Evaluating the ability of the NLHA2 and artificial neural network models to predict COVID-19 severity, and comparing them with the four existing scoring systems

Yue Dong [a,d,1], Kai Wang [a,1], Xu Zou [b], Xiaoping Tan [c], Yi Zang [d], Xinyu Li [e], Xiaoting Ren [a], Desheng Xie [f], Zhijun Jie [a], Xiaohua Chen [g], Yingying Zeng [a,h,**], Jindong Shi [a,*]

[a] Department of Pulmonary and Critical Care Medicine, Shanghai Fifth People's Hospital, Fudan University, 801 Heqing Road, Minhang District, Shanghai, 200240, China
[b] Intensive Care Unit, Guangzhou Hospital of Traditional Chinese Medicine, 111 Dade Road, Yuexiu District, Guangzhou, Guangdong Province, 510000, China
[c] Department of Respiratory Medicine, Hubei Jingzhou Jiangling People's Hospital, 29 Chujiang Dadao, Jiangling County, Jingzhou City, Hubei Province, 434101, China
[d] Lingang Laboratory, Shanghai, 200031, China
[e] Department of Infectious Disease, Gongli Hospital of Shanghai Pudong New Area, 219 Miaopu Road, Pudong New Area, Shanghai, 200120, China
[f] Department of Infectious Disease, Shanghai Fifth People's Hospital, Fudan University, 801 Heqing Road, Minhang District, Shanghai, 200240, China
[g] Department of Infectious Diseases, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, 600 Yishan Road, Xuhui District, Shanghai, 200030, China
[h] Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Xuhui District, Shanghai, 200032, China

ARTICLE INFO

ABSTRACT

To improve the identification and subsequent intervention of COVID-19 patients at risk for ICU admission, we constructed COVID-19 severity prediction models using logistic regression and artificial neural network (ANN) analysis and compared them with the four existing scoring systems (PSI, CURB-65, SMARTCOP, and MuLBSTA). In this prospective multi-center study, 296 patients with COVID-19 pneumonia were enrolled and split into the General-Ward-Care group (N = 238) and the ICU-Admission group (N = 58). The PSI model (AUC = 0.861) had the best results among the existing four scoring systems, followed by SMARTCOP (AUC = 0.770), motified-MuLBSTA (AUC = 0.761), and CURB-65 (AUC = 0.712). Data from 197 patients (training set) were analyzed for modeling. The beta coefficients from logistic regression were used to develop a severity prediction model and risk score calculator. The final model (NLHA2) included five covariates (consumes alcohol, neutrophil count, lymphocyte count, hemoglobin, and AKP). The NLHA2 model (training: AUC = 0.959; testing: AUC = 0.857) had similar results to the PSI model, but with fewer variable items. ANN analysis was used to build another complex model, which had higher accuracy (training: AUC = 1.000; testing: AUC = 0.907). Discrimination and calibration were further verified through bootstrapping (2000 replicates), Hosmer-Lemeshow goodness of fit testing, and Brier score calculation. In conclusion, the PSI model is the best existing system for predicting ICU admission among COVID-19 patients, while two newly-designed models (NLHA2 and ANN) performed better than PSI, and will provide a new approach for the development of prognostic evaluation system in a novel respiratory viral epidemic.

1. Introduction

Since December 8th, 2019, several cases of pneumonia in combination with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have been reported in Wuhan, Hubei Province, China [1,2]. The early clinical symptoms of Coronavirus Disease 2019 (COVID-19) include fever, cough, and fatigue, which is similar to the manifestation of community-acquired pneumonia (CAP) [1,3]. Moreover, COVID-19 is a highly infectious respiratory tract disease that lacks effective therapy

[3]. In the absence of timely intervention, some patients will rapidly develop severe pneumonia, acute respiratory distress syndrome (ARDS),

COVID-19 is limited since their safety has yet to be confirmed [7]. Recent studies have evaluated the application of certain scoring systems

Abbreviations			
NLHA2	A prediction model designed by our team with five covariates: neutrophil count, lymphocyte count, hemoglobin, alkaline phosphatase, and alcohol consumption	**EPV**	Event per variable
		HR	Heart rate
		RR	Respiratory rate
		RBC	Red blood cell
		Hb	Hemoglobin
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2	**HCT**	Hematocrit
COVID-19	Coronavirus Disease 2019	**PLT**	Blood platelets
ARDS	Acute respiratory distress syndrome	**ESR**	Erythrocyte sedimentation rate
CAP	Community-acquired pneumonia	**CRP**	C-reactive protein
ROC	Receiver operating characteristic	**PCT**	Procalcitonin
AUC	Area under curve	**SAA**	Serum amyloid A
LR	Likelihood ratio	**ALT**	Alanine amino transferase
KNN	K-nearest neighbor	**AST**	Aspartate amino transferase
ANN	Artificial neural network	**GGT**	Gamma-glutamyl transpeptidase
		AKP	Alkaline phosphatase

Fig. 1. Flowchart of the COVID-19 patients-screening process.

and may even die [3]. Furthermore, inappropriate intervention as well as delayed diagnosis or admission to the ICU may escalate the risk of severe pneumonia and worsen prognosis [4]. Therefore, an effective severity scoring model would benefit the treatment of patients with COVID-19.

The pneumonia severity index (PSI) and CURB-65 are prediction models for the prognosis of CAP patients, and thus assist in directing treatment strategy [5,6]. However, the utilization of these models for

(PSI, CURB-65, HNC-LL, A-DROP, SMARTCOP, MuLBSTA, NEWS2 and qSOFA) to COVID-19 [7–10]. However, the conclusions of these studies contradict one another. The purpose of this study was to identify a suitable scoring system to assist with the management and treatment of patients with COVID-19 while ensuring safety.

Previous studies have found that the existing pneumonia prediction models may not be ideal for screening severe COVID-19 [7–10]. In addition to traditional biological methods, advances in computer

Table 1
Baseline characteristics, complications, and scores for patients with COVID-19 pneumonia.

	General Ward Care (N = 238)	ICU Admission (N = 58)	P Value
Median Age (IQR) – yrs	49.0 (37.3–59.0)	66.0 (56.0–76.0)	<0.0001
Gender – count (%)			
Male	117 (49.2)	36 (62.1)	0.078
Female	121 (50.8)	22 (37.9)	
Smoker – count (%)	5 (2.1)	7 (12.1)	<0.01
Consumes Alcohol – count (%)	20 (8.4)	20 (34.5)	<0.0001
Tumor – count (%)	8 (3.4)	6 (10.3)	0.057
Chronic Liver Disease – count (%)	8 (3.4)	10 (17.2)	<0.001
Congestive Heart Failure – count (%)	15 (6.3)	10 (17.2)	<0.05
Cerebrovascular Disease – count (%)	17 (7.1)	18 (31.0)	<0.0001
Chronic Renal Disease – count (%)	11 (4.6)	16 (27.6)	<0.0001
Hypertension – count (%)	71 (29.8)	25 (43.1)	0.053
Diabetes – count (%)	30 (12.6)	22 (37.9)	<0.0001
Nursing Home Resident – count (%)	1 (0.4)	4 (6.9)	<0.01
PSI Score (IQR) – points	72 (0.0–91.0)	119.5 (102.5–136.8)	<0.0001
PSI Grade – count (%)			<0.0001
1	112 (47.1)	3 (5.2)	
2	61 (25.6)	0 (0.0)	
3	48 (20.2)	3 (5.2)	
4	13 (5.5)	30 (51.7)	
5	4 (1.7)	22 (37.9)	
CURB-65 – count (%)			<0.0001
0-1	191 (80.3)	22 (37.9)	
2	37 (15.5)	29 (50.0)	
3	9 (3.8)	6 (10.3)	
4-5	1 (0.4)	1 (1.7)	
SMARTCOP – count (%)			<0.0001
0	37 (15.5)	0 (0.0)	
1	151 (63.4)	17 (29.3)	
2	37 (15.5)	31 (53.4)	
3	10 (4.2)	7 (12.1)	
4-6	3 (1.3)	3 (5.2)	
Modified MuLBSTA (IQR) – points	3.0 (3.0–8.0)	8.0 (6.3–9.0)	<0.0001

IQR: interquartile range.

science have made it easier to carry out disease research through protein structure prediction, target-based drug prediction, system biology, and classifier construction with machine learning [11–15]. In this study, we attempted to establish a prediction model specific to COVID-19 using machine learning algorithms. Some studies provide evidence in favor of the excellent classifier effect of machine learning [16,17]. As a traditional machine learning algorithm, logistic regression is often applied to predict outcomes based on patient characteristics [18]. Artificial neural network (ANN) models can precisely predict complex relationships between high-dimensional datasets and outcomes, which is beyond the reach of traditional methods [19].

In this study, we used both our models and existing models to investigate patients with confirmed COVID-19 diagnoses. We compared the efficiency and predictive accuracy of four established scoring systems (PSI, CURB-65, SMARTCOP and Modified-MuLBSTA) in differentiating severity [20–22]. Based on patient data, we constructed our own prediction models using logistic regression or ANN. These models may help to improve the management and treatment of patients with severe COVID-19.

2. Materials and methods

2.1. Study design and setting

This multi-center prospective observational study was a collaborative effort between Wuhan Leishenshan Hospital (the designated hospital for COVID-19 in Wuhan), Fudan University Affiliated Shanghai Fifth People's Hospital, and Hubei Jingzhou Jiangling People's Hospital. Patients who participated in the study were required to meet the following criteria: (1) age≥18 years; and (2) definitive diagnosis of SARS-CoV-2 infection. All patients enrolled in this research were diagnosed with COVID-19 from February 21, 2020 to October 1, 2020, as according to the guidelines published by the World Health Organization (WHO) and National Health Commission of the People's Republic of China. Local health authorities performed laboratory confirmation of SARS-CoV-2 infection as previously described [23,24]. Once the diagnosis was confirmed, each patient was assessed using the four established scoring systems. Each patient was classified into one of the following two groups based on outcomes: the ICU-Admission group, which included patients who were treated in the ICU or who died; and the General-Ward-Care group, which included all other patients.

The ethics commissions of Fudan University Affiliated Shanghai Fifth People's Hospital and Hubei Jingzhou Jiangling People's Hospital approved this study. All patients provided informed consent before participating in the study. The patient screening process is outlined in Fig. 1.

2.2. Data collection

Infections were confirmed using nucleic acid amplification testing (NAAT), such as reverse transcription polymerase chain reaction (RT-PCR). Data on demographic characteristics, pre-existing comorbidities, initial vital signs, routine laboratory test results, admission diagnoses, and variables necessary to determine the PSI, CURB-65, SMARTCOP and modified-MuLBSTA scores were prospectively collected for all eligible patents [21,22,25–27].

2.3. Missing data imputation

Missing data is common in clinical studies. Deleting missing data is a simple way to deal with it, but can cause information loss and sample size reduction. In this study, we imputed missing data with probable values predicted from available data to preserve all cases [28]. Imputation was performed using the KNN (K-Nearest Neighbor) algorithm with k = 10. The 'DmWR2' R package was used to conduct this work.

2.4. Splitting data into training and testing sets

To conduct self-validation of the constructed prediction models, we randomly split the dataset into two subgroups: the training and testing sets. The training dataset was used for model building, while the testing set was used for model effect validation. Patients were divided between the groups at a 2:1 ratio, respectively. Supplementary Table 2 summarizes the basic distribution of demographic data.

2.5. Logistic regression model

To ensure practical operability of the scoring system, we selected easily accessible data as target variables, including demographic information, past medical history, smoking and drinking history, and blood indexes. First, we converted these variables into categorical variables. We performed univariate logistic regression for each variable to identify factors that predict overall outcome. Variables with $P < 0.1$ were further investigated using multivariate modeling. Multivariate regression analysis was carried out using a stepwise procedure, and the number of variables in the final model was determined by referring to event-per-
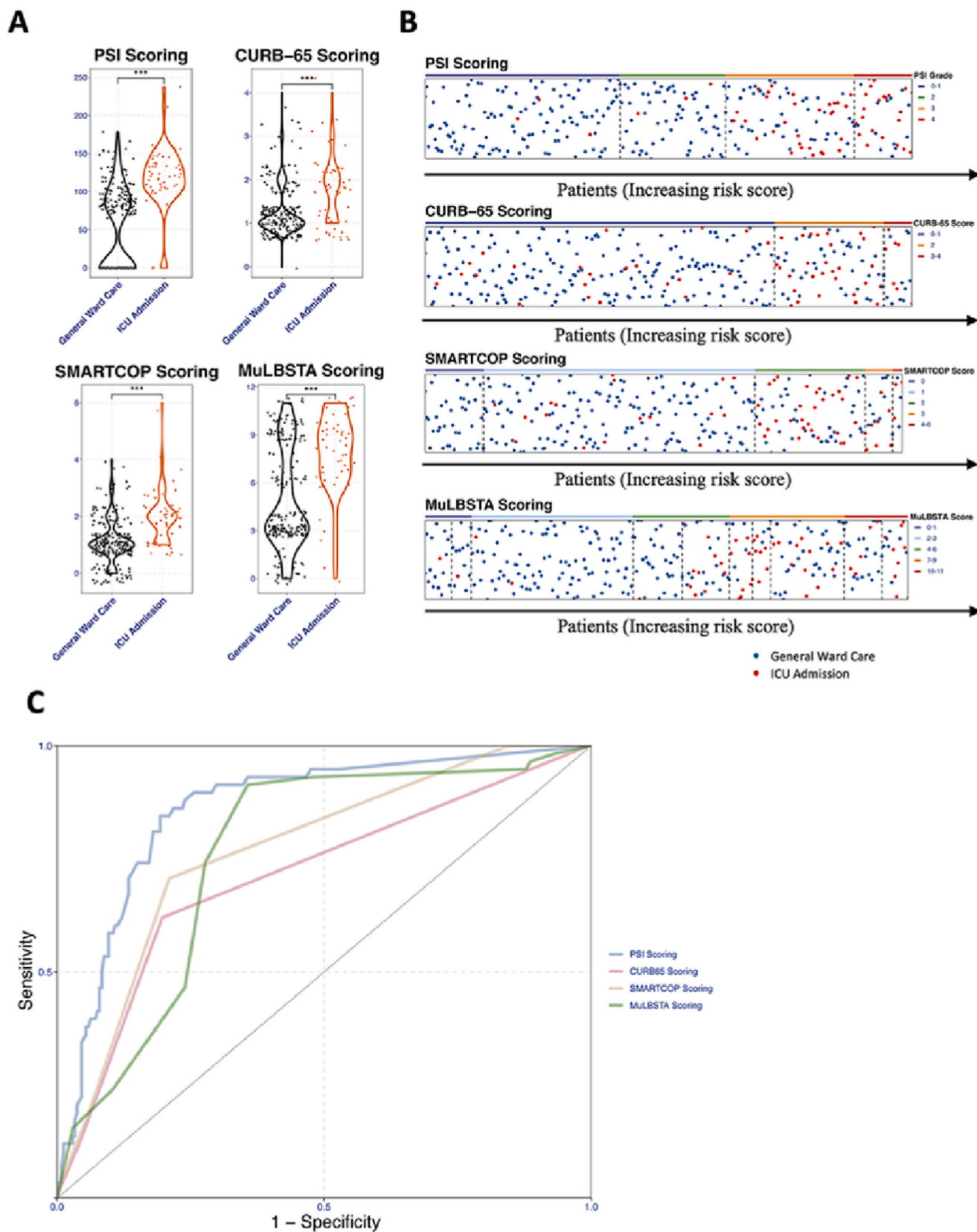
Fig. 2. The PSI scoring system had the best prediction effect. (A) Violin plots illustrating the four scores of each patient in both groups. Differences between groups were assessed using the Wilcoxon Rank Sum Test. *P* < 0.05 was considered statistically significant. (B) Distribution of risk scores in the General-Ward and ICU cohorts. Dotted lines represent dividing boundaries for scoring groups. Patients were sorted on the horizontal axis according to their scores. (C) ROC curves verifying the predictive performance of the four risk scoring systems in the COVID-19 cohort.

Table 2
ROC comparisons between PSI, CURB-65, SMARTCOP, and MuLBSTA.

	AUC	95% CI	-LR	+LR	*P* Value	Sensitivity (%)	Specificity (%)
PSI	0.861	0.816–0.898	0.19	4.37	<0.0001	84.48	80.67
CURB-65	0.712	0.657–0.763	0.47	3.14	<0.0001	62.07	80.25
SMARTCOP	0.770	0.718–0.817	0.37	3.36	<0.0001	70.69	78.99
MuLBSTA	0.761	0.708–0.808	0.13	2.56	<0.0001	91.38	64.29

ROC: receiver operating characteristic curve; AUC: area under curve; 95% CI: 95% confidence interval; LR: likelihood ratio.

variables (EPV). The range of EPV was limited to 5–10 to reduce the risk of over-fitting [29,30]. Each significant variable was assigned a weighted point based on its β coefficient value.

2.6. Artificial neural network model

Significant variables from univariate regression were selected for artificial neural network modeling. The 'RSNNS' R package was used to build a multilayer perceptron (MLP) model. Modeling parameters were confirmed empirically. The importance of each variable was calculated using the connection weighted algorithm published by Olden et al. [31]. The Olden calculation was performed with the 'NeuralNetTools' R package.

2.7. Statistical analysis

For categorical variables in our study, the Chi-Square Test (with or without Yate's correction) or Fisher's Exact Test was used for univariate comparison. To describe the distribution of numerical variables, we applied either the mean ± SD (standard deviation) or median ± IQR (interquartile range) depending on the results of normality testing. We compared the differences among numerical variables using the ANOVA test, Welch's Test, or Wilcoxon Rank-Sum Test depending on the results of normality and homogeneity of variance testing.

The predictive values of PSI, CURB-65, SMARTCOP and MuLBSTA were assessed in two dimensions of discrimination. Discrimination was determined using the area under the curve (AUC) with a 95% CI. The best cut-off value in each scoring system was confirmed by the Youden index. The cut-off value was used to divide cases into the low-risk and high-risk groups. To further assess the discrimination of a scoring model deemed to be effective, the bootstrap method (2000 replicates) is recommended to obtain a better estimate of AUC. Brier score calculation and Hosmer-Lemeshow goodness of fit testing were performed for calibration assessment.

All tests were two-tailed. $P < 0.05$ was considered statistically significant. Computations were carried out with R version 4.0.2.

3. Results

3.1. Patient population

A total of 329 patients with COVID-19 met the inclusion criteria for our cohort study. The enrolled cases consisted of 258 patients from Wuhan Leishenshan Hospital, 37 patients from Shanghai Fifth People's Hospital, and 34 patiets from Hubei Jingzhou Jiangling People's Hospital. Of the total 329 patients, 33 patients withdrew from this study because they didn't meet our inclusion criteria. Additionally, 58 patients were transferred to the ICU because of worsening situation (Fig. 1).

3.2. Patient characteristics

The characteristics of the 296 consecutive admissions are summarized in Table 1. The percentage of male subjects was 49.2% in the General-Ward-Care group and 62.1% in the ICU group ($P = 0.078$). The median age of patients in the ICU-Admission group (66 years) was significantly higher than that of the General-Ward group (49 years) ($P <$

0.0001). Rates of cigarette smoking ($P < 0.01$) and drinking ($P < 0.0001$) were higher in the ICU-Admission group, and patients who originally lived in nursing homes were more likely to develop severe COVID-19 ($P < 0.01$). These three characteristics suggest that patients with severe COVID-19 may have had poor health before the onset of the pandemic. Furthermore, patients in the severe group had higher rates of liver diseases (17.2% versus 3.4%, $P < 0.001$), congestive heart failure (17.2% versus 6.3%, $P < 0.05$), diabetes (37.9% versus 12.6%, $P < 0.0001$), chronic renal diseases (27.6% versus 4.6%, $P < 0.0001$), and cerebrovascular diseases (31.0% versus 7.1%, $P < 0.0001$). No other characteristics significantly differed between the two groups.

3.3. Statistical evaluation of the PSI, CURB-65, SMARTCOP, and MuLBSTA scoring systems

Table 1 briefly describes the overall distribution of the scores and differences between groups. Patients in the ICU group had higher PSI scores (119.5 versus 72, $P < 0.0001$) and higher PSI grades ($P < 0.0001$). The CURB-65, SMARTCOP, and modified MuLBSTA scoring systems shared the same conclusion ($P < 0.0001$). Violin plots were used to visualize the scores of the two groups (Fig. 2A).

The predictive accuracies of the PSI, CURB-65, SMARTCOP, and MuLBSTA scoring systems are shown in Fig. 2B. We noticed that all four systems were able to identify patients admitted to the ICU to some degree, but none were able to accurately distinguish between the two groups of patients.

We performed ROC analysis and calculated AUC values to assess the discriminative power of the four systems (Fig. 2C, Table 2, Supplementary Table 1). The AUC was highest for the PSI scoring system (AUC = 0.861; 95% CI: 0.816–0.898), followed by the MuLBSTA system (AUC = 0.761, 95% CI: 0.708–0.808), SMARTCOP system (AUC = 0.770, 95% CI: 0.718–0.817), and CURB-65 system (AUC = 0.712, 95% CI: 0.657–0.763). Ultimately, the PSI scoring system was considered the best choice for discriminating the severity of outcomes among patients according to our ROC analysis (Fig. 2C). The sensitivity and specificity values corresponding to each cut-off value are shown in Supplementary Table 1. The PSI scoring system with a cut-off value of 97 points yielded the best sensitivity (82.35%) and specificity (89.67%). The corresponding highest positive likelihood ratio of 4.37 and lowest negative likelihood ratio of 0.19 further validated the sensitivity and specificity of this system. Detailed data on the four scoring systems are included in Supplementary Table 1.

3.4. Logistic regression model: NLHA2 scoring

We split the original dataset into the training set (N = 197) and the testing set (N = 99). Then, we compared baseline data between the two groups. Median age, gender ratio, smoker count, and other variables did not significantly differ between the two groups, suggesting that grouping was objective and random (Supplementary Table 2). Univariate regression analysis was performed for every variable, which resulted in 35 pre-selected variables with $P < 0.1$ (Fig. 3, Table 3). We used stepwise multivariate regression to reanalyze the above variables and build a prediction model. The number of events was defined as the number of ICU admissions in the training set (N = 32), which limited the number of selected variables to less than six (EPV: 5–10). The final
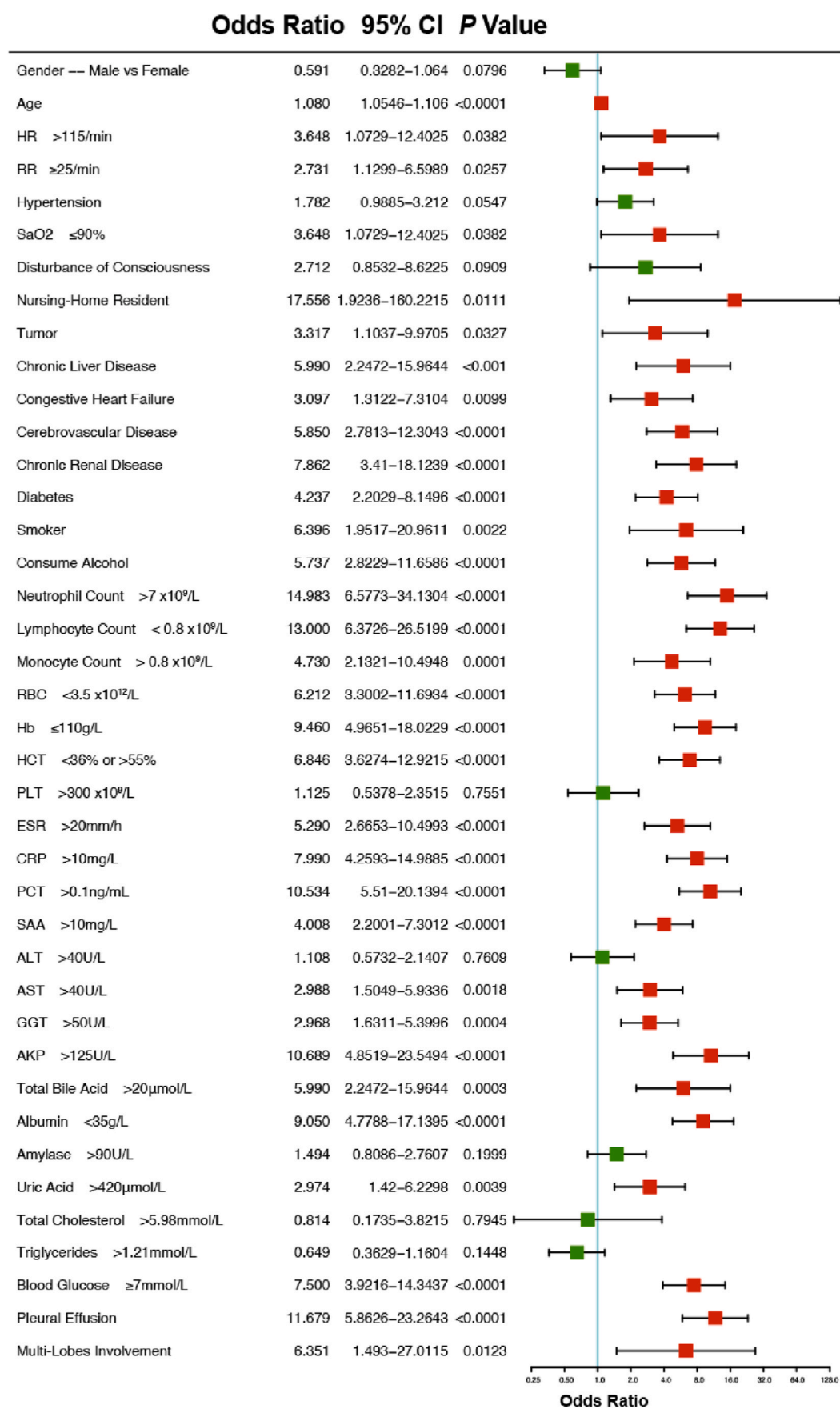
	Odds Ratio	95% CI	P Value
Gender — Male vs Female	0.591	0.3282–1.064	0.0796
Age	1.080	1.0546–1.106	<0.0001
HR >115/min	3.648	1.0729–12.4025	0.0382
RR ≥25/min	2.731	1.1299–6.5989	0.0257
Hypertension	1.782	0.9885–3.212	0.0547
SaO2 ≤90%	3.648	1.0729–12.4025	0.0382
Disturbance of Consciousness	2.712	0.8532–8.6225	0.0909
Nursing-Home Resident	17.556	1.9236–160.2215	0.0111
Tumor	3.317	1.1037–9.9705	0.0327
Chronic Liver Disease	5.990	2.2472–15.9644	<0.001
Congestive Heart Failure	3.097	1.3122–7.3104	0.0099
Cerebrovascular Disease	5.850	2.7813–12.3043	<0.0001
Chronic Renal Disease	7.862	3.41–18.1239	<0.0001
Diabetes	4.237	2.2029–8.1496	<0.0001
Smoker	6.396	1.9517–20.9611	0.0022
Consume Alcohol	5.737	2.8229–11.6586	<0.0001
Neutrophil Count >7 x10⁹/L	14.983	6.5773–34.1304	<0.0001
Lymphocyte Count < 0.8 x10⁹/L	13.000	6.3726–26.5199	<0.0001
Monocyte Count > 0.8 x10⁹/L	4.730	2.1321–10.4948	0.0001
RBC <3.5 x10¹²/L	6.212	3.3002–11.6934	<0.0001
Hb ≤110g/L	9.460	4.9651–18.0229	<0.0001
HCT <36% or >55%	6.846	3.6274–12.9215	<0.0001
PLT >300 x10⁹/L	1.125	0.5378–2.3515	0.7551
ESR >20mm/h	5.290	2.6653–10.4993	<0.0001
CRP >10mg/L	7.990	4.2593–14.9885	<0.0001
PCT >0.1ng/mL	10.534	5.51–20.1394	<0.0001
SAA >10mg/L	4.008	2.2001–7.3012	<0.0001
ALT >40U/L	1.108	0.5732–2.1407	0.7609
AST >40U/L	2.988	1.5049–5.9336	0.0018
GGT >50U/L	2.968	1.6311–5.3996	0.0004
AKP >125U/L	10.689	4.8519–23.5494	<0.0001
Total Bile Acid >20µmol/L	5.990	2.2472–15.9644	0.0003
Albumin <35g/L	9.050	4.7788–17.1395	<0.0001
Amylase >90U/L	1.494	0.8086–2.7607	0.1999
Uric Acid >420µmol/L	2.974	1.42–6.2298	0.0039
Total Cholesterol >5.98mmol/L	0.814	0.1735–3.8215	0.7945
Triglycerides >1.21mmol/L	0.649	0.3629–1.1604	0.1448
Blood Glucose ≥7mmol/L	7.500	3.9216–14.3437	<0.0001
Pleural Effusion	11.679	5.8626–23.2643	<0.0001
Multi-Lobes Involvement	6.351	1.493–27.0115	0.0123

Fig. 3. Results of univariate logistic regression. Risk factors are colored red whereas protection factors are colored blue. The green box represents nonsignificant differences.

Table 3
Univariate logistic regression analysis.

Variable	Odds Ratio	95% CI	P Value
Gender – Male vs Female	0.5909	0.3282–1.0640	0.0796[#]
Age	1.0800	1.0546–1.1060	<0.0001****
HR > 115/min	3.6478	1.0729–12.4025	0.0382*
RR ≥ 25/min	2.7306	1.1299–6.5989	0.0257*
SaO2 ≤ 90%	3.6478	1.0729–12.4025	0.0382*
Disturbance of Consciousness	2.7123	0.8532–8.6225	0.0909[#]
Nursing-Home Resident	17.5556	1.9236–160.2215	0.0111*
Hypertension	1.7819	0.9885–3.2120	0.0547[#]
Tumor	3.3173	1.1037–9.9705	0.0327*
Chronic Liver Disease	5.9896	2.2472–15.9644	<0.001***
Congestive Heart Failure	3.0972	1.3122–7.3104	0.0099**
Cerebrovascular Disease	5.8500	2.7813–12.3043	<0.0001****
Chronic Renal Disease	7.8615	3.4100–18.1239	<0.0001****
Diabetes	4.2370	2.2029–8.1496	<0.0001****
Smoker	6.3961	1.9517–20.9611	0.0022**
Consumes Alcohol	5.7368	2.8229–11.6586	<0.0001****
Neutrophil Count > 7×10^9/L	14.9829	6.5773–34.1304	<0.0001****
Lymphocyte Count <0.8×10^9/L	13.0000	6.3726–26.5199	<0.0001****
Monocyte Count >0.8×10^9/L	4.7303	2.1321–10.4948	0.0001***
RBC <3.5×10^{12}/L	6.2121	3.3002–11.6934	<0.0001****
Hb ≤ 110 g/L	9.4597	4.9651–18.0229	<0.0001****
HCT <36% or >55%	6.8463	3.6274–12.9215	<0.0001****
PLT >300×10^9/L	1.1245	0.5378–2.3515	0.7551
ESR >20 mm/h	5.2900	2.6653–10.4993	<0.0001****
CRP >10 mg/L	7.9900	4.2593–14.9885	<0.0001****
PCT >0.1 ng/mL	10.5341	5.5100–20.1394	<0.0001****
SAA >10 mg/L	4.0079	2.2001–7.3012	<0.0001****
ALT >40U/L	1.1077	0.5732–2.1407	0.7609
AST >40U/L	2.9882	1.5049–5.9336	0.0018**
GGT >50U/L	2.9677	1.6311–5.3996	0.0004***
AKP >125U/L	10.6892	4.8519–23.5494	<0.0001****
Total Bile Acid >20 μmol/L	5.9896	2.2472–15.9644	0.0003***
Albumin <35 g/L	9.0502	4.7788–17.1395	<0.0001****
Amylase >90U/L	1.4941	0.8086–2.7607	0.1999
Uric Acid >420 μmol/L	2.9743	1.4200–6.2298	0.0039**
Total Cholesterol >5.98 mmol/L	0.8143	0.1735–3.8215	0.7945
Triglycerides >1.21 mmol/L	0.6490	0.3629–1.1604	0.1448
Blood Glucose ≥ 7 mmol/L	7.5000	3.9216–14.3437	<0.0001****
Pleural Effusion	11.6786	5.8626–23.2643	<0.0001****
Multi-Lobe Involvement	6.3505	1.493–27.0115	0.0123

HR: heart rate; RR: respiratory rate; RBC: red blood cell; Hb: hemoglobin; HCT: haematocrit; PLT: blood platelets; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; PCT: procalcitonin; SAA: serum amyloid A; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transpeptidase; AKP: alkaline phosphatase.
****: <0.0001, ***: <0.001, **: <0.01, *: <0.05, #: <0.1.

model included five variables: Consumes alcohol, neutrophil count (>7 x10^9/L), lymphocyte count (<0.8 × 10^9/L), hemoglobin (≤110 g/L), and AKP (>125 U/L) (Fig. 4A, Table 4). The model was named 'NLHA2' based on the variable names. The risk score was calculated as follows: Risk Score = 3*[Drink]+4*[Neutrophil Count]+2.5*[Lymphocyte Count]+2.5*[Hb]+2.5*[AKP] (Table 4). All patients were assigned scores and accordingly stratified into the high-risk or low-risk group depending on the cut-off value. The NLHA2 scores of ICU-Admission patients were significantly higher than those for General-Ward-Care patients in both datasets (Fig. 4B). The predictive performance of the risk score for COVID-19 severity was evaluated through ROC curve analysis, where the AUC reached 0.959 in the training set and 0.857 in the testing set (Fig. 4D, Table 5). The NLHA2 model had ideal values for sensitivity (training: 96.87%; testing: 73.08%) and specificity (training: 83.03%; testing: 91.78%). It is worth noting that greater sensitivity was taken into account in the determination of the cut-off value (Table 5, Supplementary Table 1). Our scatter plot also supports the model's excellent distinguishing ability (Fig. 4C). Furthermore, the discrimination and calibration of the model verify its effectiveness. The results of Hosmer-Lemeshow goodness of fit testing suggested a better fitting

effect (training: *P* = 0.991; testing: *P* = 0.152). Additionally, Brier score analysis revealed a screening effect (training: 0.048; testing: 0.136) (Table 6).

3.5. Artificial neural network prediction model

A total of 42 variables were included in the input layer of the network. We set the number of hidden layers to two and the number of nodes to 10 and eight, according to the results of several previous training trials (Fig. 5A). The importance of each variable was calculated using Olden's method (Fig. 5B, Supplementary Table 3). Excitingly, our model demonstrated fantastic prediction effects (Fig. 5C and D). The AUC values for the training and testing sets were 1.000 and 0.907, respectively. The specificity of the model was 100% for the training set and 88.46% for the testing set (Fig. 5D, Table 5). The AUC values (training: 1.000; testing: 0.907) and bootstrap AUC values (training: 1.000; testing: 0.957) were significantly higher than those for the NLHA2 scoring system. Furthermore, the model also performed better in terms of both Hosmer-Lemeshow goodness of fit testing and Brier score analysis (training: *P* = 1.000; testing: *P* = 1.000) (Table 6).

4. Discussion

This study enrolled a total of 296 patients who were diagnosed with COVID-19 from Feb. 21, 2020 to Dec. 28, 2021. The primary COVID-19 outbreak involving so many patients occurred in December of 2019 and spread globally to many different countries [32,33]. Compared to CAP, COVID-19 shows rapid progression, high infectivity, and significant mortality in the advanced stage of the disease [34,35]. Since no specialized medication has been discovered to treat SARS-CoV-2 infection, our healthcare system is in great need of a valid scoring metric for doctors to determine whether to provide aggressive medical intervention. This study aimed to develop an efficient, safe, and dependable model to discriminate the severity and predict the outcomes of patients with COVID-19.

In this study, we restricted enrollment to patients aged 18 years or older because the WHO provides different recommendations on COVID-19 severity classification and management for adults versus children. In addition, the application of several scoring systems (e.g. PSI or CURB-65) in adult patients with community-acquired pneumonia is widely recognized. However, factors such as advanced age, tumor, hypertension and other chronic diseases hardly exist in pediatric pneumonia patients, which engendered controversy when those scoring systems applied to children. Considering the above reasons, we only included adult patients for analysis. We first summarized baseline data and found that the General-Ward-Care and ICU-Admission groups were significantly different in terms of many variables: Age, Smoking, Consumes Alcohol, Nursing home residence, and several comorbidities. This finding would suggest that the underlying health status of patients may influence the inflammation intensity and severity of COVID-19 after infection. Previous studies have also found that patients with underlying diseases are more susceptible to SARS-CoV-2 infection, which can result in severe and even fatal respiratory diseases [36]. The immune status of the patient may also influence the development of disease after infection. However, we did not consider the impact of immunodeficiency on the clinical manifestations of COVID-19 in this study since all participants were free of AIDS and had no long-term history of immunosuppressive drug use.

In our study, we identified PSI as the best existing system to determine a patient's risk. The PSI model had excellent discrimination and calibration. Patients with a PSI grade >2 are considered to have a stronger probability of admission to the ICU and higher mortality. The PSI scale, derived from the PORT (Pneumonia Patient Outcomes Research Team) Project, is a prediction system based on data from more than 10,000 adults with CAP and categorizes patients into five classes in terms of death risk [37]. Recently, some articles have described the
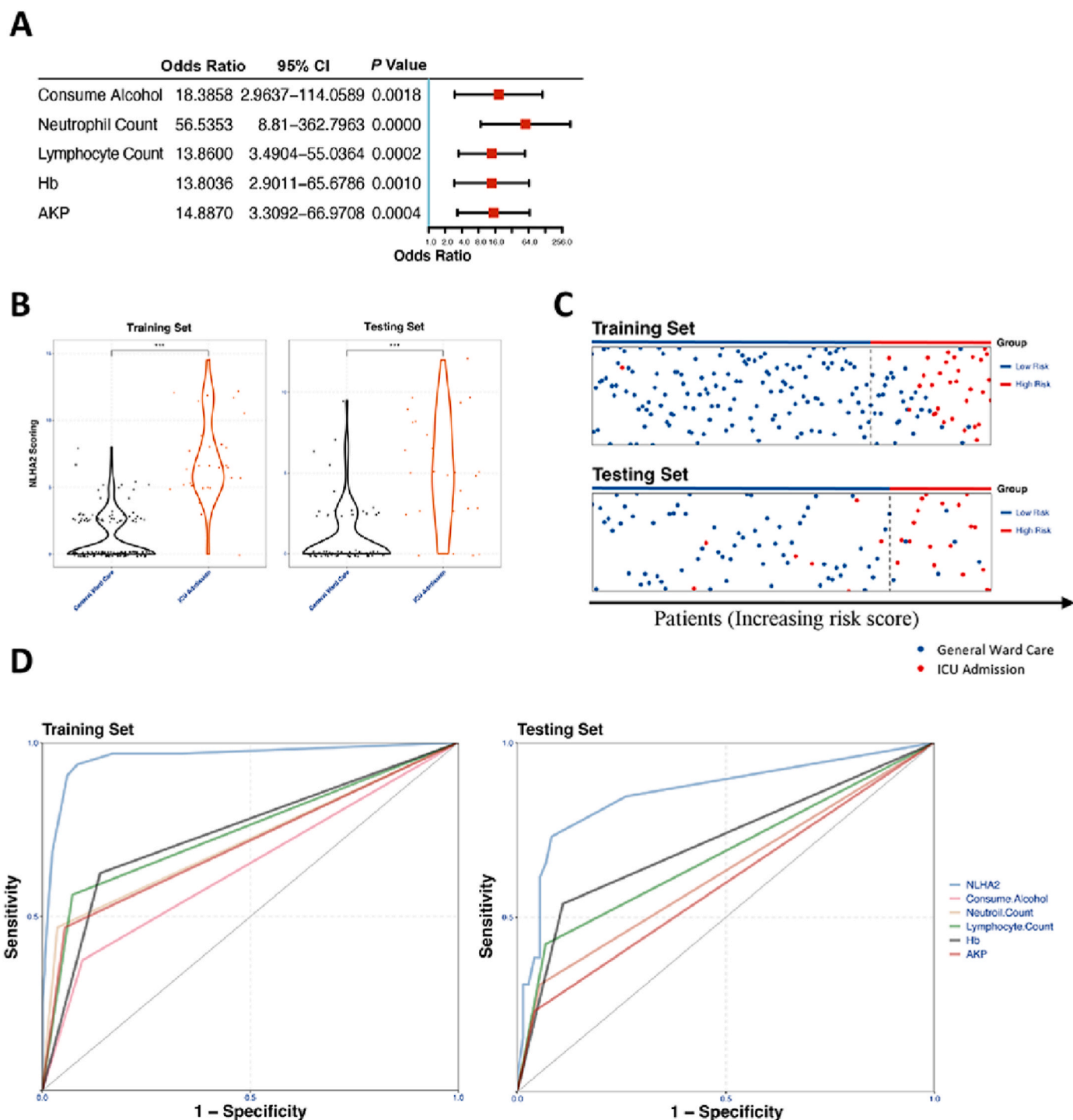
Fig. 4. The NLHA2 model built using multivariate regression effectively predicts who is likely to be admitted to the ICU. (A) Forest plot for multivariate regression. (B) Violin plots showing the NLHA2 scores of the patients (training or testing dataset) in both groups. Differences between groups were assessed using the Wilcoxon Rank Sum Test. $P < 0.05$ was considered statistically significant. (C) Distribution of the NLHA2 scores in the General-Ward and ICU cohorts. The dotted line represents the cut-off value of 2.5 points. (D) ROC curve analysis verifying the prediction performance of the NLHA2 scoring system in the COVID-19 cohort. ROC analysis was also performed for each individual variable.

value of PSI scoring in predicting COVID-19 severity. For example, García Clemente et al. determined that PSI is the best predictor for mortality (AUC = 0.874) among the four scoring systems, but not for risk of ICU admission (AUC = 0.620) [9]. In Fan's research, PSI and CURB-65 were found to share an AUC of 0.85 during clinical comparison, which was greater than the AUC values of the other scoring systems [8]. By British Thoracic Society (BTS) guidelines, patients with CURB-65 scores of 0 are recommended to receive outpatient treatment [38]. According

to a comparison of two systems (PSI and CURB-65) performed by Aujesky et al., CURB-65 is becoming a potential alternative to PSI due to its ease of use. However, the effectiveness and safety of CURB-65 are ambiguous, which makes its utilization less valuable [5]. Both our research and the article published by Nguyen et al. support this conclusion under the conditions of either CAP or COVID-19 [7]. In our study, there were no deaths in the low-risk group – this supports the safety of the PSI classification system, especially considering that the

Table 4

Multivariate logistic regression analysis.

Variable	Odds Ratio	95% CI	P Value
Consumes Alcohol	18.3858	2.9637–114.0589	0.0018**
Neutrophil Count $> 7 \times 10^9$/L	56.5353	8.8100–362.7963	<0.0001****
Lymphocyte Count $<0.8 \times 10^9$/L	13.8600	3.4904–55.0364	0.0002***
Hb $\leq$ 110 g/L	13.8036	2.9011–65.6786	0.001**
AKP >125U/L	14.8870	3.3092–66.9708	0.0004***

Risk Score = 3*[Drink]+4*[Neutrophil Count]+2.5*[Lymphocyte Count]+2.5* [Hb]+2.5*[AKP].

CURB-65 system was accompanied by 21.1% mortality in the low-risk class (CURB-65 = 0 or 1) in the research from Nguyen et al. [7]. In our study, 213 patients had a CURB-65 score equal to 0 or 1, of whom 21 were admitted to the ICU or transferred to superior hospitals. The safety of the CURB-65 system was not satisfactory, and the results of ROC analysis (AUC = 0.712, Youden Index = 0.42) indicated weaker discrimination.

In this study, we also analyzed the two other scoring systems: SMARTCOP and MuLBSTA. The SMARTCOP method is useful for predicting the risk of intensive respiratory or vasopressor support (IRVS) among patients with CAP, where IRVS is one of the criteria for ICU admission. However, our research concluded that SMARTCOP is not applicable to COVID-19 severity (AUC = 0.770, Youden Index = 0.50), which is similar to the findings of García Clemente et al. (AUC = 0.749). The MuLBSTA system was designed as a mortality prediction tool for patients with viral pneumonia, and researchers have recently considered its value of application to COVID-19 [39]. The study from Ma et al. indicated that MuLBSTA has a good early warning effect for patients with severe COVID-19 (AUC = 0.927) [40], while Iijima et al. obtained an AUC value of 0.87 in their work [27]. Additionally, the study from Xu et al. found that the MuLBSTA scoring system was extremely effective in predicting mortality (AUC = 0.956) and ICU admission (AUC = 0.875) [10]. However, similar to our results, the AUC values were only 0.773 for morality and 0.777 for ICU admission in the paper from García Clemente et al. [9]. In our research, the modified-MuLBSTA scoring system also did not perform well (AUC = 0.761). The traditional version of MuLBSTA includes an item on bacterial infection testing, which is rarely used in patients with COVID-19. As a result, Iijima et al. created the modified version of the model, which substitutes CRP level for bacterial infection status. Further studies or meta-analyses are required to better understand the value of applying MuLBSTA to the prediction of severe outcomes among patients with COVID-19.

Given the unsatisfactory status of the existing scoring systems, we attempted to construct new models for predicting COVID-19 severity through logistic regression and ANN. Our logistic regression model has simple and feasible application to clinical practice, considering it only contains five variables. Furthermore, the internal validation of this dataset showed excellent prediction results (AUC = 0.857, Youden Index = 0.65). The five items in the NLHA2 model are easy to acquire, which is conductive to the model's clinical application and external validation.

In contrast to the NLHA2 model, the ANN model more accurately predicted ICU admission. It correctly classified every patient in the training set and more than 90% of patients in the testing set (testing set: AUC = 0.907; Youden Index = 0.84). However, it was difficult to apply the ANN model because of the black box portion of the network (a complicated relationship network that cannot be completely understood). With this in mind, we may need to assess each patient's status using the R environment until we develop a visualized version of the scoring system. Despite the model's high predictive accuracy, its complexity limits its clinical application value.

To validate the models, we carried out Hosmer-Lemeshow goodness of fit testing and Brier score analysis, which are important tests but generally neglected in other studies. In summary, our results indicate that the NLHA2 and ANN models both have the potential to predict COVID-19 severity, but more research is required to confirm our conclusions.

This study has several limitations that future studies should address. First, we hope to dynamically monitor the scores of hospitalized patients to explore the characteristics of their scores regarding the course of disease progression. Second, we plan to compare our prediction results with physician judgment to investigate the stability and utility of the models. Finally, and most importantly, the current virus strain is significantly different from the former versions. However, the emergence of mutant strains does not render past models completely worthless. Since the data we used to build our scoring models were based on clinical manifestations of the patient rather than the characteristics of the virus variants, it should be widely applicable to different strains. This is especially likely considering that we found the MuLBSTA system performed well in predicting COVID-19 (Fig. 2A–C) even though it was originally designed to predict mortality among patients with unspecific viral pneumonia [39,40]. Furthermore, the SARS-CoV transcriptome shows high similarity to the transcriptome of SARS-CoV-2, and the same is true for different strains of SARS-CoV-2 [12,41]. Therefore, although our results were mainly based on the analysis of wild-type strains, they may also be of value to the Omicron variant. Given these points, future studies are required for the external validation of our models.

5. Conclusion

In conclusion, we identified the PSI scoring system as the best existing system for predicting ICU admission among patients with wild-type SARS-COV-2 infection with COVID-19, while the two newly-

Table 6

Discrimination and calibration analysis of the two predction models.

	NLHA2 Model		ANN Model	
	Training Set	Testing Set	Training Set	Testing Set
AUC	0.959	0.857	1.000	0.907
AUC (2000 boostrap replicates)	0.967	0.911	1.000	0.957
Hosmer-Lemeshow goodness of fit test				
X-squared	1.595	11.98	0.000	0.053
P Value	0.991	0.152	1.000	1.000
Brier Score	0.048	0.136	0.000	0.071

Table 5

ROC comparison between the NLHA2 and ANN models.

	AUC	95% CI	-LR	+LR	P Value	Sensitivity (%)	Specificity (%)	Cut-off
NLHA2 – Training Set	0.959	0.920–0.982	0.04	5.71	<0.0001	96.87	83.03	>2.5
NLHA2 – Testing Set	0.857	0.772–0.919	0.29	8.89	<0.0001	73.08	91.78	
ANN – Training Set	1.000	0.981–1.000	–	–	<0.0001	100.00	100.00	-
ANN – Testing Set	0.907	0.832–0.956	0.05	8.31	<0.0001	95.89	88.46	-

NLHA2: the logistic regression model with five variables mentioned in Table 4; ANN: artificial neural network.
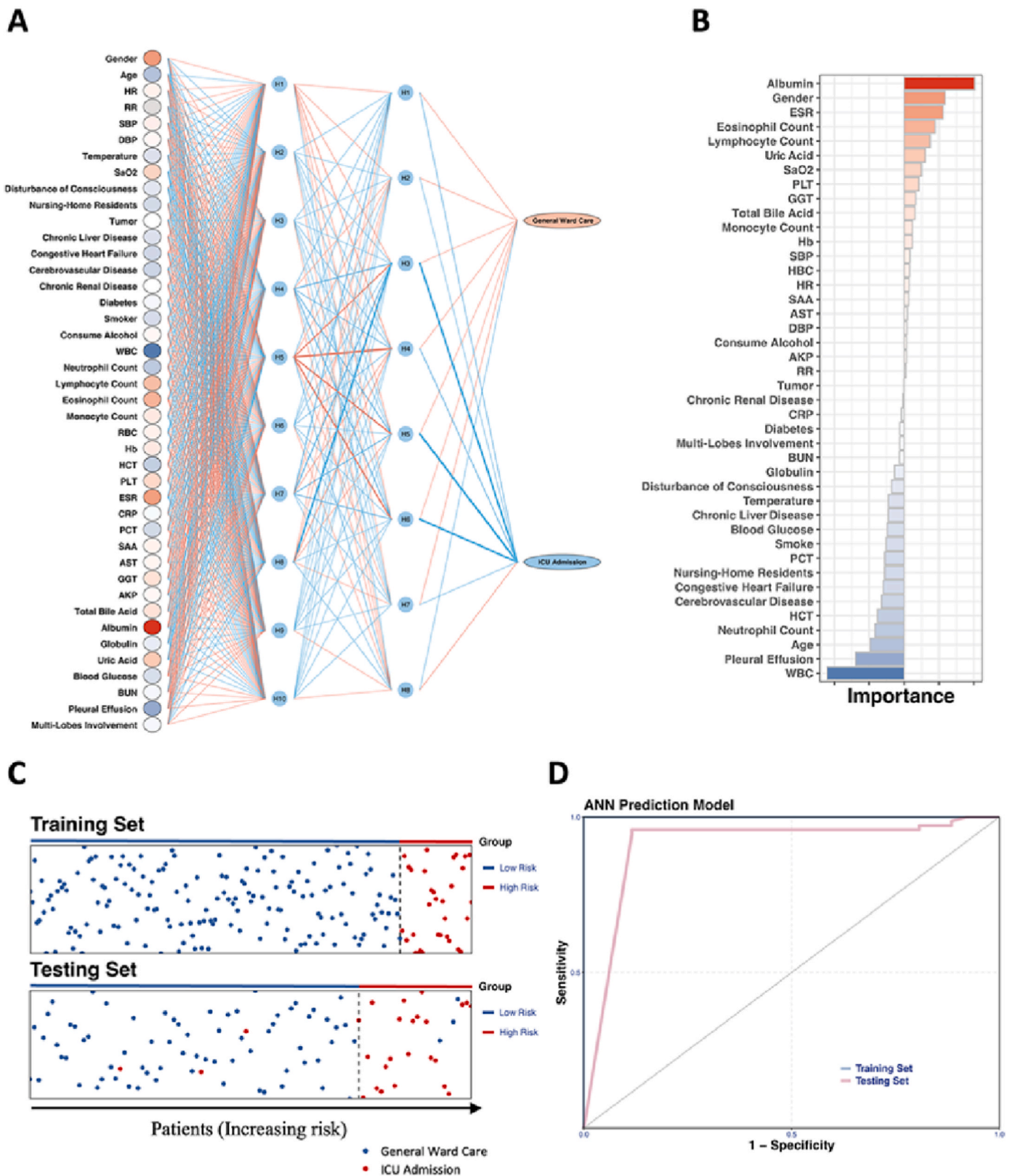
Fig. 5. The ANN model can predict the severity of COVID-19 patients with great accuracy. (A) Network for the ANN model. The colors of circles (variables) were determined by importance. The red line represents a positive correlation whereas the blue line represents a negative correlation. The thickness of the line represents the correlation index. (B) The importance of each variable was calculated using Olden's connection-weighted algorithm. (C) Distribution of the predicted ICU admissions in the General-Ward and ICU cohorts. (D) ROC curve analysis verifying the prediction performance of the ANN Model in the COVID-19 cohort.

designed models (NLHA2 and ANN) performed better and could be valuable for predicting the severity of COVID-19. This study will provide a new approach for the development of prognostic evaluation system in a novel respiratory viral epidemic.

Funding statement

CRediT authorship contribution statement

Yue Dong: Writing – original draft, Software, Methodology, Investigation, Formal analysis. **Kai Wang:** Writing – original draft, Methodology, Data curation. **Xu Zou:** Supervision, Data curation. **Xiaoping Tan:** Investigation, Data curation. **Yi Zang:** Formal analysis, Writing – review & editing. **Xinyu Li:** Data curation. **Xiaoting Ren:** Data curation. **Desheng Xie:** Supervision. **Zhijun Jie:** Supervision, Conceptualization. **Xiaohua Chen:** Data curation. **Yingying Zeng:** Writing – review & editing, Validation, Data curation, Conceptualization. **Jindong Shi:** Writing – review & editing, Validation, Supervision, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jindong Shi reports financial support was provided by the China Public Health Union.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.micpath.2022.105735.

References

[1] H. Lu, C.W. Stratton, Y.W. Tang, Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle, J. Med. Virol. 92 (2020) 401–402, https://doi.org/10.1002/jmv.25678.

[2] G.J. Cao, Z.F. Xing, L. Hua, et al., Evaluation of the diagnostic performance of panfungal polymerase chain reaction assay in invasive fungal diseases, Exp. Ther. Med. 14 (2017) 4208–4214, https://doi.org/10.3892/etm.2017.5081.

[3] X. Yang, Y. Yu, J. Xu, et al., Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study, Lancet Respir. Med. 8 (2020) 475–481, https://doi.org/10.1016/s2213-2600(20)30079-5.

[4] M.I. Restrepo, E.M. Mortensen, J. Rello, et al., Late admission to the ICU in patients with community-acquired pneumonia is associated with higher mortality, Chest 137 (2010) 552–557, https://doi.org/10.1378/chest.09-1547.

[5] D. Aujesky, M.J. Fine, The pneumonia severity index: a decade after the initial derivation and validation, Clin. Infect. Dis. 47 (Suppl 3) (2008) S133–S139, https://doi.org/10.1086/591394.

[6] A. Ilg, A. Moskowitz, V. Konanki, et al., Performance of the CURB-65 score in predicting critical care interventions in patients admitted with community-acquired pneumonia, Ann. Emerg. Med. 74 (2019) 60–68, https://doi.org/10.1016/j.annemergmed.2018.06.017.

[7] Y. Nguyen, F. Corre, V. Honsel, et al., Applicability of the CURB-65 pneumonia severity score for outpatient treatment of COVID-19, J. Infect. 81 (2020) e96–e98, https://doi.org/10.1016/j.jinf.2020.05.049.

[8] G. Fan, C. Tu, F. Zhou, et al., Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study, Eur. Respir. J. 56 (2020), https://doi.org/10.1183/13993003.02113-2020.

[9] M.M. García Clemente, J. Herrero Huertas, A. Fernández Fernández, et al., Assessment of risk scores in Covid-19, Int. J. Clin. Pract. (2020), e13705, https://doi.org/10.1111/ijcp.13705.

[10] R. Xu, K. Hou, K. Zhang, et al., Performance of two risk-stratification models in hospitalized patients with coronavirus disease, Front. Med. 7 (2020) 518, https://doi.org/10.3389/fmed.2020.00518.

[11] R. Zolfaghari Emameh, R. Falak, E. Bahreini, Application of system biology to explore the association of Neprilysin, angiotensin-converting enzyme 2 (ACE2), and carbonic anhydrase (CA) in pathogenesis of SARS-CoV-2, Biol. Proced. Online 22 (2020) 11, https://doi.org/10.1186/s12575-020-00124-6.

[12] R. Zolfaghari Emameh, H. Nosrati, M. Eftekhari, et al., Expansion of single cell transcriptomics data of SARS-CoV infection in human bronchial epithelial cells to COVID-19, Biol. Proced. Online 22 (2020) 16, https://doi.org/10.1186/s12575-020-00127-3.

[13] A. Goulas, F. Damicelli, C.C. Hilgetag, Bio-instantiated recurrent neural networks: integrating neurobiology-based network topology in artificial networks, Neural Network. 142 (2021) 608–618, https://doi.org/10.1016/j.neunet.2021.07.011.

[14] S. Shan, S. Luo, Z. Yang, et al., Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization, Proc. Natl. Acad. Sci. U. S. A. 119 (2022), e2122954119, https://doi.org/10.1073/pnas.2122954119.

[15] W. Jin, J.M. Stokes, R.T. Eastman, et al., Deep learning identifies synergistic drug combinations for treating COVID-19, Proc. Natl. Acad. Sci. U. S. A. 118 (2021), https://doi.org/10.1073/pnas.2105070118.

[16] Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, et al., Do we need hundreds of classifiers to solve real world classification problems? J. Mach. Learn. Res. 15 (2014) 3133–3181, https://doi.org/10.5555/2627435.2697065.

[17] Jacques Wainer, Comparison of 14 different families of classification algorithms on 115 binary datasets, arXiv preprint (2016), https://doi.org/10.48550/arXiv.1606.00930. 1606.00930.

[18] W.J. Meurer, J. Tolles, Logistic regression diagnostics: understanding how well a model predicts outcomes, JAMA 317 (2017) 1068–1069, https://doi.org/10.1001/jama.2016.20441.

[19] V. Renganathan, Overview of artificial neural network models in the biomedical domain, Bratisl. Lek. Listy 120 (2019) 536–540, https://doi.org/10.4149/bll_2019_087.

[20] J.L. Liu, F. Xu, H. Zhou, et al., Expanded CURB-65: a new score system predicts severity of community-acquired pneumonia with superior efficiency, Sci. Rep. 6 (2016), 22911, https://doi.org/10.1038/srep22911.

[21] C. Gonzalez, T. Johnson, K. Rolston, et al., Predicting pneumonia mortality using CURB-65, PSI, and patient characteristics in patients presenting to the emergency department of a comprehensive cancer center, Cancer Med. 3 (2014) 962–970, https://doi.org/10.1002/cam4.240.

[22] Y. Shen, T. Yang, X. Cao, et al., Conjugation of DM1 to anti-CD30 antibody has potential antitumor activity in CD30-positive hematological malignancies with lower systemic toxicity, mAbs 11 (2019) 1149–1161, https://doi.org/10.1080/19420862.2019.1618674.

[23] C. Huang, Y. Wang, X. Li, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506, https://doi.org/10.1016/s0140-6736(20)30183-5.

[24] World Health Organization, Clinical Management of COVID-19: Living Guideline, 2022.

[25] J.H. Chen, S.S. Chang, J.J. Liu, et al., Comparison of clinical characteristics and performance of pneumonia severity score and CURB-65 among younger adults, elderly and very old subjects, Thorax 65 (2010) 971–977, https://doi.org/10.1136/thx.2009.129627.

[26] P.G. Charles, R. Wolfe, M. Whitby, et al., SMART-COP: a tool for predicting the need for intensive respiratory or vasopressor support in community-acquired pneumonia, Clin. Infect. Dis. 47 (2008) 375–384, https://doi.org/10.1086/589754.

[27] Y. Iijima, T. Okamoto, T. Shirai, et al., MuLBSTA score is a useful tool for predicting COVID-19 disease behavior, J. Infect. Chemother. 27 (2021) 284–290, https://doi.org/10.1016/j.jiac.2020.10.013.

[28] H. Kang, The prevention and handling of the missing data, Korean J Anesthesiol 64 (2013) 402–406, https://doi.org/10.4097/kjae.2013.64.5.402.

[29] E.W. Steyerberg, F.E. Harrell Jr., G.J. Borsboom, et al., Internal validation of predictive models: efficiency of some procedures for logistic regression analysis, J. Clin. Epidemiol. 54 (2001) 774–781, https://doi.org/10.1016/s0895-4356(01)00341-9.

[30] M. van Smeden, K.G. Moons, J.A. de Groot, et al., Sample size for binary logistic prediction models: beyond events per variable criteria, Stat. Methods Med. Res. 28 (2019) 2455–2474, https://doi.org/10.1177/0962280218784726.

[31] Julian D. Olden, Michael K. Joy, Russell G. Death, An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, Ecol. Model. 178 (2004) 389–397, https://doi.org/10.1016/j.ecolmodel.2004.03.013.

[32] J. Sun, W.T. He, L. Wang, et al., COVID-19: epidemiology, evolution, and cross-disciplinary perspectives, Trends Mol. Med. 26 (2020) 483–495, https://doi.org/10.1016/j.molmed.2020.02.008.

[33] G. Pascarella, A. Strumia, C. Piliego, et al., COVID-19 diagnosis and management: a comprehensive review, J. Intern. Med. 288 (2020) 192–206, https://doi.org/10.1111/joim.13091.

[34] E.G. Price-Haywood, J. Burton, D. Fort, et al., Hospitalization and mortality among black patients and white patients with covid-19, N. Engl. J. Med. 382 (2020) 2534–2543, https://doi.org/10.1056/NEJMsa2011686.

[35] J. Zhang, X. Wang, X. Jia, et al., Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China, Clin. Microbiol. Infect. 26 (2020) 767–772, https://doi.org/10.1016/j.cmi.2020.04.012.

[36] N. Chen, M. Zhou, X. Dong, et al., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, Lancet 395 (2020) 507–513, https://doi.org/10.1016/s0140-6736(20)30211-7.

[37] M.J. Fine, T.E. Auble, D.M. Yealy, et al., A prediction rule to identify low-risk patients with community-acquired pneumonia, N. Engl. J. Med. 336 (1997) 243–250, https://doi.org/10.1056/nejm199701233360402.

[38] W.S. Lim, S.V. Baudouin, R.C. George, et al., BTS guidelines for the management of community acquired pneumonia in adults: update 2009, Thorax 64 (Suppl 3) (2009), https://doi.org/10.1136/thx.2009.121434 iii1-55.

[39] L. Guo, D. Wei, X. Zhang, et al., Clinical features predicting mortality risk in patients with viral pneumonia: the MuLBSTA score, Front. Microbiol. 10 (2019) 2752, https://doi.org/10.3389/fmicb.2019.02752.

[40] B. Ma, J. Gong, Y. Yang, et al., Applicability of MuLBSTA scoring system as diagnostic and prognostic role in early warning of severe COVID-19, Microb. Pathog. 150 (2021), 104706, https://doi.org/10.1016/j.micpath.2020.104706.

[41] A. Stukalov, V. Girault, V. Grass, et al., Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV, Nature 594 (2021) 246–252, https://doi.org/10.1038/s41586-021-03493-4.