

Development and internal verification of the hospitalization risk score to triage COVID-19 patients upon admission: a multicenter nested case-control study.

Hongbing Peng

Department of Pulmonary and Critical Care Medicine,Loudi Central Hospital

Chao Hu

Department of Pulmonary and Critical Care Medicine, Xiangtan Central Hospital;

Wusheng Deng

Department of Pulmonary and Critical Care Medicine, Shaoyang Central Hospital

Lingmei Huang

Department of Pulmonary and Critical Care Medicine, the First People's Hospital of YueYang,

Baowei Luo

Department of Respiratory Medicine,Shuangfeng County people's Hospital

Xingxing Wang

Department of Infection, Loudi Central Hospital,

Xiaodan Long

Department of Urology Surgery,Loudi Central Hospital

Honghui Li (✉ redcat8851@163.com)

Medical Rehabilitation Center, Loudi Central Hospital

Xiaoying Huang (✉ 1921596923@qq.com)

Department of Pulmonary and Critical Care Medicine,Loudi Central Hospital

Research Article

Keywords: Risk Score, Prediction model, COVID-19, nested case-control, incubation period, semi-quantitative CT scores

Posted Date: March 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-252154/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Novel coronavirus disease 2019(COVID-19)has a large population base of infection, and only worsening cases required hospitalization. A hospitalization risk score of COVID-19(HRS-COVID-19) to identify and treat worsening patients early thus seems crucial.

Method: We conducted a multicenter nested case-control study with 200 cases enrolling confirmed symptomatic COVID-19 patients from four designated hospitals before worsened. Least Absolute Shrinkage and Selection Operator(LASSO), Directed Acyclic Graph(DAG)and Change-in-Estimate(CIE) screened out independent risk factors for HRS-COVID-19 from demographic, clinical and imaging data of the cases. The HRS-COVID-19 was evaluated by the area under curve(AUC), calibration curve, decision curve and clinical impact curve; internal validation by the bootstrap-resampling(boot=1000); and through a nomogram and a network calculator to show.

Results: In the nested case cohort, 50 patients reached the compound endpoint(requiring hospitalization) and 150 patients were able to avoid hospitalization. Dyspnea, incubation period, age, lymphocyte count, C-reactive protein and semi-quantitative CT scores were included in HRS-COVID-19. The HRS-COVID-19 has good fitting(Hosmer-Lemeshow goodness, $P=0.45$); high discrimination(AUC 0.980,95%CI, 0.965-0.996); with excellent calibration and clinical benefits. We make a free online risk calculator (<https://hospitalization-risk-score-of-covid-19.shinyapps.io/DynNomapp/>).

Conclusion: In the study, the HRS-COVID-19 based on the clinical characteristics at admission is helpful for early identification and active treatment of aggravated high-risk COVID-19 patients.

Introduction

Since December 2019,The global pandemic of COVID-19 caused by the severe acute respiratory syndrome coronavirus (SARS-CoV-2) has continued more than one year[1, 2]. As of December 21, 2020, the WHO reported that more than 77 million people worldwide had been infected and more than 1.7 million had died. The pandemic poses a huge threat to global health.

As all know, SARS-CoV-2 is highly contagious and susceptible to infection, resulting in widespread epidemic among people[3-5].The fatality rate in the early stage of the epidemic is more than 7.0%[6, 7].The clinical outcome of different severity was heterogeneous, and the mild or common cases often rely on their own immune ability to recover, and only medical observation[8, 9]. However, most patients with severe or critical COVID-19 have mild initial symptoms, and the median time from onset to sepsis is 10.0 days (IQR7.0-14.0)[5].Among them, the time from symptom onset to first medical treatment, and from symptom attack to hospitalization of dead cases are longer than those of recovered cases[7]. Therefore, early active treatment is possible to improve the clinical outcome.

Due to the large population base of COVID-19and the majority of mild or common patients, it would be devastating and wasteful to treat all positive cases equally and actively. Therefore, early identification of

COVID-19 requiring hospitalization and active treatment is the most effective disease management. Many predictive models about COVID-19 had been published, focusing on diagnosis and prognosis[8, 10–12], and few models had been used to predict the hospitalization risk to triage patients. Jehi,L and colleagues[13] developed and validated a hospital risk prediction model with 24 variables; however, all the variables screened by LASSO regression were included in the model, and there is a risk of over-fitting. In addition, the variables are too complex for clinicians. Facing many confirmed patients, a simple and effective prediction model which can early predict aggravation and triage is particularly important.

In this study, we develop a simple hospitalization risk score(HRS-COVID-19) with six variables including incubation period, dyspnea, age, lymphocyte count, C-reactive protein(CRP) and semi-quantitative chest CT scores(CT scores). Nomogram or network calculator for HRS-COVID-19 is used to calculate the risk value for each COVID-19 at first visitor admission; clinicians can determine whether the patient is hospitalized or only for medical isolation observation based on the risk value and local medical conditions.

Method

Participants and design

We established a case cohort from the four hospitals (Loudi Central Hospital, Xiangtan Central Hospital, Yueyang First People's Hospital and Shaoyang Central Hospital)in Hunan Province between January 11 and February 28, 2020. This study was approved by the ethics committee of the Loudi central hospital. Considering the particularity of COVID-19 disease, all the data were collected anonymously and retrospectively and did not invade the privacy of patients, and the Ethics Committee abandoned the written informed consent for the new infectious diseases.

According to the guidelines for the diagnosis and treatment of COVID-19 issued by the Chinese Health Commission (1st-7th edition)[14], we enrolled 209 patients who were diagnosed by pharyngeal or nasopharyngeal swabs. And for diagnostic grading, mild patients refer to patients with mild clinical symptoms and no exudative lesions on chest imaging. Common patients refer to the symptoms such as fever and respiratory tract, and abnormalities on chest X-ray or CT. The criteria for severe diagnosis are any of the following: Shortness of breath, respiratory frequency \geq 30 beats/min; resting state, finger oxygen saturation \leq 93%; Oxygen partial pressure (PaO₂)/oxygen absorption concentration (FiO₂) \leq 300mmHg (1mmHg = 0.133kPa);lung imaging showed the lesion progressed more than 50% within 24–48 hours. Critically ill patients meet any of the following criteria: respiratory failure, mechanical ventilation, shock, and other organ failure that needs intensive care.

ThreeCOVID-19patients who were asymptomatic since onset and three who were diagnosed as severe on admission were excluded. The remaining 203 patients were followed up for 30 days, and 50 patients worsened and reached the compound endpoint(progression to severe, critical, or persistent symptoms with continuous second progression of chest X-ray),and were included in the hospitalization group. The

progression-free patients were matched with the hospitalization group at 1:3 according to gender, and the remaining 2 males and 1 female failed to match. So 150 patients were classified as non-hospitalization group(Fig. 1).

Data collection

The cases data included demographics, comorbidities, clinical symptoms, laboratory results and imaging data, all of which were collected and cross-verified by two experienced physicians in each COVID-19 treatment center from the electronic medical records.

The demographic include: age, sex, Wuhan origin (living in Wuhan, traveling or taking public transport through Wuhan), incubation period (The time from the first exposure to the onset time); Comorbidities information: NO. of comorbidities, coronary heart disease, hypertension, endocrine system disease(Diabetes, Obesity, Hyperlipidemia and Hyperthyroidism), chronic lung disease (chronic obstructive pulmonary disease, asthma, chronic bronchitis, emphysema and bronchiectasis), malignant tumor and chronic digestive system diseases (viral hepatitis, liver cirrhosis, fatty liver, and drug-induced liver injury);Symptoms at admission: fever (the highest body temperature before admission), cough, dyspnea, headache, dizziness, muscle pain, fatigue, gastrointestinal symptoms; Laboratory data: White blood cell count, neutrophil count, lymphocyte count, neutrophil/lymphocyte ratio, platelet count, hemoglobin, D-dimer, albumin, total bilirubin, direct bilirubin, creatine kinase, creatine kinase isoenzyme, lactate dehydrogenase, myoglobin, urea nitrogen, creatinine, blood glucose, C-reactive protein, procalcitonin. Imaging data: A semi-quantitative scoring system was used to evaluate the scores of each affected lobe, 0 means no involvement; 1, less than 5%; 2, 5%-25%; 3, 26%-49%; 4, 50%-75%; and 5, more than 75%; and the total pulmonary involvement score was obtained by adding them[15–17]. The image analysis was independently scored by two radiologists according to the semi-quantitative score system, and then the average value was taken.

Statistical methods and Variable Selection

In the baseline data table, Continuous variables are presented as median (inter quartile range, IQR) and categorical variables as n (%).We used the Mann-Whitney U test, Chi-square test, or Fisher's exact test to compare the differences between the non-hospitalization group and the hospitalization group when appropriate. R Software (version 3.6.3, R Foundation) performs all statistical analysis. All cases were included in variable selection and risk score development. All variables from admission or the first outpatient COVID-19 patients must collect disease information, and no missing value. The Least Absolute Shrinkage and Selection Operator(LASSO) regression was applied to data dimensionality reduction by L1 penalized to avoid potential collinearity and over fitting among variables. The LASSO regression adopted 10-fold cross-validation to select the best lambda value. Under the minimum lambada (lambada.min) compression, the variables with small regression coefficients are directly compressed to 0, to eliminate the corresponding variables. Finally, only the strongest predictors are retained in the regression model. The glmnet packet completes LASSO regression[18].

Risk score development and internal validation

Variables selected by LASSO regression were first analyzed by single-factor regression. According to the characteristic variables of published articles on severe or critical cases[7, 19], Directed Acyclic Graph(DAG)[20] and Change-in-Estimate(CIE)[21]were used to screen out the variable of binary logistic regression. The package caret[22] was used to establish abinary logistic regression score. We used Hosmer-lemeshow to evaluate the Logistic regression model, and perform internal validation with 1000 times of bootstrap-resampling; and use nomogram and network calculator to visualize the predictive ability of HRS-COVID-19 by the package DynNom[23].

Evaluation risk score

The discriminating ability of HRS-COVID-19 was assessed using area under the Receiver Operating Characteristic(ROC) Curve(AUC)by pROC packets[24]. The AUC was used to compare HRS-COVID-19 with the CURB-65 model[25], which has been used as the admission criteria for community-acquired pneumonia. With 1000 bootstrap-resampling methods, Calibration curve of the nomogram visually demonstrated the consistency between the predicted results and the real results of HRS-COVID-19(rms packets)[26].To assess the clinical usefulness of HRS-COVID-19, we set different decision thresholds (Pt 10%-95%) to make a decision curve based on the equation

$$\text{Netbenefit} = \frac{\text{True Positive Count}}{n} - \frac{\text{False Positive Count}}{n} \left(\frac{Pt}{1-Pt} \right)$$

[27]to evaluate the net benefit of HRS-COVID-19. Clinical impact curve shows clinical practical value.

Results

Patients Demographic and Characteristics

200 COVID-19 patients were matched at the 1:3 ratio by sex(hospitalization risk group: non-hospitalization risk group) a straining set to build scores. Of these, 50 (25%) reached the study endpoint (requiring hospitalization), including 40 (20%) progressed to severe, 5 (2.5%) progressed to critical, 2 (1.0%) eventually died, and the remaining patients were discharged. Age(median 57.1vs41.1);Slightly more men (28/50);Patients with fever (≥ 38.0 , 64.0%vs 39.97%), dyspnea (72% vs 8%), cough (86.0% vs 70.0%) and fatigue (48.0% vs 21.3%) were at relatively high risk of admission.

Similarly, patients in the hospitalization risk group were more likely to have comorbidities (70% vs 24.7%), especially endocrine and metabolic diseases (32.0% vs 8.67%), hypertension (36.6% vs 9.33%), respiratory diseases (16.0% vs 4.67%) and coronary heart disease (12.0% vs 2.67%).There was only one case of malignant tumor. In the study, we noted that patients with a short incubation period also had a higher risk of hospitalization (5daysvs7days).All the basic feature are shown in Table 1.

Table 1

Demographic and clinical characteristics of hospitalized or non-hospitalized patients.

	All patients N = 200	non-Hospitalization N = 150	Hospitalization N = 50	p-value
Characteristics				
Age [median(SD)]	45.1 (16.4)	41.1 (15.0)	57.1 (14.7)	< 0.001
< 35	58(29.0%)	54(36.0%)	4(8.00%)	< 0.001
35–45	42(21.0%)	34(22.7%)	8(16.0%)	
45–55	54(27.0%)	41(27.3%)	13(26.0%)	
> 55	46(23.0%)	21(14.0%)	25(50.0%)	
Gender				0.870
female	88 (44.0%)	66 (44.0%)	22 (44.0%)	
male	112 (56.0%)	84 (56.0%)	28 (56.0%)	
Imput case(Close contact with Wuhan)				
No	120 (60.0%)	91 (60.7%)	29 (58.0%)	
Yes	80 (40.0%)	59 (39.3%)	21 (42.0%)	
Frequent breathing	20.0[20.0–21.0]	20.0 [19.2–20.0]	20.5[20.0–22.0]	< 0.001
Blood pressure on admission				
Systolic pressure	123 (13.4)	123 (12.9)	125 (14.8)	0.332
diastolic pressure	80.0 [70.0–84.0]	80.0 [70.0–83.0]	80.0[74.5–85.8]	0.133
Symptoms on admission				
Temperature(°C)	38.0 [36.8–38.5]	37.8 [36.6–38.4]	38.3 [38.0-38.6]	< 0.001
< 37.3	56(28.0%)	53(35.3%)	3(6.00%)	< 0.001
37.3–38.0	52(26.0%)	37(24.7%)	15(30.0%)	
38.0–39.0	83(41.5%)	56(37.3%)	27(54.0%)	
> 39.0	8(4.50%)	4(2.67%)	5(10.0%)	
Dyspnea				< 0.001

Continuous variable data are presented as median (inter quartile ranges, IQR). Classified variable data are presented as n(%). Unless otherwise stated, the Mann-Whitney U test is used for the continuous variable, the χ^2 test or the Fisher's exact test for the categorical variable.

	All patients	non-Hospitalization	Hospitalization	p-value
	N = 200	N = 150	N = 50	
No	152 (76.0%)	138 (92.0%)	14 (28.0%)	
Yes	48 (24.0%)	12 (8.00%)	36 (72.0%)	
Cough				0.041
No	52 (26.0%)	45 (30.0%)	7 (14.0%)	
Yes	148 (74.0%)	105 (70.0%)	43 (86.0%)	
Headache				1.000
No	175 (87.9%)	131 (87.9%)	44 (88.0%)	
Yes	24 (12.1%)	18 (12.1%)	6 (12.0%)	
Fatigue				0.001
No	144 (72.0%)	118 (78.7%)	26 (52.0%)	
Yes	56 (28.0%)	32 (21.3%)	24 (48.0%)	
Muscle soreness				0.237
No	184 (92.0%)	140 (93.3%)	44 (88.0%)	
Yes	16 (8.00%)	10 (6.67%)	6 (12.0%)	
Gastrointestinal symptoms(Nausea, Vomiting, Diarrhea)				0.145
No	174 (87.0%)	134 (89.3%)	40 (80.0%)	
Yes	26 (13.0%)	16 (10.7%)	10 (20.0%)	
Incubation period	7.0 [5.0–9.0]	7.0 [5.0–10.0]	5.0 [3.25-7.0]	< 0.001
Comorbidities				< 0.001
No	128 (64.0%)	113 (75.3%)	15 (30.0%)	
Yes	72 (36.0%)	37 (24.7%)	35 (70.0%)	
No.of comorbidities	0[0–1]	0 [0–1]	1[1–2]	< 0.001
Respiratory disease				0.014
(COPD, Chronic bronchitis, Emphysema, Asthma and Bronchiectasis)				
No	185 (92.%)	143 (95.3%)	42 (84.0%)	

Continuous variable data are presented as median (inter quartile ranges, IQR). Classified variable data are presented as n(%). Unless otherwise stated, the Mann-Whitney U test is used for the continuous variable, the χ^2 test or the Fisher's exact test for the categorical variable.

	All patients	non-Hospitalization	Hospitalization	p-value
	N = 200	N = 150	N = 50	
Yes	15 (7.50%)	7 (4.67%)	8 (16.0%)	
Cardiovascular diseases				0.017
No	190 (95.0%)	146 (97.3%)	44 (88.0%)	
Yes	10 (5.00%)	4 (2.67%)	6 (12.0%)	
Endocrine System disease and Metabolic related diseases (Diabetes, Obesity, Hyperlipidemia and Hyperthyroidism)				< 0.001
No	171 (85.5%)	137 (91.3%)	34 (68.0%)	
Yes	29 (14.5%)	13 (8.67%)	16 (32.0%)	
Hypertension				< 0.001
No	168 (84.0%)	136 (90.7%)	32 (64.0%)	
Yes	32 (16.0%)	14 (9.33%)	18 (36.0%)	
Malignant tumour				1.000
No	199 (99.5%)	149 (99.3%)	50 (100%)	
Yes	1 (0.50%)	1 (0.70%)	0 (0.00%)	
Digestive disease(Hepatitis B, Drug-induced hepatitis, Fatty liver)				0.217
No	187 (93.5%)	142 (94.7%)	45 (90.0%)	
Yes	13 (6.50%)	8 (5.30%)	5 (10.0%)	
Continuous variable data are presented as median (inter quartile ranges, IQR). Classified variable data are presented as n(%). Unless otherwise stated, the Mann-Whitney U test is used for the continuous variable, the χ^2 test or the Fisher's exact test for the categorical variable.				

Laboratory outcomes

From the laboratory data, we observed there were significant differences in the several blood tests between the hospitalized group and the non-hospitalized group. For example, NLR(5.55vs 2.88), Myoglobin(104vs56.8), Lactate dehydrogenase (LDH, 255vs204), D-dimer(0.3vs0.22), C-reactive protein(CRP29.6vs5.54), and CT score(11.2vs3)increased in hospitalization group. Other, lymphocyte count (0.70vs1.14), platelet count (155vs208), and albumin (35.9vs39.3) decreased(Table 2).

Table 2
Laboratory findings of all patients on admission.

	All patients	non-Hospitalization	Hospitalization	p-value
	N = 200	N = 150	N = 50	
White cell count, × 10³/L	4.96[3.80–6.59]	4.98 [3.80–6.46]	4.94[3.69–7.20]	0.671
Neutrophil count, × 10³/L	3.65 [2.60–6.64]	3.55 [2.54–6.03]	4.14[2.77–7.71]	0.114
Lymphocyte count, ×10³/L	1.00 [0.73–1.51]	1.14 [0.84–1.62]	0.70[0.47–0.82]	< 0.001
NLR*	3.30 [2.08–6.69]	2.88 [1.88–5.29]	5.55[3.38–9.24]	< 0.001
Platelet count, × 10³/L	196 [146–246]	208 [154–248]	155 [126–209]	0.003
Haemoglobin, g/L	132 [122–146]	135 [123–147]	130 [120–144]	0.106
D-dimer, mg/L	0.23 [0.13–0.36]	0.22 [0.13–0.33]	0.30[0.13–0.57]	0.013
Albumin, g/L	39.0 [35.5–41.0]	39.3 [37.0-42.7]	35.9[33.0-38.9]	< 0.001
Myoglobin, ng/mL	62.8 [35.4–112]	56.8 [34.3–107]	104 [50.1–134]	< 0.001
C-reactive protein(mg/L)	7.55 [2.97-25.0]	5.54 [1.90–10.4]	29.6[12.3–63.3]	< 0.001
Creatine kinase, U/L	77.0 [49.8–109]	73.0 [48.5–97.0]	87.5 [53.0-160]	0.021
Creatine kinase MB, U/L	13.6 [9.52–17.8]	13.0 [9.00–17.0]	15.1[12.0-19.8]	0.016
Lactate dehydrogenase, U/L	214 [163–268]	204 [160–248]	255 [211–329]	< 0.001
Blood urea nitrogen, mmol/L	4.10 [3.10–5.10]	3.90 [3.00-4.89]	4.87[3.92–5.47]	< 0.001
Creatinine, µmol/L	71.0 [58.8–80.0]	70.3 [56.2–79.1]	75.7[63.0-81.2]	0.062

Continuous variable data are presented as median (inter quartile ranges, IQR). Unless otherwise stated, the Mann-Whitney U test is used for the continuous variable, the χ^2 test or the Fisher's exact test for the categorical variable.

*NLR : Neutrophil-to-lymphocyte ratio.

	All patients	non-Hospitalization	Hospitalization	p-value
	N = 200	N = 150	N = 50	
Procalcitonin, ng/mL	0.04 [0.04–0.10]	0.04 [0.04–0.10]	0.08[0.03–0.10]	0.045
Blood Glucose,mmol/L	5.80 [5.20–7.78]	5.60 [4.93–6.72]	7.02[5.62–11.1]	<0.001
Total bilirubin, mmol/L	9.90 [7.80–15.9]	9.77 [7.80–16.0]	10.4[8.17–15.8]	0.530
Direct bilirubin, mmol/L	4.00 [3.21–5.68]	4.00 [3.16–5.69]	3.91[3.46–5.52]	0.873
Semi-quantitative chest CT Score	4.00 [2.00–8.00]	3.00 [2.00–5.00]	11.2[8.00-14.8]	<0.001

Continuous variable data are presented as median (inter quartile ranges, IQR). Unless otherwise stated, the Mann-Whitney U test is used for the continuous variable, the χ^2 test or the Fisher's exact test for the categorical variable.

※NLR : Neutrophil-to-lymphocyte ratio.

Predictor Selection

The 39 variables measured at admission (Tables 1 and 2, excluding gender and imported cases) were included in LASSO regression (eFigure1 in the supplement), and 13 variables with non-zero coefficients were obtained, and including incubation period, dyspnea, age, NO. Of comorbidities, respiratory diseases, hypertension, lymphocyte count, D-dimer, CRP Procalcitonin, urea nitrogen, Blood Glucose and CT Score. In univariate regression analysis(Table 3), all variables were independent risk factors for hospitalization ($P < 0.05$).When all variables were included in the Logistic regression model, only the CT score was correlated with the risk of admission ($P < 0.001$).Next, based on DAG(eFigure2 in the supplement), age on the risk of hospitalization, it was more pronounced after not adjusting for intermediate variables, NO. of comorbidities, respiratory disease and hypertension. According to CIE(eTable1 in the supplement), combined with clinical and published studies, incubation period, dyspnea, age, lymphocyte count, CRP and CT score were selected as risk score variables.

Table 3
Univariable and multivariable Logistics regression models on the risk of hospitalization.

Variables	Univariable		multivariable	
	Odds ratio(95%CI)	P value	Odds ratio(95%CI)	P value
Dyspnea	29.58(12.59–69.45)	< 0.001	3.96(1.04–15.34)	0.04
Respiratory system disease	3.89(1.33–11.36)	0.01		
Hypertension	5.46(2.46–12.13)	< 0.001		
Incubation period	0.79(0.70–0.90)	< 0.001	0.82(0.67–0.99)	0.049
No of comorbidities	2.61(1.83–3.73)	< 0.001		
Age	1.07(1.05–1.10)	< 0.001	1.08(1.02–1.14)	0.01
Lymphocyte count	0.01(0.00-0.06)	< 0.001	0.12(0.01–0.81)	0.045
D-dimer	7.68(2.19–26.94)	0.001		
C-reactive protein	1.04(1.02–1.06)	< 0.001	1.02(1-1.05)	0.07
Blood urea nitrogen	1.34(1.11–1.62)	0.002		
Blood Glucose	1.16(1.06–1.26)	< 0.001		
Procalcitonin	56.15(2.76-1144.22)	0.01		
Semi-quantitative chest CT score	1.82(1.52–2.18)	< 0.001	1.59(1.32–2.10)	< 0.001

Construction of hospitalization risk score

In multivariate analysis(Table 3), the logistic regression model identified the correlation between 6 variables and HRS-COVID-19. Incubation period(OR, 0.82;95% CI,0.67–0.99; $P= 0.049$) and lymphocyte count(OR, 0.12;95%CI,0.01–0.81; $P= 0.045$) were negatively correlated. Age(OR, 1.08;95%CI,1.02-14; $P= 0.01$), dyspnea (OR, 3.96;95%CI,1.04–16.34; $P= 0.04$),CRP(OR,1.02;95%CI,1-1.05; $P= 0.07$) and CT score (OR 1.59;95%CI,1.32–2.10; $P< 0.001$)were positively correlated. Hosmer-lemeshow test showed a lack of significance in the regression model ($P= 0.45$). Similarly, the model verified by bootstrap-resampling 1000 times, R2 is 0.61, indicating that the prediction accuracy of HRS-COVID-19 is good. A nomogram for HRS-COVID-19 containing Incubation period, dyspnea, age, CRP, lymphocyte count, and CT score was constructed(Fig. 2a).An online calculator based on nomogram has been developed, which allows clinicians to automatically calculate the hospitalization risk of COVID-19 patients (95%CI)(Fig. 2b),

(<https://hospitalization-risk-score-of-covid-19.shinyapps.io/ DynNomapp/>).

Risk score evaluation and internal validation

The AUC of HRS-COVID-19 was 0.980(95%CI, 0.965–0.996, specificity 0.953, sensitivity 0.940), showing excellent differentiation, outstanding sensitivity and specificity(Fig. 3a). The optimal cut-off value of ROC curve is 0.353. Interestingly, the calibration curve showed the predicted risk of hospitalization was greater than 0.30, the predicted value was consistent with the ideal value(Fig. 3b). The decision curve shows good net benefits, when the threshold probability is between 0.0 and 0.95(Fig. 3c).The clinical impact curve of HRS-COVID-19,among 1000 patients, visually shows the number of high-risk hospitalization(solid red line) versus the actual number of hospitalizations required (dotted blue line, Fig. 3d).We carried out internal validation of the nomogram and obtained the C statistic 0.977 through bootstrap-resampling for 1000 times.

Comparison of two scoring systems

The hospitalization standard of community-acquired pneumonia was that the CURB-65 score was greater than 2. The AUC value of CURB-65 is 0.833(95% CI 0.7703–0.8953), and the AUC value of HRS-COVID-19 is 0.980(95% CI 0.965–0.9958). The two scoring systems were compared by delong method ($P < 0.001$) (Fig. 3a).

Discussion

In the study, we constructed a nomogram of hospitalization risk in patients with COIVD-19. The internal verification indicates that the HRS-COVID-19 fits well. Our results demonstratedIncubation period, dyspnea, age, CRP, lymphocyte count, and CT score were independent risk factors for predicting exacerbation of COVID-19 patients requiring hospitalization. The network calculator for nomogram provides data reference for whether symptomatic COVID-19 patients need to be hospitalized or not, whichfacilitates the clinicians operation.

Based on the AUC value, HRS-COVID-19(AUC = 0.980) may be more appropriate as a standard of admission for COVID-19 patients with symptoms than CURB-65 (AUC = 0.833). The cutoff of the ROC curve is 0.353. Interestingly, the calibration curves of the nomogram indicated the predicted risk is more than 0.3, and the predicted value is consistent with the ideal result. Therefore, When a single symptomatic COVID-19 patient was hospitalized according to the predicted risk of more than 0.3, the predictive specificity was 0.953 and the sensitivity was 0.940. If the predicted risk of COVID-19 patients is above 0.3, they need to be hospitalized; when the predicted risk is about 0.2, the choice should be made according to the conditions of the local hospital. Clinical decision-making curve and clinical impact curve show better clinical benefits than "full hospitalization" or "non-hospitalization", when HRS-COVID-19 is used to determine whether to be hospitalized or not.

Variables commonly found in published literature, such as body temperature ($> 38.0^{\circ}\text{C}$)[12], comorbidities[28] or NO. of comorbidities[8, 29], neutrophilic-lymphocyte ratio (NLR)[30, 31], urea nitrogen[10], NT-proBNP[11], D-dimer[32], and lactate deoxyenzyme[12, 33], were not included in our risk score. It is related to the time node of data acquisition. Our study used baseline data from the onset stage

of COVID-19 patients, so there were no coagulation abnormalities, renal failure or heart involvement associated with critical illness.

We found the longer incubation period, the lower risk of developing into severe or critical, and similar results have been published[34]. Respiratory viruses induce the immune response through inflammatory mediators and cytokines leading to clinical symptoms, thus determining the incubation period[35]. Pneumonia caused by SARS-CoV-2 is also related to immune response. The shorter incubation period was associated with more pulmonary exudate lesions[34], and the greater risk of hospitalization. Therefore, the incubation period is an independent risk factor for HRS-COVID-19.

Another most prominent variable was CT scores. The CT scores of hospitalized group (mean value \pm SD: 11.4 ± 4.66) was significantly higher than non-hospitalized group ($3.47 \pm 2.60, P < 0.0001$). Francone, M. team[17] study showed the CT scores of the critical group(20.3 ± 3) and the severe group(17.4 ± 3.1) were significantly higher than those of the mild group($8.7 \pm 4, P < 0.0001$). However, there was no statistically significant difference between severe and critical conditions ($P = 0.7921$). In our study, the CT scores of the hospitalized group was between severe and mild, with some overlap, which was related to the grouping mode of our cases cohort and the goal of the study.

Limitations

Potential limitations of the study include the study pairs only collected partial data from one province, without external verification; the control of confounding factors and the elimination of intermediate variables still have the risk of misjudgment; although the nesting case study controlled for some confounding factors, it may also miss other important information. For example, the cases imported from Wuhan were more serious.

Conclusions

In the study, hospitalization risk score and web-based calculator were developed. Using HRS-COVID-19 to triage each confirmed patients with symptoms (medical observation or hospitalization) can obtain better clinical benefits, achieve the goal of optimizing medical resources and alleviating the pressure.

Declarations

Acknowledgment

Special thanks to Yuechun, Hu Prof. and Wen, Mao Prof. Radiology Department of Loudi Central Hospital, they help in lung CT score. Thanks to all the medical workers working in the front line of coronary pneumonia.

Funding/Support

This study is supported by the Special Topic Project for Pneumonia Epidemic Infected by New Coronavirus in Loudi City (award number: Lou Caijiao [2020] No. 43).

References

1. Cucinotta D, Vanelli M: **WHO Declares COVID-19 a Pandemic.** *Acta Biomed* 2020, **91**(1):157-160.
2. Viruses. CSGO: **The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.** *NAT MICROBIOL* 2020, **5**(4):536-544.
3. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y *et al:* **Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.** *LANCET* 2020, **395**(10223):507-513.
4. Argenziano MG, Bruce SL, Slater CL, Tiao JR, Baldwin MR, Barr RG, Chang BP, Chau KH, Choi JJ, Gavin N *et al:* **Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series.** *BMJ* 2020, **369**:m1996.
5. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X *et al:* **Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study.** *The Lancet* 2020, **395**(10229):1054-1062.
6. Onder G, Rezza G, Brusaferro S: **Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy.** *JAMA* 2020, **323**(18):1775-1776.
7. Deng X, Yang J, Wang W, Wang X, Zhou J, Chen Z, Li J, Chen Y, Yan H, Zhang J *et al:* **Case fatality risk of the first pandemic wave of novel coronavirus disease 2019 (COVID-19) in China.** *CLIN INFECT DIS* 2020.
8. 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.
9. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, Tobin KA, Cerfolio RJ, Francois F, Horwitz LI: **Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study.** *BMJ* 2020, **369**:m1966.
10. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, Dunning J, Fairfield CJ, Gamble C, Green CA *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**:m3339.
11. Dong YM, Sun J, Li YX, Chen Q, Liu QQ, Sun Z, Pang R, Chen F, Xu BY, Manyande A *et al:* **Development and Validation of a Nomogram for Assessing Survival in Patients with COVID-19 Pneumonia.** *CLIN INFECT DIS* 2020.
12. Bartoletti M, Giannella M, Scudeller L, Tedeschi S, Rinaldi M, Bussini L, Fornaro G, Pascale R, Pancaldi L, Pasquini Z *et al:* **Development and validation of a prediction model for severe respiratory**

- failure in hospitalized patients with SARS-CoV-2 infection: a multicentre cohort study (PREDICO study). *Clin Microbiol Infect* 2020.
13. Jehi L, Ji X, Milinovich A, Erzurum S, Merlino A, Gordon S, Young JB, Kattan MW: **Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19.** *PLOS ONE* 2020, **15**(8):e237419.
14. NHCoPsRo C: **Guideline for the diagnosis and treatment of COVID-19 infections (version 1–7).** 2020(http://www.nhc.gov.cn/xcs/zhengcwj/list_gzbd_3.shtml. Accessed 3 Mar 2020.).
15. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, Zheng D, Wang J, Hesketh RL, Yang L *et al:* **Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19).** *RADIOLOGY* 2020, **295**(3):715-721.
16. Chang YC, Yu CJ, Chang SC, Galvin JR, Liu HM, Hsiao CH, Kuo PH, Chen KY, Franks TJ, Huang KM *et al:* **Pulmonary sequelae in convalescent patients after severe acute respiratory syndrome: evaluation with thin-section CT.** *RADIOLOGY* 2005, **236**(3):1067-1075.
17. Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, Panebianco V, Andreoli C, Colaiacomo MC, Zingaropoli MA *et al:* **Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis.** *EUR RADOL* 2020:1-10.
18. Friedman J, Hastie T, Tibshirani R: **Regularization Paths for Generalized Linear Models via Coordinate Descent.** *J STAT SOFTW* 2010, **33**(1):1-22.
19. Xu PP, Tian RH, Luo S, Zu ZY, Fan B, Wang XM, Xu K, Wang JT, Zhu J, Shi JC *et al:* **Risk factors for adverse clinical outcomes with COVID-19 in China: a multicenter, retrospective, observational study.** *THERANOSTICS* 2020, **10**(14):6372-6383.
20. Guohua W, Wuxuan P, Xingchen H, Rui W, Huangke C: **Configuring differential evolution adaptively via path search in a directed acyclic graph for data clustering.** *SWARM EVOL COMPUT* 2020, **55**.
21. Lee PH: **Is a cutoff of 10% appropriate for the change-in-estimate criterion of confounder identification?** *J EPIDEMIOL* 2014, **24**(2):161-167.
22. Kuhn M: **caret: Classification and Regression Training.** *R package version 6.0-86* 2020.
23. Amirhossein Jalali DRAA: **DynNom: Visualising Statistical Models using Dynamic Nomograms.** *R package version 5.0.1* 2019.
24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M: **pROC: an open-source package for R and S+ to analyze and compare ROC curves.** *BMC BIOINFORMATICS* 2011, **12**:77.
25. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, Lewis SA, Macfarlane JT: **Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study.** *THORAX* 2003, **58**(5):377-382.
26. Frank E HJ: **rms: Regression Modeling Strategies.** *R package version 6.0-1* 2020.
27. Vickers AJ, Elkin EB: **Decision curve analysis: a novel method for evaluating prediction models.** *MED DECIS MAKING* 2006, **26**(6):565-574.

28. Gao S, Jiang F, Jin W, Shi Y, Yang L, Xia Y, Jia L, Wang B, Lin H, Cai Y *et al*: **Risk factors influencing the prognosis of elderly patients infected with COVID-19: a clinical retrospective study in Wuhan, China.** *Aging (Albany NY)* 2020, **12**(13):12504-12516.
29. Ye C, Zhang S, Zhang X, Cai H, Gu J, Lian J, Lu Y, Jia H, Hu J, Jin C *et al*: **Impact of comorbidities on patients with COVID-19: A large retrospective study in Zhejiang, China.** *J MED VIROL* 2020.
30. Yang AP, Liu JP, Tao WQ, Li HM: **The diagnostic and predictive role of NLR, d-NLR and PLR in COVID-19 patients.** *INT IMMUNOPHARMACOL* 2020, **84**:106504.
31. Liu YP, Li GM, He J, Liu Y, Li M, Zhang R, Li YL, Wu YZ, Diao B: **Combined use of the neutrophil-to-lymphocyte ratio and CRP to predict 7-day disease severity in 84 hospitalized patients with COVID-19 pneumonia: a retrospective cohort study.** *Ann Transl Med* 2020, **8**(10):635.
32. McRae MP, Dapkins IP, Sharif I, Anderman J, Fenyo D, Sinokrot O, Kang SK, Christodoulides NJ, Vurmaz D, Simmons GW *et al*: **Managing COVID-19 With a Clinical Decision Support Tool in a Community Health Network: Algorithm Development and Validation.** *J MED INTERNET RES* 2020, **22**(8):e22033.
33. Zhang S, Guo M, Duan L, Wu F, Hu G, Wang Z, Huang Q, Liao T, Xu J, Ma Y *et al*: **Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study.** *CRIT CARE* 2020, **24**(1):438.
34. Lai C, Yu R, Wang M, Xian W, Zhao X, Tang Q, Chen R, Zhou X, Li X, Li Z *et al*: **Shorter incubation period is associated with severe disease progression in patientswith COVID-19.** *VIRULENCE* 2020, **11**(1):1443-1452.
35. Hermesh T, Moltedo B, López CB, Moran TM: **Buying time-the immune system determinants of the incubation period to respiratory viruses.** *Viruses* 2010, **2**(11):2541-2558.

Figures

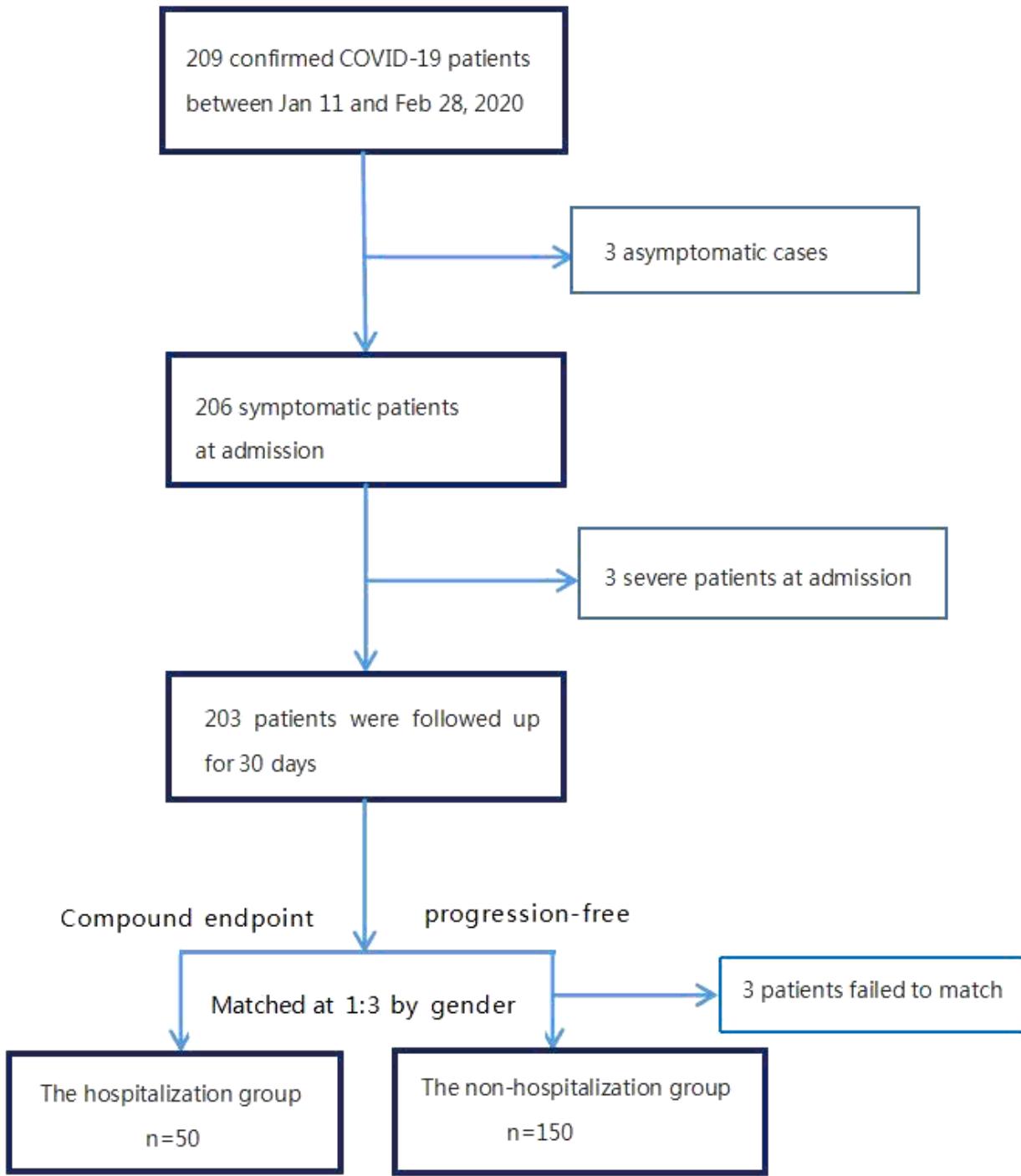
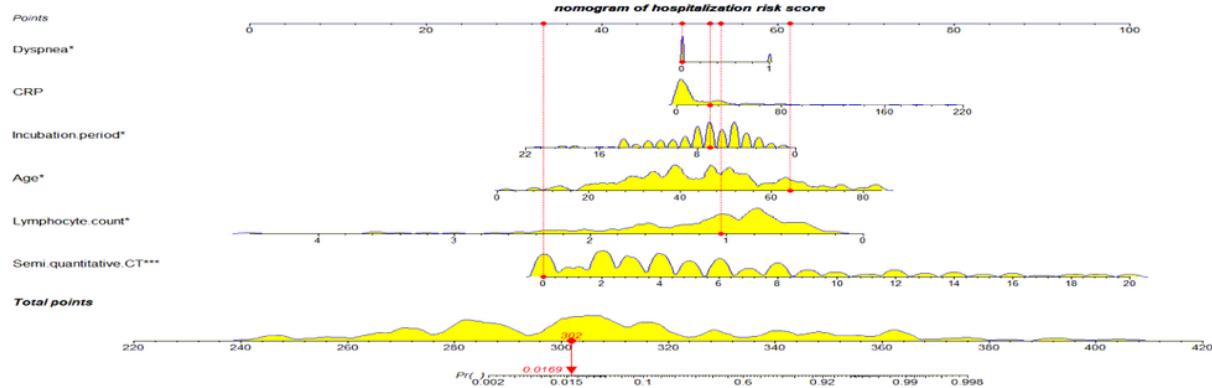


Figure 1

The flow chart of this study.

a**b**

Dynamic Nomogram of Hospitalization Risk Score

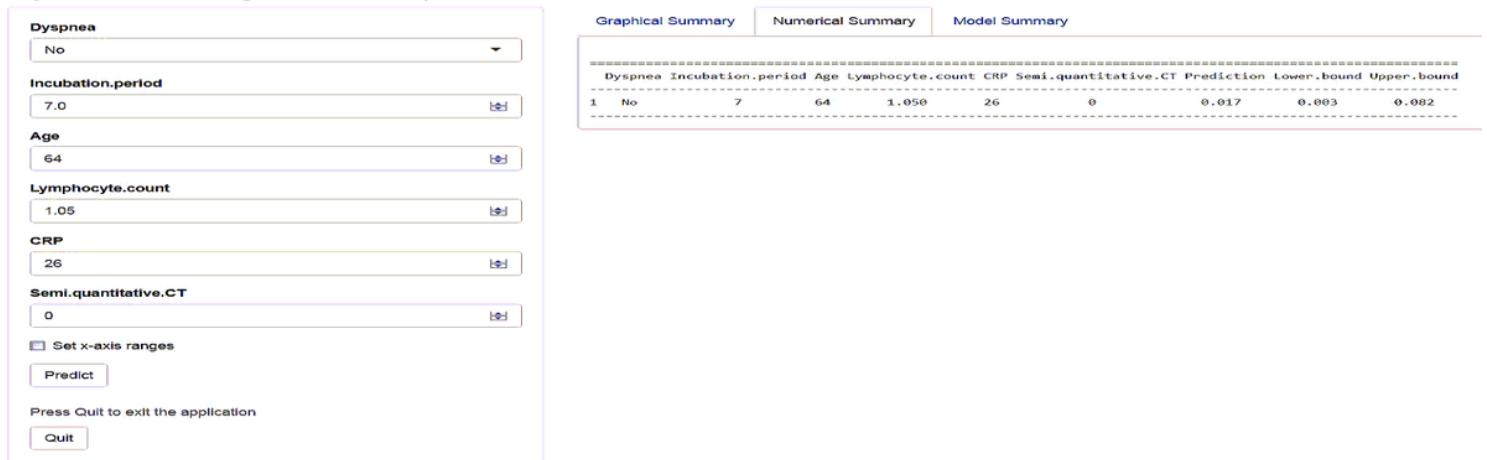


Figure 2

a. Nomogram for predicting hospitalization. Each variable is projected onto line 1 to obtain the score. The sum of the scores is marked in the penultimate line, and a vertical line to the bottom line represents a final prediction from the risk score count. The case shown here is 64 years old with no dyspnea, no CT lesions in the lungs, lymphocytic count $1.05 \times 10^9/L$, Incubation period 7 day and CRP 26mg/L, final score 302, predicted risk value 0.0169, low risk of hospitalization. b. The case calculated a close prediction value of 0.017 (95%CI 0.003-0.082) according to the variable through the network counter.

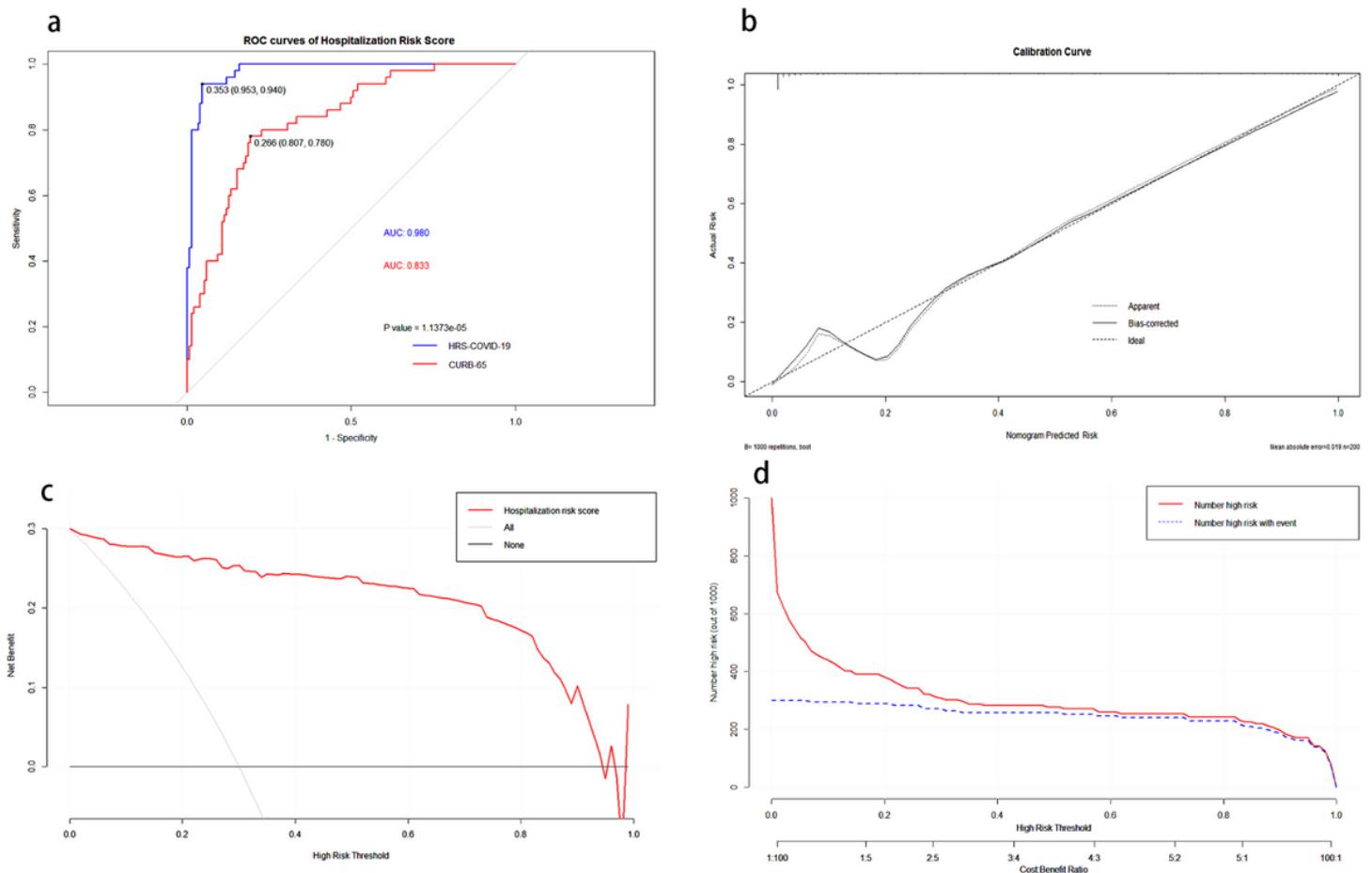


Figure 3

a. Receiver-operating characteristic curve of the nomogram for the hospitalization risk score of COVID-19(HRS-COVID-19), The discriminant ability of hospitalization risk score and CURB-65 scoring system was compared(0.980 vs 0.833). The two scoring systems were compared by delong method ($P < 0.001$). b. Calibration curves of the nomogram for HRS-COVID-19. If the risk score is above 30%, the probability of prediction, reality and ideal is the same; At 20%, the predicted rate is higher than the measured rate; at 10%, the predicted rate is lower than the measured rate. In the actual operation process, missed diagnosis and over treatment can be reduced. c. Decision curves of the nomogram for HRS-COVID-19; Red line: risk score. Light line: assume all patients are hospitalized. Horizontal thin line: assumed no patient is hospitalized. The chart shows the expected net benefit of each patient relative to the non-hospitalization of any patient. The farther away the curve is from the horizontal thin line, the greater the clinical benefit. d. Clinical impact curve of the nomogram for HRS-COVID-19. Of 1000 patients, the solid red line represents the total number of hospitalizations considered high-risk at each risk threshold. The blue dotted line shows how many people actually need to be hospitalized.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryOnlineContent.docx
- SupplementaryOnlineContenteTable1.docx