


Article

Lung Ultrasound, Clinical and Analytic Scoring Systems as Prognostic Tools in SARS-CoV-2 Pneumonia: A Validating Cohort

Jaime Gil-Rodríguez ¹, Michel Martos-Ruiz ¹, José-Antonio Peregrina-Rivas ², Pablo Aranda-Laserna ¹, Alberto Benavente-Fernández ¹, Juan Melchor ^{3,4,5,*}  and Emilio Guirao-Arrabal ^{2,*}

- ¹ Internal Medicine Unit, San Cecilio University Hospital, 18012 Granada, Spain; jaigilro@gmail.com (J.G.-R.); michelmartosruiz@gmail.com (M.M.-R.); pablo.aranlas@gmail.com (P.A.-L.); alberto.benavente.sspa@juntadeandalucia.es (A.B.-F.)
- ² Infectious Diseases Unit, San Cecilio University Hospital, 18012 Granada, Spain; joseperegrina91@gmail.com
- ³ Department of Statistics and Operations Research, University of Granada, 18011 Granada, Spain
- ⁴ Biomechanics Group (TEC-12), Instituto de Investigación Biosanitaria (IBS), 18012 Granada, Spain
- ⁵ Research Unit “Modelling Nature” (MNat), University of Granada, 18011 Granada, Spain
- * Correspondence: jmelchor@ugr.es (J.M.); emilio.guirao@gmail.com (E.G.-A.)

Abstract: At the moment, several COVID-19 scoring systems have been developed. It is necessary to determine which one better predicts a poor outcome of the disease. We conducted a single-center prospective cohort study to validate four COVID-19 prognosis scores in adult patients with confirmed infection at ward. These are National Early Warning Score (NEWS) 2, Lung Ultrasound Score (LUS), COVID-19 Worsening Score (COWS), and Spanish Society of Infectious Diseases and Clinical Microbiology score (SEIMC Score). Our outcomes were the combined variable “poor outcome” (non-invasive mechanical ventilation, intubation, intensive care unit admission, and death at 28 days) and death at 28 days. Scores were analysed using univariate logistic regression models, receiver operating characteristic curves, and areas under the curve. Eighty-one patients were included, from which 21 had a poor outcome, and 9 died. We found a statistically significant correlation between poor outcome and NEWS2, LUS > 15, and COWS. Death at 28 days was statistically correlated with NEWS2 and SEIMC Score although COWS also performs well. NEWS2, LUS, and COWS accurately predict poor outcome; and NEWS2, SEIMC Score, and COWS are useful for anticipating death at 28 days. Lung ultrasound is a diagnostic tool that should be included in COVID-19 patients evaluation.



Citation: Gil-Rodríguez, J.; Martos-Ruiz, M.; Peregrina-Rivas, J.-A.; Aranda-Laserna, P.; Benavente-Fernández, A.; Melchor, J.; Guirao-Arrabal, E. Lung Ultrasound, Clinical and Analytic Scoring Systems as Prognostic Tools in SARS-CoV-2 Pneumonia: A Validating Cohort. *Diagnostics* **2021**, *11*, 2211. <https://doi.org/10.3390/diagnostics11122211>

Academic Editor: Po-Hsiang Tsui

Received: 13 October 2021

Accepted: 22 November 2021

Published: 26 November 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: SARS-CoV-2; COVID-19; lung; ultrasound; early warning score

1. Introduction

Pneumonia caused by SARS-CoV-2 (COVID-19) has been an unprecedented health disruption in recent decades. In December 2019, a lung disease outbreak due to a previously unknown pathogen was announced in the city of Wuhan, China. Since then, according to the Coronavirus Resource Center at Johns Hopkins University, as of 1 June 2021, more than 170 million cases and 3.5 million deaths have been reported all around the world [1]. Of these data, more than 50 million cases correspond to Europe and, in particular, more than 3.5 million to Spain [2]. This has resulted in 250,000 hospital admissions in Spain, from which 25,000 went to intensive care units [3].

One of the difficulties in assessment and management lies on knowing which patients will have a more unfavourable evolution. Acute respiratory distress syndrome (ARDS) is the worse consequence of SARS-CoV-2 inflammatory process in the lungs, needing frequent high-flow nasal oxygen therapy or noninvasive mechanical ventilation (NIMV), invasive mechanical ventilation (IMV), or extracorporeal membrane oxygenation (ECMO) in critical cases [4]. This leads to an overuse of intensive care unit resources, extended hospital stays, and a high mortality rate of this population.

Several scoring systems have tried to classify patients at admission [5–9]. Classifying patients has the utility of trying to predict which of them may have a worse outcome in order to provide them an earlier administration of ventilatory support or drugs that change the patient's prognosis. These scores are usually based on clinical and/or analytical data.

However, lung ultrasound has been reported to be a promising tool for the diagnosis and follow-up of pneumonia, interstitial diseases, and ARDS [10,11]. Lung damage originated from COVID-19 is typically patchy and subpleural, which makes COVID-19 specially interesting to be evaluated by lung ultrasound [12]. Its role in the pandemic has so far been secondary, but it has proven to be a useful tool for the stratification of severity and risk of death in these patients [13], as supported by a strong correlation between lung ultrasound and CT, with greater accessibility and speed at lower cost [14].

The aim of our study is to evaluate some of these scoring systems in order to validate them and assess which one of them better predicts a worse outcome in our cohort.

2. Materials and Methods

2.1. Study Design and Setting

We conducted a single-center prospective cohort study to validate four COVID-19 prognosis scores. The study protocol was approved by the regional ethics committee with the code 0259-N-21. This study was written in accordance with the TRIPOD statement for risk prediction models [15]. It was conducted at the San Cecilio University Hospital in Granada (Spain), with a reference population of almost 500,000 people.

2.2. Selection of Participants

This study included consecutive adult patients (age 18 years or older) by quota sampling, with infection confirmed by real-time reverse transcription–polymerase chain reaction (RT–PCR) to be SARS-CoV-2 positive [16], in the first 24 h of admission to the COVID-19 inpatient ward from 1 January 2021 to 20 March 2021.

Patients with suspected or confirmed bacterial pulmonary superinfection, those unable to cooperate with the examination (for physical or cognitive reasons), and those who voluntarily refused to participate were excluded from the study. All patients were given a study summary and signed an informed consent form with the possibility of revocation.

2.3. Measurements

Registration data, anthropometric variables, symptoms and onset date, comorbidities, vital signs, radiological severity, analytical data, and ultrasound lung involvement were compiled. The researchers on the COVID-19 transitional hospitalisation ward collected those clinical variables at the time of evaluation. The Emergency Department requested the analytical data at the first hospital contact, whereas the missing data were completed during the examination.

Ultrasound imaging was performed with an ultrasound scanner PHILIPS® SPARQ (Koninklijke Philips N. V.; Amsterdam, The Netherlands). The predefined lung programme was established, using a convex transducer 6–2 MHz. The Clinical Care Ultrasound Group of the hospital trained 6 Internal Medicine physicians in lung ultrasound, who obtained the images. At least 2 examiners were present when the ultrasound images were acquired in order to reduce the inherent inter-individual variability in this technique. We used a 12 lung fields protocol, as it is the protocol that has shown the best balance between accuracy and acquisition time [17]. The clavicular midline and paravertebral line divided each hemithorax into anterior, lateral, and posterior fields. These fields were further subdivided into superior and posterior fields, resulting in 6 fields in each lung (Figure 1). Prior examination of the patients medical history was avoided to prevent possible biases in image acquisition.

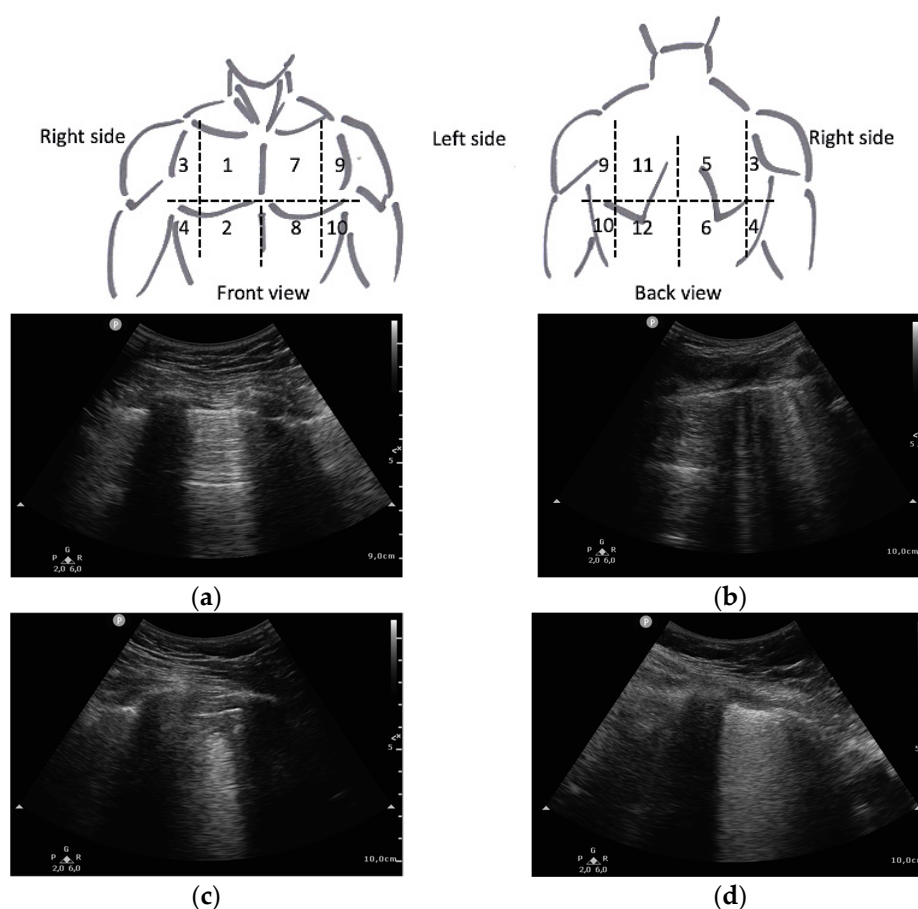


Figure 1. Lung areas distribution and LUS score as described by Soldati et al. [7]: (a) 0 points (normal A-lines pattern); (b) 1 point (>3 B-lines); (c) 2 points (subpleural consolidation); (d) 3 points (pulmonary consolidation or “with lung”).

2.4. Outcomes

Outcomes collected were non-invasive mechanical ventilation or high-flow nasal oxygen therapy (NIMV), intubation (IMV), intensive care unit (ICU) admission, and death at 28 days. The combined outcome variables NIMV or IMV and NIMV or IMV or ICU admission or death at 28 days were also elaborated. Our primary variable was the combined variable poor outcome (NIMV or IMV or ICU admission or death at 28 days) and the secondary variable death at 28 days.

2.5. Scores Selection

The studied scores cover clinical, analytical, and imaging variables associated with poor prognosis in the evolution of COVID-19.

National Early Warning Score (NEWS) 2 [5], recognised in previous studies as a useful prognostic tool and superior to other scales in COVID 19 for both severe illness and in-hospital mortality [18], was used to analyse the prognostic value of clinical data. This takes into account vital signs, such as heart rate, systolic blood pressure, respiratory rate, peripheral oxygen saturation (SaO₂), and body temperature; and it also includes the need for oxygen supply and the level of consciousness.

Furthermore, the most widely recognised international index for lung ultrasound in COVID-19 is the Lung Ultrasound Score (LUS) [11]. It scores from 0 to 3 each of the fields analysed, depending on the severity of the findings in each of them, as shown in Figure 1. The final sum of the score for all fields is used as a prognostic value. It was developed only correlating the score with clinical severity [7].

In addition to the scores mentioned above, two other indexes that include two or more of these categories were analysed. The COVID-19 Worsening Score (COWS) measures both LUS >15 points and clinical and analytical variables of the previously validated COVID-GRAM score [19], including dyspnea, symptom days, arterial oxygen pressure to oxygen inspired fraction ratio, and number of co-morbidities, including dementia, chronic obstructive pulmonary disease (COPD), chronic renal disease (CRD), diabetes mellitus (DM), hypertension, obesity, neoplasm, hepatitis B virus, cerebrovascular disease, cardiovascular disease, heart failure, and immunosuppression. The COWS measures the likelihood of developing severe disease in hospitalised patients [8]. Another prognostic tool is the COVID-19 Spanish Society of Infectious Diseases and Clinical Microbiology score (SEIMC Score), which studies clinical and analytical variables, such as age, low age-adjusted SaO₂, neutrophil to lymphocyte ratio, estimated glomerular filtration rate, dyspnea, and sex in hospitalised patients and correlates it with 30-day mortality [9].

2.6. Analysis

Based on the data available at the beginning of the study, an estimated sample calculation of 22 patients was made to perform mean differences. We performed a simple logistic regression assuming an alpha of 0.05, a beta of 0.8, a probability of the event in the null hypothesis of 0.3, and a probability of the event in the alternative hypothesis of 0.7, with an estimated proportion of the population with the alternative hypothesis of 0.3. We set >15 as the severity cutoff point in the LUS, as used by COWS [8].

Descriptive statistics on the patients subjected to the ultrasonography exploration were performed by means and ranges for discrete variables, median, and inter-quartile ranges for continuous variables. For this analysis, normality was verified using the Shapiro–Wilk test. Non-normal variables means differences were tested using the Mann–Whitney test and the chi-square test for categorical variables. The main statistical analyses were derived in order to explore the relationship between our principal and second outcomes with several indices detailed above, such as NEWS2, LUS, COWS, or SEIMC Score. The methods consisted of univariate logistic regression models, estimating the risk of low index versus high index per unit depending on the classification. Receiver operating characteristic (ROC) curves were calculated to evaluate the predictive performance of these five scores for our primary and secondary outcomes. Areas under the curve (AUC) and 95% confidence intervals (CI) were also calculated. The cut-off point for the AUC was set at 0.7. This value indicates a favourable result, especially for the validation of a new score [20].

All tests were two-sided, and we determined a *p*-value < 0.05 and an AUC > 0.7 as significant. Data derivation procedure was calculated using SSPS software, version 24 (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY, USA: IBM Corp.).

3. Results

3.1. Characteristics of Study Subjects

Between 1 January and 20 March, a total of 769 patients were admitted to the COVID-19 hospitalisation ward of the San Cecilio University Hospital with a positive SARS-CoV-2 PCR, and 656 of those were admitted diagnosed of COVID-19 pneumonia. Of the remaining, 198 had one or more of the exclusion criteria described above. Finally, of the 458 eligible patients, 81 successive patients could be included in the study depending on the availability of human and material resources, as described in Figure 2.

The baseline characteristics of the participants are listed in Appendix A. The median age was 62 years (interquartile range 52–72), being significantly higher in dead patients (73) than in alive patients (59). The majority of participants were male (66.7%) although among the deceased, there was an equal number of males and females (4). Some comorbidities were more frequent in poor outcome and dead patients, such as COPD, CRD, immunosuppression (these three were statistically significant in death outcome), obesity, neoplasia, and heart failure. Radiologically, the median chest X-ray severity according to RALE

was 4 (statistically significant for both main outcomes) and the incidence of pulmonary thromboembolism 8.6%, more frequent in favourable outcome (10% vs. 5%) and death (11.1% vs. 8.3%) without reaching statistical significance. In clinical variables, the median number of symptom days was higher in the living (8 vs. 5) and RR in poor outcome (26 vs. 21, $p < 0.01$), while in poor outcome and death, altered level of consciousness (5% vs. 3.3% and 22.2% vs. 1.3% with $p < 0.01$) and supplementary O₂ (85% vs. 56.7% with $p < 0.05$ and 77.8% vs. 62.5%) were more frequent. Analytically, poor outcome and dead patients presented a higher frequency of levels above the established cut-off points in LDH and CRP scores although lower in D-dimer and neutrophil-lymphocyte ratio. They also had lower PaO₂-FiO₂ ratio and median eGFR ($p < 0.01$ in death outcome), lymphocyte, and platelet values ($p < 0.05$ in both outcomes) while higher NT-proBNP levels ($p < 0.05$ in death outcome).

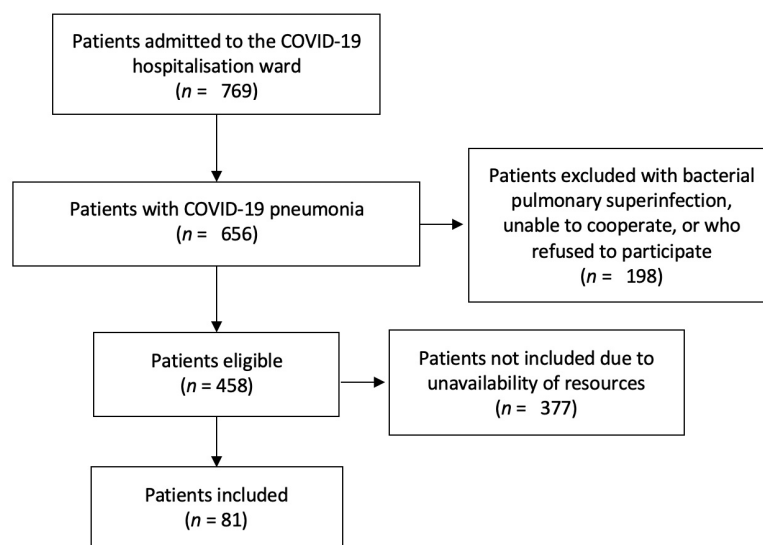


Figure 2. Patient selection flow chart.

3.2. Main Results

Differences between the subgroups for the two main outcomes under study are described in Table 1. In total, 21 patients in our cohort had a poor outcome, and nine patients died. Univariate analysis of each of the scores showed a statistically significant correlation between poor outcome and NEWS2 (OR 1.611, $p < 0.001$), LUS > 15 (OR 3.5, $p = 0.019$), and COWS (OR 3.968, $p = 0.019$). The study of death at 28 days showed statistically significant correlations in NEWS2 (OR 1.315, $p = 0.041$) and SEIMC Score (OR 1.190, $p = 0.008$).

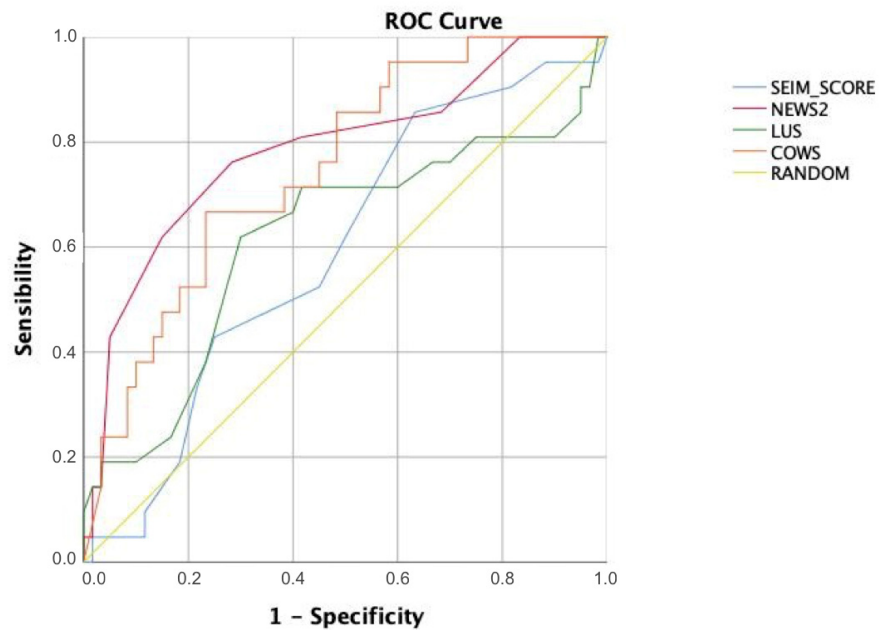
Table 1. Differences between favourable and poor outcome and between alive and dead patients.

	Vs. Poor Outcome		Vs. Dead at 28 Days	
	<i>p</i>	OR (CI)	<i>p</i>	OR (CI)
NEWS2	0.001 **	1.611 (1.228–2.113)	0.041 *	1.315 (1.012–1.710)
LUS > 15	0.019 *	3.5 (1.192–10.275)	0.271	2.235 (0.519–9.636)
SEIMC Score	0.516	1.034 (0.935–1.143)	0.008 **	1.190 (1.046–1.354)
COWS	0.019 *	3.968 (1.251–12.573)	0.243	2.552 (0.530–12.283)

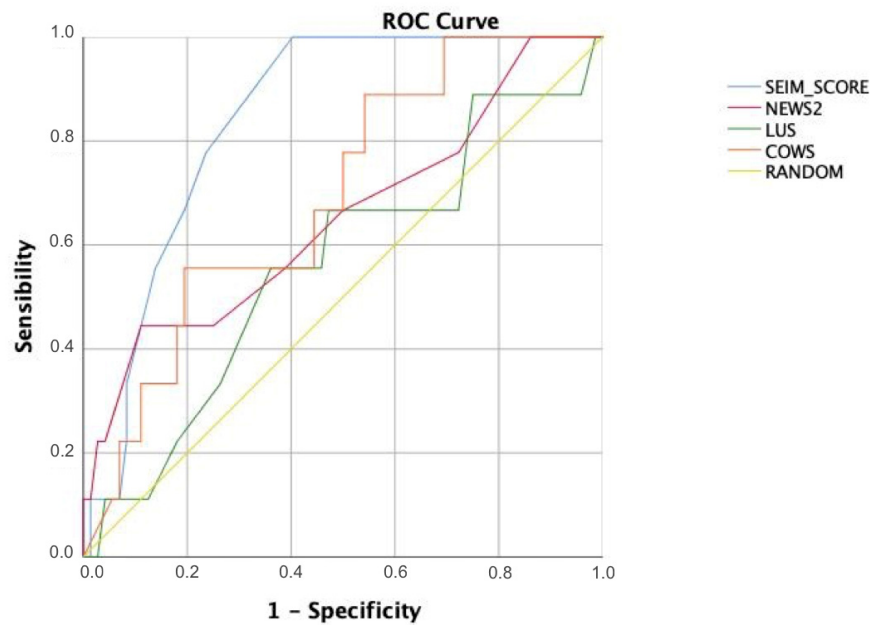
* $p < 0.05$; ** $p < 0.01$.

The performance of each of the scores was represented in the pooled ROC curves (Figure 3), detailing their AUC, optimal cut-point value, sensitivity, and specificity in Table 2. As in the univariate analysis, NEWS (AUC 0.785), LUS > 15 (0.617), and COWS (AUC 0.751) showed the best performance at poor outcome. It also coincides in death at

28 days, with NEWS (AUC 0.654) and SEIMC Score (AUC 0.854) showing the best results although COWS (AUS 0.690) also performs well. Statistical differences between the AUC of the scores were also studied according to DeLong's non-parametric method [21], as shown in Appendix B. As a result, NEWS ($p = 0.028$), SEIMC ($p = 0.004$), and COWS ($p < 0.001$) are superior to LUS as a continuous variable for predicting poor outcome, while COWS is also superior to LUS for predicting death at 28 days ($p = 0.007$). There is also significant superiority of NEWS over SEIMC for predicting poor outcome ($p = 0.049$).



(a)



(b)

Figure 3. Pooled ROC curves for (a) poor outcome and (b) dead at 28 days.

Table 2. Performance of prognostic scores for (a) poor outcome and (b) dead at 28 days.

(a)				
Poor Outcome	AUC	Optimal Cut-Off Point Value	Sensitivity	Specificity
NEWS2	0.785 *	>5	0.619	0.850
LUS	0.617	>17	0.619	0.700
SEIMC Score	0.593	>9	0.333	0.783
COWS	0.751 *	≥0.1007	0.857	0.617
(b)				
Dead 28 Days	AUC	Optimal Cut-Off Point Value	Sensitivity	Specificity
NEWS2	0.654	>5	0.444	0.887
LUS	0.560	>17	0.556	0.634
SEIMC Score	0.840 *	>9	0.667	0.803
COWS	0.690	≥0.1007	0.889	0.431

* AUC > 0.7.

4. Discussion

Most physicians agree on the importance of clinical findings of haemodynamic instability [22] and analytical findings of hyperinflammation syndrome [23] in the management of patients with COVID-19, supported by the classical radiological studies of chest X-ray and chest CT [24]. However, weighing each of these findings separately as a better predictor of the disease outcome proves to be very complicated. In this context, different scores that group several of these variables and assign a weight to them have emerged continuously up to the present time [25–27] as well as alternatives to conventional diagnostic tests, such as lung ultrasound [28]. However, a new question arises when trying to use these indices in clinical practice, where effectiveness in the management of time and resources is essential in a pandemic context. For this reason, we conducted this comparative study between representative scores of the most commonly used variables in medical practice (clinical, analytical, and imaging), avoiding some that have raised doubts about their usefulness [29].

Although several studies have attempted to deal with this problem, some of them have focused on validating previous indices not specific to COVID-19 [30,31], comparing them with new indices developed particularly for this disease [32,33] or comparing scores that evaluate one or two categories of the variables used to assess the severity of the disease [34]. However, none of them have validated COVID-specific indices taking into account the new diagnostic alternative of lung ultrasound or scores derived from it. More and more clinicians are becoming proficient in this technique, and it is no longer a diagnostic test offered by the radiology department but an increasingly accessible tool in the emergency department and on the hospital ward.

The results described above validate many of the original studies that developed the scores analysed. NEWS2 was confirmed as a useful tool both for predicting severe disease and for studying in-hospital death, showing significant results for both outcomes in our cohort. LUS findings had already shown good correlation with disease severity, repeating these results in this study with a cut-off point of 15. The SEIMC Score data for 30-day mortality were also confirmed, being significantly associated as well with 28-day death in our study. Finally, the COWS was developed to measure the probability of developing severe disease and, in our study, not only confirmed this result but also showed an acceptable diagnostic ability for 28-day mortality according to the AUC (however, it did not reach the pre-determined statistical significance), with a high sensitivity for both outcomes.

From the detailed analysis of the results, we appreciate the importance of the cut-off points. Setting the same cut-off point for the two prognostic variables (death and poor prognosis) may imply decreasing the utility, predicting one of the variables in favour of the other. In the case of LUS, on the other hand, it seems that its usefulness as a discrete

variable is less than as an ordinal variable. Categorising into mild and severe groups with a cut-off point of 15 (as used by the COWS) has greater statistical significance in our sample than its AUC. The NEWS is a clinical scale, which measures the impact of COVID-19 at the time of the evaluation. Perhaps for this reason it is a better predictor of poor outcome, as clinical variables are established criteria for starting mechanical ventilation. However, its usefulness in predicting death as an outcome is less in the AUC analysis although it is possible that with greater sample power it would reach statistical significance, as it does in simple logistic regression. The COWS has a similar issue, in which a larger sample size would possibly show statistical significance in the analysis of death as an outcome (AUC 0.69, having used 0.7 as the cut-off point for statistical significance). Finally, the SEIMC Score appears to be an index developed specifically to predict death as an outcome, including the variables that, in the cohort analysed, best predicted this outcome. Specifically, the inclusion of age and eGFR as variables in their calculation may be two of the reasons for this. Our study provides external validity to these results but confirms that it is not a useful score for predicting poor prognosis.

Among the strengths of our study, as an underused clinical tool equal to other more commonly used techniques such as blood tests, it is worth noting the inclusion of lung ultrasound and the scores derived from it. Finally, the outcomes measured included both poor outcome and death at 28 days, in contrast to other studies that assessed only one or the other, and there were been any missing data among participants.

The study was designed with the purpose to reduce the limitation of being an operator-dependent technique, with two or more operators always required to be present during image acquisition and interpretation. However, the biases in lung ultrasound performance by our team of operators, although mitigated by the presence of other colleagues, were still present and cannot be completely ruled out. Potential limitations of the study could arise from the sample size. Although the initial calculation showed us that the sample of 24 patients was sufficient, probably with a larger sample, we could perform a multivariate analysis, including joint clinical and ultrasound parameters, and confirm the clear tendency of some variables to reach statistical significance without reaching it. Participants included among the eligible patients were recruited on a randomised basis, so we believe that the patients included are a representative sample of the total population admitted to the inpatient ward of our hospital and therefore do not represent a bias in itself. In addition, some parameters, such as D-dimer or cardiovascular disease, are more frequent in favourable outcomes than in other studies, which could probably change with more cases favouring regression to the mean. Though a careful selection was made from among the many existing scores, selecting representative examples of the variables most commonly used in clinical practice, the continuous proliferation of these indices makes it impossible to compare all of them. At last, the exclusion of patients with potential bacterial superinfection in the lung may be a bias that eliminates patients with more severe evolution, while the inclusion of lung ultrasound implies an operator-dependent technique that may make reproducibility in other studies difficult.

This study raises some issues that will need to be addressed in further studies. Which type of transducer (convex or linear) is most suitable for lung ultrasound should be defined and demonstrated. Most of these scores have also been developed and validated in the hospitalised population, and studies are needed to test their usefulness in the out-of-hospital setting. Our study indicates the necessity of establishing LUS cut-off points for mild, moderate, and severe disease in future studies. We also consider it necessary to establish guidelines in the elaboration of these scores and to combine those that measure the same variables into indices with greater statistical power.

5. Conclusions

In summary, NEWS2, LUS, and COWS are good scores for predicting poor outcome of COVID-19, while NEWS2, SEIMC Score, and COWS are useful scores for predicting death at 28 days. Both COWS and NEWS are the scores that best predict the evolution of COVID

patients in our cohort. Lung ultrasound is a diagnostic tool that should be included in the evaluation of COVID-19 patients.

Author Contributions: Conceptualization, J.G.-R., J.-A.P.-R. and E.G.-A.; methodology, J.G.-R., J.M. and E.G.-A.; formal analysis, J.M.; investigation, J.G.-R., M.M.-R., J.-A.P.-R., P.A.-L., A.B.-F. and E.G.-A.; data curation, J.M.; writing—original draft preparation, J.G.-R.; writing—review and editing, J.G.-R., M.M.-R., P.A.-L., E.G.-A.; visualization, J.G.-R., M.M.-R., P.A.-L.; supervision, J.G.-R., E.G.-A.; project administration, J.G.-R., J.-A.P.-R., E.G.-A.; funding acquisition, J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by regional ethics committee of CEI/CEIM of Granada (protocol code 0259-N-21, 20 February 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Not applicable.

Acknowledgments: To Valme Sanchez Cabrera, for her support in managing the database, and to Alejandra Ciria García, for her assistance in text edition.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Clinical, analytical and imaging characteristics of participants.

	All Patients (n = 81)	Poor Outcome (n = 21)	Favourable Outcome (n = 60)	p-Value	Dead 28 Days (n = 9)	Alive 28 Days (n = 72)	p-Value
Age (median, IQR)	62 (52–72)	61 (47.5–72.3)	62 (52.8–71.3)	0.550	73 (72–78)	59 (51–70)	0.001 **
Sex							
Male (%)	66.7	75	63.9	0.362	55.6	68.1	0.292
Female (%)	33.3	25	36.1		44.4	31.9	
Weight [kg] (mean, SD)	84 (1.9)	84 (3.3)	84 (2.3)	0.852	83.9 (7.3)	84 (1.9)	0.733
Height [cm] (mean, SD)	168.7 (1.1)	168 (2.1)	168.9 (1.3)	0.653	165 (3.7)	169 (1.2)	0.868
Dementia (%)	2.5	0	3.3	0.412	0	2.7	0.635
COPD¹ (%)	4.9	10	3.3	0.229	33.3	1.4	0.001 **
Chronic renal disease (%)	12.4	15	11.7	0.678	33.3	9.7	0.023 *
Diabetes mellitus (%)	28.4	25	30	0.698	55.6	25	0.153
Hypertension (%)	46.9	30	53.3	0.081	55.6	45.8	0.854
Obesity (%)	34.6	40	33.3	0.600	44.4	33.3	0.505
Neoplasia (%)	15.2	20	13.6	0.414	33.3	12.9	0.064
Hepatitis B virus (%)	0	0	0	-	0	0	-
Cerebrovascular disease (%)	5.1	5	5.1	0.964	0	5.7	0.491

Table A1. Cont.

	All Patients (n = 81)	Poor Outcome (n = 21)	Favourable Outcome (n = 60)	p-Value	Dead 28 Days (n = 9)	Alive 28 Days (n = 72)	p-Value
Cardiovascular disease (%)	8.9	0	11.9	0.119	0	10	0.352
Heart failure (%)	7.6	15	5.1	0.122	22.2	5.7	0.050
Immunosuppression (%)	6.4	10	5.1	0.400	22.2	4.3	0.023 *
Number of comorbidities (median, IQR)	1 (0–2)	1 (0–2)	1 (0–2)	0.402	2 (2–3)	1 (0–2)	0.006 **
Days of symptoms (median, IQR)	8 (5–10)	7 (4.75–9)	9 (6–10)	0.071	5 (4–6)	8 (7–10)	0.057
Dyspnea (%)	69.1	75	66.7	0.513	66.7	69.4	0.705
Cough (%)	76.5	75	76.7	0.851	66.7	77.8	0.914
Altered consciousness (%)	3.7	5	3.3	0.724	22.2	1.3	0.001 **
Fever (%)	44.4	50	41.7	0.466	44.4	44.4	0.598
SaO₂² [%] (mean, SD)	96.2 (0.2)	95.2 (0.6)	96.6 (0.2)	0.012 *	95 (0.9)	96.4 (0.24)	0.014 *
Supplementary O₂ (%)	64.2	85	56.7	0.025 *	77.8	62.5	0.502
RR³ [breaths/min] (median, IQR)	24 (20–28)	26 (24–28)	21 (18–24.5)	0.009 **	24 (22–28)	24 (21–28)	0.245
HR⁴ [beats/min] (median, IQR)	81 (70–90)	85 (79.5–96)	76 (67.8–89)	0.030 *	82 (73–86)	81 (70–90)	0.692
SBP⁵ (mmHg) (mean, SD)	127.3 (1.7)	127.5 (2.9)	127.2 (2.1)	0.939	130.9 (7.3)	126.8 (1.7)	0.893
Rx. thorax [RALE⁶] (median, IQR)	4 (3–5)	4 (3–5)	4 (3–5)	0.019 *	5 (5–8)	4 (3–5)	0.001 **
CT PTE⁷ (%)	8,6	5	10,0	0.504	11.1	8.3	0.682
PaO₂-FiO₂⁸ ratio (mean, SD)	366.3 (10.8)	306.2 (23)	388.3 (11.1)	0.001 **	289.9 (30.8)	373.1 (10.8)	0.015 *
Glucose [mg/dL] (mean, SD)	147.2 (7.5)	126.75 (9.5)	154 (9.4)	0.139	160.7 (22.2)	145.5 (8)	0.862
eGFR⁹ [ml/min/1.73 m ²] (mean, SD)	80 (2.8)	75 (5.9)	81 (3.1)	0.330	52.6 (7.4)	83.4 (2.7)	0.001 **
AST¹⁰ [U/L] (mean, SD)	52.6 (6.3)	66 (17.7)	48.6 (6.1)	0.920	39.2 (6.7)	47.6 (6.5)	0.837
Ferritin [ng/mL] (mean, SD)	824.6 (80.7)	745.8 (132.7)	848.4 (99.3)	0.814	681.5 (183.3)	832.5 (87.7)	0.914
>700 ng/mL (%)	44.4	45	43.3	0.852	33.3	45.2	0.731
LDH¹¹ [U/L] (mean, SD)	337.2 (13.3)	365.9 (33.4)	329.6 (14)	0.738	339.8 (44.9)	336.9 (14)	0.849
>400 U/L (%)	21	30	18.3	0.254	22.2	20.8	0.769
CRP¹² [mg/L] (mean, SD)	91.5 (6.8)	84.2 (14.1)	92.4 (7.7)	0.473	91.5 (23)	91.5 (7.1)	0.681
>15 mg/mL (%)	95.1	100	93.3	0.240	100	94.4	0.497
Lymphocyte [10³/μL] (mean, SD)	0.87 (0.05)	0.83 (0.1)	0.88 (0.06)	0.763	0.7 (0.15)	0.89 (0.05)	0.101
Neutrophil-lymphocyte ratio (mean, SD)	6.4 (0.46)	7.3 (0.8)	6.1 (0.6)	0.540	8.6 (2.1)	6.1 (0.4)	0.264
>10 (%)	17.3	7.31	16.7	0.711	11.1	18.1	0.706

Table A1. Cont.

	All Patients (n = 81)	Poor Outcome (n = 21)	Favourable Outcome (n = 60)	p-Value	Dead 28 Days (n = 9)	Alive 28 Days (n = 72)	p-Value
Platelets [$10^3/\mu\text{L}$] (mean, SD)	200.6 (8.3)	163.4 (12.7)	214 (9.9)	0.008 **	152.6 (20.6)	206.6 (8.8)	0.020 *
D dimer [mg/L] (mean, SD)	1.68 (0.5)	1.1 (0.3)	1.9 (0.6)	0.144	4.4 (3.9)	1.3 (0.2)	0.168
>15 mg/mL (%)	19.7	10	23.3	0.538	11.1	20.8	0.587
NT-proBNP ¹³ [pg/mL] (mean, SD)	815 (171.2)	1007.7 (372.5)	741.8 (188.7)	0.180	2086.9 (731.9)	593.8 (136.6)	0.038 *
NEWS2 (median, IQR)	4 (2–6)	6 (5–7)	3 (2–5)	0.001 **	5 (3–7)	3 (2–6)	0.041 *
ECO LUS total (median, IQR)	15 (10–19)	18 (13.5–19)	14 (9.8–18.3)	0.177	18 (10–19)	15 (10–19)	0.541
>15 (%)	49.4	70	43.3	0.019 *	66.7	47.2	0.271
SEIMC Score (median, IQR)	6 (4–10)	6 (5–10)	6 (4–9)	0.516	11 (9–14)	5 (4–8)	0.008 **
COWS (mean, SD)	0.52 (0.05)	0.72 (0.10)	0.47 (0.06)	0.019 *	0.71 (0.15)	0.5 (0.06)	0.243

¹ COPD: Chronic obstructive pulmonary disease; ² SaO₂: Peripheral arterial oxygen saturation; ³ RR: Respiration rate; ⁴ HR: Heart rate; ⁵ SBP: Systolic blood pressure; ⁶ RALE: Radiographic Assessment of Lung (O)Edema; ⁷ CT PTE: Computer Tomography Pulmonary Thromboembolism; ⁸ PaO₂-FiO₂: Arterial oxygen pressure to oxygen inspired fraction ratio; ⁹ eGFR: Estimated glomerular filtration rate calculated by the CKD-EPI; ¹⁰ AST: Aspartate aminotransferase; ¹¹ LDH: Lactate dehydrogenase; ¹² CRP: C-reactive protein; ¹³ NT-proBNP: N-terminal pro-brain natriuretic peptide. (* $p < 0.05$ and ** $p < 0.01$).

Appendix B

Table A2. Statistical differences between the AUC of the scores according to DeLong's non-parametric method [21].

	Poor Outcome p-Value (CI)	Dead 28 Days p-Value (CI)
LUS vs. NEWS2	0.028 * (−0.319–0.018)	0.507 (−0.357–0.176)
LUS vs. SEIMC	0.004 ** (−0.471–0.090)	0.812 (−0.172–0.220)
LUS vs. COWS	< 0.001 ** (−0.201–0.067)	0.007 ** (−0.226–0.036)
COWS vs. NEWS2	0.590 (−0.160–0.233)	0.676 (−0.151–0.091)
COWS vs. SEIMC	0.059 (−0.006–0.322)	0.081 (−0.318–0.018)
NEWS2 vs. SEIMC	0.049 * (0.0003–0.385)	0.126 (−0.435–0.054)

* $p < 0.05$; ** $p < 0.01$.

References

- COVID-19 Map—Johns Hopkins Coronavirus Resource Center. Available online: <https://coronavirus.jhu.edu/map.html> (accessed on 1 June 2021).
- Ministerio de Sanidad, Consumo y Bienestar Social—Profesionales—Situación Actual Coronavirus. Available online: <https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/situacionActual.htm> (accessed on 10 June 2021).
- Instituto de Salud Carlos III. Informe No^o 81. In *Situación de COVID-19 en España a 2 de Junio de 2021*; Ministerio de Ciencia. e Innovación: Madrid, Spain, 2021; Volume 1.

4. Budinger, G.R.S.; Misharin, A.V.; Ridge, K.M.; Singer, B.D.; Wunderink, R.G. Distinctive Features of Severe SARS-CoV-2 Pneumonia. *J. Clin. Investig.* **2021**, *131*, e149412. [[CrossRef](#)]
5. Royal College of Physicians. *National Early Warning Score (NEWS) 2; Standardising the Assessment of Acute-Illness Severity in the NHS Updated Report of a Working Party Executive Summary and Recommendations*; Royal College of Physicians: London, UK, 2017; pp. 1–15.
6. Webb, B.J.; Peltan, I.D.; Jensen, P.; Hoda, D.; Hunter, B.; Silver, A.; Starr, N.; Buckel, W.; Grisel, N.; Hummel, E.; et al. Clinical Criteria for COVID-19—Associated Hyperinflammatory Syndrome: A Cohort Study. *Lancet Rheumatol.* **2020**, *2*, 754–763. [[CrossRef](#)]
7. Soldati, G.; Smargiassi, A.; Inchingolo, R.; Buonsenso, D.; Perrone, T.; Briganti, D.F.; Perlini, S.; Torri, E.; Mariani, A.; Mossolani, E.E.; et al. Proposal for International Standardization of the Use of Lung Ultrasound for Patients with COVID-19. *J. Ultrasound Med.* **2020**, *39*, 1413–1419. [[CrossRef](#)] [[PubMed](#)]
8. Boero, E.; Rovida, S.; Schreiber, A.; Berchiolla, P.; Charrier, L.; Cravino, M.M.; Converso, M.; Gollini, P.; Puppo, M.; Gravina, A.; et al. The COVID-19 Worsening Score (COWS)—A Predictive Bedside Tool for Critical Illness. *Echocardiography* **2021**, *38*, 207–216. [[CrossRef](#)]
9. Berenguer, J.; Borobia, A.M.; Ryan, P.; Rodríguez-Banõ, J.; Bellón, J.M.; Jarrín, I.; Carratalà, J.; Pachón, J.; Carcas, A.J.; Yllescas, M.; et al. Development and Validation of a Prediction Model for 30-Day Mortality in Hospitalised Patients with COVID-19: The COVID-19 SEIMC Score. *Thorax* **2021**, *76*, 1–10. [[CrossRef](#)]
10. Dietrich, C.F.; Mathis, G.; Blaiwas, M.; Volpicelli, G.; Seibel, A.; Atkinson, N.S.; Cui, X.W.; Mei, F.; Schreiber-Dietrich, D.; Yi, D. Lung Artefacts and Their Use. *Med. Ultrason.* **2016**, *18*, 488–499. [[CrossRef](#)] [[PubMed](#)]
11. Allinovi, M.; Parise, A.; Giacalone, M.; Amerio, A.; Delsante, M.; Odone, A.; Franci, A.; Gigliotti, F.; Amadasi, S.; Delmonte, D.; et al. Lung Ultrasound May Support Diagnosis and Monitoring of COVID-19 Pneumonia. *Ultrasound Med. Biol.* **2020**, *46*, 2908–2917. [[CrossRef](#)]
12. Casella, F.; Barchiesi, M.; Leidi, F.; Russo, G.; Casazza, G.; Valerio, G.; Torzillo, D.; Ceriani, E.; Medico, M.D.; Brambilla, A.M.; et al. Lung Ultrasonography: A Prognostic Tool in Non-ICU Hospitalized Patients with COVID-19 Pneumonia. *Eur. J. Intern. Med.* **2021**, *85*, 34–40. [[CrossRef](#)]
13. Song, G.; Qiao, W.; Wang, X.; Yu, X. Association of Lung Ultrasound Score with Mortality and Severity of COVID-19: A Meta-Analysis and Trial Sequential Analysis. *Int. J. Infect. Dis.* **2021**, 603–609. [[CrossRef](#)]
14. Nouvenne, A.; Zani, M.D.; Milanese, G.; Parise, A.; Baciarello, M.; Bignami, E.G.; Odone, A.; Sverzellati, N.; Meschi, T.; Ticinesi, A. Lung Ultrasound in COVID-19 Pneumonia: Correlations with Chest CT on Hospital Admission. *Respiration* **2020**, *99*, 617–624. [[CrossRef](#)]
15. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G.M. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *J. Br. Surg.* **2015**, *102*, 148–158. [[CrossRef](#)]
16. Kevadiya, B.D.; Machhi, J.; Herskovitz, J.; Oleynikov, M.D.; Blomberg, W.R.; Bajwa, N.; Soni, D.; Das, S.; Hasan, M.; Patel, M.; et al. Diagnostics for SARS-CoV-2 Infections. *Nat. Mater.* **2021**, *20*, 593–605. [[CrossRef](#)] [[PubMed](#)]
17. Mento, F.; Perrone, T.; Macioce, V.N.; Tursi, F.; Buonsenso, D.; Torri, E.; Smargiassi, A.; Inchingolo, R.; Soldati, G.; Demi, L. On the Impact of Different Lung Ultrasound Imaging Protocols in the Evaluation of Patients Affected by Coronavirus Disease 2019. *J. Ultrasound Med.* **2021**, *40*, 2235–2238. [[CrossRef](#)]
18. Myrstad, M.; Ihle-Hansen, H.; Tveita, A.A.; Andersen, E.L.; Nygård, S.; Tveit, A.; Berge, T. National Early Warning Score 2 (NEWS2) on Admission Predicts Severe Disease and in-Hospital Mortality from COVID-19—A Prospective Cohort Study. *Scand. J. Trauma Resusc. Emerg. Med.* **2020**, *28*, 66. [[CrossRef](#)] [[PubMed](#)]
19. Liang, W.; Liang, H.; Ou, L.; Chen, B.; Chen, A.; Li, C.; Li, Y.; Guan, W.; Sang, L.; Lu, J.; et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients with COVID-19. *JAMA Intern. Med.* **2020**, *180*, 1081–1089. [[CrossRef](#)] [[PubMed](#)]
20. Swets, J.A. Measuring the Accuracy of Diagnostic Systems. *Science* **1988**, *240*, 1285–1293. [[CrossRef](#)]
21. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837. [[CrossRef](#)]
22. Jasiński, T.; Stefaniak, J. COVID-19 and Haemodynamic Failure: A Point of View on Mechanisms and Treatment. *Anaesthesiol. Intensive Ther.* **2020**, *52*, 409–417. [[CrossRef](#)]
23. Manson, J.J.; Crooks, C.; Naja, M.; Ledlie, A.; Goulden, B.; Liddle, T.; Khan, E.; Mehta, P.; Martin-Gutierrez, L.; Waddington, K.E.; et al. COVID-19-Associated Hyperinflammation and Escalation of Patient Care: A Retrospective Longitudinal Cohort Study. *Lancet Rheumatol.* **2020**, *2*, 594–602. [[CrossRef](#)]
24. Schalekamp, S.; Bleeker-Rovers, C.P.; Beenen, L.F.M.; Quarles van Ufford, H.M.E.; Gietema, H.A.; Lauran Stöger, J.; Harris, V.; Reijers, M.H.E.; Rahamat-Langendoen, J.; Korevaar, D.A.; et al. Chest CT in the Emergency Department for Diagnosis of COVID-19 Pneumonia: Dutch Experience. *Radiology* **2021**, *298*, 98–106. [[CrossRef](#)]
25. Liu, H.; Chen, J.; Yang, Q.; Lei, F.; Zhang, C.; Qin, J.-J.; Chen, Z.; Zhu, L.; Song, X.; Bai, L.; et al. Development and Validation of a Risk Score Using Complete Blood Count to Predict In-Hospital Mortality in COVID-19 Patients. *Med* **2021**, *2*, 435–447. [[CrossRef](#)]
26. Grifoni, E.; Valoriani, A.; Cei, F.; Vannucchi, V.; Moroni, F.; Pelagatti, L.; Tarquini, R.; Landini, G.; Masotti, L. The CALL Score for Predicting Outcomes in Patients With COVID-19. *Clin. Infect. Dis.* **2021**, *72*, 182–183. [[CrossRef](#)]

27. Alfaro-Martínez, J.J.; Calbo Mayo, J.; Molina Cifuentes, M.; Abizanda Soler, P.; Guillén Martínez, S.; Rodríguez Marín, Y.; Sirvent Segovia, A.E.; Nuñez Ares, A.; Alcaraz Barcelona, M.; Paterna Mellinas, G.; et al. Generation and Validation of In-Hospital Mortality Prediction Score in COVID-19 Patients: Alba-Score. *Curr. Med. Res. Opin.* **2021**, *37*, 719–726. [[CrossRef](#)] [[PubMed](#)]
28. Sorlini, C.; Femia, M.; Nattino, G.; Bellone, P.; Gesu, E.; Francione, P.; Paternò, M.; Grillo, P.; Ruffino, A.; Bertolini, G.; et al. The Role of Lung Ultrasound as a Frontline Diagnostic Tool in the Era of COVID-19 Outbreak. *Intern. Emerg. Med.* **2021**, *16*, 749–756. [[CrossRef](#)] [[PubMed](#)]
29. Yoshioka, T.; Funada, S.; Luo, Y. The Inherent Problems with the Generalizability of the CALL Score: Towards Reliable Clinical Prediction Models for Coronavirus Disease 2019 (COVID-19). *Clin. Infect. Dis.* **2020**, *73*, e2818. [[CrossRef](#)] [[PubMed](#)]
30. García Clemente, M.M.; Herrero Huertas, J.; Fernández Fernández, A.; De La Escosura Muñoz, C.; Enríquez Rodríguez, A.I.; Pérez Martínez, L.; Gómez Mañas, S.; Iscar Urrutia, M.; López González, F.J.; Madrid Carbajal, C.J.; et al. Assessment of Risk Scores in COVID-19. *Int. J. Clin. Pract.* **2020**, e13705. [[CrossRef](#)]
31. Saleh, A.; Ahmed, M.; Abdel-lateif, A. Comparison of the Mortality Prediction of Different ICU Scoring Systems (APACHE II and III, SAPS II, and SOFA) in Acute Respiratory Distress Syndrome Patients. *Chest* **2016**, *149*, A147. [[CrossRef](#)]
32. Shi, Y.; Pandita, A.; Hardesty, A.; McCarthy, M.; Aridi, J.; Weiss, Z.F.; Beckwith, C.G.; Farmakiotis, D. Validation of Pneumonia Prognostic Scores in a Statewide Cohort of Hospitalised Patients with COVID-19. *Int. J. Clin. Pract.* **2021**, *75*, e13926. [[CrossRef](#)]
33. Holten, A.R.; Nore, K.G.; Tveiten, C.E.V.W.K.; Olasveengen, T.M.; Tonby, K. Predicting Severe COVID-19 in the Emergency Department. *Resusc. Plus* **2020**, *4*, 100042. [[CrossRef](#)]
34. Kamran, S.M.; Mirza, Z.-H.; Moeed, H.A.; Naseem, A.; Hussain, M.; Fazal, I.; Saeed, F.; Alamgir, W.; Saleem, S.; Riaz, S. CALL Score and RAS Score as Predictive Models for Coronavirus Disease 2019. *Cureus* **2020**, *12*, 11368. [[CrossRef](#)]