



Stratification of patients admitted for SARS-CoV2 infection: prognostic scores in the first and second wave of the pandemic

F. Innocenti¹ · A. De Paris¹ · A. Lagomarsini¹ · L. Pelagatti¹ · L. Casalini¹ · A. Gianni¹ · M. Montuori¹ · P. Bernardini¹ · F. Caldi¹ · I. Tassinari¹ · R. Pini¹

Received: 11 February 2022 / Accepted: 23 May 2022 / Published online: 22 June 2022
© The Author(s), under exclusive licence to Società Italiana di Medicina Interna (SIMI) 2022

Abstract

To test the prognostic performance of different scores, both specifically designed for patients with COVID-19 and generic, in predicting in-hospital mortality and the need for mechanical ventilation (MV). We retrospectively collected clinical data of patients admitted to the Emergency Department of the University Hospital AOU Careggi, Florence, Italy, between February 2020 and January 2021, with a confirmed infection by SARS-CoV2. We calculated the following scores: Sequential Organ Failure Assessment (SOFA) score, CALL score, 4C Mortality score, QUICK score, CURB-65 and MuLBSTA score. The end-points were in-hospital mortality and the need for MV. We included 1208 patients, mean age 60 ± 17 years, 57% male sex. Compared to survivors, non-survivors showed significantly higher values of all the prognostic scores (4C: 13 [10–15] vs 8 [4–10]; CALL: 11 [10–12] vs 9 [7–11]; QUICK: 4 [1–6] vs 0 [0–3]; SOFA: 5 [4–6] vs 4 [4–5]; CURB: 2 [1–3] vs 1 [0–1]; MuLBSTA: 11 [9–13] vs 9 [7–11], all $p < 0.001$). Discriminative ability evaluated by the Receiver Operating Curve analysis showed the following values of the Area under the Curve: 0.83 for 4C, 0.74 for CALL, 0.70 for QUICK, 0.68 for SOFA, 0.76 for CURB and 0.64 for MuLBSTA. The mortality rate significantly increased in increasing quartiles of 4C and CALL score (respectively, 2, 8, 24 and 54% for the 4C score and 1, 17, 33 and 68% for the CALL score, both $p < 0.001$). 4C and CALL score allowed an early and good prognostic stratification of patients admitted for pneumonia induced by SARS-CoV2.

Keywords COVID-19 · Prognostic scores · Mortality rate · Respiratory failure

Introduction

In December 2019, the SARS-CoV2 virus, a member of the coronavirus family, responsible for the global COVID-19 pandemic, was isolated. SARS-CoV2 is a highly contagious respiratory virus, which is able to cause disease with a broad spectrum of severity, from asymptomatic infection to ARDS and MOF [1].

The disease caused by SARS-CoV2 can present a rapid evolution, within hours or days, even in paucisymptomatic patients. It is, therefore, essential to identify, right from the access to the Emergency Department, those patients at risk to develop a severe form of the disease and a pattern of ARDS [2].

Viral and bacterial pneumonia represent a common cause of ED presentation and it carries an adverse outcome, especially in old patients and in those with previous medical conditions [3]. In past years, several scores have been developed for the prognostic stratification of patients with community-acquired pneumonia, which demonstrated a good prognostic stratification ability. These tools, combined with the clinical evaluation, allow the physician to predict the disease evolution and to make appropriate therapeutic and care choices [4].

Pneumonia caused by SARS-CoV2 presents special features from several points of view, including symptoms of presentation, arterial blood gas parameters and bio-humoral findings [5]. A systematic review published in November 2020, identified a long list of clinical parameters, with an established prognostic value in patients with COVID-19 [6]. Thereafter, clinicians involved in the management of these patients elaborated several scores, specifically targeted for this disease. Most of them have been validated in small populations, with a specific ethnic characterization, and do

✉ F. Innocenti
innocenti.fra66@gmail.com

¹ High-Dependency Unit, Department of Clinical and Experimental Medicine, Careggi University Hospital, Lg. Brambilla 3, 50134 Florence, Firenze, Italy

not appear to be effectively generalizable among different contexts [7–9]. 4Cscore represents the only specific score validated on a large European population [10].

We selected two established scores for patients with pneumonia, CURB65 [11] and MuLBSTA [7], and we compared them to those newly designed for patients with COVID-19, including 4C mortality score [10], the CALL score [8] and the Quick score [9]. We added Sequential Organ Failure Assessment (SOFA) score [12], as it is the most widespread tool to evaluate the degree of organ dysfunction in critically ill patients, with a good prognostic stratification ability.

Our aim was to compare the prognostic performance of generic and specifically designed scores in predicting in-hospital mortality and the need for invasive ventilation among patients hospitalized for COVID-19.

Methods

Patients selection

This is a retrospective study, which included all patients hospitalized for pneumonia caused by SARS-CoV2 from the Emergency Department of the Careggi University Hospital in the period March 2020–February 2021. We considered the period from March 2020 to July 2020 as the first wave of the pandemic and the following period as the second wave. The Area Vasta Centro Local Ethical Committee approved this study (NO. 17104). Inclusion criteria were: age equal to or above 18 years, a positive nasopharyngeal swab for SARS-CoV2 and the need for hospitalization. All these patients were admitted primarily for COVID-19 and they all presented radiological findings diagnostic for pneumonia. Discharged patients with a positive swab for SARS-CoV2 were not included.

A database including all patients, with a confirmed diagnosis of SARS-COVID 19 in the ED of this University Hospital, was filled. For each patient, basic demographic data, previous medical conditions, vital signs and laboratory parameters were collected from medical records using a standardized collection template. From the database, we extracted data regarding the ED stay and the in-hospital outcome of the included patients. The primary end-point was in-hospital mortality, while the need for endotracheal intubation (ETI) and mechanical ventilation (MV) was the secondary end-point. Patients not considered eligible for ETI and MV for age and/or comorbidities were excluded from the analysis for the secondary end-point.

Scores calculation

We calculated the following scores: 4C mortality score [10], CALL score [8], MuLBSTA score [7], Quick Severity Index

[9], CURB-65 [11] and SOFA score [12]. The scores were calculated as shown in Table 1. The Quick severity score was calculated only for patients, who required low-flow oxygen therapy in the ED (less than or equal to 6 L/min). We followed the calculation modality of the original scores, except for the 4C mortality score, for which we substituted BUN with creatinine. In fact, BUN was available in a minority of patients, while creatinine was evaluated in all subjects. The points assigned to different creatinine values are listed in Table 1. All scores were evaluated as continuous values and after dichotomization based on the calculation of quartiles (Table 1). SOFA score was calculated based on the original score and was dichotomized as follows: 1st quartile if lowest thru 3, 2nd quartile if 4, 3rd quartile if 5, 4th quartile if 6 thru Highest.

Statistical analysis

Categorical data were reported as proportions and counts. Continuous parameters were reported as mean \pm standard deviation or median and interquartile range based on their distribution. For continuous variables, the null hypothesis was tested using the Student's t test for independent groups; Fisher exact test was used to compare counts in cross tables. The discrimination ability of each prognostic scoring system was evaluated by the receiver-operating-characteristic analysis (ROC analysis). A p value < 0.05 was accepted as statistically significant. The analyses were computed using Statistical Package for Social Sciences 27.0.1.0 (SPSS Statistics, IBM Corporation, Chicago, Ill, USA).

Results

We included 1208 patients, mean age was 69 ± 17 years, 57% of male sex. As reported in Table 2, the most common previous medical conditions were hypertension, diabetes mellitus, coronary artery disease and atrial fibrillation. The most common clinical presentation was dyspnea, followed by cough and weakness.

Most of the patients ($n = 960$, 80%) have been initially transferred to ordinary wards, while 87 (7%) were admitted to the High-Dependency Unit (HDU) and 161 (13%) to the Intensive Care Unit (ICU).

One-hundred-thirteen patients (9%) underwent ETI and MV, with a mean latency of 7 ± 8 days from ED admission and among them mortality rate was 50% ($n = 57$). The overall mortality rate was 21% ($n = 248$ patients), with a mean latency of 14 ± 14 days. One-hundred-sixty-seven patients were not considered eligible for ETI and MV and were excluded from the analysis for the secondary end-point. We compared the prevalence of the two end-points between the first and the second wave: the mortality rate decreased from

Table 1 Prognostic scores

CALL score	4C mortality	MuLBSTA	Quick	CURB-65
Comorbidities No = 1 Yes = 4	Age (years) < 50 = 0 50–59 = 2 60–69 = 4 70–79 = 6 ≥ 80 = 7	Multilobe infiltrate No = 0 Yes = 5	RR (breaths/min) ≤ 22 = 0 23–28 = 1 > 28 = 2	Mental confusion No = 0 Yes = 1
Age (years) ≤ 60 = 1 < 60 = 3	Sex Female = 0 Male = 1	Lymphocyte count ≤ 0.8 × 10 ⁹ /L No = 0 Yes = 4	SO ₂ (%) > 92 = 0 89–92 = 2 ≤ 88 = 5	RR ≥ 30 b/min No = 0 Yes = 1
Lymphocytes (10 ⁹ /L) > 1,0 = 1 ≤ 1,0 = 3	Number of comorbidities 0 = 0 1 = 1 ≥ 2 = 2	Bacterial coinfection No = 0 Yes = 4	O ₂ flow rate (L/min) ≤ 2 = 0 3–4 = 4 5–6 = 5	SPB < 90 mmHg or DBP < 60 mmHg No = 0 Yes = 1
LDH (U/L) ≤ 250 = 1 251–500 = 2 > 500 = 3	RR (breath/min) < 20 = 0 20–29 = 1 ≥ 30 = 2 SO ₂ on room air (%) ≥ 92 = 0 < 92 = 2 Glasgow coma scale 15 = 0 < 15 = 2 Creatinine (mg/L) < 1.199 = 0 1.2–1.9 = 1 > 1.9 = 3 CRP (mg/L) < 50 = 0 50–99 = 1 ≥ 100 = 2	Smoking history Non-smoker = 0 Prior smoker = 2 Active smoker = 3 History of hypertension No = 0 Yes = 2 Age ≥ 60 years No = 0 Yes = 2		Age ≥ 65 years No = 0 Yes = 1
Dichotomization in quartiles				
1st: lowest thru 6	1st: lowest thru 4	1st: lowest thru 6	1st: 0	1st: 0
2nd: 7 thru 9 = 2	2nd: 5 thru 8 = 2	2nd: 7 thru 8 = 2	2nd: 1	2nd: 1
3th: 10 thru 11 = 3	3th: 9 thru 11 = 3	3th: 9 thru 11 = 3	3th: 2 thru 3	3th: 2
4th: 12 thru highest	4th: 12 thru highest	4th: 12 thru highest	4th: 4 thru highest	4th: 3 thru highest

25 to 18% ($p=0.006$) and the need for MV from 14 to 7% ($p<0.001$). Among survivors, mean length of hospital stay was 15 ± 14 days.

Upon admission to the ED, non-survivors were older and were more likely to have previous medical conditions (Table 2). They showed lower peripheral oxygen saturation and PaO₂/FiO₂ ratio, systolic and diastolic blood pressure, and higher heart rate (HR) and respiratory rate (RR). Among laboratory parameters, they had higher White Blood Cell (WBC) counts, Neutrophil/Lymphocyte ratio, C Reactive Protein, procalcitonin, creatinine, IL-6 and glucose level (Table 3).

In Table 4, we reported the values of the scores in the whole study population and based on in-hospital mortality

and need for MV. All the scores were significantly higher in non-survivors compared to survivors and in those who needed mechanical ventilation. The prognostic stratification ability was further evaluated by ROC Curves analysis. 4C-score showed the best predictive ability for in-hospital mortality, followed by the CALL score (respectively, 0.837 and 0.743). SOFA, QUICK and MuLBSTA showed a fair discriminative ability (Fig. 1). We evaluated separately the discrimination ability of the scores in the first and the second wave and we did not find significant differences (SOFA score: AUC 0.652 vs 0.691; MuLBSTA AUC 0.664 vs 0.671; QUICK AUC 0.663 vs 0.701; 4C score AUC 0.816 vs 0.846; CALL score AUC 0.738 vs 0.751, all $p<0.05$).

Table 2 Characteristics of the study population as a whole and based on in-hospital mortality

	All patients (n = 1208)	Non survivors (n = 248)	Survivors (n = 960)	<i>p</i>
Age (years)	69 ± 17	81 ± 11	65 ± 17	<0.001
Male sex (%)	690 (57%)	145 (59%)	545 (57%)	0.654
Comorbidities				
Hypertension (%)	637 (54%)	158 (65%)	479 (51%)	<0.001
Diabetes (%)	235 (20%)	62 (25%)	173 (19%)	0.019
CAD (%)	177 (15%)	62 (25%)	115 (12%)	<0.001
COPD (%)	107 (9%)	33 (14%)	74 (8%)	0.008
CHF (%)	106 (9%)	41 (17%)	65 (7%)	<0.001
CKF (%)	106 (9%)	43 (18%)	63 (7%)	<0.001
Atrial fibrillation (%)	146 (12%)	54 (22%)	92 (10%)	<0.001
Active cancer (%)	60 (5%)	28 (12%)	32 (3%)	<0.001
Number of comorbidities				<0.001
0	362 (30%)	16 (7%)	346 (36%)	
1–3	686 (57%)	164 (66%)	522 (54%)	
> 3	160 (13%)	68 (27%)	92 (10%)	
Symptoms				
Cough (%)	460 (39%)	65 (27%)	395 (42%)	<0.001
Dyspnea (%)	647 (54%)	164 (67%)	483 (51%)	<0.001
Diarrhea (%)	103 (9%)	13 (5%)	90 (10%)	0.041
Syncope (%)	72 (6%)	11 (5%)	61 (7%)	0.294
Vomit (%)	66 (6%)	12 (5%)	54 (6%)	0.754
Weakness (%)	221 (19%)	36 (15%)	185 (20%)	0.08

CAD coronary artery disease, COPD chronic obstructive pulmonary disease, CHF chronic heart failure, CKF chronic kidney failure

When we considered the secondary end-point, the discriminative ability of all the scores was fair (SOFA: AUC 0.60, 95% CI 0.54–0.66, $p = 0.030$; CURB65: 0.60, 95% CI 0.53–0.67, $p = 0.005$; QUICK: 0.66, 95% CI 0.59–0.74, $p = 0.040$; CALL: AUC 0.68, 95% CI 0.62–0.74, $p = 0.032$; MulBSTA: AUC 0.64, 95% CI 0.58–0.70, $p = 0.031$; 4C score: AUC 0.65, 95% CI 0.60–0.70, $p = 0.024$).

We finally evaluated the mortality rate in the different quartiles of all the scores (Fig. 2). The prevalence of in-hospital mortality significantly increased from the first to the fourth quartile of the 4C mortality score, being virtually absent in the first quartile and disproportionately high in the fourth one. The same behavior was observed for the CALL score and the CURB score. With the SOFA score, only patients in the fourth quartile showed significantly higher mortality, compared to those in the lower quartiles. For homogeneity, we adopted the subdivision in quartiles for all scores. However, in the original paper for validation of the 4C score, the authors proposed a division into 4 risk classes: low-risk (points 0–3), intermediate risk (points 4–8), high risk (points 9–14) and very high risk (points ≥ 15). We tested this classification in our study population and we confirmed the results obtained with

the division in quartiles (mortality in different subgroups, respectively, 1, 6, 27 and 69%, all $p < 0.001$, except for low and intermediate-risk subgroups $p = 0.05$).

Discussion

Among a large population of patients admitted to the hospital with SARS-COV2 infection, the prognostic scores specifically designed for COVID patients demonstrated a good prognostic stratification ability for in-hospital mortality. Among the scores usually applied to patients with bacterial or viral pneumonia, only the CURB-65 score showed a good prognostic performance. The SOFA score did not allow a useful prognostic stratification in our study population. These data were confirmed for both the first and the second wave of the pandemic. An adjunctive strength of CURB, CALL and 4C scores was that the mortality rate significantly increased in increasing quartiles. The novelty of the present study is represented by the evaluation of all the considered scores by parameters collected at the moment of ED admission, allowing an early prognostic stratification. The increased mortality in

Table 3 Vital signs, arterial blood gas and laboratoristic values in the whole population and based on in-hospital mortality

	All patients (n = 1208)	Non survivors (n = 248)	Survivors (n = 960)	p
Vital signs				
SBP (mmHg)	132 ± 21	130 ± 25	133 ± 20	<0.001
Heart rate (b/min)	88 ± 18	91 ± 22	87 ± 17	<0.001
Respiratory rate (b/min)	21 ± 5	24 ± 7	20 ± 5	<0.001
SpO2 (%)	93 ± 6	90 ± 9	94 ± 5	<0.001
GCS	14.7 ± 1.2	14.2 ± 2.1	14.8 ± 0.8	<0.001
Temperature (°C)	36.7 ± 0.9	36.6 ± 0.9	36.8 ± 0.9	0.621
Arterial blood gas				
Ph	7.45 ± 0.06	7.44 ± 0.06	7.45 ± 0.06	0.88
pO2	71.7 ± 31.5	72.4 ± 41.5	71.5 ± 28.2	<0.001
pCO2	35.7 ± 7.9	36.0 ± 10.1	35.6 ± 7.2	<0.001
HCO ₃ ⁻ (meq/L)	25.3 ± 4.1	25.2 ± 4.7	25.4 ± 3.9	0.58
Lac (meq/L)	1.3 ± 1.1	1.9 ± 1.5	1.2 ± 0.9	<0.001
PaO2/FiO2	283 ± 98	219 ± 98	299 ± 91	<0.001
FiO2 (%)	31 ± 22	42 ± 29	27 ± 17	<0.001
Laboratory data				
Hb (g/dL)	13.2 ± 2	12.6 ± 2.3	13.4 ± 1.9	<0.001
WBC (*10 ⁻⁹ /L)	8.5 ± 9.9	11.2 ± 19.2	7.7 ± 4.9	<0.001
PLT (*10 ⁻⁹ /L)	220 ± 104	211 ± 96	222 ± 105	0.436
N/L	10.1 ± 16	15.6 ± 20.8	8.7 ± 14.1	<0.001
CRP (mg/L)	81 ± 74	117 ± 87	71 ± 67	<0.001
Procalcitonin (ng/ml)	5.1 ± 37.3	11.5 ± 62	3.5 ± 27	0.006
Creatinine (mg/dl)	1.2 ± 1.1	1.6 ± 1.4	1.1 ± 1.0	<0.001
Urea (g/L)	1.1 ± 0.6	1.2 ± 0.6	1.1 ± 0.6	0.321
IL-6 (pg/ml)	165 ± 694	379 ± 1154	105 ± 479	<0.001
Bilirubin (mg/dl)	0.5 ± 0.4	0.6 ± 0.4	0.5 ± 0.3	0.008
GPT (U/L)	40 ± 151	32 ± 32	42 ± 168	0.247
Glucose (mg/dl)	138 ± 63	158 ± 91	132 ± 53	<0.001
LDH (U/L)	364 ± 233	469 ± 325	334 ± 190	<0.001
INR	1.3 ± 0.6	1.4 ± 0.8	1.2 ± 0.5	<0.001
Fibrinogen (mg/dl)	603 ± 211	573 ± 203	610 ± 213	0.423

SBP systolic blood pressure, GCS glasgow coma scale, Lac lactate, Hb haemoglobin, WBC white blood cells, PLT platelet count, N/L neutrophil-to-lymphocyte ratio, CRP C-reactive protein, GPT glutamic pyruvic transaminase, LDH lactate dehydrogenase, INR international normalized ratio

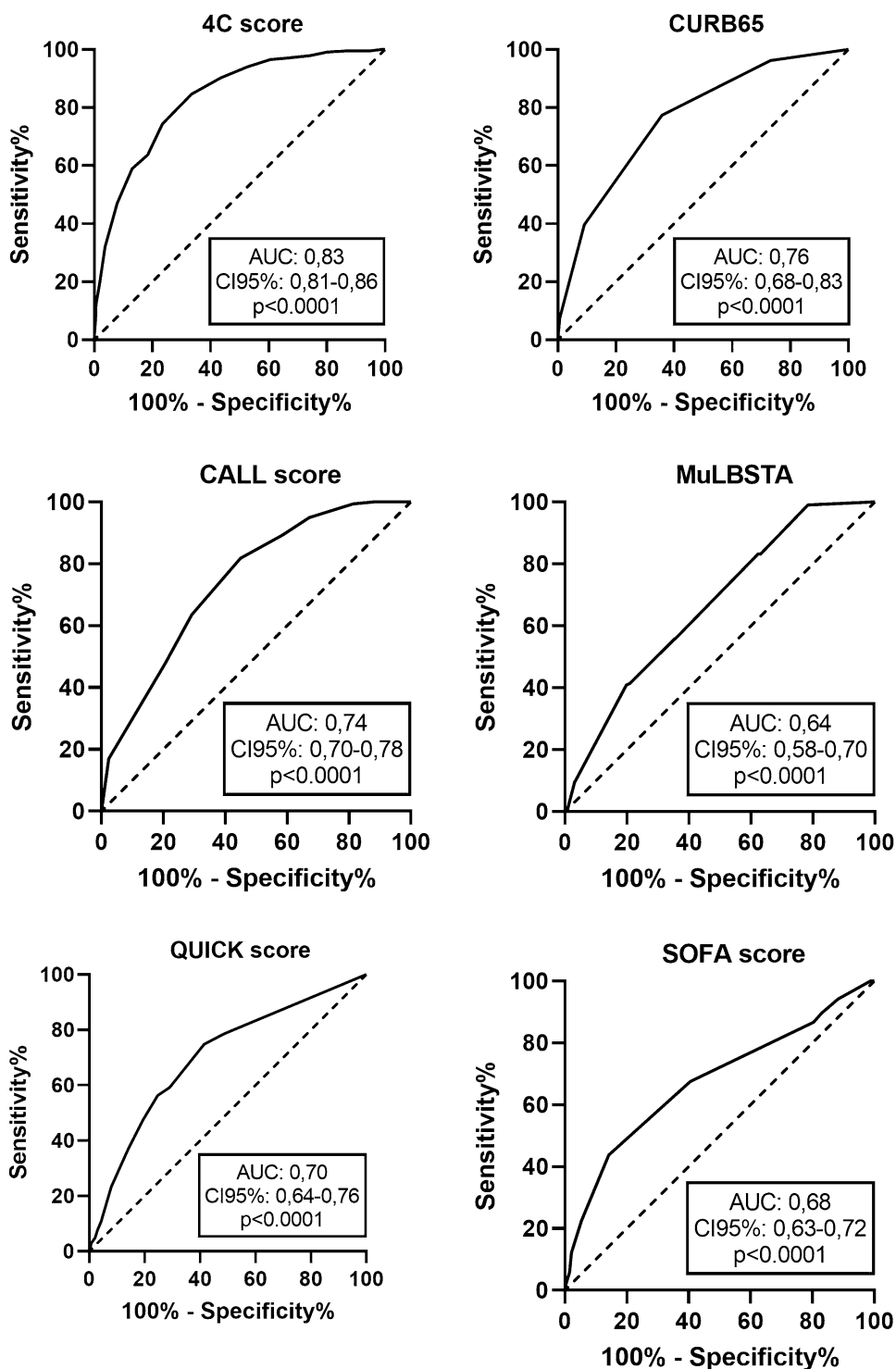
Table 4 Scores values in the whole population and based on the primary outcomes

	All patients (n = 1208)	In-hospital mortality		p	ETI+MV		p
		Survivors (n = 960)	Non-survivors (n = 248)		Effective NIV (n = 928)	Failed NIV (n = 113)	
SOFA	4 [4–5]	4 [4–5]	5 [4–6]	<0.001	4 [4–5]	5 [4–6]	<0.001
CURB	0 [0–1]	1 [0–1]	2 [1–3]	<0.001	1 [0–1]	1 [1–2]	0.006
MuLBSTA	9 [7–12]	9 [7–11]	11 [9–13]	<0.001	9 [7–11]	11 [9–13]	<0.001
4C-score	9 [5–12]	8 [4–10]	13 [10–15]	<0.001	13 [12–15]	14 [11–17]	<0.001
CALL	10 [7–12]	9 [7–11]	11 [10–12]	<0.001	9 [7–11]	11.5 [10–12]	<0.001
QUICK	1 [0–4]	0 [0–3]	4 [1–6]	<0.001	0 [0–3]	3 [1.5–6]	<0.001

higher quartiles can give the clinician a definite reference for prognostic stratification. This is of utmost relevance, as during a pandemic, in the presence of a shortage of

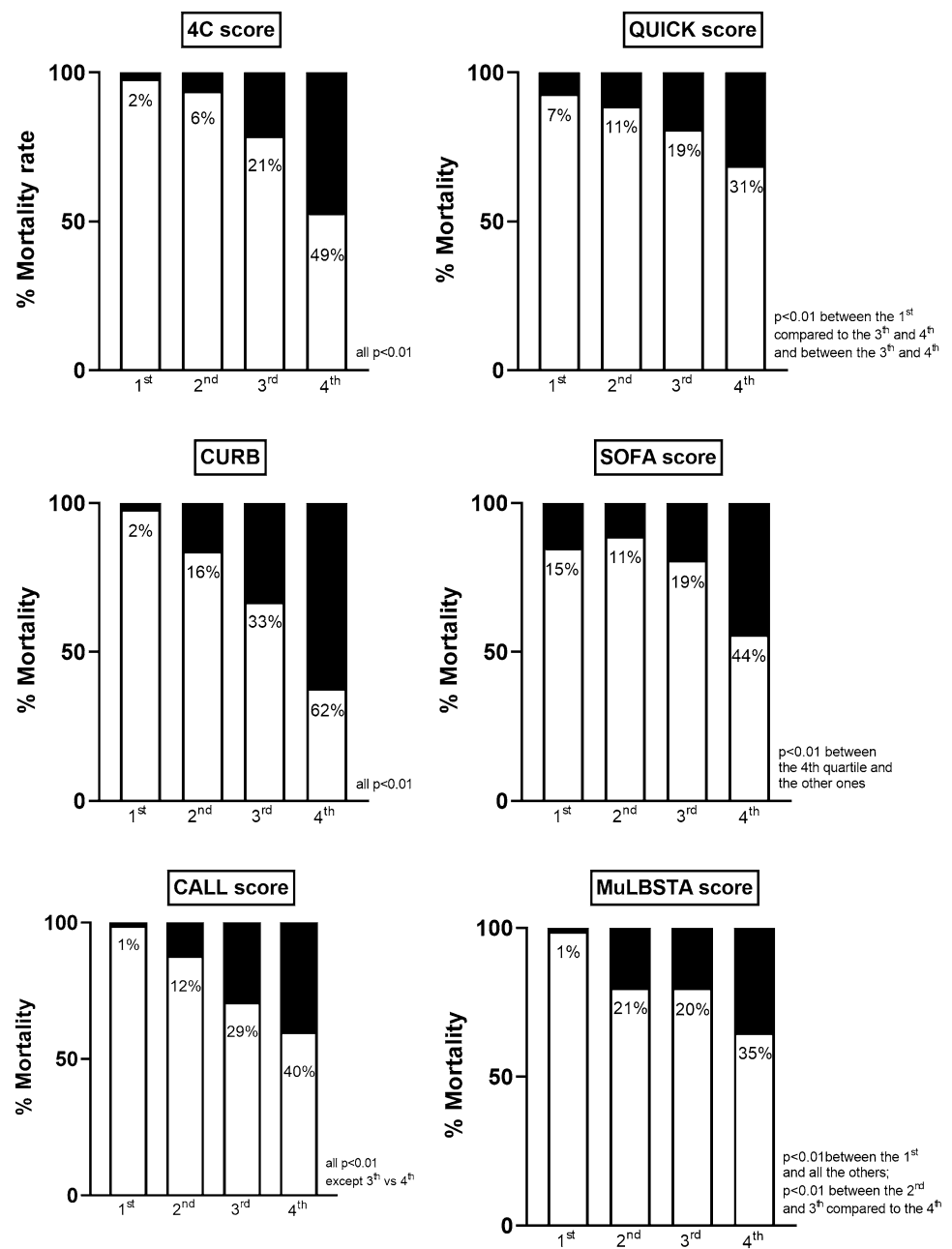
resources, the early identification of patients at risk of an adverse prognosis allows an appropriate utilization of hospital facilities [13]. As relevant changes in the admission

Fig. 1 Prognostic stratification ability of the examined scores evaluated by ROC curves analysis



policies and treatments between the first and the second wave of the pandemic occurred, we could not take for granted a similar diagnostic accuracy in the whole study period. However, our analysis did not show significant differences and, to the best of our knowledge, a similar comparison has not been reported by previous papers on this topic.

We tested the SOFA score as it represents one of the most widespread prognostic scores in the medical community and it plays a pivotal role in the diagnosis and prognostic stratification of septic patients [14]. However, in the present as well as in other populations of patients with pneumonia caused by COVID-19, its prognostic value was fair [9]. A possible explanation could be that these patients

Fig. 2 Mortality rate in the quartiles of the examined scores

do not have significant extrapulmonary organ damage at the beginning of the disease when they usually come to the ED. Multiorgan failure is a late complication, in those who develop severe respiratory failure [15–17]. Therefore, the SOFA score evaluated upon ED admission cannot capture this possible evolution.

Among scores specifically designed for patients with COVID-19, we selected the CALL score, for its ease to use and for the inclusion of parameters routinely evaluated in the ED. It was validated in a small Chinese population [8] and, when we applied it in our study group, the

performance was good in predicting mortality, while it was less satisfying for the prediction of respiratory deterioration. To the best of our knowledge, this study represents the first attempt at external validation of the score.

On the other side, the prognostic stratification ability of the 4C-score, validated in a large European population, has already been confirmed among patients evaluated in the Emergency Department [18, 19] and among those admitted to the Intensive Care Unit [20] or in the general ward [21], in different geographic areas. In our study group, which is one of the largest among the previously mentioned papers, 4C score

outperformed all the other tools. The need to substitute BUN with creatinine, as it was not available in the vast majority of our patients, did not reduce its prognostic stratification ability. Both 4C and CALL score included different laboratory parameters, which were identified by previous studies as prognosticators in patients with a severe form of infection by COVID-19. In the same way, an increased creatinine value was associated with a more severe form of the disease. This could be the reason why the CURB-score, which includes demographics and parameters of renal function, demonstrated a good prognostic performance. Conversely, the MuLBSTA score did not allow a useful prognostic stratification, as some of the included parameters are not discriminative in patients with pneumonia induced by COVID-19. In fact, a bacterial coinfection is a late event, while multilobar infiltration is a ubiquitous finding in these patients [22].

When we considered the need for mechanical ventilation, the prognostic performance of all the scores was fair, consistently with similar study populations [21]. 4C score was developed to predict mortality, while the CALL score was tested to predict both mortality and disease progression, but the ability to early identify patients at risk of developing severe respiratory failure was fair. A possible explanation for this finding is that there was a considerable latency between the hospital admission and the event of intubation. Probably, patients with good and bad prognoses present similar characteristics upon ED admission, but the response to the early treatment plays a pivotal role in determining the prognosis. Therefore, scores calculated at the very beginning of the illness do not capture the difference.

A separate mention has to be made for the QUICK score, which was tested in the present study as it was conceived for patients who required low-flow oxygen supplementation in the ED. It is based on simple parameters, immediately available at the bedside, and could be especially useful in the ED, where patients with these characteristics represent a significant proportion. However, its prognostic ability was fair, limiting its applicability. Therefore, we were not able to confirm the results reported by Rodriguez-Nava et al. [23], who enrolled a significantly smaller population compared to the present one. A relevant difference between the QUICK score and the other ones specifically designed for patients with COVID-19 is that it does not take into account the presence of comorbidities. Our patients were all admitted for pneumonia caused by COVID-19, but we are aware that most of all had previous medical conditions, which definitively affected their outcome. Therefore, considering their presence allowed the 4Ca and CALL score to obtain a better prognostic stratification than the QUICK score.

Limitations

The retrospective and single-center design represents a significant limitation of the present study. In fact, these results may

not be generalizable in the presence of different local admission and management policies. However, the use of a standardized template for data collection limited the possibility of a subjective interpretation of data.

We decided to test several scores, including both those specifically designed for patients with COVID and commonly used scores in patients with bacterial or viral pneumonia. The choice was suggested by the need to confirm the discrimination ability of the first ones as well as to ascertain whether familiar scores could be applied to these patients. We confirmed that specifically designed scores demonstrated a good discrimination ability in this large European population in the following waves. On the other side, scores for patients do not find a relevant role among these patients.

Conclusions

In conclusion, scores specifically designed for patients with pneumonia induced by SARS-CoV2 allowed a good prognostic stratification in the ED, especially for the risk of in-hospital mortality. They maintained their discrimination ability despite the changes in the clinical features of the pandemic in the following waves. The prevalence of an adverse prognosis significantly increased in higher quartiles of the scores. This finding may give clinicians a definite reference to define the risk, with the possibility of early identified patients at high risk of an adverse prognosis.

Author contributions FI and ADeP gave substantial contributions to the conception and design of the work. LP, LC, AG, MM and PB gave substantial contributions to the acquisition, analysis, or interpretation of data for the work. FI, AL and LP drafted the manuscript; ITeFC revised it critically for important intellectual content. RP gave the final approval of the version to be published.

Funding No funding source to declare.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest No conflict of interest to declare.

Ethical approval The study protocol was approved by the “Toscana–Area Vasta-Centro” inter-institutional ethic committee (registration number NO. 17104) and was conducted in accordance with the Helsinki Declaration of 1964 (revised 2008).

Human and animal rights and informed consent The Area Vasta Centro Local Ethical Committee approved this study (NO. 17104). The study was conducted in agreement with the principles of Helsinki declaration.

References

- Chen N, Zhou M, Dong X, Qu J et al (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 395:507–513. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)
- Duca A, Memaj I, Zanardi F, Preti C et al (2020) Severity of respiratory failure and outcome of patients needing a ventilatory support in the emergency department during Italian novel coronavirus SARS-CoV2 outbreak: preliminary data on the role of helmet CPAP and non-invasive positive pressure ventilation. *EClinicalMedicine* 24:100419. <https://doi.org/10.1016/j.eclim.2020.100419>
- Chandra A, Nicks B, Maniago E, Nouh A et al (2010) A multicenter analysis of the ED diagnosis of pneumonia. *Am J Emerg Med* 28:862–865. <https://doi.org/10.1016/j.ajem.2009.04.014>
- Brabrand M, Henriksen DP (2018) CURB-65 score is equal to NEWS for identifying mortality risk of pneumonia patients: an observational study. *Lung* 196:359–361. <https://doi.org/10.1007/s00408-018-0105-y>
- Guo LY, Wang YG, Liu J, Yuan HP et al (2020) Studies on viral pneumonia related to novel coronavirus SARS-CoV-2, SARS-CoV, and MERS-CoV: a literature review. *APMIS* 128:423–432. <https://doi.org/10.1111/apm.13047>
- Izcovich A, Ragusa MA, Tortosa F, Lavena Marzio MA et al (2020) Prognostic factors for severity and mortality in patients infected with COVID-19: a systematic review. *PLoS ONE* 15:e0241955. <https://doi.org/10.1371/journal.pone.0241955>
- Guo L, Wei D, Zhang X, Wu Y et al (2019) Clinical features predicting mortality risk in patients with viral pneumonia: the MuLBSTA score. *Front Microbiol* 10:2752. <https://doi.org/10.3389/fmicb.2019.02752>
- Ji D, Zhang D, Xu J, Chen Z et al (2020) Prediction for progression risk in patients with COVID-19 pneumonia: the call score. *Clin Infect Dis* 71:1393–1399. <https://doi.org/10.1093/cid/ciaa14>
- Haimovich AD, Ravindra NG, Stoytchev S, Young HP et al (2020) Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. *Ann Emerg Med* 76:442–453. <https://doi.org/10.1016/j.annemergmed.2020.07.022>
- Knight SR, Ho A, Pius R, Buchan I et al (2020) 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:m3339. <https://doi.org/10.1136/bmj.m3339>
- Lim WS, van der Eerden MM, Laing R, Boersma WG et al (2003) Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 58:377–382. <https://doi.org/10.1136/thorax.58.5.377>
- Vincent JL, Moreno R, Takala J, Willatts S et al (1996) The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. on behalf of the working group on sepsis-related problems of the European society of intensive care medicine. *Intensive Care Med* 22:707–710. <https://doi.org/10.1007/BF01709751>
- Jang JG, Hur J, Hong KS, Lee W et al (2020) Prognostic accuracy of the SIRS, qSOFA, and NEWS for early detection of clinical deterioration in SARS-CoV-2 infected patients. *J Korean Med Sci* 35:e234. <https://doi.org/10.3346/jkms.2020.35.e234>
- Singer M, Deutschman CS, Seymour CW, Shankar-Hari M et al (2016) The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 315:801–810. <https://doi.org/10.1001/jama.2016.0287>
- Wang D, Hu B, Hu C, Zhu F et al (2020) Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA*. <https://doi.org/10.1001/jama.2020.1585>
- Huang C, Wang Y, Li X, Ren L et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Guan WJ, Ni ZY, Hu Y, Liang WH et al (2020) Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 382:1708–1720. <https://doi.org/10.1056/NEJMoa2002032>
- Jones A, Pitre T, Junek M, Kapralik J et al (2021) External validation of the 4C mortality score among COVID-19 patients admitted to hospital in Ontario, Canada: a retrospective study. *Sci Rep* 11:18638. <https://doi.org/10.1038/s41598-021-97332-1>
- Covino M, Sandroni C, Santoro M, Sabia L et al (2020) Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores. *Resuscitation* 156:84–91. <https://doi.org/10.1016/j.resuscitation.2020.08.124>
- Aletreby WT, Mumtaz SA, Shahzad SA, Ahmed I et al (2022) External validation of 4C ISARIC mortality score in critically ill COVID-19 patients from Saudi Arabia. *Saudi J Med Sci* 10:19–24. https://doi.org/10.4103/sjmms.sjmms_480_21
- Yildiz H, Castanares-Zapatero D, Hanneke C, Vandermeersch D et al (2021) Prospective validation and comparison of COVID-GRAM, NEWS2, 4C mortality score, CURB-65 for the prediction of critical illness in COVID-19 patients. *Infect Dis (Lond)* 53:640–642. <https://doi.org/10.1080/23744235.2021.1896777>
- Lansbury L, Lim B, Baskaran V, Lim WS (2020) Co-infections in people with COVID-19: a systematic review and meta-analysis. *J Infect* 81:266–275. <https://doi.org/10.1016/j.jinf.2020.05.046>
- Rodriguez-Nava G, Yanez-Bello MA, Trelles-Garcia DP, Chung CW et al (2021) Performance of the quick COVID-19 severity index and the brescia-COVID respiratory severity scale in hospitalized patients with COVID-19 in a community hospital setting. *Int J Infect Dis* 102:571–576. <https://doi.org/10.1016/j.ijid.2020.11.003>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.