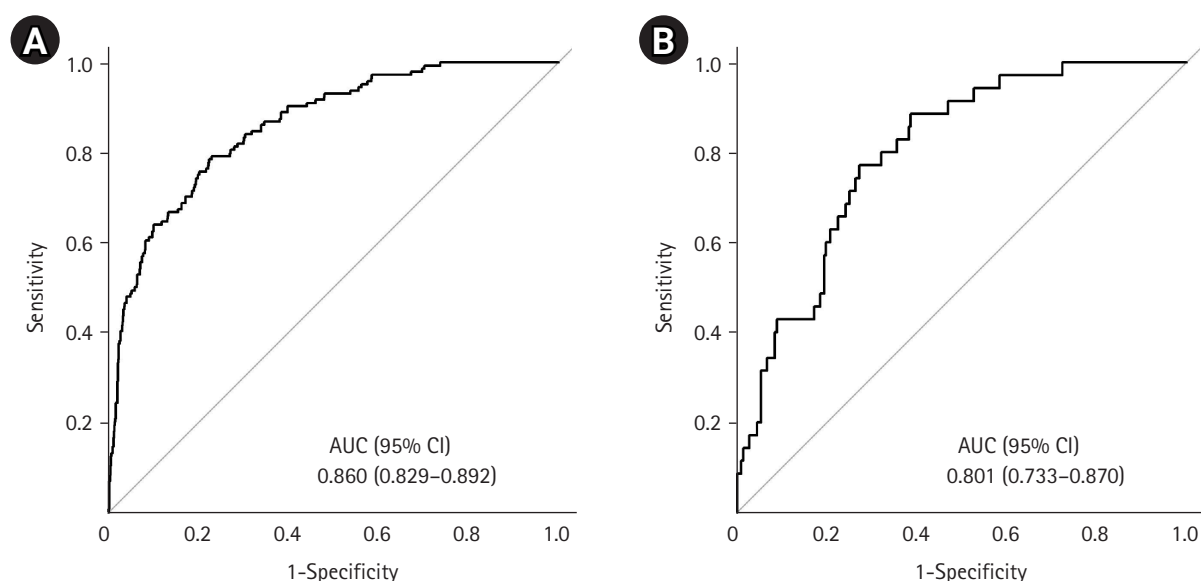
A few studies have documented the variation of laboratory biomarker measurements during the course of the disease. In



**Table 4.** Multivariable logistic regression model for predicting the requirement of mechanical ventilation in 1,650 patients hospitalized for COVID-19 in Sao Paulo, Brazil

Variable	Odds ratio (95% CI)	P-value
Sex (male vs. female)	1.810 (1.110–2.940)	0.018
Age (per year)	1.030 (1.020–1.050)	<0.001
Obesity (yes vs. no)	2.560 (1.570–4.150)	<0.001
Duration from symptom onset to hospital admission (per day)	0.910 (0.850–0.980)	0.011
SpO <sub>2</sub> (%)	0.950 (0.920–0.990)	0.012
C-reactive protein (mg/dl)	1.005 (1.002–1.008)	<0.001
NLR	1.046 (1.005–1.089)	0.029
LDH (U/L)	1.005 (1.003–1.007)	<0.001

COVID-19: coronavirus disease 2019; CI: confidence interval; SpO<sub>2</sub>: arterial oxygen saturation; NLR: neutrophil-to-lymphocyte ratio; LDH: lactate dehydrogenase.

**Figure 5.** Receiver operating characteristic (ROC) curves for the multivariable logistic regression model to identify predictors of mechanical ventilation. (A) Development cohort. (B) Validation cohort. AUC: area under the ROC curve; CI: confidence interval.

a cohort of 75 patients, Li et al. [21] observed that survivors had lower neutrophil, d-dimer and CRP levels and higher lymphocyte count at day 5 to 8 as compared to baseline. Other investigators found an early rise of CRP concentration during hospital stay, a finding that resembles the trajectory observed in our population [22,23]. In one study, patients who were critically ill, who died, or developed acute kidney injury had CRP levels significantly higher over time [23]. Furthermore, a retrospective cohort reported a similar trend towards greater levels of inflammatory biomarkers and lower lymphocyte count among patients who required a higher level of respiratory support or who died [24].

There are some similarities between predictors found in our algorithm and studies addressing pneumonia from non-COVID-19 cause [25,26]. Nevertheless, severity scoring sys-

tems validated for community acquired pneumonia, such as CURB-65 (6-point pneumonia severity score; 1 point for each of Confusion, Urea >7 mmol/L, respiratory rate >30/min, low systolic [<90 mm Hg] or diastolic [<60 mm Hg] blood pressure, age <65 years) [27] or PSI [28], address the mortality risk instead of the risk for IMV. Furthermore, they involve clinical findings not included in our algorithm such as altered mental status, hypotension, tachycardia and tachypnea, along with low oxygen saturation values and comorbidities. Also importantly, significant predictors in our model are not present in pneumonia severity scoring systems such as obesity, duration from symptom onset to hospital admission and serum levels of CRP and NLR at admission. Additional factors are implicated in the distinction between COVID-19 pneumonia from other causes of pneumonia. COVID-19 patients normally have

more prolonged duration of mechanical ventilation and use of sedation, more profound hypoxemia, extensive use of prone positioning and higher risk of pulmonary embolism [29-31]. Moreover, the pandemic-related overcrowding may have compromised the usual level of care due to resource shortage [32].

Different predictive models have been developed in patients with COVID-19 to assist medical decision making [6]. Overall, studies reported a good predictive performance assessed by ROC curves, nevertheless, a wide range of predictors were evaluated and the outcome definition was not consistent across algorithms. In the CALL (comorbidity, age, lymphocyte, and LDH) score model, a multivariate Cox regression assessed factors associated with disease progression and identified age and LDH as significant predictors, variables that were also relevant in our model [33]. In another study using machine learning algorithms to predict severe outcomes, one of which was requirement of IMV, it was found that age, serum creatinine and CPR were significant predictors [34]. Furthermore, in a retrospective multicenter study performed in Michigan, an XGBoost algorithm was built to predict mechanical ventilation and the most important predictors were age and oxygen saturation [35]. In a multicenter investigation performed in Wuhan, a Machine Learning approach was implemented to predict severe disease [36]. Age was the most important feature, but the model also included CRP and LDH measurements. Notably, age and male sex were relevant predictors in other large multicenter cohorts, two of them were performed in South Korea [37-40]. Moreover, a retrospective study also in South Korea showed greater risk of disease severity with increased age and higher likelihood for critically ill symptoms in male patients [41]. Finally, a multicenter Italian study aiming to predict severe respiratory failure observed that age, obesity, creatinine and LDH were significant variables [42]. Importantly, we found that lower duration from symptom onset to hospital admission was a predictor of IMV requirement, a variable not commonly addressed by other investigators. Such inverse relationship between time from symptom onset and outcome was documented in a retrospective study assessing mortality [43].

Our study adds to the literature by providing data on epidemiological aspects, laboratory biomarker pattern and prognosis among patients hospitalized for COVID-19. Furthermore, Sao Paulo was one of the regions severely affected by the pandemic waves and faced an increased demand for ICU beds and shortage of medical equipment and healthcare staff. In addition, patients who require mechanical ventilation may not present with clinical signs of respiratory failure upon hospital

admission, but frequently undergo rapid deterioration over the course of disease. Therefore, our findings may assist health systems to allocate resources and enable early intervention. On a patient level, those who are at high risk of receiving IMV may be considered to therapies demonstrated to shorten time to disease recovery. Moreover, our model seems to be feasible as the included parameters are widely available across different settings.

Our study has some limitations. Firstly, the retrospective nature of the study design may pose a greater risk of information bias as relevant clinical data may not have been well documented during the pandemic surge. However, our study subgroups were divided by unequivocal events (ICU admission and mechanical ventilation) and the model outcome was well defined (requirement of mechanical ventilation). Secondly, the timeframe of our data collection comprised different phases of the pandemic. As a result, patients may have been treated differently as initially clinical evidence was being accumulated and protocols were still under review. Nevertheless, the sample size was representative as we included a large number of consecutive patients admitted to our hospital network. Thirdly, we did not account for patients who had a do-not-intubate orders in place. As the evidence on disease prognosis were still being accrued in a such a complex and dynamic scenario, end of life decisions might have been affected and not consistently applied across the entire population. Finally, our model was developed and tested on patients from a single health system. Thus, future prospective studies implemented in different healthcare systems will provide external validity for our model.

Our study demonstrated that COVID-19 patients who developed worse outcomes had different epidemiological and laboratory biomarker characteristics accessible as early as at hospital admission. Also, our multivariable regression model showed good ability to predict the requirement of mechanical ventilation by including parameters widely available in clinical practice.

## CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

## FUNDING

None.

## ACKNOWLEDGMENTS

The authors express their appreciation for the statistical support provided by Ms. Luciana M Ishihara and for the data collection undertaken by the nursing team of Imed Group (Renata Melo, Selma Bezerra, Milena P Sanches and Mariana Pecoraro).

## ORCID

Eduardo Atsushi Osawa <https://orcid.org/0000-0002-1590-8658>  
 Alexandre Toledo Maciel <https://orcid.org/0000-0001-6043-0967>

## AUTHOR CONTRIBUTIONS

Conceptualization: ATM. Data curation: all authors. Formal analysis: all authors. Methodology: ATM. Project administration: ATM. Visualization: EAO. Writing—original draft: all authors. Writing—review & editing: all authors.

## REFERENCES

1. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, 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.
2. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020;323:1061-9.
3. Liu J, Li H, Luo M, Liu J, Wu L, Lin X, et al. Lymphopenia predicted illness severity and recovery in patients with COVID-19: a single-center, retrospective study. *PLoS One* 2020;15:e0241659.
4. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708-20.
5. Kiss S, Gede N, Hegyi P, Németh D, Földi M, Dembrowsky F, et al. Early changes in laboratory parameters are predictors of mortality and ICU admission in patients with COVID-19: a systematic review and meta-analysis. *Med Microbiol Immunol* 2021;210:33-47.
6. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328.
7. RECOVERY Collaborative Group, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in hospitalized patients with COVID-19. *N Engl J Med* 2021;384:693-704.
8. Cavalcanti AB, Zampieri FG, Rosa RG, Azevedo LC, Veiga VC, Avezum A, et al. Hydroxychloroquine with or without azithromycin in mild-to-moderate COVID-19. *N Engl J Med* 2020;383:2041-52.
9. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med* 2020;46:846-8.
10. Chen W, Zheng KI, Liu S, Yan Z, Xu C, Qiao Z. Plasma CRP level is positively associated with the severity of COVID-19. *Ann Clin Microbiol Antimicrob* 2020;19:18.
11. Herold T, Jurinovic V, Arnreich C, Lipworth BJ, Hellmuth JC, von Bergwelt-Baildon M, et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J Allergy Clin Immunol* 2020;146:128-36.
12. Luo X, Zhou W, Yan X, Guo T, Wang B, Xia H, et al. Prognostic value of C-reactive protein in patients with coronavirus 2019. *Clin Infect Dis* 2020;71:2174-9.
13. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497-506.
14. Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang YQ, et al. Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther* 2020;5:33.
15. Wagner J, DuPont A, Larson S, Cash B, Farooq A. Absolute lymphocyte count is a prognostic marker in COVID-19: a retrospective cohort review. *Int J Lab Hematol* 2020;42:761-5.
16. Ziadi A, Hachimi A, Admou B, Hazime R, Brahim I, Douirek F, et al. Lymphopenia in critically ill COVID-19 patients: a predictor factor of severity and mortality. *Int J Lab Hematol* 2021;43:e38-40.
17. Martha JW, Wibowo A, Pranata R. Prognostic value of elevated lactate dehydrogenase in patients with COVID-19: a systematic review and meta-analysis. *Postgrad Med J* 2022;98:422-7.
18. Szarpak L, Ruetzler K, Safiejko K, Hampel M, Pruc M, Kanczuga-Koda L, et al. Lactate dehydrogenase level as a COVID-19 severity marker. *Am J Emerg Med* 2021;45:638-9.
19. Chen XY, Huang MY, Xiao ZW, Yang S, Chen XQ. Lactate dehydrogenase elevations is associated with severity of COVID-19: a meta-analysis. *Crit Care* 2020;24:459.
20. Mo P, Xing Y, Xiao Y, Deng L, Zhao Q, Wang H, et al. Clinical characteristics of refractory coronavirus disease 2019 in Wuhan, China. *Clin Infect Dis* 2021;73:e4208-13.

21. Li T, Wang X, Zhuang X, Wang H, Li A, Huang L, et al. Baseline characteristics and changes of biomarkers in disease course predict prognosis of patients with COVID-19. *Intern Emerg Med* 2021;16:1165-72.
22. Mueller AA, Tamura T, Crowley CP, DeGrado JR, Haider H, Jezmir JL, et al. Inflammatory biomarker trends predict respiratory decline in COVID-19 patients. *Cell Rep Med* 2020;1:100144.
23. Smilowitz NR, Kunichoff D, Garshick M, Shah B, Pillinger M, Hochman JS, et al. C-reactive protein and clinical outcomes in patients with COVID-19. *Eur Heart J* 2021;42:2270-9.
24. Manson JJ, Crooks C, Naja M, Ledlie A, Goulden B, Liddle T, et al. COVID-19-associated hyperinflammation and escalation of patient care: a retrospective longitudinal cohort study. *Lancet Rheumatol* 2020;2:e594-602.
25. Kohno S, Seki M, Takehara K, Yamada Y, Kubo K, Ishizaka A, et al. Prediction of requirement for mechanical ventilation in community-acquired pneumonia with acute respiratory failure: a multicenter prospective study. *Respiration* 2013;85:27-35.
26. Ferrer M, Traverso C, Cilloniz C, Gabarrus A, Ranzani OT, Polverino E, et al. Severe community-acquired pneumonia: characteristics and prognostic factors in ventilated and non-ventilated patients. *PLoS One* 2018;13:e0191721.
27. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58:377-82.
28. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997;336:243-50.
29. Tian J, Xu Q, Liu S, Mao L, Wang M, Hou X. Comparison of clinical characteristics between coronavirus disease 2019 pneumonia and community-acquired pneumonia. *Curr Med Res Opin* 2020;36:1747-52.
30. Bain W, Yang H, Shah FA, Suber T, Drohan C, Al-Yousif N, et al. COVID-19 versus non-COVID-19 acute respiratory distress syndrome: comparison of demographics, physiologic parameters, inflammatory biomarkers, and clinical outcomes. *Ann Am Thorac Soc* 2021;18:1202-10.
31. Di Mitri C, Arcoleo G, Mazzuca E, Camarda G, Farinella EM, Soresi M, et al. COVID-19 and non-COVID-19 pneumonia: a comparison. *Ann Med* 2021;53:2321-31.
32. Wicky PH, Niedermann MS, Timsit JE. Ventilator-associated pneumonia in the era of COVID-19 pandemic: how common and what is the impact? *Crit Care* 2021;25:153.
33. Ji D, Zhang D, Xu J, Chen Z, Yang T, Zhao P, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clin Infect Dis* 2020;71:1393-9.
34. Fernandes FT, de Oliveira TA, Teixeira CE, Batista AF, Dalla Costa G, Chiavegatto Filho AD. A multipurpose machine learning approach to predict COVID-19 negative prognosis in São Paulo, Brazil. *Sci Rep* 2021;11:3343.
35. Yu L, Halalau A, Dalal B, Abbas AE, Ivascu F, Amin M, et al. Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. *PLoS One* 2021;16:e0249285.
36. Wu G, Yang P, Xie Y, Woodruff HC, Rao X, Guiot J, et al. Development of a clinical decision support system for severity risk prediction and triage of COVID-19 patients at hospital admission: an international multicentre study. *Eur Respir J* 2020;56:2001104.
37. Heo J, Han D, Kim HJ, Kim D, Lee YK, Lim D, et al. Prediction of patients requiring intensive care for COVID-19: development and validation of an integer-based score using data from Centers for Disease Control and Prevention of South Korea. *J Intensive Care* 2021;9:16.
38. Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, et al. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020;180:1345-55.
39. Gupta S, Hayek SS, Wang W, Chan L, Mathews KS, Melamed ML, et al. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020;180:1436-47.
40. Song J, Park DW, Cha JH, Seok H, Kim JY, Park J, et al. Clinical course and risk factors of fatal adverse outcomes in COVID-19 patients in Korea: a nationwide retrospective cohort study. *Sci Rep* 2021;11:10066.
41. Park HY, Lee JH, Lim NK, Lim DS, Hong SO, Park MJ, et al. Presenting characteristics and clinical outcome of patients with COVID-19 in South Korea: a nationwide retrospective observational study. *Lancet Reg Health West Pac* 2020;5:100061.
42. Bartoletti M, Giannella M, Scudeller L, Tedeschi S, Rinaldi M, Bussini L, et al. Development and validation of a prediction model for severe respiratory failure in hospitalized patients with SARS-CoV-2 infection: a multicentre cohort study (PREDI-CO study). *Clin Microbiol Infect* 2020;26:1545-53.
43. Azoulay E, Fartoukh M, Darmon M, Géri G, Voiriot G, Dupont T, et al. Increased mortality in patients with severe SARS-CoV-2 infection admitted within seven days of disease onset. *Intensive Care Med* 2020;46:1714-22.