Original Research

Clinical prediction models in hospitalized patients with COVID-19: A multicenter cohort study

Maria Cristina Vedovati [a,*], Greta Barbieri [b,c], Chiara Urbini [a], Erika D'Agostini [a,d], Simone Vanni [e], Chiara Papalini [f], Giacomo Pucci [g], Ludovica Anna Cimini [a], Alessandro Valentino [d], Lorenzo Ghiadoni [h,c], Cecilia Becattini [a]

[a] Internal, Vascular and Emergency Medicine – Stroke Unit, University of Perugia, Perugia, Italy
[b] Department of Surgical, Medical, Molecular and Critical Area Pathology, University of Pisa, Italy
[c] Emergency Medicine Unit, Pisa University Hospital, Italy
[d] Emergency Department, "M. Bufalini" Hospital, Cesena, Italy
[e] Emergency Department, Empoli Hospital, Empoli, Italy
[f] Infectious Diseases Clinic, University of Perugia, Perugia, Italy
[g] Department of Medicine and Surgery, University of Perugia, Unit of Internal Medicine, "Santa Maria" Terni University Hospital, Terni, Italy
[h] Department of Clinical and Experimental Medicine, Pisa University Hospital, Pisa, Italy

ARTICLE INFO

ABSTRACT

Background: Clinical spectrum of novel coronavirus disease (COVID-19) ranges from asymptomatic infection to severe respiratory failure that may result in death. We aimed at validating and potentially improve existing clinical models to predict prognosis in hospitalized patients with acute COVID-19.
Methods: Consecutive patients with acute confirmed COVID-19 pneumonia hospitalized at 5 Italian non-intensive care unit centers during the 2020 outbreak were included in the study. Twelve validated prognostic scores for pneumonia and/or sepsis and specific COVID-19 scores were calculated for each study patient and their accuracy was compared in predicting in-hospital death at 30 days and the composite of death and orotracheal intubation.
Results: During hospital stay, 302 of 1044 included patients presented critical illness (28.9%), and 226 died (21.6%). Nine out of 34 items included in different prognostic scores were independent predictors of all-cause-death. The discrimination was acceptable for the majority of scores (APACHE II, COVID-GRAM, REMS, CURB-65, NEWS II, ROX-index, 4C, SOFA) to predict in-hospital death at 30 days and poor for the rest. A high negative predictive value was observed for REMS (100.0%) and 4C (98.7%) scores; the positive predictive value was poor overall, ROX-index having the best value (75.0%).
Conclusions: Despite the growing interest in prognostic models, their performance in patients with COVID-19 is modest. The 4C, REMS and ROX-index may have a role to select high and low risk patients at admission. However, simple predictors as age and PaO2/FiO2 ratio can also be useful as standalone predictors to inform decision making.

1. Introduction

The spectrum of disease in patients affected by SARS-CoV-2 infection ranges from asymptomatic infection to critical illness. Patients with coronavirus disease 19 (COVID-19) admitted to the hospital require assessment for short-term prognosis to optimize clinical management and resources allocation. During the first wave, in-hospital mortality ranged from 21% to 29%, and admission to intensive care units (ICUs) from 14% to 16% [1–4]. It should be noted that the efficacy of immunomodulatory and antiviral agents in reducing death or adverse outcome is still debated in COVID-19 patients, mainly when used in advanced phases of the disease [5–7]. Despite vaccination and spreading of several new variants may have impacted on severity of the disease, caution is required in inferring a reduced severity of COVID-19 infection [8,9]. In this view, the opportunity of early identification of patients that will progress to severe disease would allow tailored treatment strategies.

Clinical prediction models have been specifically developed to predict prognosis in patients with COVID-19; in addition, several scores

---

validated in patients with no-COVID-19-related lung failure have been applied to patients with COVID-19 [10,11]. Unfortunately, currently available evidence showed poor performance for all these models to predict prognosis in patients with COVID-19 and no incremental value was observed for scores in comparison to simple univariable predictors [3,10]. Among univariable predictors, oxygen saturation on room air and patient age have been found to be strong predictors of deterioration and of mortality, respectively [10]. In addition, many of the existing COVID-19 clinical prediction models have a high risk of bias due to poor reporting, over-estimation of predictive performance, and lack of external validation [4,12].

For these reasons, further external validation is required before moving these scores to the clinical use.

We performed a multicenter cohort study to identify predictors of 30- day in-hospital death or critical illness in patients admitted to hospital for COVID-19 pneumoniae and to evaluate the accuracy of available clinical prediction models in this setting.

2. Materials and methods

2.1. Patients and study design

Data from retrospectively collected cohorts of patients enrolled at 5 Italian non-ICU centers (Perugia, Pisa, Cesena, Empoli and Terni) were merged in a collaborative database. Consecutive patients to the study centers (3 Emergency Departments, 2 Internal Medicine wards, 1 Infectious Disease unit) with confirmed COVID-19 pneumoniae from March 3rd’ 2020 to March 16th’ 2021 were evaluated for inclusion in the

Table 1
Parameters included in the prognostic scores.

	Score											
	APACHE-II	COVID-Gram	CSS	CURB-65	4C	HACOR	MEWS	NEWS-II	qSOFA	*REMS*	*ROX index*	*SOFA*
Age	✓	✓	✓	✓	✓					✓		
Gender					✓							
Body temperature	✓						✓	✓				
History of severe organ failure or immunocompromise	√§											
Coronary heart disease			✓									
Cancer		✓										
Number of comorbidities		√°			√^							
Clinical presentation												
Shortness of breath		✓										
Hemoptysis		✓										
Systolic blood pressure				✓			✓	✓	✓			
Mean blood pressure	✓									✓		√*
Heart rate	✓					✓	✓	✓		✓		
Respiratory rate	✓			✓	✓	✓	✓	✓	✓	✓	✓	
SpO2					✓					✓	✓	
Need for oxygen								✓				
Mechanical ventilation												✓
FiO2											✓	✓
GCS	✓				✓	✓			✓	✓		✓
AVPU score						✓						
Unconsciousness		✓						✓				
Confusion				✓								
Laboratory findings												
Platelet, mmc												✓
White blood cells, mmc	✓											
Lymphocyte count			✓									
N/L ratio		✓										
Hematocrit	✓											
Urea				✓	✓							
Creatinine	✓											✓
Lactate deidrogenase		✓										
Bilirubin		✓										✓
pH	✓					✓						
PaO2	√+											✓
PaO2/FiO2						✓						
AaDO₂	√+											
Hypercapnia								✓				
Sodium	✓											
Potassium	✓											
C-RP					✓							
Procalcitonin			✓									
D-dimer			✓									
Chest Xray abnormalities		✓										

AaDO2 = alveolar-arterial oxygen gradient; AVPU = Alert, Verbal, Pain, Unresponsive scale; C-RP=C reactive protein; FiO2 = fraction of inspired oxygen; GCS = Glasgow Coma Scale; N/L = neutrophil to lymphocyte ratio; PaO2 = arterial partial pressure of oxygen; SpO2 = oxygen saturation.
§ Heart failure class IV, cirrhosis, chronic lung disease, or dialysis-dependent.
+ AaDO₂ or PaO₂ (for FiO₂≥0.5 or <0.5, respectively).
° chronic obstructive pulmonary disease (COPD), hypertension, diabetes, coronary heart disease, chronic heart disease, chronic kidney disease, cancer, cerebral vascular disease, hepatitis B, and immunodeficiency.
^ Chronic cardiac disease, chronic respiratory disease -excluding asthma-, chronic renal disease, mild to severe liver disease, dementia, chronic neurological conditions, connective tissue disease, diabetes mellitus, HIV or AIDS, and malignancy.
* Mean arterial pressure OR administration of vasoactive agents required.

study.

Inclusion criteria were:

- age 18 years or older;
- need for oxygen support or COVID-19 pneumonia (clinical or imaging diagnosed);
- SARS-CoV2 infection confirmed by RT-PCR throat/nasal swab.

Exclusion criteria were:

- patients intubated before hospital arrival or intubated in the Emergency Department;
- insufficient data to calculate at least one risk stratification score.

The study was approved by the Ethical Committee and/or Institutional Review Boards of the participating centers.

Informed consent was obtained in accordance with current regulations for observational studies.

2.1.1. Study outcome

The study primary outcome was all-cause-death occurring within 30 days from hospital admission. Secondary outcome was critical illness defined as the composite of all-cause-death or orotracheal intubation. Study outcomes were considered occurring within hospital stay or at a maximum of 30 days whichever comes first.

2.2. Measurements

The following risk stratification scores were calculated for each patient at admission: APACHE II, COVID-GRAM, CSS, CURB-65, 4C, HACOR, MEWS, NEWS II, qSOFA, REMS, ROX-index, SOFA (Table 1) [13–22].

2.3. Data collection

For included patients the following data were collected: age, gender, comorbidities (cardiovascular disease [coronary or peripheral artery disease], type 2 diabetes, chronic obstructive pulmonary disease [COPD], cancer, kidney disease), number of comorbidities, clinical presentation at admission (hemoptysis, shortness of breath, body temperature, systolic and diastolic blood pressure [mmHg], heart rate [beats per minute], respiratory rate [breaths per minute], oxygen saturation, fraction of inspired oxygen [FiO_2], respiratory index [oxygen saturation to respiratory rate ratio] [23], Glasgow coma scale [GCS], confusion), chest x-ray abnormalities, laboratory findings (platelet count, white blood cells count, neutrophils and lymphocytes count, haematocrit, urea, creatinine, bilirubin, lactate dehydrogenase, electrolytes, d-dimer, C-reactive protein, and arterial partial pressure of oxygen [PaO_2] values), date of hospital admission, date of discharge, date of death, date of endotracheal intubation.

2.4. Statistical analysis

Main baseline characteristics of patients were reported as frequencies for categorical data and as mean ± standard deviation (SD) for continuous data. The predictive value for in-hospital death at 30 days and for critical illness of the individual items included in prognostic scores and models was assessed by Cox proportional Hazard model. Separate multivariable analyses were performed - to verify the independent prognostic value of the included items; results were reported as hazard ratio (HR) at 95% confidence interval (CI).

Information on candidate predictors was collected at admission; however, where information on physiological or laboratory measures was not available on the day of admission, measures recorded up to 24 h after vedmission were used. Patients were excluded from analyses when candidate predictors were missing.

Irrespective of these results, the performance of each risk stratification score (APACHE II, COVID-GRAM, CSS, CURB-65, 4C, HACOR, MEWS, NEWS II, qSOFA, REMS, ROX-index, and SOFA) at the established cut-off value was assessed.

In addition to validating existing scores, independent predictors of in-hospital death at 30 days at multivariate analysis were combined to derive a new prognostic model. According to the presence of a significant association with study outcomes and to the magnitude of the effect (β-coefficient), candidate variables were arranged in new candidate models. The final model was chosen according to the best performance. The most discriminant cut-off was determined by calculating the Youden's index. Bootstrapping analysis (2000 samples, bias corrected and accelerated 95% CIs) of the new models was performed.

The following parameters were evaluated to assess performance of clinical models and scores: discrimination by calculating the area under the receiver operating characteristic curve (AUC) at 95% CI, calibration by applying the Hosmer-Lemeshow test (p > 0.05 showed no significant differences between observed and predicted values), sensitivity, specificity, positive and negative predictive value. The risk for in-hospital death at 30 days and for critical illness by risk categories according to individual models/score was calculated using Cox proportional hazard model. Survival was reported by Kaplan-Meier curves and log-rank test.

All tests were 2-sided and statistical significance was accepted if p value < 0.05.

We followed the TRIPOD statement for reporting [24].

Statistical analysis was performed with SPSS software (version 25).

3. Results

Overall, 1047 patients were evaluated and 3 excluded for lack of baseline information. Thus, 1044 patients were included in the study: mean age 68.3 ± 15.6 years, 62.1% were males. Characteristics of the included population at admission are reported in Table 2.

During a mean hospital stay of 15.5 days, in-hospital death at 30 days

Table 2
Baseline characteristics of included patients.

Baseline characteristics	Population (n = 1044)	
Age, years mean, SD	68.3	15.6
Male, n/N, %	648/1044	62.1
Main comorbidities		
Cardiovascular disease, n/N, %	331/1044	31.7
Type 2 diabetes, n/N, %	187/1044	17.9
COPD, n/N, %	123/1044	11.8
Cancer, n/N, %	140/1043	13.4
Clinical presentation		
Shortness of breath, n/N, %	472/839	56.3
Hemoptysis, n/N, %	10/839	1.2
Fever, n/N, %	274/751	36.5
Systolic blood pressure, mmHg mean, SD	131	19.6
Diastolic blood pressure, mmHg mean, SD	76	12.0
Heart rate, beats per minute mean, SD	86	16.9
Respiratory rate, breaths per minute mean, SD	22	6.0
PaO2/FiO2, mmHg mean, SD	272	103
SatO2, % mean, SD	92.3	6.4
PaO2, mmHg mean, SD	73.9	31.0
Respiratory index, mean, SD	4.7	1.3
Laboratory findings		
Platelet, mmc mean, SD	212689	96818
White blood cells, mmc mean, SD	7917	4477
N/L ratio, mean, SD	8.6	10.8
Hematocrit, % mean, SD	38.6	5.4
Urea, mg/dl mean, SD	45.9	36.1
Creatinine, mg/dl mean, SD	1.17	0.91
D-dimer, ng/ml mean, SD	2091	4827
Lactate deidrogenase, U/l mean, SD	377	242

COPD = chronic obstructive pulmonary disease; PaO2/FiO2 = ratio of the arterial partial pressure of oxygen to the fraction of inspired oxygen; PaO2 = arterial partial pressure of oxygen; SpO2 = oxygen saturation; respiratory index = ratio of the oxygen saturation to the respiratory rate.

occurred in 226 patients (21.6%) (Fig. 1 A) and critical illness in 302 (28.9%).

Predictors of in-hospital death at 30 days at univariate analysis are reported in e-Table 1. Association of age or respiratory rate at increasing cutoffs and of PaO2, PaO2 to FiO2 ratio, and oxygen saturation at decreasing cutoffs with death are reported in Fig. 1 B–F.

Out of 34 items included in the 12 assessed scores, age $\geq$60 years (HR 4.13, 95% CI 1.49 to 11.43), the presence of at least 2 comorbidities (HR 2.43, 95% CI 1.57 to 3.76), GCS lower than 15 (HR 1.94, 95% CI 1.07 to 3.54), mean blood pressure lower than 70 mmHg (HR 4.19, 95% CI 1.50 to 11.71), respiratory rate higher than 20 Bpm (1.58, 95% CI 1.00 to 2.50), PaO$_2$ to FiO$_2$ ratio lower than 200 mmHg (HR 1.88, 95% CI 1.22 to 2.89), PaO$_2$ lower than 60 mmHg (HR 1.60, 95% CI 1.05 to 2.45), oxygen saturation lower than 90% (HR 1.73, 95% CI 1.12 to 2.69), and respiratory index lower or equal to 3.8 (HR 1.82, 95% CI 1.19 to 2.80) were identified as independent predictors of in-hospital death at 30 days (Table 3, e-Table 2).

Age $\geq$60 years (HR 2.57, 95% CI 1.27 to 5.21), the presence of at least 2 comorbidities (HR 1.93, 95% CI 1.32 to 2.83), respiratory rate higher than 20 Bpm (HR 1.63, 95% CI 1.07–2.46), PaO$_2$ lower than 60 mmHg (HR 1.67, 95% CI 1.15 to 2.43), oxygen saturation lower than 92% (HR 1.83, 95% CI 1.26 to 2.67), and respiratory index lower or equal to 3.8 (HR 1.72, 95% CI 1.19 to 2.50) were independent predictors of 30-day critical illness (e-Table 3 and e-Table 4).

3.1. Performance of existing risk stratification scores

The ability to discriminate categories of patients with different risk of in-hospital death at 30 days (discrimination) was acceptable for nine out of the 12 assessed scores (Table 4 and Fig. 2). The APACHE-II, 4C and REMS scores showed the highest AUC values, while the HACOR, MEWS and qSOFA the lowest. The established cut-offs of the individual scores showed a high sensitivity for REMS (100.0%) and 4C (98.9%) scores, and a high specificity for ROX-index (98.3%) and SOFA (96.6%) scores. The negative predictive value was high for REMS (100.0%) and 4C (98.7%), the positive predicted value overall was poor with the highest value for the ROX-index (75.0%).

However, small proportions of patients were categorized at high risk for death according to the qSOFA, SOFA score and the ROX-index; the REMS score identified only a minority of patients as being at low risk. The established cut-offs were significantly associated with the risk for in-hospital death at 30 days as shown in Table 4 and e-Fig. 1. The number of patients included in the high-risk categories varied widely across scores from 5.0% in the ROX-index to 91.4% in the REMS. Similarly, in-hospital death at 30 days in high-risk category ranged from 27.9% in the REMS to 75.0% in the ROX-index and in the low-risk category from 0 in the REMS to 21.5% in the ROX-index.

No significant differences between observed and predicted events were found for CSS, 4C, HACOR, NEWS-II, q-SOFA, REMS, ROX-index and SOFA scores, while calibration was poor for the APACHE-II, COVID-GRAM, CURB-65 and MEWS scores.

The 4C score was confirmed to have the highest discrimination in predicting critical illness: AUC 0.770, 95% CI 0.720 to 0.820 (e-Table 5). Similarly, the 4C and the REMS showed a high negative predictive value: 96.2% and 97.4%, respectively. ROX-index showed the best specificity and positive predictive value: 99.1% and 87.5%, respectively. Calibration was good for ROX-index, CSS, NEWS-II, SOFA, qSOFA, CURB-65, REMS, HACOR, and 4C scores.

3.2. Optimization of existing risk stratification scores

The performance of the 4C and REMS scores was re-evaluated by recalculating SpO2 and respiratory rate at a different cut-off point. No advantage was found for the 4C score with SpO2 lower than 90% (AUC 0.802, 95% CI 0.755 to 0.849), however, discrimination improved after REMS score recalculation for respiratory rate higher than 20 Bpm (AUC 0.800, 95% CI 0.755 to 0.845).

3.3. Derivation of new predicting models

Variables identified as independent predictors for in-hospital death at 30 days were combined in new candidate models. All the examined new models showed a good discrimination and high negative predicted values (e-Table 6, Table 4). All the new models showed a good
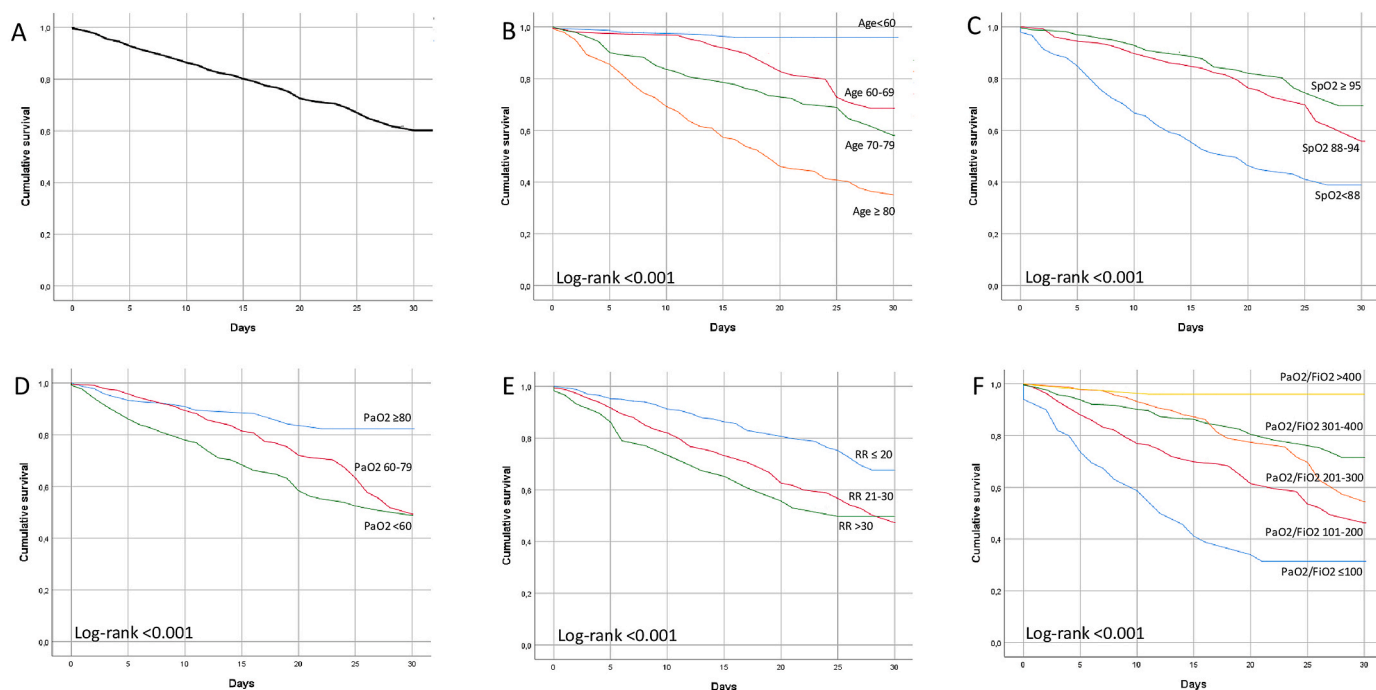

Fig. 1. Cumulative survival in the overall study population (A), according to age categories (B), according to oxygen saturation categories (C), according to PaO2 categories (D), according to respiratory rate categories (E), according to PaO2/FiO2 categories (F).

Table 3
Performance of independent predictors in the assessment of in-hospital death at 30 days.

	AUC	95% CI	Cut-off	Patients	In-hospital death N (%)	SE	SP	NPV	PPV	HR[a] (95% CI)
Age	0.778	0.747–0.810	≥60	718 (68.8)	215 (29.9)	95.1	38.5	96.6	30.0	8.00 (4.24–15.10)
			<60	326 (31.2)	11 (3.4)					
Number of comorbidities	0.679	0.639–0.719	≥2	306 (29.3)	119 (38.9)	52.7	77.1	85.5	38.9	2.93 (2.25–3.81)
			<2	738 (70.7)	107 (14.5)					
GCS	0.538	0.479–0.598	<15	32 (6.1)	15 (46.9)	11.9	95.8	77.6	46.9	2.80 (1.63–4.80)
			≥15	495 (93.9)	111 (22.4)					
Mean blood pressure	0.578	0.530–0.627	<70	24 (3.0)	12 (50.0)	6.4	98.1	77.7	50.0	3.11 (1.73–5.58)
			≥70	784 (97.0)	175 (22.3)					
Respiratory rate	0.663	0.620–0.706	>20	394 (43.6)	127 (32.2)	64.8	62.2	32.2	86.4	2.12 (1.57–2.86)
			≤20	509 (56.4)	69 (13.6)					
PaO2/FiO2	0.722	0.682–0.762	<200	246 (25.1)	108 (43.9)	51.9	82.1	86.4	43.9	2.63 (1.99–3.47)
			≥200	734 (74.9)	100 (13.6)					
SpO2	0.682	0.632–0.732	<90	129 (19.0)	60 (46.5)	38.5	86.8	82.6	46.5	2.42 (1.75–3.36)
			≥90	550 (81.0)	96 (17.5)					
Respiratory index	0.691	0.640–0.742	≤3.8	173 (29.4)	70 (40.5)	50.4	77.1	83.4	40.5	2.14 (1.52–2.99)
			>3.8	415 (70.6)	69 (16.6)					

[a] Results of the univariate analysis.

Table 4
Performance of risk stratification scores in the prediction of in-hospital death at 30 days.

Scores	N	%	AUC	95% CI	Cutoff	Patients n (%)	In-hospital Death n (%)	SE	SP	NPV	PPV	HR	95% CI
APACHE-II	607	58.1	0.785	0.745–0.825	≥10	229 (37.7)	94 (41.0)	68.1	71.2	88.4	41.0	3.54	2.46–5.10
					<10	378 (62.3)	44 (11.6)						
COVID-GRAM	401	38.4	0.711	0.659–0.763	≥139	194 (48.4)	75 (38.7)	73.5	60.2	87.0	38.7	2.88	1.85–4.48
					<139	207 (51.6)	27 (13.0)						
CSS	590	56.5	0.733	0.687–0.778	≥3	419 (71.0)	120 (28.6)	93.8	35.3	**95.3**	28.6	5.49	2.55–11.76
					<3	171 (29.0)	8 (4.7)						
CURB-65	704	67.4	0.705	0.662–0.749	≥2	491 (69.7)	145 (29.5)	93.5	37.0	**95.3**	29.5	5.42	2.76–10.64
					<2	213 (30.3)	10 (4.7)						
4C	340	32.6	0.803	0.756–0.851	≥9	262 (77.1)	90 (34.4)	**98.9**	30.9	**98.7**	34.4	16.76	2.33–120.49
					<9	78 (22.9)	1 (1.3)						
HACOR	427	40.9	0.675	0.615–0.736	≥6	69 (16.2)	33 (47.8)	30.6	88.7	79.1	47.8	1.76	1.15–2.68
					<6	358 (83.8)	75 (20.9)						
MEWS	476	45.6	0.632	0.575–0.690	≥5	131 (27.5)	46 (35.1)	40.7	76.6	80.6	35.1	1.53	1.04–2.23
					<5	345 (72.5)	67 (19.4)						
NEWS-II	721	69.1	0.736	0.691–0.781	≥5	335 (46.5)	113 (33.7)	74.3	61.0	89.9	33.7	2.66	1.83–3.85
					<5	386 (53.5)	39 (10.1)						
QSOFA	826	79.1	0.668	0.623–0.713	≥2	59 (7.1)	30 (50.8)	17.3	95.6	81.4	50.8	3.32	2.24–4.93
					<2	767 (92.9)	143 (18.6)						
REMS	440	42.1	0.793	0.747–0.839	≥3	402 (91.4)	112 (27.9)	100.0	11.6	**100.0**	27.9	22.01	0.45–1081.74
					<3	38 (8.6)	0						
ROX-index	479	45.9	0.712	0.655–0.769	<4.88	24 (5.0)	18 (75.0)	15.5	98.3	78.5	75.0	4.46	2.69–7.41
					≥4.88	455 (95.0)	98 (21.5)						
SOFA	780	74.7	0.776	0.738–0.814	≥6	59 (7.6)	39 (66.1)	20.9	96.6	79.5	66.1	4.31	3.02–6.17
					<6	721 (92.4)	148 (20.5)						
New Model 5	434	41.6	0.809	0.765–0.853	≥1.5	341 (78.6)	104 (30.5)	99.1	28.0	98.9	30.5	26.31	(2.30–301.56)
					<1.5	93 (21.4)	1 (1.1)						

New Model 5: age≥60 years (+1.5), number of comorbidities≥2 (+1), GCS<15 (+1), mean blood pressure <70 mmHg (+1.5), respiratory rate>20 Bpm (+1), PaO2/FiO2 <200 mmHg (+1).

calibration.

Among the candidate models, a good performance in predicting in-hospital death at 30 days for New Model 5 (age 60 years or older, at least two comorbidities, GCS<15, mean blood pressure <70 mmHg, respiratory rate over 20 Bpm, PaO2/FiO2 <200 mmHg; AUC 0.809, 95% CI, 0.765 to 0.853) was observed (e-Fig. 2). The results of the bootstrapping analysis on the predictive value of the New Model 5 were consistent with those of the main analysis.

4. Discussion

Our multicenter cohort study shows that, among 12 existing models for risk stratification of acutely ill patients, the 4C and REMS scores had good accuracy to predict in-hospital death at 30 days in patients hospitalized with COVID-19, with a sensitivity of 99% and 100%, respectively, and negative predictive values of 99% and 100%, respectively. Among the 34 items included in the 12 models, age ≥60 years, having at least two comorbidities, GCS lower than 15, mean blood pressure lower than 70 mmHg, respiratory rate higher than 20 Bpm, PaO2 to FiO2 ratio lower than 200 mmHg, PaO2 at admission lower than 60 mmHg, oxygen saturation lower than 90%, or respiratory index lower or equal to 3.8 were identified as independent predictors of in-hospital death at 30 days.

Similar findings were observed for the prediction of critical illness.

COVID-19 continues to be a major health problem still causing high number of deaths worldwide [25,26], despite the understanding of its pathophysiology continues to grow exponentially [27].

The experience from the initial phase of the pandemic along with the overwhelming amount of scientific evidence on mortality risk factors, respiratory support measures and treatments, allowed greater expertise in facing upcoming pandemic waves. Notwithstanding, the mortality among critical COVID-19 patients remains unacceptably high [28], and it is unclear if improvements in patients' care have truly improved clinical outcomes [29]. In this view, a better understanding of
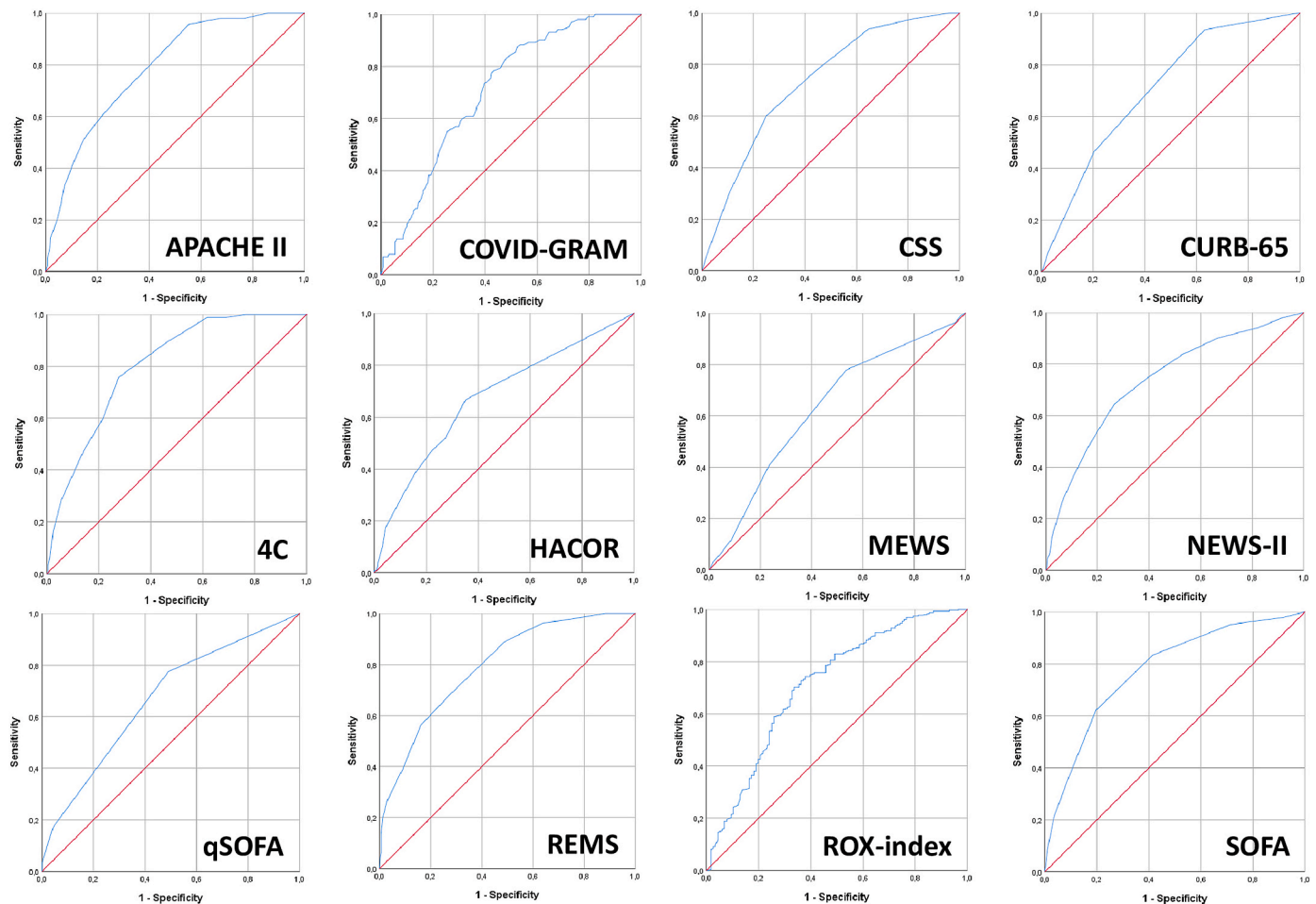
Fig. 2. ROC curves for the prediction of in-hospital death at 30 days according to clinical prediction model.

determinants of prognosis could improve clinical management and also resources allocation for current COVID-19 patients and for potential new COVID-19 waves. These data can certainly serve to implement the management of non-COVID related acute respiratory distress syndromes.

Our study shows that currently available prognostic models recommended for risk stratification in different setting of critically ill patients (sepsis, pneumonia, …) have modest performance in assessing mortality and risk of critical illness in hospitalized patients with COVID-19. The 4C model that has been specifically derived and validated to risk stratify patients with COVID-19 and the REMS score, derived and validated to predict mortality and length of stay in nonsurgical patients attending the emergency department, showed good sensitivity and negative predictive values. These results have biological plausibility as COVID-19 infection rarely causes hemodynamic impairment rather severe respiratory failure potentially leading to acute respiratory distress syndrome as the final cause of death. Thus, clinical models with high performance in patients with sepsis or bacterial pneumonia can have limited value in patients with COVID-19. The ROX-index, derived to assess the risk of mechanical ventilation in pneumonia patients with hypoxemic acute respiratory failure treated with high-flow nasal cannula, revealed to be useful to identify COVID-19 at increased risk for mortality and for critical illness (positive predictive value 75% and 88%, respectively). Therefore, 4C and REMS could be used to identify patients at low-risk of death (1.3% and 0% for patients classified at high risk) and ROX index to identify patients at high-risk (75% for patients classified at high risk). Of note, the 4C could be preferred to the REMS score for the higher prevalence of the low-risk category (22.9% for 4C vs 8.6% for REMS). Of note, only 5%

of patients were categorized in the high-risk category of the ROX-index.

In the analysis of individual predictors, age, number of comorbidities, GCS, mean blood pressure, respiratory rate, PaO2 to FiO2 ratio, PaO2 value, oxygen saturation, or respiratory index were identified as independently associated with in-hospital death at 30 days. Age was associated with both, in-hospital death and critical illness. In particular, age≥60 years showed high sensitivity (95%) and negative predictive value (97%). According to these results, age itself has similar accuracy to 4C and REMS for prognostic stratification. These results are in line with data from a recent study in 411 patients with COVID-19 [10]. In this study, none of the evaluated prognostic models offered incremental value for patient stratification with respect to univariable predictors such as oxygen saturation on room air at admission (AUC 0.76, 95% CI 0.71 to 0.81 for in-hospital deterioration) and patient age (AUC 0.76, 95% CI 0.71 to 0.81 for in-hospital mortality). However, in our study the combination of age to other predictors in the 4C and REMS scores allowed to improve negative predictive value up to 99 and 100%.

In addition to offer external validation of existing models, we built new models from the combination of independent predictors. The performance of model 5 including age≥60 years, having two or more comorbidities, GCS<15, mean blood pressure <70 mmHg, respiratory rate>20 Bpm and PaO2/FiO2 <200 mmHg was similar to that of 4C and REMS scores with the advantage of being simpler. Bootstrapping analysis on the predictive value of the new model 5 showed consistent results with those of the main analysis. However, as for all new scores, this model requires external validation.

Study limitations include the observational study design, thus, owing to the limitations of routinely collected data, predictor variables were

available for a varying number of patients for each model. Therefore, heterogeneity in model performance cannot be excluded. Second, we used admission data as predictors because most prognostic scores are intended to predict outcomes at the point of hospital admission. However, some scores are designed for dynamic inpatient monitoring. We included patients from the first and second waves of COVID-19. In the current COVID-19 patients, vaccination could also influence clinical outcome. Moreover, patients with extremely poor prognosis could be included in the study, this may have influenced study outcomes. At last, since patients hospitalized in non-intensive care unit were included, results cannot be generalized to the whole population of COVID-19.

Our study has some strengths the multicenter design and the completeness of follow-up allowed to retain consistent sample for all the study analyses.

5. Conclusion

Despite the interest and widespread use of prognostic models for COVID-19 patients, the performance of the COVID-19-specific models is modest. The 4C, REMS and ROX-index may have a role to select high and low risk patients at admission. Age and PaO2/FiO2 ratio can also be useful as standalone predictors to inform decision making, despite their accuracy is lower with respect to models.

Fundings

CRediT authorship contribution statement

Maria Cristina Vedovati: Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft. **Greta Barbieri:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Chiara Urbini:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Erika D'Agostini:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Simone Vanni:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Chiara Papalini:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Giacomo Pucci:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Ludovica Anna Cimini:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft. **Alessandro Valentino:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Lorenzo Ghiadoni:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Cecilia Becattini:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft.

Declaration of competing interest

None of the authors have conflict of interest to declare for this study.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rmed.2022.106954.

References

[1] S. Richardson, J.S. Hirsch, M. Narasimhan, et al., Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area, JAMA 323 (2020) 2052–2059.

[2] G. Grasselli, A. Pesenti, M. Cecconi, Critical care utilization for the COVID-19 outbreak in Lombardy, Italy, JAMA 323 (2020) 1545.

[3] P.N. Perez-Guzman, A. Daunt, S. Mukherjee, P. Crook, R. Forlano, M.D. Kont, A. Løchen, M. Vollmer, P. Middleton, R. Judge, C. Harlow, A. Soubieres, G. Cooke, P.J. White, T.B. Hallett, P. Aylin, N. Ferguson, K. Hauck, M.R. Thursz, S. Nayagam, Clinical characteristics and predictors of outcomes of hospitalized patients with coronavirus disease 2019 in a multiethnic london national health service trust: a retrospective cohort study, Clin. Infect. Dis. 73 (11) (2021 Dec 6) e4047–e4057, https://doi.org/10.1093/cid/ciaa1091. PMID: 32766823; PMCID: PMC7454410.

[4] L. Wynants, B. Van Calster, G.S. Collins, R.D. Riley, G. Heinze, E. Schuit, M.M. J. Bonten, D.L. Dahly, J.A.A. Damen, T.P.A. Debray, V.M.T. de Jong, M. De Vos, P. Dhiman, M.C. Haller, M.O. Harhay, L. Henckaerts, P. Heus, M. Kammer, N. Kreuzberger, A. Lohmann, K. Luijken, J. Ma, G.P. Martin, D.J. McLernon, C. L. Andaur Navarro, J.B. Reitsma, J.C. Sergeant, C. Shi, N. Skoetz, L.J.M. Smits, K.I. E. Snell, M. Sperrin, R. Spijker, E.W. Steyerberg, T. Takada, I. Tzoulaki, S.M.J. van Kuijk, B. van Bussel, I.C.C. van der Horst, F.S. van Royen, J.Y. Verbakel, C. Wallisch, J. Wilkinson, R. Wolff, L. Hooft, K.G.M. Moons, M. van Smeden, Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal, BMJ 369 (2020 Apr 7) m1328, https://doi.org/10.1136/bmj.m1328. Update in: BMJ. 2021 Feb 3;372:n236. Erratum in: BMJ. 2020 Jun 3;369: m2204. PMID: 32265220; PMCID: PMC7222643.

[5] A.C. Kalil, T.F. Patterson, A.K. Mehta, K.M. Tomashek, C.R. Wolfe, V. Ghazaryan, V.C. Marconi, G.M. Ruiz-Palacios, L. Hsieh, S. Kline, V. Tapson, N.M. Iovine, M. K. Jain, D.A. Sweeney, H.M. El Sahly, A.R. Branche, J. Regalado Pineda, D.C. Lye, U. Sandkovsky, A.F. Luetkemeyer, S.H. Cohen, R.W. Finberg, P.E.H. Jackson, B. Taiwo, C.I. Paules, H. Arguinchona, N. Erdmann, N. Ahuja, M. Frank, M.D. Oh, E.S. Kim, S.Y. Tan, R.A. Mularski, H. Nielsen, P.O. Ponce, B.S. Taylor, L. Larson, N. G. Rouphael, Y. Saklawi, V.D. Cantos, E.R. Ko, J.J. Engemann, A.N. Amin, M. Watanabe, J. Billings, M.C. Elie, R.T. Davey, T.H. Burgess, J. Ferreira, M. Green, M. Makowski, A. Cardoso, S. de Bono, T. Bonnett, M. Proschan, G.A. Deye, W. Dempsey, S.U. Nayak, L.E. Dodd, J.H. Beigel, ACTT-2 Study Group Members, Baricitinib plus remdesivir for hospitalized adults with covid-19, N. Engl. J. Med. 384 (9) (2021 Mar 4) 795–807, https://doi.org/10.1056/NEJMoa2031994. Epub.2020.Dec.11. PMID: 33306283; PMCID: PMC7745180.

[6] Q. Cheng, G. Zhao, J. Chen, Q. Jia, Z. Fang, Efficacy and safety of current treatment interventions for patients with severe COVID-19 infection: a network meta-analysis of randomized controlled trials, J. Med. Virol. (2021 Dec 9), https://doi.org/10.1002/jmv.27512. Epub ahead of print. PMID: 34882805.

[7] WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group, M. Shankar-Hari, C.L. Vale, P.J. Godolphin, D. Fisher, J.P.T. Higgins, F. Spiga, J. Savovic, J. Tierney, G. Baron, J.S. Benbenishty, L.R. Berry, N. Broman, A. B. Cavalcanti, R. Colman, S.L. De Buyser, L.P.G. Derde, P. Domingo, S.F. Omar, A. Fernandez-Cruz, T. Feuth, F. Garcia, R. Garcia-Vicuna, I. Gonzalez-Alvaro, A. C. Gordon, R. Haynes, O. Hermine, P.W. Horby, N.K. Horick, K. Kumar, B. N. Lambrecht, M.J. Landray, L. Leal, D.J. Lederer, E. Lorenzi, X. Mariette, N. Merchante, N.A. Misnan, S.V. Mohan, M.C. Nivens, J. Oksi, J.A. Perez-Molina, R. Pizov, R. Porcher, S. Postma, R. Rajasuriar, A.V. Ramanan, P. Ravaud, P.D. Reid, A. Rutgers, A. Sancho-Lopez, T.B. Seto, S. Sivapalasingam, A.S. Soin, N. Staplin, J. H. Stone, G.W. Strohbehn, J. Sunden-Cullberg, J. Torre-Cisneros, L.W. Tsai, H. van Hoogstraten, T. van Meerten, V.C. Veiga, P.E. Westerweel, S. Murthy, J.V. Diaz, J. C. Marshall, J.A.C. Sterne, Association between administration of IL-6 antagonists and mortality among patients hospitalized for COVID-19: a meta-analysis, JAMA 326 (6) (2021 Aug 10) 499–518, https://doi.org/10.1001/jama.2021.11330. PMID: 34228774; PMCID: PMC8261689.

[8] https://www.cdc.gov/coronavirus/2019-ncov/variants/about-variants.html. (Accessed 7 February 2022).

[9] M. Salehi-Vaziri, M. Fazlalipour, S.M. Seyed Khorrami, K. Azadmanesh, M. H. Pouriayevali, T. Jalali, Z. Shoja, A. Maleki, The ins and outs of SARS-CoV-2 variants of concern (VOCs), Arch. Virol. (2022 Jan 28) 1–18, https://doi.org/10.1007/s00705-022-05365-2. Epub ahead of print. PMID: 35089389; PMCID: PMC8795292.

[10] R.K. Gupta, M. Marks, T.H.A. Samuels, A. Luintel, T. Rampling, H. Chowdhury, M. Quartagno, A. Nair, M. Lipman, I. Abubakar, M. van Smeden, W.K. Wong, B. Williams, M. Noursadeghi, UCLH COVID-19 Reporting Group, Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study, Eur. Respir. J. 56 (6) (2020 Dec 24), 2003498, https://doi.org/10.1183/13993003.03498-2020. PMID: 32978307; PMCID: PMC7518075.

[11] M. Luo, J. Liu, W. Jiang, S. Yue, H. Liu, S. Wei, IL-6 and CD8+ T cell counts combined are an early predictor of in-hospital mortality of patients with COVID-19, JCI Insight 5 (13) (2020 Jul 9), e139024, https://doi.org/10.1172/jci.insight.139024. PMID: 32544099; PMCID: PMC7406244.

[12] G.P. Martin, M. Sperrin, G. Sotgiu, Performance of prediction models for COVID-19: the Caudine Forks of the external validation, Eur. Respir. J. 56 (6) (2020 Dec

24), 2003728, https://doi.org/10.1183/13993003.03728-2020. PMID: 33060155; PMCID: PMC7562696.

[13] W.A. Knaus, E.A. Draper, D.P. Wagner, J.E. Zimmerman, Apache II: a severity of disease classification system, Crit. Care Med. 13 (10) (1985 Oct) 818–829. PMID: 3928249.

[14] C. Armiñanzas, F. Arnaiz de Las Revillas, M. Gutiérrez Cuadra, A. Arnaiz, M. Fernández Sampedro, C. González-Rico, D. Ferrer, V. Mora, B. Suberviola, M. Latorre, J. Calvo, J.M. Olmos, J.M. Cifrián, M.C. Fariñas, Usefulness of the COVID-GRAM and CURB-65 scores for predicting severity in patients with COVID-19, Int. J. Infect. Dis. 108 (2021 Jul) 282–288, https://doi.org/10.1016/j.ijid.2021.05.048. Epub 2021 May 24. PMID: 34044145; PMCID: PMC8142713.

[15] Y. Shang, T. Liu, Y. Wei, J. Li, L. Shao, M. Liu, Y. Zhang, Z. Zhao, H. Xu, Z. Peng, F. Zhou, X. Wang, Scoring systems for predicting mortality for severe patients with COVID-19, EClinicalMedicine 24 (2020 Jul 3), 100426, https://doi.org/10.1016/j.eclinm.2020.100426. PMID: 32766541; PMCID: PMC7332889.

[16] S.R. Knight, A. Ho, R. Pius, I. Buchan, G. Carson, T.M. Drake, J. Dunning, C. J. Fairfield, C. Gamble, C.A. Green, R. Gupta, S. Halpin, H.E. Hardwick, K. A. Holden, P.W. Horby, C. Jackson, K.A. Mclean, L. Merson, J.S. Nguyen-Van-Tam, L. Norman, M. Noursadeghi, P.L. Olliaro, M.G. Pritchard, C.D. Russell, C.A. Shaw, A. Sheikh, T. Solomon, C. Sudlow, O.V. Swann, L.C. Turtle, P.J. Openshaw, J. K. Baillie, M.G. Semple, A.B. Docherty, E.M. Harrison, ISARIC4C investigators. 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 370 (2020 Sep 9) m3339, https://doi.org/10.1136/bmj.m3339. Erratum in: BMJ. 2020 Nov 13;371:m4334. PMID: 32907855; PMCID: PMC7116472.

[17] J. Duan, S. Wang, P. Liu, X. Han, Y. Tian, F. Gao, J. Zhou, J. Mou, Q. Qin, J. Yu, L. Bai, L. Zhou, R. Zhang, Early prediction of noninvasive ventilation failure in COPD patients: derivation, internal validation, and external validation of a simple risk score, Ann. Intensive Care 9 (1) (2019 Sep 30) 108, https://doi.org/10.1186/s13613-019-0585-9. PMID: 31565779; PMCID: PMC6766459.

[18] C.P. Subbe, M. Kruger, P. Rutherford, L. Gemmel, Validation of a modified early warning score in medical admissions, QJM 94 (10) (2001 Oct) 521–526, https://doi.org/10.1093/qjmed/94.10.521.PMID:11588210.

[19] G.B. Smith, D.R. Prytherch, P. Meredith, P.E. Schmidt, P.I. Featherstone, The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death, Resuscitation 84 (4) (2013 Apr) 465–470, https://doi.org/10.1016/j.resuscitation.2012.12.016. Epub 2013 Jan 4. PMID: 23295778.

[20] T. Olsson, A. Terent, L. Lind, Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients,

[21] J. Intern. Med. 255 (5) (2004 May) 579–587, https://doi.org/10.1111/j.1365-2796.2004.01321.x.PMID:15078500.

[21] O. Roca, J. Messika, B. Caralt, M. García-de-Acilu, B. Sztrymf, J.D. Ricard, J. R. Masclans, Predicting success of high-flow nasal cannula in pneumonia patients with hypoxemic respiratory failure: the utility of the ROX index, J. Crit. Care 35 (2016 Oct) 200–205, https://doi.org/10.1016/j.jcrc.2016.05.022. Epub 2016 May 31. PMID: 27481760.

[22] F.L. Ferreira, D.P. Bota, A. Bross, C. Mélot, J.L. Vincent, Serial evaluation of the SOFA score to predict outcome in critically ill patients, JAMA 286 (14) (2001 Oct 10) 1754–1758, https://doi.org/10.1001/jama.286.14.1754.PMID:11594901.

[23] M.C. Vedovati, L.A. Cimini, L. Pierpaoli, S. Vanni, M. Cotugno, P. Pruszczyk, F. Di Filippo, V. Stefanone, L.G. Torrecillas, M. Kozlowska, M.G. De Natale, F. Mannucci, G. Agnelli, C. Becattini, Prognostic value of respiratory index in haemodynamically stable patients with acute pulmonary embolism: the Respiratory Index model study, Eur. Heart J. Acute Cardiovasc. Care 9 (4) (2020 Jun) 286–292, https://doi.org/10.1177/2048872620913849.Epub.2020.Mar.13. PMID: 32166955.

[24] G.S. Collins, J.B. Reitsma, D.G. Altman, et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement, BMJ 350 (2015) g7594.

[25] A.D. Iuliano, J.M. Brunkard, T.K. Boehmer, et al., Trends in disease severity and health care utilization during the early omicron variant period compared with previous SARS-CoV-2 high transmission periods — United States, december 2020–january 2022, MMWR Morb. Mortal. Wkly. Rep. 71 (2022) 146–152, https://doi.org/10.15585/mmwr.mm7104e4.

[26] World Health Organization, WHO coronavirus (COVID-19) dashboard. https://covid19.who.int. (Accessed 12 February 2022).

[27] A. Parasher, COVID-19: current understanding of its pathophysiology, clinical presentation and treatment, Postgrad. Med. 97 (2021) 312–320.

[28] R.A. Armstrong, A.D. Kane, E. Kursumovic, F.C. Oglesby, T.M. Cook, Mortality in patients admitted to intensive care with COVID-19: an updated systematic review and meta-analysis of observational studies, Anaesthesia 76 (2021) 537–548.

[29] R. Carbonell, S. Urgelés, A. Rodríguez, M. Bodí, I. Martín-Loeches, J. Solé-Violán, E. Díaz, J. Gómez, S. Trefler, M. Vallverdú, J. Murcia, A. Albaya, A. Loza, L. Socias, J.C. Ballesteros, E. Papiol, L. Viña, S. Sancho, M. Nieto, M.D.C. Lorente, O. Badallo, V. Fraile, F. Arméstar, A. Estella, L. Sanchez, I. Sancho, A. Margarit, G. Moreno, COVID-19 SEMICYUC Working Group, Mortality comparison between the first and second/third waves among 3,795 critical COVID-19 patients with pneumonia admitted to the ICU: a multicentre retrospective cohort study, Lancet Reg. Health Eur. 11 (2021 Dec), 100243, https://doi.org/10.1016/j.lanepe.2021.100243. Epub 2021 Nov 4. PMID: 34751263; PMCID: PMC8566166.