Scores based on neutrophil percentage and lactate dehydrogenase with or without oxygen saturation predict the risk of hospital mortality in severe COVID-19 patients

Xiude Fan  
First Affiliated Hospital of Xi’an Jiaotong University

Bin Zhu  
Capital Medical University

Masoud Nouri-Vaskeh  
Tabriz University of Medical Sciences

Chunguo Jiang  
Capital Medical University

Xiaokai Feng  
Capital Medical University

Kyle Poulsen  
Cleveland Clinic

Behzad Baradaran  
Tabriz University of Medical Sciences

Jiansong Fang  
Guangzhou University of Chinese Medicine

Erfan Ahmadi Ade  
Tabriz University of Medical Sciences

Akbar Sharifi  
Tabriz University of Medical Sciences

Zhigang Zhao  
Capital Medical University

Yong Zhang  
Tongji Medical College

Liming Zhang  
Capital Medical University

Zhengwen Liu (✉ liuzhengwen113@xjtu.edu.cn)  
First Affiliated Hospital of Xi’an Jiaotong University
Abstract

**Background.** Risk scores are urgently needed to assist clinicians in predicting the risk of death in severe patients with SARS-CoV-2 infection in the context of millions of people infected, rapid disease progression, and shortage of medical resources.

**Method.** A total of 139 severe patients with SARS-CoV-2 from China and Iran were included. Using data from China (training dataset, n = 96), prediction models were developed based on logistic regression models, nomogram and risk scoring system for simplification. Leave-one-out cross validation was used for internal validation and data from Iran (test dataset, n = 43) for external validation.

**Results.** The NSL model (Area under the curve (AUC) 0.932) and NL model (AUC 0.903) were developed based on neutrophil percentage (NE), lactate dehydrogenase (LDH) with or without oxygen saturation (SaO₂) using the training dataset. Compared with the training dataset, the predictability of NSL model (AUC 0.910) and NL model (AUC 0.871) were similar in the test dataset. The risk scoring systems corresponding to these two models were established for clinical application. The AUCs of the NSL and NL scores were 0.928 and 0.901 in the training dataset, respectively. At the optimal cut-off value of NSL score, the sensitivity was 94% and specificity was 82%. In addition, for NL score, the sensitivity and specificity were 94% and 75%, respectively.

**Conclusion.** NSL and NL score are straightforward means for clinicians to predict the risk of death in severe patients. NL score could be used in selected regions where patients’ SaO₂ cannot be tested.

Introduction

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which is also called coronavirus infectious disease 2019 (COVID-19), a highly contagious and fast-spreading infectious disease, is only four months old but is already spreading in many countries with millions of people being infected [1].

The clinical spectrum of COVID-19 ranges from mild to critically ill cases according to the largest cohort study (44,672 persons with COVID-19) from China [2]. This disease can progress rapidly into acute respiratory distress syndrome (ARDS), multiorgan failure, and even death during the later stages in some severe cases [2-6]. Clinicians should be aware that some serious patients may deteriorate rapidly after admission.

Since the outbreak of COVID-19, researchers and clinicians are acting quickly, but perhaps not fast enough to as compared to the rate of this disease. Previous studies had identified that lymphopenia, neutrophilia, elevated serum alanine aminotransferase (ALT), aspartate aminotransferase levels (AST), lactate dehydrogenase (LDH), D-dimer and C-reactive protein (CRP) may be associated with disease progression and death [3-5, 7, 8], however, there is no easy-to-use risk-scoring system for the risk of death in severe patients. Currently, clinicians urgently need a convenient clinical risk assessment tool to assist
them in predicting the risk of hospital mortality, in order to select the time and method of medical intervention and to evaluate the effectiveness of treatment strategies.

Therefore, in the current study we utilized data from severe patients with confirmed COVID-19 who were admitted to hospitals in China and Iran to establish straightforward and user-friendly prediction models for clinicians to predict the risk of in-hospital death in severe patients with COVID-19.

**Patients And Methods**

**Patient population**

This multicentric retrospective observational study was based on two datasets of severe patients with confirmed SARS-CoV-2 infection selected by the same criteria [9] from 2 medical centers (West Branch of Union Hospital affiliated to Tongji Medical College of Huazhong University of Science and Technology) in China and (Tabriz University of Medical Sciences) Iran. The patients’ data from China was used as the training dataset to establish models in predicting the risk of hospital mortality, whereas the patients’ data from Iran was used for external validation of the prediction models. As shown in the Figure 1. All severe patients with confirmed SARS-CoV-2 infection in training and test datasets were included if they were adults. Pregnant patients and patients with human immunodeficiency virus infection were excluded.

This study was approved by the Ethics Committees of all participating hospitals in China and Iran.

**Data Collection**

We reviewed clinical medical records, nursing records, laboratory examinations for all severe patients with laboratory-confirmed SARS-CoV-2 infection. The severity of disease was classified according to Chinese Clinical Guidance for COVID-19 Pneumonia Diagnosis and Treatment (7th edition) [9]. We collected admission data of these patients on age, sex, symptoms (fever, cough, sputum, fatigue, shortness of breath, headache, and diarrhea), medical histories (chronic cardiovascular disease, chronic pulmonary disease, cerebrovascular disease, diabetes, malignancy, chronic liver and kidney disease and smoking history), signs (heart rate, respiratory rate, and oxygen saturation (SaO\textsubscript{2})), laboratory indexes (white blood cells (WBC), neutrophil percentage (NE), lymphocyte percentage (LY), hemoglobin (HGB), hematocrit (HCT), Platelets (PLT), LDH, total bilirubin (Tbil), direct bilirubin (Dbil), ALT, AST, total protein, albumin (ALB), activated partial thromboplastin time (APTT), prothrombin time (PT), D-dimer, CRP, blood urea nitrogen (BUN), serum creatinine (Cr), creatinine clearance (CCr), blood glucose, creatine kinase isoenzymes (CKMB), high density lipoprotein (HDL), low density lipoprotein (LDL), total cholesterol (TC), triglyceride (TG), Lipoprotein, Apolipoprotein A (ApoA), Apolipoprotein B (ApoB), serum potassium (K), and serum sodium (Na)). HDL, LDL, TC, TG, Lipoprotein, ApoA, ApoB, HGB, and HCT were not collected in the Iranian population. Information about treatment during hospitalization (antiviral therapy, antibacterial therapy, corticosteroids, and immunoglobulin therapy) and outcome (in-hospital death) were also collected.
Statistical analysis

Continuous variables are reported as means±standard error (SE). Unpaired t-test or the Mann-Whitney test was used to compare two groups of data. Categorical variables are expressed as counts and percentages; Chi-square or Fisher’s exact tests were used for comparisons of categorical factors. Feature selection was performed to select the suitable variables to establish the prognostic model using the information gain method. Information gain was calculated by comparing the entropy of the data before and after transformation [10]. Factors with attributes of variables >0.2 were selected for modeling. The establishment of death risk models were based on multivariable logistic regression models using training dataset. The predictive accuracy for the prognostic accuracy of hospital mortality of severe patients was calculated using receiver operating characteristic (ROC) curves. When the sensitivity, specificity and area under the curve (AUC) were basically similar between different models, we selected models for further analysis based on the premise of minimizing the number of factors included in the model. Validity assessment of the predictive models was conducted using internal and external validation. We used leave-one-out cross-validation method for internal validation to limit model over-fitting and to assess predictive potential [11]. In external validation, models developed in the training dataset were applied on the test dataset to assess the predictive performance of models. We used calibration plots to show the goodness-of-fit of models and plotted nomograms to facilitate the clinical application of both models. In order to simplify the computation of in-hospital death risk estimate, we develop risk scores based on the points system from the Framingham Heart Study methodology [12]. All statistical analyses were performed using STATA (Version 13.0, IBM, New York, USA) and Orange (Version 3.24.1, USA).

Result

Characteristics of the study population

There were 96 patients from China in the training dataset and 43 patients from Iran in the test dataset. The mean age of patients in the training and test datasets were 63.47 and 63.37 years, respectively. The patients in the two datasets differ in several characteristics at the time of admission (Table 1). In total, 49 (51%) male patients in the training and 30 (69.8%) male patients in the test dataset (P=0.039). There were more patients with fever (89.6% versus 46.5%), fatigue (89.6% versus 42.2%) and diarrhea (20.8% versus 2.3%) in the training dataset compared to those in test dataset. in addition, patients in the training dataset had faster respiratory rates (27.24 versus 22.76) than those in the test dataset. The proportion of deaths in the two data sets (32.3% versus 30.2%) was roughly the same.

Feature selection

Figure 2 shows the results from information gain ranking, the top 8 of the available 60 variables (LDH, NE, SaO2, LY, NLR, CKMB, D-dimer, and CRP) were selected for modeling according to the criteria (information gain > 0.2). As shown in Supplementary Figure 1A, LDH, NE, SaO2, NLR, CKMB, D-dimer, and CRP were significantly higher and LY was lower in the severe patients who died during hospitalization compared to patients who did not die.
Derivation and validation of NSL model and NL model

When used individually to predict the risk of death, AUCs of top 8 ranked variables range from 0.763 to 0.880, sensitivities range from 73% to 100%, and specificities range from 51% to 88% (Table 2). Each of these indicators had a good prediction ability for the risk of death, but there were some exceptions, such as some patients with normal indicators who also died during hospitalization, so integrated prediction models were needed to reduce the defects of a single indicator in predicting death risk.

In the modeling, we tried to use as few variables as possible to facilitate clinical application. Because the NE and LY had a reciprocal relationship and integrated models were based on the logistic regression method, we established three model groups depending on whether the NE, LY, or neutrophils/lymphocytes ratio (NLR) was added to the model. AUCs of all integrated models range from 0.903 to 0.948, sensitivities range from 77% to 97%, and specificities range from 77% to 97% (Table 2). Integrated model, combining all top 8 variables (AUC 0.945; sensitivity 97% and specificity 83%), the NSL model, combining top 3 variables (AUC 0.932; sensitivity 97% and specificity 78%; Supplementary Figure 1B), and NL model combining NE and LDH (AUC 0.903; sensitivity 94% and specificity 82%; Supplementary Figure 1B) had High sensitivity and specificity in predicting the risk of death. Considering the need for convenient clinical application and the regions with backward medical care level, we selected the NSL model and NL model for validation in the test dataset. NL model could be used in regions where patients’ SaO$_2$ cannot be tested.

Compared with the training dataset, NSL model (AUC 0.910; sensitivity 92% and specificity 96%) and NL model (AUC 0.871; sensitivity 92% and specificity 82%) provided similarly accurate predictability of in-hospital death in the test dataset (Table 2 and Supplementary Figure 1C).

Nomogram prediction for in-hospital death of severe patients

In order for clinicians to easily calculate the risk of mortality using the NSL model or NL model, we created two nomograms to provide graphical depictions of all indicators in the NSL model and NL model, respectively (Figure 3A,B). In both the training and test datasets, the calibration plots of nomograms were consistent between the predicted risk and the observed probability of death (Figure 3C-F). The Hosmer–Lemeshow tests for NSL model and NL model were not significant (P=0.47 and P=0.45), suggesting the NSL model and NL model were correctly specified for the prediction of in-hospital death from COVID-19.

Development of risk scoring system for predicting in-hospital death

In addition to providing a nomogram to help clinicians predict the mortality risk of severe patients, we also developed two risk scoring systems based on NSL model and NL model. As shown in Table 3, simple point systems were developed based on the logistic regression coefficients (Supplementary Table 1) and reference values for each significant risk factor (Table 3). The NSL risk score included NE (16 points), SaO$_2$ (9 points), and LDH (9 points). The total points ranged from 0 to 34. With an increasing total points, the risk of death increased. Points of 0–13 were associated with a less than 10% risk of death and points
of 14-20 with a 10–50% risk of death. Finally, points above 20 were associated with an extremely high risk of death over 50%. The cut-off of the NSL risk score for the prediction of death in training dataset is 15 (sensitivity 94% and specificity 82%, Supplementary Table 2). The AUCs of the NSL risk score were 0.928 and 0.901 in the training and test dataset, respectively. In addition, the NL risk score included NE (16 points) and LDH (9 points). The score ranged from 0 to 25. The AUCs of the NL risk score were 0.895 and 0.857 in the training and test dataset, respectively. Points of 0–9 were associated with a less than 10% risk of death, points of 10-15 with a 10–50% risk of death, and points above 16 were associated with an extremely high risk of death over 50%. The cut-off of the NL risk score for the prediction of death in training dataset is 12 (sensitivity 94% and specificity 75%, Supplementary Table 2). In clinical practice, clinicians can calculate the risk scores of each patient at admission based on the points provided in Table 3 and Table 4.

Discussion

To our knowledge, this is the first study to develop in-hospital death risk scoring systems in severe patients with COVID-19 from China and Iran. The NSL score and NL score described in this study are easy to understand and to use. These two risk scores make it easy for clinicians to predict the risk of death in severe patients and avoid the influence of personal bias in the course of evaluation. In some regions where medical resources are scarce, the NL score enables medical staffs to predict the risk of death of severe patients with only NE and LDH at the time of admission, which will greatly improve the efficiency of medical resource allocation. The NSL score and NL score developed in a dataset of Chinese patients and was validated in another dataset of Iranian patients. There were several differences in the clinical characteristics of the severe patients in the training and test datasets, but this enhances the reliability of our risk scores, which provides similar predictability across different patient populations.

Lymphopenia, neutrophilia, LDH, D-dimer and CRP may be related to the progression of the disease according to previous studies [3-5, 7, 8]. Among these factors, elevated D-dimer and lymphopenia have been reported to be associated with death [3, 4, 7]. An SaO₂ rate below 93% (normal range is 95% to 100%) has long been considered a sign of underlying hypoxia and impending organ failure [13, 14]. For COVID 19, SaO₂ is also a good indicator for the disease progression [15], which was also confirmed in our study. Previous study found that higher SOFA score, older age, and D-dimer greater than 1 μg/mL at admission were associated with increased risk of death, which could help medical staffs to assess the prognosis of patients [3]. In addition, Ji et al. established a risk score (CALL) based on patients’ age, lymphocyte count, serum LDH levels and comorbidities at admission, which could help medical staffs to identify patients with a high risk of disease progression [5]. Outside of the CALL risk score to quantitatively predict risk of disease progression, clinicians lack a relevant scoring system to quantitatively predict the risk of death in severe patients. This may lead to an underestimation of the risk of death in some severe patients, resulting in delays in treatment and unnecessary mortality.

Here, in the establishment of the predictive models, we utilized the feature selection method of machine learning and also considered the needs of clinicians. We established two risk scores (NSL score and NL
score) only based on NE, SaO2 with and without LDH concentration at admission. An NSL score ≤ 11 is associated with a risk of death is less than 5%, whereas NSL score > 15 and particularly > 20 indicated an increased risk of death, requiring urgent support symptomatic treatment, and careful surveillance for these patients. In particular, the cut-off point of 20 in NSL score offered 71% sensitivity and 94% specificity for death risk prediction in training datasets and 92% sensitivity and 82% specificity in the test dataset. For some regions without appropriate access to tests for LDH concentrations in patients, the NL score can also be used to predict the risk of death with high risk prediction accuracy. NL score ≤ 8 is associated with a risk of death is less than 5%, whereas NL score > 9 and NL score > 14 indicated the risk of death exceeded 10% and 40%, respectively.

Our study has a few limitations. Firstly, the machines and methods used in China and Iran to detect serum LDH concentrations are different, so the normal range of LDH concentrations is slightly different. In China, LABOSPECT 008 α Hitachi Automatic Analyzer (Hitachi High-Technologies Corporation, Japan) was used to detect serum LDH concentrations, while in Iran, LDH Cytotoxicity Detection Kit (Roche, Germany) was applied to detect serum LDH. Secondly, the sample size is relatively small, especially the test data from Iran. Finally, due to the limitations of data, we did not analyze the effects of different medical interventions on prognosis.

In summary, The NSL score and NL score are straightforward means for clinicians to predict the risk of death in severe patients and avoid the influence of human factors in the course of evaluation.

Declarations

This retrospective study was approved by the Ethics Committees of all participating hospitals in China and Iran. We use the De-identified medical record for establishing the prediction model.

Acknowledgements. We thank the all medical staffs and patients involved in the study.

Financial support. None.

Potential conflict of Interests. All authors declared there were no conflicts of interest.

References

1. World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report – 92.. Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200421-sitrep-92-covid19.pdf. Accessed 21 April 2020

2. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. JAMA 2020.
3. Zhou F, Yu T, Du R, 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): 1054-62.

4. Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. JAMA Intern Med 2020.

5. 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.

6. Grein J, Ohmagari N, Shin D, et al. Compassionate Use of Remdesivir for Patients with Severe Covid-19. N Engl J Med 2020.

7. Yu B, Li X, Chen J, et al. Evaluation of variation in D-dimer levels among COVID-19 and bacterial pneumonia: a retrospective analysis. 2020.

8. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020.

9. Commission CNH. Chinese Clinical Guidance for COVID-19 Pneumonia Diagnosis and Treatment (7th edition). Available at: http://kjfy.meetingchina.org/msite/news/show/cn/3337.html. Accessed 4 March 2020.

10. Alhaj TA, Siraj MM, Zainal A, Elshoush HT, Elhaj F. Feature Selection Using Information Gain for Improved Structural-Based Alert Correlation. PLoS One 2016; 11(11): e0166017.

11. Leave-One-Out Cross-Validation. In: Sammut C, Webb GI. Encyclopedia of Machine Learning. Boston, MA: Springer US, 2010:600-1.

12. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004; 23(10): 1631-60.

13. Moussa MD, Durand A, Leroy G, et al. Central venous-to-arterial PCO2 difference, arteriovenous oxygen content and outcome after adult cardiac surgery with cardiopulmonary bypass: A prospective observational study. Eur J Anaesthesiol 2019; 36(4): 279-89.

14. Kazune S, Caica A, Luksevics E, Volceka K, Grabovskis A. Impact of increased mean arterial pressure on skin microcirculatory oxygenation in vasopressor-requiring septic patients: an interventional study. Ann Intensive Care 2019; 9(1): 97.

15. Colombi D, Bodini FC, Petrini M, et al. Well-aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia. Radiology 2020: 201433.

Tables

Table 1. Clinical characteristics of the severe patients with COVID-19
| Variables                               | Training dataset (China data) n=96 | Test dataset (Iran data) n=43 | P value |
|----------------------------------------|----------------------------------|-------------------------------|---------|
| Age(years), mean(SE)                   | 63.47(1.36)                      | 63.37(2.70)                   | 0.972   |
| Male, n(%)                             | 49(51.0)                         | 30(69.8)                      | 0.039   |
| Smoking history, n(%)                  | 1(1.0)                           | 2(4.7)                        | 0.064   |
| Symptoms on admission, n(%)            |                                  |                               |         |
| Fever                                  | 86(89.6)                         | 20(46.5)                      | <0.001  |
| Cough                                  | 78(81.3)                         | 23(53.5)                      | 0.222   |
| Fatigue                                | 86(89.6)                         | 27(42.2)                      | <0.001  |
| Shortness of breath                    | 70(72.9)                         | 24(55.8)                      | 0.983   |
| Headache                               | 17(17.7)                         | 4(9.3)                        | 0.453   |
| Diarrhea                               | 20(20.8)                         | 1(2.3)                        | 0.017   |
| Coexisting disorder, n(%)              |                                  |                               |         |
| Hypertension                           | 33(34.4)                         | 12(27.9)                      | 0.748   |
| Diabetes                               | 16(16.7)                         | 8(18.6)                       | 0.296   |
| Chronic obstructive pulmonary disease  | 3(3.1)                           | 2(4.7)                        | 0.444   |
| Cerebral infarction                    | 1(1.0)                           | 1(2.3)                        | 0.42    |
| Coronary heart disease                 | 10(10.4)                         | 0                              | 0.057   |
| Chronic kidney disease                 | 3(3.1)                           | 1(2.3)                        | 0.768   |
| Chronic liver disease                  | 0                                | 0                              | —       |
| Respiratory rate, mean(SE)             | 27.24(0.57)                      | 22.76(1.09)                   | <0.001  |
| Heart rate, mean(SE)                   | 92.89(1.74)                      | 90.34(1.88)                   | 0.323   |
| In-hospital deaths, n(%)               | 31(32.3)                         | 13(30.2)                      | 0.809   |

Table 2. Predictive capacity of the factors and integrated models for the risk of hospital mortality in severe patients with COVID-19
| Variable                  | AUC  | 95%CI         | SE  | Sensitivity(%) | Specificity(%) |
|--------------------------|------|---------------|-----|----------------|----------------|
| **Training data from China cohort** |      |               |     |                |                |
| LDH(U/L)                 | 0.880| 0.813-0.948   | 0.034| 97            | 71             |
| NE(%)                    | 0.879| 0.812-0.946   | 0.034| 84            | 82             |
| SpO₂(%)                  | 0.849| 0.758-0.940   | 0.046| 87            | 78             |
| LY(%)                    | 0.852| 0.776-0.929   | 0.039| 77            | 80             |
| NLR                      | 0.858| 0.783-0.933   | 0.038| 81            | 82             |
| CKMB(U/L)                | 0.829| 0.746-0.912   | 0.042| 87            | 69             |
| D-dimer(μg/mL)           | 0.763| 0.641-0.885   | 0.062| 73            | 88             |
| CRP(μg/mL)               | 0.807| 0.723-0.892   | 0.043| 100           | 51             |
| **Integrated models**    |      |               |     |                |                |
| **All variables with information gain > 0.2** |      |               |     |                |                |
| LDH+NE+SpO₂+LY+NLR+CKMB+D-dimer+CRP | 0.945| 0.897-0.992   | 0.024| 97            | 83             |
| **NE was selected for modeling** |      |               |     |                |                |
| LDH+NE+SpO₂+CKMB+D-dimer+CRP | 0.945| 0.900-0.989   | 0.023| 93            | 84             |
| LDH+NE+SpO₂+CKMB+D-dimer | 0.942| 0.898-0.987   | 0.023| 97            | 78             |
| LDH+NE+SpO₂+CKMB         | 0.937| 0.887-0.988   | 0.026| 83            | 94             |
| LDH+NE+SpO₂ (NSL risk score) | 0.932| 0.884-0.981   | 0.025| 97            | 78             |
| LDH+NE (NL risk score)   | 0.903| 0.843-0.963   | 0.031| 94            | 82             |
| **LY was selected for modeling** |      |               |     |                |                |
| LDH+SpO₂+LY+CKMB+D-dimer+CRP | 0.948| 0.904-0.992   | 0.022| 97            | 84             |
| Model                                      | AUC   | Lower 95% CI | Upper 95% CI | p-value | Hosmer-Lemeshow Goodness of Fit Test p-value |
|--------------------------------------------|-------|--------------|--------------|---------|---------------------------------------------|
| LDH+SpO₂+LY+CKMB+D-dimer                  | 0.944 | 0.901-0.987  | 0.022        | 86      | 88                                          |
| LDH+SpO₂+LY+CKMB                          | 0.932 | 0.880-0.984  | 0.026        | 97      | 77                                          |
| LDH+SpO₂+LY                               | 0.934 | 0.886-0.982  | 0.025        | 90      | 88                                          |
| LDH+LY                                     | 0.903 | 0.843-0.964  | 0.031        | 90      | 82                                          |

**NLR was selected for modeling**

| Model                                      | AUC   | Lower 95% CI | Upper 95% CI | p-value | Hosmer-Lemeshow Goodness of Fit Test p-value |
|--------------------------------------------|-------|--------------|--------------|---------|---------------------------------------------|
| LDH+SpO₂+NLR+CKMB+D-dimer+CRP              | 0.930 | 0.866-0.995  | 0.033        | 83      | 95                                          |
| LDH+SpO₂+NLR+CKMB+D-dimer                 | 0.945 | 0.901-0.989  | 0.022        | 79      | 95                                          |
| LDH+SpO₂+NLR+CKMB                         | 0.933 | 0.882-0.983  | 0.026        | 77      | 97                                          |
| LDH+SpO₂+NLR                              | 0.933 | 0.883-0.971  | 0.025        | 87      | 88                                          |
| LDH+NLR                                    | 0.919 | 0.866-0.971  | 0.027        | 90      | 82                                          |

**LDH wasn't selected for modeling**

| Model                                      | AUC   | Lower 95% CI | Upper 95% CI | p-value | Hosmer-Lemeshow Goodness of Fit Test p-value |
|--------------------------------------------|-------|--------------|--------------|---------|---------------------------------------------|
| NE+SpO₂                                    | 0.919 | 0.865-0.972  | 0.027        | 97      | 78                                          |

**Test data from Iran cohort**

**Combined models**

| Model                                      | AUC   | Lower 95% CI | Upper 95% CI | p-value | Hosmer-Lemeshow Goodness of Fit Test p-value |
|--------------------------------------------|-------|--------------|--------------|---------|---------------------------------------------|
| LDH+NE+SpO₂ (NSL risk score)               | 0.910 | 0.758-1.000  | 0.077        | 92      | 96                                          |
| LDH+NE (NL risk score)                     | 0.871 | 0.734-1.000  | 0.071        | 92      | 82                                          |

Abbreviation: LDH, lactate dehydrogenase; NE, neutrophil percentage; SpO₂, oxygen saturation; LY, lymphocyte percentage; NLR, neutrophils/lymphocytes ratio; CKMB, creatine kinase myocardial bound; CRP, C-reactive protein; AUC, area under the curve.

**Table 3. Algorithm to estimate risk for hospital mortality using total points for risk scores with logistic regression analysis in the severe patients with COVID-19 from training dataset.**
| Variables | Categories | Reference value | Bi | Regression units | Points assigned |
|-----------|------------|-----------------|----|------------------|----------------|
|           |            | ($W_{ij}$)      | $B_i$ | ($W_{ij} - W_{iREF}$) | $B_i (W_{ij} - W_{iREF})/B$ |
| **NSL risk score (NE+SaO2+LDH)** | | | | | |
| NE(%)     | ≤60        | 55 ($W_{iREF}$) | 0.000 | 0 |
|           | 60.1-70    | 65              | 1.270 | 4 |
|           | 70.1-80    | 75              | 2.540 | 8 |
|           | 80.1-90    | 85              | 3.810 | 12 |
|           | ≥90.1      | 95              | 5.080 | 16 |
| SaO$_2$(%)| 100-96     | 98 ($W_{iREF}$) | 0.000 | 0 |
|           | 95-91      | 93              | 0.875 | 3 |
|           | 90-86      | 88              | 1.750 | 6 |
|           | ≤86        | 83              | 2.625 | 9 |
| LDH(U/L)  | ≤221       | 171 ($W_{iREF}$) | 0.000 | 0 |
|           | 222-321    | 271             | 0.300 | 1 |
|           | 322-421    | 371             | 0.600 | 2 |
|           | 422-521    | 471             | 0.900 | 3 |
|           | 522-621    | 571             | 1.200 | 4 |
|           | 622-721    | 671             | 1.500 | 5 |
|           | 722-821    | 771             | 1.800 | 6 |
|           | 822-921    | 871             | 2.100 | 7 |
|           | 922-1021   | 971             | 2.400 | 8 |
|           | ≥1022      | 1071            | 2.700 | 9 |
| **NL risk score (NE+LDH)** | | | | | |
| NE(%)     | ≤60        | 55 ($W_{iREF}$) | 0.000 | 0 |
|       |       |       |       |       |
|-------|-------|-------|-------|-------|
| 60.1-70 | 65    | 1.580 | 4     |
| 70.1-80 | 75    | 3.160 | 8     |
| 80.1-90 | 85    | 4.740 | 12    |
| ≥90.1  | 95    | 6.320 | 16    |

| LDH(U/L) |       |
|----------|-------|
| 0.004    |       |

|       |       |       |
|-------|-------|-------|
| ≤221  | 171 (WiREF) | 0    |
| 222-321 | 271   | 0.400 |
| 322-421 | 371   | 0.800 |
| 422-521 | 471   | 1.200 |
| 522-621 | 571   | 1.600 |
| 622-721 | 671   | 2.000 |
| 722-821 | 771   | 2.400 |
| 822-921 | 871   | 2.800 |
| 922-1021 | 971  | 3.200 |
| ≥1022  | 1071  | 3.600 |

Wij, reference value for each category of risk factors in risk score; WiREF, the base category for each risk factor was used as the basic value for that factor and assigned 0 point. Bi, the regression coefficient of each risk factor from logistic regression; B, the smallest regression units or the smallest units divided by some constant (B=0.3 for NSL risk score and B=0.4 for NL risk score).

**Table 4 The risk of in-hospital death corresponding to the sum of points obtained from integrated models**
| NSL risk score (NE+SaO₂+LDH) | NL risk score (NE+ LDH) |
|-------------------------------|-------------------------|
| **Point of total** | **Estimate of risk of hospital mortality (%)** | **Point of total** | **Estimate of risk of hospital mortality (%)** |
| 0 | 0.18 | 0 | 0.22 |
| 1 | 0.25 | 1 | 0.33 |
| 2 | 0.33 | 2 | 0.49 |
| 3 | 0.45 | 3 | 0.74 |
| 4 | 0.61 | 4 | 1.09 |
| 5 | 0.82 | 5 | 1.62 |
| 6 | 1.10 | 6 | 2.40 |
| 7 | 1.48 | 7 | 3.54 |
| 8 | 1.99 | 8 | 5.20 |
| 9 | 2.67 | 9 | 7.56 |
| 10 | 3.57 | 10 | 10.87 |
| 11 | 4.76 | 11 | 15.39 |
| 12 | 6.32 | 12 | 21.35 |
| 13 | 8.34 | 13 | 28.82 |
| 14 | 10.94 | 14 | 37.66 |
| 15 | 14.22 | 15 | 47.40 |
| 16 | 18.29 | 16 | 57.35 |
| 17 | 23.20 | 17 | 66.73 |
| 18 | 28.97 | 18 | 74.95 |
| 19 | 35.50 | 19 | 81.70 |
| 20 | 42.63 | 20 | 86.94 |
| 21 | 50.07 | 21 | 90.85 |
| 22 | 57.52 | 22 | 93.68 |
| 23 | 64.63 | 23 | 95.67 |
| 24 | 71.16 | 24 | 97.06 |
| 25 | 76.91 | 25 | 98.01 |
|    |    |
|----|----|
| 26 | 81.80 |
| 27 | 85.85 |
| 28 | 89.12 |
| 29 | 91.71 |
| 30 | 93.72 |
| 31 | 95.27 |
| 32 | 96.45 |
| 33 | 97.35 |
| 34 | 98.02 |

Figures
Figure 1

Flow chart of the study population selection and analysis method
Figure 2

Feature selection to find variables with respect to the hospital mortality of severe patients. Abbreviation: SaO2, oxygen saturation; WBC, white blood cells; NE, neutrophil percentage; LY, lymphocyte percentage; NLR, neutrophils/lymphocytes ratio; HGB, hemoglobin; HCT, hematocrit; PLT, platelets; LDH, lactate dehydrogenase; Tbil, total bilirubin; Dbil, direct bilirubin; ALT, alanine aminotransferase; AST, aspartate amino transferase; ALB, albumin; APTT, activated partial thromboplastin time; PT, prothrombin time; CRP,
C-reactive protein; BUN, blood urea nitrogen; Cr, serum creatinine; CCr, creatinine clearance; CKMB, creatine kinase isoenzymes; HDL, high density lipoprotein; LDL, low density lipoprotein; TC, total cholesterol; TG, triglyceride; ApoA, Apolipoprotein A; ApoB, Apolipoprotein B; K, serum potassium; Na, serum sodium.

Figure 3

Nomograms for integrated models to predict hospital mortality and d the corresponding calibration plots Nomgrams of the NSL model (A) and NL model (B) to estimate the risk of death in severe patients with...
COVID-19, respectively. Calibration plots for NSL model in training (C) and test dataset (D), respectively. Calibration plots for NL model in training (E) and test dataset (F), respectively.

**Supplementary Files**

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

- Supplementalmaterial.docx