pagepress

# Usefulness of CURB-65, pneumonia severity index and MuLBSTA in predicting COVID-19 mortality

**Carlo Preti[1,2], Roberta Biza[2,3], Luca Novelli[3], Arianna Ghirardi[4], Caterina Conti[3], Chiara Galimberti[2], Lorenzo Della Bella[1,2], Irdi Memaj[1,2], Fabiano Di Marco[2,3], Roberto Cosentini[1]**

*[1]Emergency Department, ASST Papa Giovanni XXIII, Bergamo; [2]University of Milan; [3]Pulmonary Medicine Unit, ASST Papa Giovanni XXIII, Bergamo; [4]Fondazione per la Ricerca Ospedale di Bergamo (FROM), Bergamo, Italy*

## Abstract

The aim of our study is to evaluate the accuracy of CURB-65 and Pneumonia Severity Index (PSI), the most widely used scores for community acquired pneumonia, and MuLBSTA, a viral pneumonia score, in predicting 28-day mortality in Coronavirus Disease 2019 (COVID-19) pneumonia. We retrospectively collected clinical data of consecutive patients with laboratory-confirmed COVID-19 pneumonia admitted at Papa Giovanni XXIII Hospital from February 23rd to March 14th, 2020. We calculated at Emergency Department (ED) presentation CURB-65, PSI and MuLBSTA and we compared their performances in discriminating between survivors and non-survivors at 28 days. Among 431 hospitalized patients, the majority presented with hypoxic respiratory failure: median (interquartile range, IQR) $PaO_2/FiO_2$ ratio at admission was 228.6 (142.0-278.1). In the first 24 hours, 111 (27%) patients were administered low-flow oxygen cannula, 50 (12%) Venturi Mask, 95 (23%) non-rebreather mask, 106 (26%) non-invasive ventilation, 12 (3%) mechanical ventilation and 41 (9%) were not administered oxygen therapy. Mortality rate at 28-day was 35% (150/431). Between survivors and non-survivors, median (IQR) scores were, respectively, 1.0 (1.0-2.0) and 2.0 (2.0-3.0) for CURB-65 (p<0.001); 90.5 (76.0-105.5) and 115.0 (100.0-129.0) for PSI (p<0.001); 7.0 (5.0-10.0) and 11.0 (9.0-13.0) for MuLBSTA (p<0.001). Areas under the receiver operating characteristic curve (AUCs) for each score were, respectively, 0.725 (0.662-0.787), 0.776 (0.693-0.859) and 0.743 (0.680-0.806) (p>0,05). PSI and MuLBSTA did not show a better performance when compared to CURB-65. Although CURB-65, PSI and MuLBSTA scores are useful tools to discriminate between survivors and non-survivors in COVID-19 pneumonia, their diagnostic accuracy in discriminating 28-day mortality in COVID-19 pneumonia is moderate, as confirmed by AUCs <0.80, and there is a potential underestimation of disease severity in the low-risk classes. For this reason, they should not be recommended in ED to decide between inpatient and outpatient management in patients affected by COVID-19 pneumonia.

Correspondence: Luca Novelli MD, Pulmonary Medicine Unit, ASST Papa Giovanni XXIII, 24127 Bergamo, Italy.
Tel. +39.035.29673456, Mobile: +39.339.4443927.
E-mail: lnovelli@asst-pg23.it

Key words: COVID-19; CURB-65; pneumonia severity index; MuLBSTA; respiratory failure.

Contributions: CP, LN, FD, RC, conceived the study; LN, RB, CG, supervised the conduct of the trial and data collection; IM, LDB, CP, RB, undertook patient recruitment and managed the data, including quality control; AG, provided statistical advice on study design and analyzed the data; FD, RC, chaired the data oversight committee; CP, RB, drafted the manuscript, and all authors contributed substantially to its revision and approved the final version. All the authors read and approved the manuscript. LN takes responsibility for the paper as a whole.

Conflict of interest: The authors declare that they have no competing interests, and all authors confirm accuracy.

Ethics approval and informed consent: The local Institutional Ethics Committee approved the study protocol (n. 37/2020). Informed consent was collected as requested by the local ethics board.

Received for publication: 17 August 2021.
Accepted for publication: 14 February 2022.

Publisher's note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.


*Licensee PAGEPress, Italy*
*Monaldi Archives for Chest Disease 2022; 92:2054*
*doi: 10.4081/monaldi.2022.2054*



## Introduction

Since severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection had spread around the world, emergency departments (EDs) had to face a sudden wave of patients who presented with acute hypoxemic respiratory failure, thus needing any degree of oxygen therapy and ventilatory support. This situation hardly challenged healthcare systems, and EDs in particular, in order to provide optimal care for the high burden of severely ill patients.

From the end of February to April 2020, the City of Bergamo, in Northern Italy, was one of the most hit places in the world, with a huge number of deaths and people admitted to Papa Giovanni XXIII Hospital in that limited period of time.

In this kind of overwhelming scenarios, discriminating who could be safely treated at home and who needs hospitalization becomes crucial, in order to reduce pressure on the EDs, optimize allocation of medical resources and to prevent hospital-related virus transmission. Thus, rapid assessment tools as algorithms and scores, could potentially be helpful in allocating limited resources efficiently.

Given the premises, we tried to understand if common prognostic scales for community-acquired pneumonia (CAP) could be useful in the Coronavirus Disease 2019 (COVID-19) setting. We investigated the accuracy of Pneumonia Severity Index (PSI), CURB-65 and MuLBSTA, a multivariable score for viral pneumonia, in predicting the 28-day all-cause mortality rate in a cohort of patients admitted to Papa Giovanni XXIII Hospital in Bergamo, Italy, during the first two weeks of the Italian forefront of the pandemic.

## Materials and Methods

### Study design, population criteria, data collection and outcomes

We retrospectively selected consecutive adult patients, not previously hospitalized for other conditions, who were admitted with laboratory-confirmed SARS-CoV-2 pneumonia to Papa Giovanni XXIII hospital in Bergamo, at the peak of the outbreak, from February 23rd to March 14th, 2020. Epidemiological, clinical, laboratory and radiographic data were collected from electronic medical records. The Regional Healthcare Information System (SISS, Lombardy Region, Italy) was investigated about patients' survival status: follow-up was suspended on 15th May 2020.

At the presentation, blood gas analysis, routine blood exams and chest X-ray (CXR) were performed. CXRs were performed also by portable CXR for practical reasons. The most invasive respiratory support administered during the first 24 hours was recorded. Clinical diagnosis was made upon the updated World Health Organization (WHO) guidelines and confirmed by two different real time polymerase chain reaction-based methods (GeneFinder™ COVID-19, Elitech Group, and Allplex™ 2019-nCoV Assay, Seegene Inc) on respiratory samples, according to WHO protocol. All patients were treated according to the updated hospital protocol in force at that time, with antibiotics, hydroxychloroquine, lopinavir/ritonavir, enoxaparin and steroids. Primary outcome was all-cause mortality at 28-day, either during hospital stay or after discharge.

The local institutional ethics committee approved the study protocol (n. 37/2020).

### Statistical analysis

Descriptive statistics were used to summarize the baseline characteristics of COVID-19 patients. Continuous variables were expressed as mean and standard deviation (SD) or as median and interquartile range (IQR). Categorical variables were expressed as absolute counts and percentages. CURB-65, PSI and MuLBSTA were calculated only for the enrolled patients for which scores' variables were available. All the variables needed to calculate the scores were obtained at admission to the ED.

Receiver operating characteristic (ROC) analysis was used to evaluate the performance of the scores in discriminating between survivors and non-survivors at 28 days using area under the curve (AUC) value. Sensitivity and specificity values were obtained for each score at the optimal cut-off identified following Youden's index. The DeLong *et al*. method [1] for correlated samples was used to compare AUCs for the three scores. Kaplan-Meier 28-day survival curves were reported stratifying patients according to existing score cut-offs (if already defined) or according to the distribution tertiles. The corresponding curves were compared using the log-rank test.

To overcome the constraint of biased results due to missing data, multiple imputation by chained equation (MICE) was used to assign the missing covariates in each score (number of imputations = 10). Statistical analysis was performed using STATA software, release 16 (StataCorp LP, College Station TX, USA). All tests were two-sided and a $p<0.05$ was considered significant.

## Results

In the first three weeks of the pandemic, 431 adult patients with laboratory-confirmed SARS-CoV-2 pneumonia were admitted to our hospital; 119 (27.6%) were female. Caucasian ethnicity was predominant (98.6%). The mean (± SD) age was 67.6±13 years, with no significant differences between female and male (p=0.54). The most common comorbidities were systemic arterial hypertension (55.8%) and diabetes (19.8%); 27% of the patients had a BMI >30.

## Clinical, biochemical and radiological features at presentation

Fever was the most frequent symptom of onset at home (90%), followed by dyspnea (59%) and cough (50%). The median (IQR) interval between symptoms onset and ER admission was 7.0 days (5.0-10.0).

Most of the patients presented with normal state of consciousness, blood pressure and heart rate and a hypoxic respiratory failure: at arterial blood gas analysis, median (IQR) $PaO_2/FiO_2$ ratio was 228.6 (142.0-278.1) without statistically significant difference between male and female (p=0.28), median pH was 7.47 (7.44-7.50), median $PaCO_2$ was 33 (30-35) mmHg and median $HCO_3^-$ was 24.1 (22.0-26.0) mmol/L.

At the chest X-ray, bilateral pneumonia was the most prevalent abnormality detected at presentation (74%).

Blood count was normal, except lymphopenia which was very common at presentation with a median (IQR) count of 891 (582.7-1234.1) x $10^9$/L. Inflammatory markers, including C-reactive protein and procalcitonin, were generally elevated 113 (56-162) mg/L and 0.49 (0.12-1.90) mg/L, respectively. Urea and creatinine were 45.0 (34.0-66.0) and 0.92 (0.77-1.23).

## Respiratory support in the first 24 hours

Among all the analyzed patients, 301 (70%) were admitted to the general clinical wards, 84 (19%) to the Intensive Care Unit and 46 (11%) to Semi-intensive Care Unit.

During the first 24 hour in ED, 111 (27%) patients were administered low-flow oxygen cannula, 50 (12%) Venturi Mask, 95 (23%) non-rebreather mask, 106 (26%) non-invasive ventilation, 12 (3%) mechanical ventilation and 41 (9%) did not need any respiratory support.

## Clinical outcomes

The overall 28-day mortality rate was 35% (150/431). When stratified by age, death occurred in 9 (8%) patients who were ≤ 59 years old, 23 (22%) patients who were 60-69 years old, 51 (45%) 70-77 years old patients and 67 (64%) patients who were ≥78 years old. No gender differences were noted for 28-day mortality (p=0.091).

## CURB-65, PSI and MuLBSTA

It was possible to calculate CURB-65 for 226 of 431 patients involved in the study. The median (IQR) CURB-65 score was 2.0 (1.0-2.0). There was a statistically significant difference between non-survivors and survivors in median (IQR) [2.0 (2.0-3.0) *vs* 1.0 (1.0-2.0)] scores; p-value was <0.001. One hundred and ten (110) patients were considered at low risk (defined as CURB-65 score 0 or 1) and 19 (17%) of them died; 116 patients were in the high-risk group (defined as CURB-65 score 2 or more) and 59 (51%) patients died.

PSI was calculated for 122 patients with complete data. Median (IQR) score was 100.5 (85.0-118.0). As for CURB-65, there was a statistically significant difference in median (IQR) scores between non-survivors [115.0 (100.0-129.0)] and survivors [90.5 (76.0-105.5)], with a p<0.001. Among 14 patients who were in risk class I, 1 patient died.

MuLBSTA was calculated for 259 patients. Median (IQR) score was 9.0 (6.0-12.0). Once again, the two groups of patients showed a statistically significant difference in the corresponding median (IQR) [11.0 (9.0-13.0) *vs* 7.0 (5.0-10.0)] scores, with a p<0.001 (Table 1).

AUC values of the three scores were 0.725 (0.662-0.787) for CURB-65, 0.776 (0.693-0.859) for PSI and 0.743 (0.680-0.806) for MuLBSTA. Comparison between the three AUCs did not reach statistical significance. Scores' stratification in classes did not change their performance: AUCs did not deviate from the previously reported results (0.735, 0.760 and 0.707, respectively). Multiple imputations were calculated to overcome the missing data, but this did not substantially change the results (Figures 1 and 2).

## Discussion

Our aim was to evaluate the performances of the CAP-related scores and MuLBSTA in predicting 28-day mortality in COVID-19 patients and their usefulness in helping the physician to decide the best management setting.

CURB-65 and PSI are the most common scores created for assessing severity and determining the appropriate site of care in CAP. Their role is mainly in the EDs, where they are used to guide clinical decisions about which is the best setting to continue patient management. A more accurate description can be found in the Supplementary Material.

Nonetheless the mentioned scores showed statistically significant differences in discriminating survivors and non survivors in our cohort, their clinical usefulness is not as much clear, as demonstrated by a moderate accuracy in assessing 28-day mortality with AUCs <0.8, and a clinically unacceptable number of deaths observed in low-risk classes.

Referring to CURB-65 and PSI, specifically created tools for

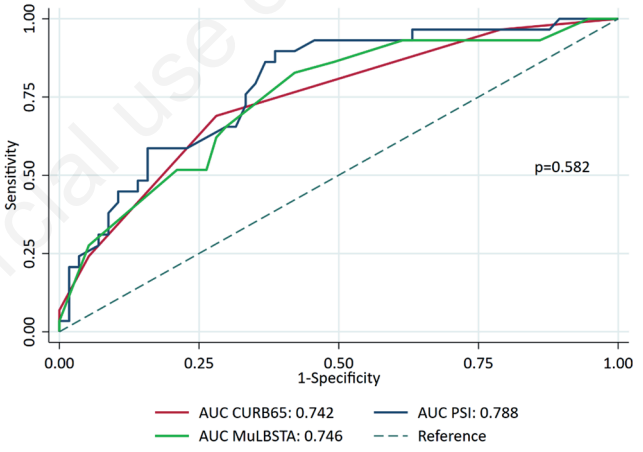


**Figure 1. Comparison of the predictive performance of CURB-65, PSI e MuLBSTA.**

**Table 1. Comparison of the scores according to 28-day mortality.**

| | n | Total n=431 | 28-day mortality No (n=281) | 28-day mortality Yes (n=150) | p | AUC (95% CI) | Best cut off | SE | SP |
|---|---|---|---|---|---|---|---|---|---|
| CURB-65*, median (IQR) | 226 | 2.0 (1.0-2.0) | 1.0 (1.0-2.0) | 2.0 (2.0-3.0) | <0.001 | 0.725 (0.662 – 0.787) | 1 | 0.76 | 0.61 |
| 0-1, n(%) | 226 | 110 (48.7) | 91 (61.5) | 19 (24.4) | <0.001 | | | | |
| 2, n(%) | | 74 (32.7) | 39 (26.4) | 35 (44.9) | | | | | |
| 3+, n(%) | | 42 (18.6) | 18 (12.2) | 24 (30.8) | | | | | |
| PSI*, median (IQR) | 122 | 100.5 (85.0-118.0) | 90.5 (76.0-105.5) | 115.0 (100.0-129.0) | <0.001 | 0.776 (0.693 – 0.859) | 105 | 0.63 | 0.80 |
| Risk class I, n(%) | 122 | 14 (11.5) | 13 (17.1) | 1 (2.2) | <0.001 | | | | |
| Risk class II, n(%) | | 29 (23.8) | 25 (32.9) | 4 (8.7) | | | | | |
| Risk class III, n(%) | | 63 (51.6) | 33 (43.4) | 30 (65.2) | | | | | |
| Risk class IV, n(%) | | 16 (13.1) | 5 (6.6) | 11 (23.9) | | | | | |
| MuLBSTA#, median (IQR) | 259 | 9.0 (6.0-12.0) | 7.0 (5.0-10.0) | 11.0 (9.0-13.0) | <0.001 | 0.743 (0.680 – 0.806) | 9 | 0.63 | 0.73 |
| <7, n(%) | 259 | 75 (29.0) | 66 (38.2) | 9 (10.5) | <0.001 | | | | |
| 7-11, n(%) | | 89 (34.4) | 64 (37.0) | 25 (29.1) | | | | | |
| 11+, n(%) | | 95 (36.7) | 43 (24.9) | 52 (60.5) | | | | | |

*Traditional categories for CURB-65 and PSI; #distribution tertiles for MuLBSTA.

CAP management, these results could be explained by the pathophysiological differences between CAP and COVID-19 pneumonia. As we know by literature, SARS-CoV-2 causes a systemic infection with a pattern of syndromes, ranging from the absence of symptoms to life-threatening acute respiratory distress syndrome (ARDS). Many patients develop an interstitial pneumonia that leads to type 1 respiratory failure and the need of respiratory support, which varies from low flow nasal cannula to endotracheal intubation with mechanical ventilation. In COVID-19 pneumonia a cytokines hyperactivation is observed, leading to a misregulation of the immune system and a systemic disease with vasculitic and pro-thrombotic features. Despite the frequent multiorgan involvement, including kidneys, liver, myocardium, central nervous system and gastrointestinal system, respiratory failure seems to be predominant in COVID-19 pneumonia. Therefore, the clinical presentation of this pneumonia differed considerably from the typical CAP, which often leads to altered mental status, renal failure, sepsis or septic shock.

Consistently, both CURB-65 and PSI scores were already investigated in the influenza virus A H1N1 2009 pandemic and were not found to be useful in estimating prognosis [2]. Moreover, PSI was more recently investigated in CAP with viral etiology and it was found that it is associated with mortality regardless of respiratory virus detection [3].

MuLBSTA was considered in our study since it is a tool previously proposed for viral pneumonia including lymphocytes count. In effect, as confirmed by literature, lymphopenia and its severity levels may serve as reliable predictive factors for COVID-19 clinical outcomes including mortality [4]. Despite that, in our cohort, MuLBSTA did not increase diagnostic accuracy in comparison to CURB-65 and PSI. This could be explained again with the pathophysiological differences between other viral pneumonia (i.e., influenza) and the more complex clinical features of SARS-CoV-2 pneumonia. Moreover, we are aware that missing data could reduce CURB-65, PSI and MuLBSTA accuracy in this study population, even if multiple imputation did not change our results.

Our data are consistent with those reported in similar studies evaluating CURB-65 or PSI in COVID-19 [5,6], even considering that our population of patients showed severe hypoxic respiratory failure and higher mortality rates during the peak of the pandemics, which could be due to a very high viral load, incidence and need of health assistance in a limited period of time, with many patients treated as outpatients as a consequence of the pressure on our healthcare system [7].

Other scores such as the CALL score [8], the 4C score [9] and the Quick COVID-19 Severity score [10] were created *ad hoc* with promising performances. These multiparameter scores include respiratory rate, peripheral oxygen saturation and a lower lymphocyte count, analyzing clinical features which are more consistent with COVID-19 pneumonia pathophysiology. Anyway, all these scores are required to be validated on wider populations.



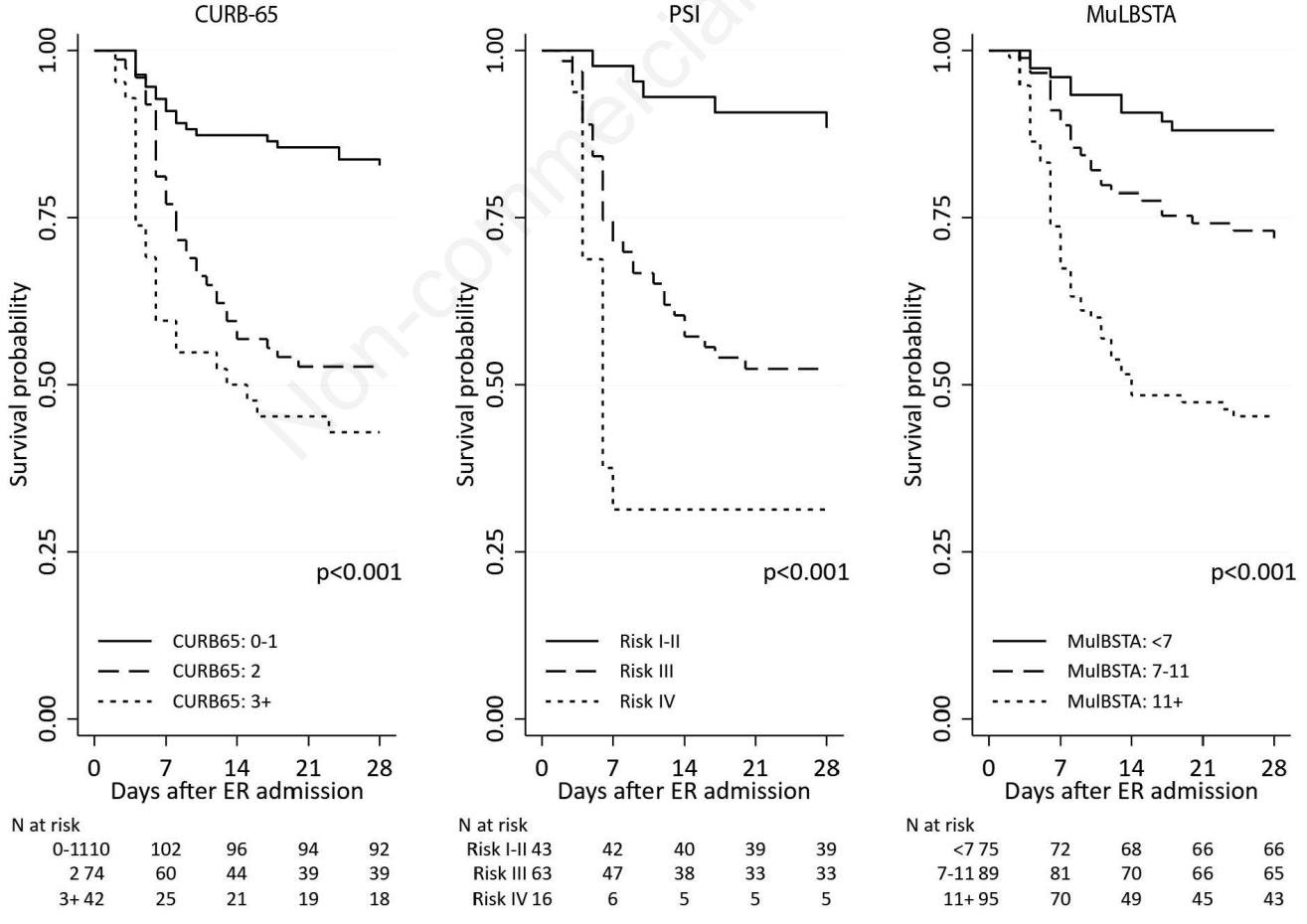


**Figure 2. Kaplan-Meier 28-day survival curves stratified by the categories of CURB-65, PSI and MuLBSTA.**

This study has many limitations, being a single-center, retrospective study with a small number of enrolled patients. Moreover, we focused on a limited, although extraordinary, period of time with a certain amount of possible expected missing data. Diagnostic imaging was performed just by using CXR because, due to the overflow of patients at the beginning of the COVID-19 pandemic, it was not possible to make all patients presented with COVID-19 pneumonia undergo CT scan.

## Conclusions

In conclusion, our aim was to assess CURB-65, PSI and MuLBSTA scores in discriminating 28-day mortality in COVID-19 pneumonia. Despite they demonstrated statistically significant differences in discriminating between survivors and non survivors, the scores were not able to assess accurately COVID-19 associated mortality, leading to a potential underestimation of the disease severity, probably as a consequence of the complex and unique features of this disease. Therefore, these scores should not be recommended in ED to discriminate inpatient *vs* outpatient management of COVID-19 patients.

## References


1. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-845.
2. Estella A. Usefulness of CURB-65 and pneumonia severity index for influenza A H1N1v pneumonia. Monaldi Arch Chest Dis 2012;77:118-21.
3. Kim MA, Park JS, Lee CW, Choi W Il. Pneumonia severity index in viral community acquired pneumonia in adults. PLoS One 2019;14:1-12.
4. Guo L, Wei D, Zhang X, et al. Clinical features predicting mortality risk in patients with viral pneumonia: The MuLBSTA score. Front Microbiol 2019;10:2752.
5. Bradley P, Frost F, Tharmaratnam K, Wootton DG. Utility of established prognostic scores in COVID-19 hospital admissions: Multicentre prospective evaluation of CURB-65, NEWS2 and qSOFA. BMJ Open Respir Res 2020;7:1-9.
6. Nguyen Y, Corre F, Honsel V, et al. Applicability of the CURB-65 pneumonia severity score for outpatient treatment of COVID-19. J Infect 2020;81:e96-e98.
7. Novelli L, Raimondi F, Ghirardi A, et al. At the peak of Covid-19 age and disease severity but not comorbidities are predictors of mortality. Covid-19 burden in Bergamo, Italy. Panminerva Med 2021;63:51-61.
8. Ji D, Zhang D, Xu J, et al. Prediction for progression risk in patients with COVID-19 pneumonia: The CALL score. Clin Infect Dis 2020;71:1393-9.
9. Knight SR, Ho A, Pius R, et al. 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 2020;370:1-13.
10. Haimovich AD, Ravindra NG, Stoytchev S, et al. Development and validation of the quick COVID-19 severity index: A prognostic tool for early clinical decompensation. Ann Emerg Med 2020;76:442-53.