Clin. Lab. 2024;70:XXX-XXX

ORIGINAL ARTICLE

A Prediction Model for Clinical Outcomes of COVID-19 Hospitalized Patients: Construction and Accuracy Assessment

Han Chen [#], Meng-jia Chen [#], Li-qun Ling [#], Jian-rong Yang, Hai-xia Huang, Jia-jing Zhou, Ning Yang, Mei-juan Zhang

[#] *These authors contributed equally and should be considered co-first authors*
Department of Laboratory Medicine, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

SUMMARY

Background: **The coronavirus disease 2019 (COVID-19) pandemic spread rapidly with considerable morbidity nationwide since China's liberalization in December 2022. Our work has focused on identifying different predictive factors from the laboratory examination of critically ill patients, and forecasting the unfavorable outcome of critically ill patients with COVID-19 through a combined diagnosis of biological markers.**
Methods: **We conducted a retrospective study at the Department of First Affiliated Hospital of Wenzhou Medical University, China, from December 24, 2022, to January 10, 2023, where 434 critically ill patients who met the inclusion criteria were involved. Machine analysis was employed to search for the parameters with the highest predictive value to calculate COVID-19 mortality by exploiting 66 typical laboratory results.**
Results: **Combined diagnosis of serum albumin (ALB), lactate dehydrogenase (LDH), direct bilirubin (Dbil), ferritin, pulse oxygen saturation (SpO_2), and neutrophil count (NEUT#) was evaluated, and the result with the highest predictive value (NEUT#) was selected as the predictor for COVID-19 mortality with a sensitivity of 89.2% and a specificity of 77.4%.**
Conclusions: **The increased levels of LDH, Dbil, ferritin, and NEUT#, along with lowered ALB and SpO_2 levels are the most decisive variables for forecasting the mortality for COVID-19 according to our machine-learning-based model. The combined diagnosis could be used to improve further diagnostic performance.**
(Clin. Lab. 2024;70:xx-xx. DOI: 10.7754/Clin.Lab.2023.230721)

Correspondence:
Mei-juan Zhang
Department of Clinical Laboratory
Key Laboratory of Clinical Laboratory Diagnosis and Translational Research of Zhejiang Province
The First Affiliated Hospital of Wenzhou Medical University
Wenzhou
Zhejiang
China
Email: 18767482298@126.com

Manuscript accepted September 19, 2023

KEYWORDS

COVID-19, machine analysis, mortality, combined diagnosis

INTRODUCTION

The COVID-19 epidemic, well-known to all, continues to be raged as the number of infections increasing exponentially since December 2019 [1]. As of February 27, 2023, there were more than 679 million confirmed COVID-19 cases worldwide, with over 6.7 million deaths reported. A retrospective study showed a striking 28-day mortality up to 27% among 379 critically ill adult patients, especially for elder patients with underlying diseases [2].

The clinical manifestations of COVID-19 range from respiratory symptoms to pneumonia, and in more severe cases, multiple organ failure and eventually death. The intensive care unit (ICU) has been overwhelmed since China's liberalization. In Iran, the first country affected by the outbreak, the ICU admission rate is about 32% of all hospitalizations, while the ICU mortality rate is approximately 39% [3]. High morbidity and mortality led to overburdened healthcare systems and massive economic losses, consequently, recognition of early biomarkers which can be associated with diagnosis, prognosis, and outcomes are of paramount importance to prevent COVID-19 from developing into severe illness, even death [4].

Most biomarkers such as C-reactive protein (CRP), interleukin (IL)-6, procalcitonin (PCT), white blood cell (WBC) count, D-dimer, prothrombin time (PT), lactate dehydrogenase (LDH), and transaminases have been investigated in COVID-19 patients [3,4]. Machine analysis can help clinical medical staff to determine the level of care for critically ill patients and avoid the increased mortality from the lack of accurate judgment. A meta-analysis showed that machine learning models can be implemented in hospital settings and predict the incidence or prevalence of COVID-19 effectively [5]. In addition, Guan X et al. also observed that the simple-tree XGBoost model can help predict death risk of hospitalized COVID-19 patients accurately with high precision and sensitivity [2].

Therefore, in this work, using machine analysis, we intend to search for the early laboratory biomarkers that could be clinically essential for predicting the prognosis and outcome in severe COVID-19 cases.

MATERIALS AND METHODS

Data resources

This study involving patients was carried out in the Department of Laboratory Medicine at the First Affiliated Hospital of Wenzhou Medical University, China. We retrospectively studied 434 patients who had stayed in ICU with severe COVID-19 infection from December 24, 2022, through January 10, 2023.

The inclusion criteria were: positive PCR result for COVID-19 (confirmed by RT-PCR test) and meeting the severe illness standard according to the tenth version of diagnosis and treatment protocol for COVID-19 issued by the National Health Committee of the People's Republic of China. All the test results were completed within 24 hours after admission.

Data collection

Seventy-nine parameters in total were evaluated for each patient, including 13 clinical characteristics and 66 laboratory biomarkers as listed in Supplemental Table S1 and S2. Demographic and baseline characteristics during hospitalization were collected, including age, gender, length of stay (in days), clinical assessment at admission, comorbidities, clinical symptoms, infection situation, treatments, and laboratory findings. Patients were divided into two groups according to the clinical outcome: 120 patients (27.6%) and 314 patients (72.4%) were assigned to the non-recovery group and recovery group, respectively.

Statistical analysis

First, measurement data was represented by mean with associated standard deviation. Second, the chi-squared (X^2) test was used to analyze categorical variables, while the independent sample t-test was applied for the normally distributed numerical values between two groups. For non-normally distributed continuous variables, the Mann-Whitney U-test was employed. Furthermore, all variables with a statistical significance level $p < 0.05$ in the analysis were included in the initial machine analysis (Random Forests). The receiver operating curve (ROC) was applied to evaluate the diagnostic price value. The logistic regression model was also used to calculate the odds ratio and 95% confidence intervals for each factor predicting the risk of recovery in COVID-19 patients of ICU admission. Differences were considered statistically significant at $p < 0.05$. Statistical analysis was performed by IBM SPSS (Version 22.0) and R 4.0.5 statistics software.

RESULTS

Descriptive study

The average age of the two groups was 76.140 ± 10.445 years old and 70.120 ± 14.190 years old for the non-recovery group and the recovery group, respectively, which was statistically significant ($p < 0.001$). The M/F gender ratio was 2.43 for non-recovery group and 1.75 for recovery group thus not statistically significant ($p = 0.161$). The mean number of hospitalization days was similar for both groups, with the non-recovery group averaging 11.660 ± 8.750 days and recovery group averaging 11.830 ± 9.115 days, which was not statistically significant.

Besides basic information such as clinical assessment at admission (Padua score system, Caprini score system and ADL score system), comorbidities and clinical symptoms were also recorded in detail.

The two major basic diseases we considered in this study were hypertension and diabetes. In non-recovery group, these conditions were present in 59.2% and 33.3% of participants, respectively, while in recovery group, they were 56.4% and 34.7%, which was not statistically significant ($p > 0.05$).

Cerebral diseases such as cerebral infarction or hemorrhage history, Parkinson's disease, schizophrenia and so on, accounted for 29.1% in the non-recovery group and 13.1% in the recovery group, respectively. This correlation was statistically significant ($p < 0.001$).

As for cardiovascular diseases (coronary heart disease, atrial fibrillation, myocardial infarction, heart failure,

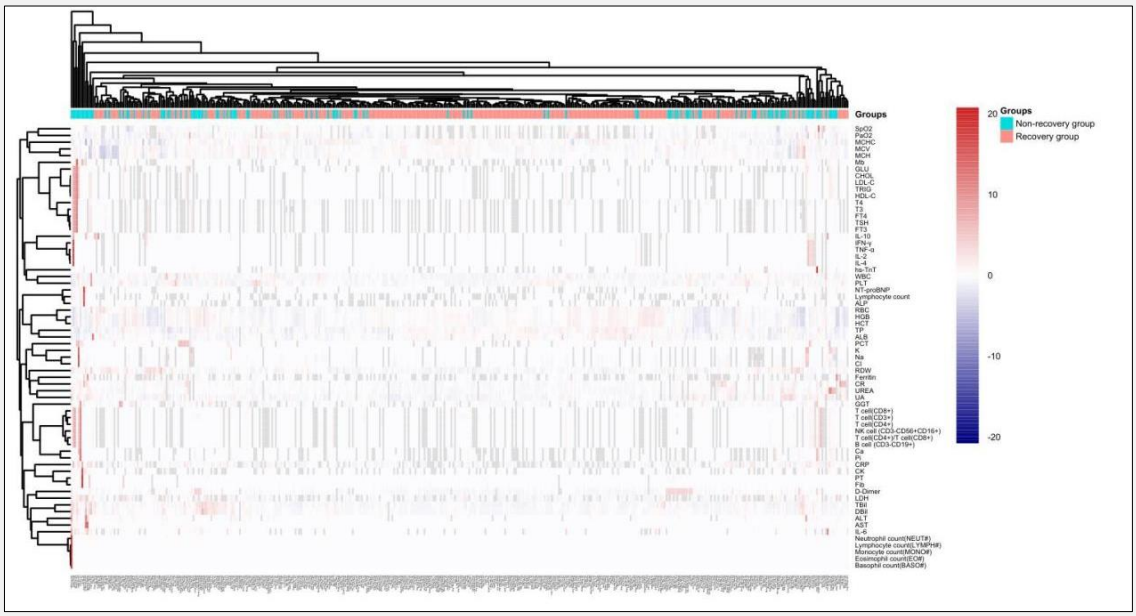

Figure 1. Heatmaps showing laboratory indexes in the non-recovery group and the recovery group.

The distribution of 66 laboratory indexes in 434 patients was represented by heat map. Patients were divided into two groups based on clinical outcomes (non-recovery group/recovery group).

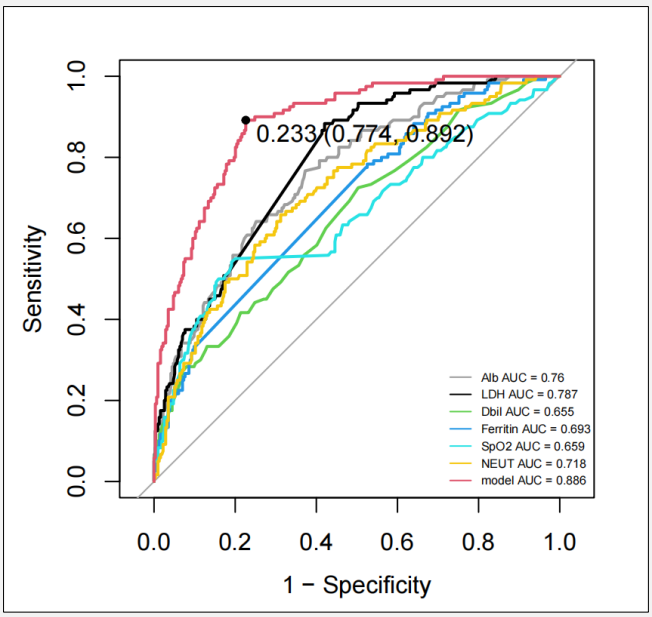

Figure 2. ROC analysis of the association between ALB, LDH, Dbil, ferritin, SpO2, NEUT# and combined diagnosis for forecasting mortality for COVID-19.

The ROC curve analysis demonstrated an AUC of 0.886 when the six biomarkers were used together, and the sensitivity and specificity were 89.2% and 77.4%, respectively. The cutoff point was 0.233.

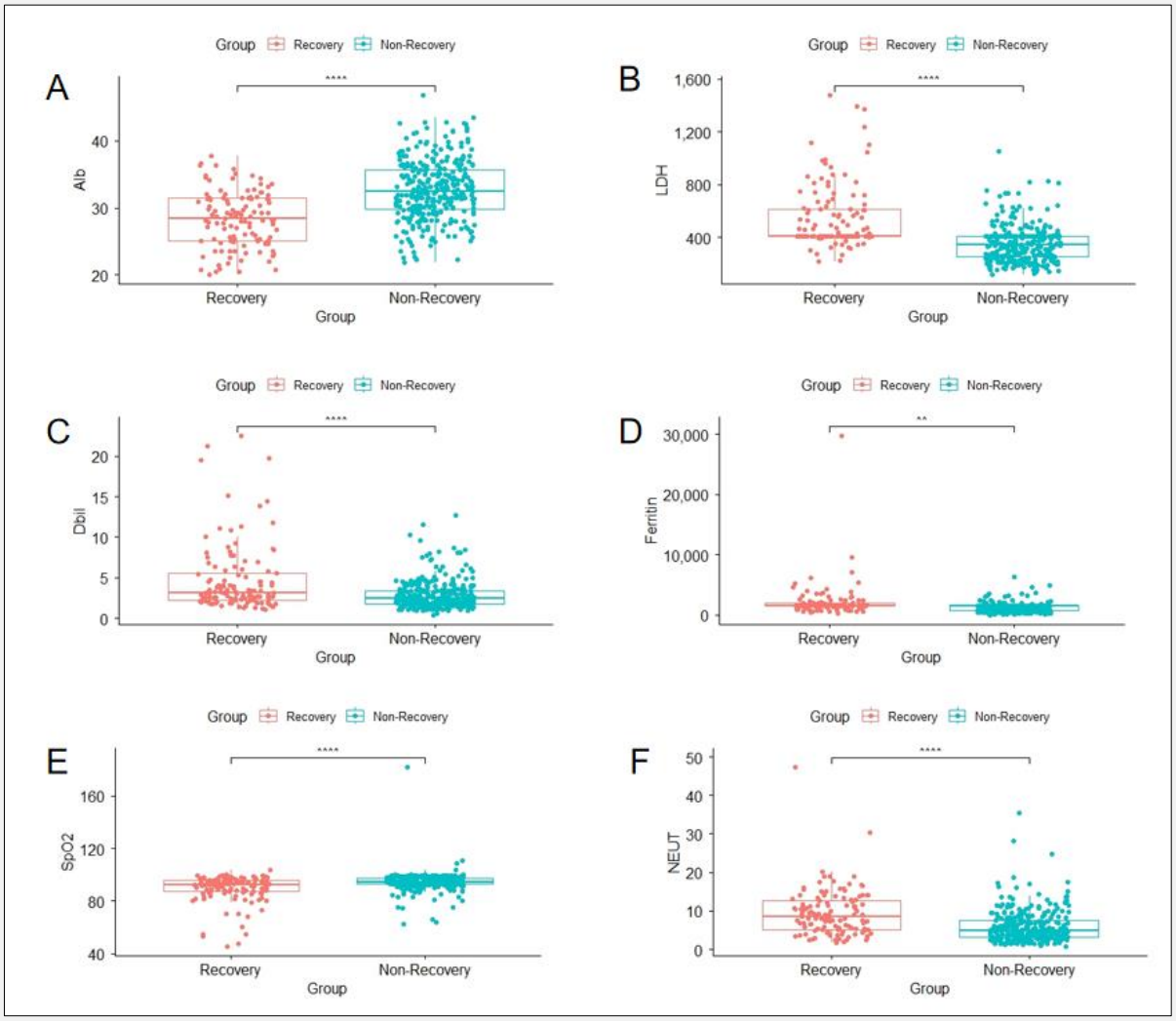

Figure 3. The comparison of the ALB, LDH, Dbil, ferritin, SpO_2 and NEUT# between the non-recovery group and the recovery group associated with critically ill patients with COVID-19.

A: The comparison of ALB between the non-recovery and the recovery group (28.350 ± 4.154 vs. 32.730 ± 4.431, p < 0.001), B: The comparison of LDH between the non-recovery and the recovery group (544.060 ± 245.287 vs. 351.960 ± 133.560, p < 0.001), C: The comparison of Dbil between the non-recovery and the recovery group (4.720 ± 4.177 vs. 2.876 ± 1.759, p < 0.001), D: The comparison of ferritin between the non-recovery and the recovery group (2,070.752 ± 2,873.237 vs. 1,185.21 ± 729.936, p = 0.001), E: The comparison of SpO_2 between the non-recovery and the recovery group (89.360 ± 10.707 vs. 94.513 ± 7.123, p < 0.001), F: The comparison of NEUT# between the non-recovery and the recovery group (9.511 ± 6.058 vs. 5.908 ± 4.183, p < 0.001).

etc.), they were represented in 16.7% of the non-recovery group and 19.4% of the recovery group, which was not statistically significant (p = 0.509). In addition, chronic kidney diseases, malignancy, pulmonary diseases, and obesity rates were also included in the study. Patients complained about a large number of symptoms of which the most observed ones were fever, cough, and dyspnea. They were found in 68.3%, 64.2%, 70.0% of the non-recovery group and 48.4%, 76.1%, 48.7% of the recovery group, respectively.

Furthermore, the infection during hospitalization was also measured. The main types of infection present, i.e. the respiratory tract infection, the bloodstream infection and the urinary infection, were 25.0%, 10.8% and 1.7% in the non-recovery group, compared to 10.5%, 1.9% and 2.2% in the recovery group. Besides, we also found that the T.SPOT-TB test in the non-recovery group was 1.7% compared to 5.1% in the recovery group. These results are shown in Supplemental Table S1.

Treatments and outcome

Treatments of COVID-19 mainly involved medication and oxygen support during hospitalization. The medication mainly included antiviral drugs, hormones, anticoagulant drugs, antibiotics, and adjuvant therapies. Oxygen support was mainly based on nasal oxygen therapy, mask oxygen therapy, and mechanical ventilation. An improvement of outcome was seen in 72.4% of cases as shown in Supplemental Table S1.

Laboratory examination

Machine analysis was used to identify the most valuable biomarkers from 66 parameters. More than 60.0% of the laboratory results were considered to be statistically significant between non-recovery group and recovery group with $p < 0.05$. The results of all laboratory examinations are shown in Supplemental Table S2.

We performed a multivariate analysis to identify different laboratory indexes associated with the risk of recovery in COVID-19 patients by the logistic regression model. Advanced levels of LDH and ferritin were found to be risk factors with p = 0.003/0.004, OR = 1.004 (1.001 - 1.007)/1.001 (1.000 - 1.001). Respiratory unstable patients with low SpO_2 levels were at higher risk with p = 0.002, OR = 0.914 (0.863 - 0.967). Meanwhile, high mortality was appreciably associated with lowered ALB with p = 0.007, OR = 0.823 (0.714 - 0.948). RBC, HGB, and MCHC were added to these indexes with p = 0.036/0.036/0.020, respectively, and OR = 0.006 (0.001 - 0.725), 1.177 (1.011-1.369), 0.916 (0.85 - 0.986), respectively, all indicating a high mortality rate. The results are shown in Supplemental Table S2.

Machine analysis revealed that six biomarkers, i.e. serum albumin (ALB), lactate dehydrogenase (LDH), direct bilirubin (Dbil), ferritin, pulse oxygen saturation (SpO_2), and neutrophil count (NEUT#), were particularly noteworthy. A ROC curve analysis was executed to provide statistical evidence for the value of the six biomarkers in forecasting the mortality of COVID-19. LDH had the strongest predictive value for mortality among all analyzed parameters (AUC = 0.787). However, ROC curve analysis found an Area Under the Curve (AUC) of 0.886 when the six biomarkers were combined. The sensitivity and specificity were up to 89.2% and 77.4%, respectively, which were higher than that of the single parameter. Figure 1 and Figure 2 list the heatmaps for all 66 laboratory parameters and the AUC of ROC for each biomarker.

We also found that the mean of LDH in the non-recovery group was 618.897 U/L vs. 335.583 U/L for the recovery group and exhibited a statistically significant correlation ($p < 0.001$). In addition, the level of ALB was 28.324 g/L and 32.773 g/L for the non-recovery group and recovery group, respectively, with a p-value < 0.001. Meanwhile, the mean of Dbil level was 4.720 μmol/L in the non-recovery group vs. 2.873 μmol/L in the recovery group, which was statistically significant ($p < 0.001$). In recovery group, the mean for ferritin levels was 1,007.636 ng/mL, compared to 2,577.564 ng/mL for the non-recovery group, which was also statistically significant with a p-value = 0.001. As for NEUT# levels, the mean in the non-recovery group was 9.533×10^9/L compared with 5.908×10^9/L in the recovery group, a correlation that was statistically significant ($p < 0.001$). Besides, the results of pulse oxygen saturation (SpO_2) were 89.329% and 94.937% in the non-recovery group and the recovery group with a p-value < 0.001. The results of the six laboratory parameters between the non-recovery group and the recovery group are presented in Figure 3.

DISCUSSION

The infection of COVID-19 has been a huge challenge worldwide since its evolution from a viral pneumonia to a systemic disease, also known as the cytokine storm syndrome [6-8]. COVID-19 is a great threat to elderly people being a vulnerable population group, especially those suffering from various comorbidities, including cardiovascular diseases, hypertension, diabetes, malignancy, cerebrovascular diseases, and chronic kidney disease [9-12]. A meta-analysis showed that patients over 70 years old have a 65% higher risk for COVID-19 [13,14] which is consistent with the results in our research. In addition to increasing age, higher levels of pro-inflammatory cytokines may further facilitate COVID-19 [15]. Thus, an appropriate therapy according to the most particular biomarkers plays an important role in reducing mortality [16].

COVID-19 infection could lead to liver injury by the direct infection of cells. Typical expressions of liver damage consist of hypoalbuminemia and the changes of liver enzymes and bilirubin [17], which is statistically significant between the non-recovery group and the recovery group during hospitalization. LDH, one of the important enzymes for glycolysis and gluconeogenesis, exists in all the body cells. A large number of studies show that, as an inflammatory marker, LDH plays an important role in acute tissue damage, cardiovascular disease, liver disease and malignancy. The acute tissue hypoxia resulting from COVID-19 infection will lead to an increase of LDH levels. Research conducted by Li et al. showed that the level of LDH at admission is an independent risk factor for disease severity and mortality, which is similar to the study by De Smet et al. [18-20]. We also found that the level of LDH in the non-recovery group was higher than that in the recovery group, and the correlation is statistically significant. It suggests that higher levels of LDH could indicate an increased mortality in COVID-19 patients.

When a virus replicates in the body, the liver needs to produce an increasing amount of protein to support the immune response. However, this can create a vicious cycle, ultimately resulting in hypoalbuminemia. Hypoproteinemia, particularly a low albumin level, could lead to longer hospitalizations and a poor prognosis, which is related to the severity of patients with severe

acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS). Additionally, it is also a risk factor for the mortality in ICU patients [21,22]. Our findings align with these previous studies and suggest that a low albumin level is a significant risk factor for adverse outcomes in COVID-19 patients.

In addition to LDH and serum albumin, we also noted an elevation in Dbil levels during hospitalization of critically ill patients with widespread hepatocyte damage. Dbil was calculated at 4.720 μmol/L in the non-recovery group vs 2.873 μmol/L in the recovery group, showing a clear statistical correlation ($p < 0.001$). Our finding suggests that patients with higher levels of DB are more likely to develop severe COVID-19.

In a comprehensive literature review conducted by Tjendra Y et al. on multiple biomarkers that predict COVID-19 severity, raised inflammatory indicators such as serum ferritin, released by damaged cells, and high neutrophil count which were associated with a higher risk of critical illness, ultimately contribute to the severity of disease and fatal consequences in COVID-19 patients [23]. During the study period, the mean of serum ferritin and NEUT# in the non-recovery group is 2,577.564 ng/mL and 9.533×10^9/L, whereas for the second group, it is 1,007.636 ng/mL and 5.908×10^9/L, an association that was statistically significant with a p-value = 0.001 and < 0.001, respectively. Hyperinflammation is considered the main cause of the critical forms of COVID-19. Based on our findings, it is reasonable to presume that serum ferritin and NEUT# can be used to predict cases that are more susceptible to progress unfavorably.

In summary, we observed by laboratory indicators during hospitalization, including elevated LDH and Dbil, decreased ALB were more widespread in critically ill patient groups. Indicators, such as serum ferritin or NEUT#, could be used to forecast disease outcome. Additionally, our findings indicated that patients with lower SpO_2 and those requiring mechanical ventilation were more likely to develop severe disease, and in some cases, even result in death.

The ROC curve analysis by combining LDH, Dbil, ferritin, NEUT#, ALB, and SpO_2 finds an optimal AUC of 0.886%. We achieved the highest sensitivity of 89.2% and specificity of 77.4%.

We ranked 66 biomarkers based on their association with COVID-19 outcomes and identified the six most promising ones: LDH, ALB, NEUT#, ferritin, SpO_2, and Dbil. As variant strains of SARS-COV-2 continue to emerge and spread to different countries, it has posed a significant threat to prevention. Therefore, monitoring changes of these laboratory indicators can aid clinicians in providing personalized treatment, delaying disease progression, and reducing COVID-19 mortality.

CONCLUSION

This is a review to use machine analysis to search for the optimal combination of LDH, ALB, NEUT#, ferritin, SpO_2, and Dbil which can forecast the outcome of critically ill patients with COVID-19 and is highly predictive of mortality.

Source of Funds:

Department of Clinical Laboratory, Key Laboratory of Clinical Laboratory Diagnosis and Translational Research of Zhejiang Province, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China. This work was funded by Key Laboratory of Clinical Laboratory Diagnosis and Translational Research of Zhejiang Province (2022E10022). The funding had no involvement in study design, collection, analysis and interpretation of data, writing of the report, and the decision to submit the article for publication.

Declaration of Interest:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References:

1. Ljungqvist O, de Boer HD, Balfour A, et al. Opportunities and Challenges for the Next Phase of Enhanced Recovery After Surgery: A Review. JAMA Surg 2021;156:775-84. (PMID: 33881466)

2. Guan X, Zhang B, Fu M, et al. Clinical and inflammatory features based machine learning model for fatal risk prediction of hospitalized COVID-19 patients: results from a retrospective cohort study. Ann Med 2021;53:257-66. (PMID: 33410720)

3. Vijenthira A, Gong IY, Fox TA, et al. Outcomes of patients with hematologic malignancies and COVID-19: a systematic review and meta-analysis of 3377 patients. Blood 2020;136:2881-92. (PMID: 33113551)

4. Terpos E, Ntanasis-Stathopoulos I, Elalamy I, et al. Hematological findings and complications of COVID-19. Am J Hematol 2020;95:834-47. (PMID: 32282949)

5. Kuo KM, Talley PC, Chang CS. The accuracy of machine learning approaches using non-image data for the prediction of COVID-19: A meta-analysis. Int J Med Inform 2022;164:104791. (PMID: 35594810)

6. Fajgenbaum DC, June CH. Cytokine Storm. N Engl J Med 2020; 383:2255-73. (PMID: 33264547)

7. Mehta P, McAuley DF, Brown M, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet 2020; 395:1033-4. (PMID: 32192578)

8. Pływaczewska-Jakubowska M, Chudzik M, Babicki M, Kapusta J, Jankowski P. Lifestyle, course of COVID-19, and risk of Long-COVID in non-hospitalized patients. Front Med (Lausanne) 2022;9:1036556. (PMID: 36353225)

9. Mazloom R. Feasibility of Therapeutic Effects of the Cholinergic Anti-Inflammatory Pathway on COVID-19 Symptoms. J Neuroimmune Pharmacol 2020;15:165-6. (PMID: 32378064)

10. Wan S, Xiang Y, Fang W, et al. Clinical features and treatment of COVID-19 patients in northeast Chongqing. J Med Virol 2020; 92:797-806. (PMID: 32198776)

11. Esme M, Koca M, Dikmeer A, et al. Older Adults With Coronavirus Disease 2019: A NationwideStudy in Turkey. J Gerontol A Biol Sci Med Sci 2021;76:e68-e75. (PMID: 32871002)

12. Wang L, Li X, Chen H, et al. Coronavirus Disease 19 Infection Does Not Result in Acute Kidney Injury: An Analysis of 116 Hospitalized Patients from Wuhan, China. Am J Nephrol 2020; 51:343-8. (PMID: 32229732)

13. Taylor EH, Marson EJ, Elhadi M, et al. Factors associated with mortality in patients with COVID-19 admitted to intensive care: a systematic review and meta-analysis. Anaesthesia 2021;76:1224-32. (PMID: 34189735)

14. Pijls BG, Jolani S, Atherley A, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ Open 2021;11:e044640. (PMID: 33431495)

15. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020;369:m1328. (PMID: 32265220)

16. Bouayed MZ, Laaribi I, Chatar CEM, et al. C-Reactive Protein (CRP): A poor prognostic biomarker in COVID-19. Front Immunol 2022;13:1040024. (PMID: 36451818)

17. Kumar MP, Mishra S, Jha DK, et al. Coronavirus disease (COVID-19) and the liver: a comprehensive systematic review and meta-analysis. Hepatol Int 2020;14:711-22. (PMID: 32623633)

18. Gálvez-Barrón C, Arroyo-Huidobro M, Miñarro A, et al. COVID-19: Clinical Presentation and Prognostic Factors of Severe Disease and Mortality in the Oldest-Old Population: A Cohort Study. Gerontology 2022;68:30-43. (PMID: 33853067)

19. Owen RK, Conroy SP, Taub N, et al. Comparing associations between frailty and mortality in hospitalised older adults with or without COVID-19 infection: a retrospective observational study using electronic health records. Age Ageing 2021;50:307-16. (PMID: 32678866)

20. Önal U, Gülhan M, Demirci N, et al. Prognostic value of neutrophile-to-lymphocyte ratio (NLR) and lactate dehydrogenase (LDH) levels for geriatric patients with COVID-19. BMC Geriatr 2022;22:362. (PMID: 35468761)

21. Zhang L, Yu W, Zhao Y, et al. Albumin Infusion May Improve the Prognosis of Critical COVID-19 Patients with Hypoalbuminemia in the Intensive Care Unit: A Retrospective Cohort Study. Infect Drug Resist 2022;15:6039-50. (PMID: 36277241)

22. Stone R, Scheib S. Advantages of, and Adaptations to, Enhanced Recovery Protocols for Perioperative Care during the COVID-19 Pandemic. J Minim Invasive Gynecol 2021;28:481-9. (PMID: 33359742)

23. Garcia-Beltran WF, Lam EC, Astudillo MG, et al. COVID-19-neutralizing antibodies predict disease severity and survival. Cell 2021;184:476-488.e411. (PMID: 33412089)

Additional material can be found online at:
http://supplementary.clin-lab-publications.com/230721/