

Machine Learning to Predict Disease Severity and Progression in Hospitalized COVID-19 Patients Using Laboratory Data on Admission

Gokhan Tazegül¹, Volkan Aydın², Elif Tükenmez Tigen³, Buket Ertürk Şengel³, Kübra Köksal^{4,5}, Buket Doğan⁴, Sait Karakurt⁶, Zehra Aysun Altıkardeş^{4,5}, Lütfiye Mülazımoğlu Durmuşoğlu³, Ali Serdar Fak^{1,5}, Abdulsamet Aktaş⁴, Uluhan Sili³, Abidin Gündoğdu¹, Fethi Gül⁷, Sena Tokay Tarhan¹, Semiha Emel Eryüksel⁶, Mümine Topçu⁵, Berrin Aysevinç⁵, Songül Çeçen Düzel⁵, Tuba Güçtekin⁸, Derya Kocakaya⁶, Beste Özben Sadıç⁸, Halil Ataş⁸, Mustafa Kürşat Tigen⁸, Ahmet Altuğ Çiçin⁸, Bülent Mutlu⁸, Alper Kepez⁸, Mehmet Baran Balcan⁹, Ayla Erdoğan¹, Emre Çapar¹, Ömer Ataç¹⁰, Beliz Bilgili⁷, İsmail Cinel⁷, Ahmet Akıcı¹¹, Rafi Haner Direskeneli¹²

¹ Marmara University Faculty of Medicine, Department of Internal Medicine, Division of General Internal Medicine, İstanbul, Turkey.

² İstanbul Medipol University International School of Medicine, Department of Medical Pharmacology, İstanbul, Turkey.

³ Marmara University Faculty of Medicine, Department of Infectious Disease and Clinical Microbiology, İstanbul, Turkey.

⁴ Marmara University Faculty of Technology, Department of Computer Engineering, İstanbul, Turkey.

⁵ Marmara University Hypertension and Atherosclerosis Research Center, İstanbul, Turkey.

⁶ Marmara University Faculty of Medicine, Department of Pulmonary and Critical Care Medicine, İstanbul, Turkey.

⁷ Marmara University Faculty of Medicine, Department of Anesthesiology and Reanimation, Division of Critical Care, İstanbul, Turkey.

⁸ Marmara University Faculty of Medicine, Department of Cardiology, İstanbul, Turkey.

⁹ Koç University Faculty of Medicine, Department of Pulmonary and Critical Care Medicine, İstanbul, Turkey.

¹⁰ İstanbul Medipol University International Faculty of Medicine, Department of Public Health, İstanbul, Turkey.

¹¹ Marmara University Faculty of Medicine, Department of Medical Pharmacology, İstanbul, Turkey.

¹² Marmara University Faculty of Medicine, Department of Internal Medicine, Division of Rheumatology, İstanbul, Turkey.

ABSTRACT

Background Herein, we aimed to develop and test machine learning (ML) models to predict disease severity and/or progression in hospitalised COVID-19 patients through baseline laboratory features.

Methods In this retrospective study of hospitalised COVID-19 patients admitted to a tertiary care centre, we evaluated routine admission data to determine the accuracy rates of different ML algorithms: k-nearest neighbour classifier, bagging classifier, random forest (RF), and decision tree. These models were compared over three outcomes: those who needed oxygen supplementation vs who did not on admission (Analysis 1, n: 180), those who later developed oxygen requirement vs those who did not (Analysis 2, n: 112), and those who needed invasive mechanical ventilation vs. those who did not during hospitalisation (Analysis 3, n: 164).

Results The median age of the patients was 55 (44-68) years, with males constituting 47.2% of the subjects. At admission, 37.8% of the patients required oxygen supplementation. During hospitalisation, 17.5% needed mechanical ventilation, and 8.3% died. For all analyses, RF had the highest accuracy in classifying the need for oxygen supplementation on admission (89.4%) or during hospitalisation (91.1%) and for invasive mechanical ventilation (92.2%). These were followed by a bagging classifier for Analysis 1 (88.3%) and Analysis 3 (91.0%) and by a decision tree for Analysis 2 (88.4%). C-reactive protein, monocyte distribution width, and high-sensitive troponin-T were the most crucial laboratory contributors to Analysis 1, Analysis 2, and Analysis 3, respectively.

Conclusion Our study showed that ML algorithms could predict the need for oxygen supplementation and mechanical ventilation during hospitalisation using baseline laboratory data, suggesting a slight superiority of RF, among others.

Turk J Int Med 2024;6(3):144-154

DOI: 10.46310/tjim.1502238

Original Article

Keywords: Accuracy, classifiers, COVID-19, inpatient, oxygen supplementation, random forest



INTRODUCTION

In March 2020, the World Health Organization (WHO) officially declared the outbreak of Coronavirus disease 2019 (COVID-19) a pandemic. By the end of March 2024, the total number of COVID-19 cases globally had surpassed a staggering 775 million, resulting in the loss of 7 million lives worldwide.¹ The pandemic was characterized by unprecedented cases that overwhelmed healthcare facilities globally², and it still poses a significant threat since presentations are heterogeneous; 15% of all infected patients deteriorate rapidly, with multiorgan damages and high fatality rates.³⁻⁵ Therefore, finding novel ways for effective triage and timely risk stratification to predict COVID-19 deterioration remains an important research area. In this context, patient progression through the healthcare system is assessed via the WHO Clinical Progression Scale (WHO-CPS), which the WHO recommends as an outcome measure.⁶ Early warning scores (EWS) that help recognize clinical deterioration in the short term have been extensively used in COVID-19 patients.⁷ Among them, the National Early Warning Score (NEWS) and Modified Early Warning Score (MEWS) have been reported to predict mortality and clinical deterioration adequately⁸⁻¹⁰; however, several recent studies on EWS models showed subpar results.¹¹⁻¹³

Artificial intelligence (AI), featuring various machine learning (ML) tools, can analyze large amounts of data and offer solutions that are not apparent. AI programs

have already been adopted as decision support systems in clinical practice, where certain ML models are known to generate better performance than traditional prediction models.^{14,15} Several studies have successfully tested ML's predictive value in COVID-19-related mortality and clinical deterioration¹⁶⁻¹⁸, with some models showing promise for possible identification of low-risk patients for early discharge.¹⁹ Nevertheless, ML studies on COVID-19 are heterogeneous, as there is a plethora of included parameters as well. Therefore, in this study, we aimed to develop and test ML models to predict WHO-CPS-oriented disease severity and/or progression in hospitalized COVID-19 patients using baseline laboratory features on hospital admission.

MATERIAL AND METHODS

Ethical considerations

This single-center retrospective study was approved by the institutional review board of the Turkish Ministry of Health's COVID-19 Scientific Research Studies, and ethical approval was obtained from Marmara University Clinical Studies Ethics Committee (Approval date: 27.04.2020, Approval number: 09.2020.487). This study was conducted by the Declaration of Helsinki and the Research and Publication Ethics, and patient data were anonymized before analysis.

Table 1. Seventy-three demographic and laboratory features are included in the dataset after preprocessing

| | | | | |
|-------------------|--------------------|---------------|------------------|-----------------|
| Gender | Creatinine | LDH | NEU#/LYM# | PT |
| Age | CRP | LYM# | NEU#/PLT# | PT,% |
| Albumin | D-dimer | LYM% | NRBC# | RBC |
| ALP | Direct bilirubin | LYM#/CRP | NRBC% | RDW |
| ALT | EOS# | Magnesium | Osmolarity | sO ₂ |
| aPTT | EOS% | MCH | pCO ₂ | Sodium |
| AST | Ferritin | MCHC | PCT | Total bilirubin |
| BAS# | Fibrinogen | MCV | PDW | Total protein |
| BAS% | GGT | MDW | pH | Troponin T-hs |
| Base Excess | Glucose | Methemoglobin | Phosphorus | Urea |
| BUN | HCO ₃ - | MON# | PLT | Uric acid |
| Calcium | HCT | MON% | PLT#/LYM# | WBC |
| Carboxyhemoglobin | HGB | MPV | pO ₂ | WDOP |
| Chloride | INR | NEU# | Potassium | |
| CK-MB (mass) | Lactate | NEU% | Procalcitonin | |

Initial and worst WHO-CPS scores were not included in this demographic and laboratory features presentation. (#) denotes counts, and (%) denotes percent. ALP: alkaline phosphatase, ALT: alanine transaminase, aPTT: activated partial thromboplastin time, AST: aspartate transaminase, BAS: basophils, BUN: blood urea nitrogen, CK-MB: creatine kinase myocardial band, CRP: C-reactive protein, EOS: eosinophils, GGT: gamma-glutamyl transferase, HCO₃-: bicarbonate, HCT: hematocrit, HGB: hemoglobin, INR: international normalized ratio, LDH: lactate dehydrogenase, LYM: lymphocytes, MCH: mean corpuscular hemoglobin, MCHC: mean corpuscular hemoglobin concentration, MCV: mean corpuscular volume, MDW: monocyte distribution width, MON: monocytes, MPV: mean platelet volume, NEU: neutrophils, PLT: platelets, NRBC: nucleated red blood cell, pCO₂: partial pressure of carbon dioxide, PCT: plateletcrit, PDW: platelet distribution width, pO₂: partial pressure of oxygen, PT: prothrombin time, RBC: red blood cell, RDW: red blood cell distribution width, sO₂: blood oxygen saturation, Troponin T-hs: troponin T-high sensitivity, WBC: white blood cells, WDOP: white cell differential optical count.

Study setting

This study evaluated WHO-CPS-oriented patient outcomes in patients hospitalized with COVID-19 infection admitted to (censored) University Training and Research Hospital between 27 April 2020 and 1 June 2020. We collected baseline data on routine clinical evaluation encompassing medical history, thorough physical examination, and initial laboratory tests, including complete blood count, biochemistry panel, and inflammatory markers. All patients were followed until death, discharge, or up to 28 days of hospital stay.

Data handling

We followed four basic stages to determine the accuracy rates of different ML algorithms: creating the dataset, preprocessing, random feature selection, and classification.

Study population, dataset, and preprocessing

The initial dataset consisted of 508 patients admitted to the COVID-19 unit within the study period and included 193 parameters, including age, sex, initial and worst WHO-CPS scores, and 189 laboratory results. Patients without a confirmed COVID-19 infection based on reverse-transcriptase polymerase chain reaction for the SARS-CoV-2 ribonucleic acid, who were rapidly treated with intubation and mechanical ventilation during initial presentation, and who were directly admitted to the intensive care unit were not included in this study. Included patients had a valid initial laboratory result obtained within the first 24 hours of admission. Parameters with substantial missing data (present in less than 50% of the cases) or duplicated (e.g., obtained from arterial and venous blood) were excluded. The final analysis was conducted on 180 patients with 75 attributes (71 laboratory parameters, age, sex, initial WHO-CPS, and worst WHO-CPS) (Table 1). Before data processing, all features are normalized to have 0 mean and unit standard deviation.

Feature selection

Feature subset selection is a critical step of data mining, where fewer parameters could achieve higher accuracy (Figure 1).²⁰ We used the random subset feature selection (RSFS) algorithm to reduce the number of features in the data set.²¹ The feature selection process is iterative. The K-Nearest Neighbor (KNN) classifier classifies the randomly selected

feature subsets, rated according to their relevance values at each step. Each subset has randomly selected features as the square root of the total number of features.²² The relative contribution and ranking of the selected features were assessed via the Correlation Attribute Eval (CA) algorithm and the Ranker method.

Classification

The targets, i.e., the outcomes of the study, were classified as to the pre-defined initial and worst WHO-CPS categories of the patients, which included reverse-transcriptase polymerase chain reaction positivity for the SARS-CoV-2 ribonucleic acid, symptomatology of patients, the need for and the severity of oxygen supplementation, and the need for non-invasive or invasive mechanical ventilation.⁶ The initial WHO-CPS category was defined as the WHO-CPS score during the initial presentation. In contrast, the worst WHO-CPS category was the highest WHO-CPS score during a patient's follow-up. In all classifications, a standardization process was performed on the dataset with the WEKA application (WEKA 3.8, Waikato, New Zealand)²³, and the model's accuracy was calculated using k-fold cross-validation.²² We used the KNN classifier, bagging classifier, random forest, and decision tree ML algorithms in the training phase.^{22,24-26} All classification results were generated using the 10-fold cross-validation technique and were evaluated according to whether there was standardization within each algorithm. The relevance value of each randomly generated subset is calculated as the difference between the performance criterion (the average recall value for the current iterations) and the expected criterion (correctly classified / correctly classified + incorrectly categorized).

Evaluation metrics

Model performances were evaluated using the accuracy value, which equals the percentage of the correctly classified positive and negative subjects: $(\text{true negatives} + \text{true positives}) / (\text{all subjects})$.¹⁶ We also calculated the F1-score, an important metric in unbalanced data sets and can be described as a weighted average of the precision and recall values. F1-score equals $2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall})$, where precision equals $\text{true positives} / (\text{true and false positives})$, and recall equals $\text{true positives} / (\text{true positives and false negatives})$.¹⁶ We only presented the F1-score of the ML model with the highest accuracy in each analysis.

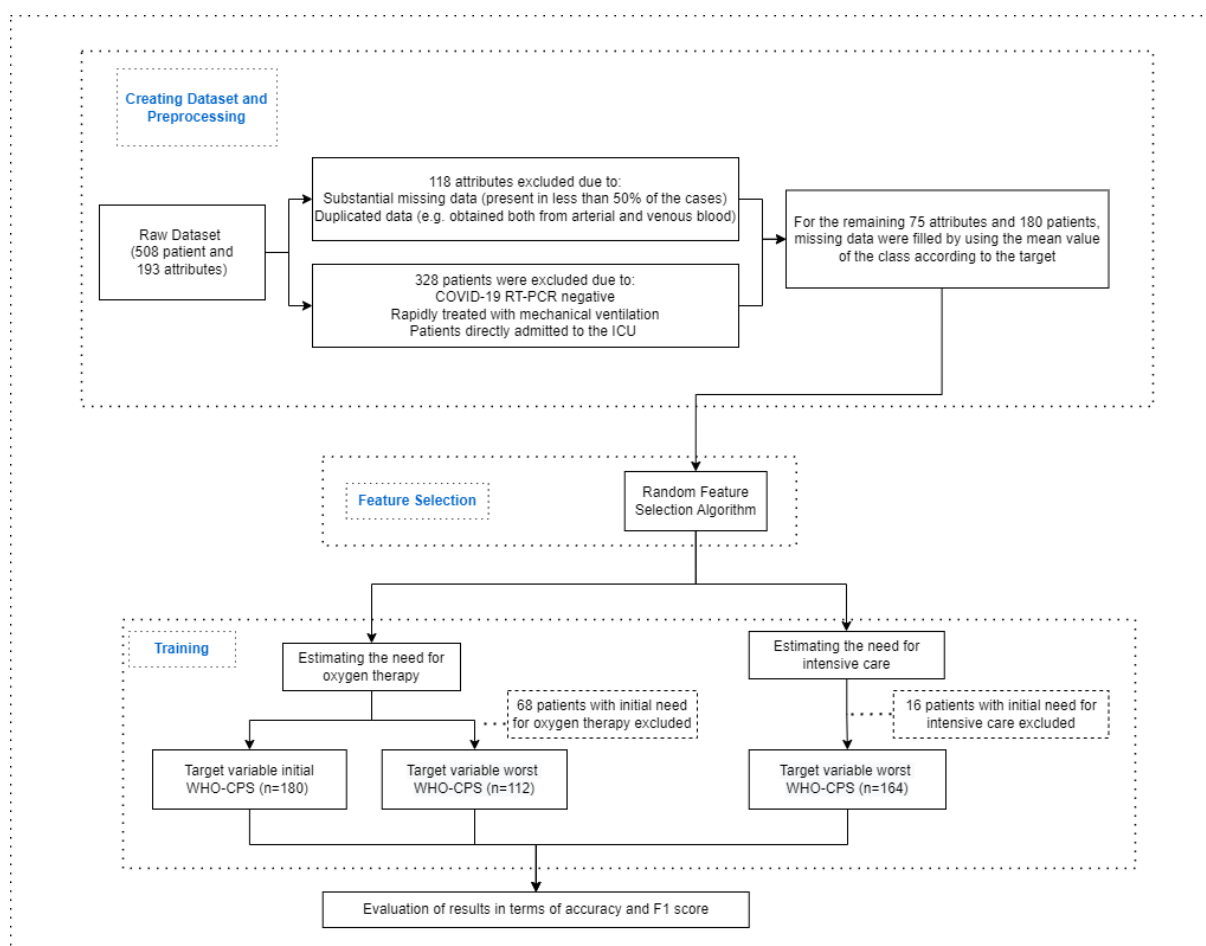


Figure 1. A model development pipeline flowchart was used in this study.

Study outcomes

Three main events were analyzed as outcomes within this study. The study outcomes were determining the accuracy rate of differentiating the subjects (i) who needed oxygen supplementation (WHO-CPS Score 5-9) from those who did not (WHO-CPS Score 1-4) on initial admission (Analysis 1, n: 180); (ii) patients who later developed the need for oxygen supplementation (WHO-CPS Score 5-9) from those who did not (WHO-CPS Score 1-4), excluding those who needed oxygen supplementation on initial admission (Analysis 2, n: 112), and (iii) who needed invasive mechanical ventilation (WHO-CPS Score 7-9) from those who did not (WHO-CPS Score 1-6) during hospitalization, excluding those who needed invasive mechanical ventilation during initial admission (Analysis 3, n: 164).

Statistical analysis

Statistical analyses were carried out using SPSS 24.0 software. Baseline categorical variables were expressed as numbers and percentages, and continuous variables were presented as medians and interquartile ranges. Each analysis's relevant target

group's features were compared through the Mann-Whitney U test. An overall 5% type-I error level was used to infer statistical significance.

RESULTS

The median age of the overall study population was 55 (44-68) years, with males constituting 47.2% of the participants. We identified comorbidities in 67.2% of the subjects, which was the most common reason for hospitalization (45.0%), followed by advanced age (37.7%), dyspnea/hypoxia (35.5%), and radiological evidence of severe pneumonia (34.4%). At admission, 62.2% did not require oxygen supplementation, while others did, with non-invasive (28.9%) or invasive mechanical ventilation (8.9%). Table 2 shows the baseline clinical characteristics of the analyzed study subgroups. During hospitalization, 29 patients (17.5%) developed the need for mechanical ventilation, and 15 patients (8.3%) died.

Within the data of 180 patients and 75 attributes, the RSFS algorithm identified 16 attributes to classify the need for oxygen supplementation on admission,

where the random forest had the highest accuracy (89.4%), followed by the bagging classifier (88.3%). The ranking of the attributes by the CA algorithm showed C-reactive protein (CRP) (0.41) and monocyte distribution width (MDW) (0.40) as the most important contributors. The F1-score for defining the patient group who did not need oxygen on admission was 0.92, whereas the F1-score for defining the group that needed oxygen on admission was 0.86 (Figure 2). All included variables in the ML algorithm had statistically significant differences between the two groups (Mann-Whitney U test, $p < 0.05$ for all pairs).

In 112 patients who did not need oxygen supplementation during admission, developing a need for oxygen during hospitalization was classified by 18 attributes, with the highest accuracy by random forest (91.1%), followed by decision tree (88.4%). Of these eighteen classifiers, MDW (0.49), high-sensitive troponin T (0.43), CRP (0.41), and calcium (0.40) were the most critical contributors to identifying the need for oxygen among those who did not require oxygen on admission. The F1 score for the patient group that did not require oxygen during the study period was 0.95. In contrast, it was 0.71 for the patient group that

had developed the need for oxygen supplementation (Figure 3). Although included in the ML algorithm, several notable variables, namely pO₂, neutrophil (NEU) and white blood cell (WBC) count, total and direct bilirubin, and white cell differential optical count (WDOC), had not shown any statistically significant difference between the two groups (Mann-Whitney U test, $p > 0.05$ for all pairs).

Twelve attributes were used to classify any need for invasive mechanical ventilation during hospitalization in patients not hospitalized in the intensive care unit on admission. Random forest achieved the highest accuracy (92.2%), followed by the bagging classifier (91.0%). Initial WHO-CPS category (0.56), high-sensitive troponin T (0.46), and CK-MB (0.42) were the most important contributors. The F1 score for developing a need for invasive mechanical ventilation during the study period was 0.73, whereas it was 0.94 for not needing invasive mechanical ventilation (Figure 4). Apart from methemoglobin levels and gender (Mann-Whitney U test, $p > 0.05$), all included variables in the ML algorithm had statistically significant differences between the two groups (Mann-Whitney U test, $p < 0.05$, Figure 4).

Table 2. Baseline characteristics of the study population.

| Characteristics | Analysis 1 (n: 180) | Analysis 2 (n: 112) | Analysis 3 (n: 164) |
|---|------------------------|------------------------|------------------------|
| Age (years) (median, IQR) | 55 (44-68) | 56 (39-68) | 56 (42-72) |
| Male n (%) | 85 (47.2) | 58 (51.8) | 82 (50) |
| Any comorbidity n (%) | 121 (67.2) | 66 (58.9) | 109 (66.4) |
| Hypertension | 54 (30.0) | 26 (23.2) | 46 (28) |
| Cardiovascular disease | 23 (12.7) | 11 (9.8) | 21 (12.8) |
| Asthma/COPD | 20 (11.0) | 7 (6.1) | 15 (9.1) |
| Diabetes mellitus | 19 (10.5) | 5 (4.4) | 12 (7.3) |
| Chronic kidney disease | 10 (5.5) | 4 (3.5) | 7 (4.2) |
| Rheumatologic/autoimmune disease | 9 (5.0) | 6 (5.3) | 8 (4.8) |
| Immunodeficiency | 8 (4.4) | 3 (2.6) | 5 (3.0) |
| Neurological disease | 8 (4.4) | 2 (1.7) | 7 (4.2) |
| Solid organ tumors | 5 (2.7) | 1 (0.8) | 3 (1.8) |
| Reasons for hospitalization n (%) | | | |
| Comorbidities | 81 (45.0) | 39 (34.8) | 69 (42) |
| Advanced age | 68 (37.7) | 44 (39.2) | 62 (37.8) |
| Dyspnea/hypoxia | 64 (35.5) | 23 (20.5) | 50 (30.4) |
| Radiological findings of severe pneumonia | 62 (34.4) | 29 (25.8) | 54 (32.9) |
| Other | 17 (9.4) | 17 (15.1) | 17 (10.3) |
| Need for oxygen supplementation on admission. | | | |
| No (WHO-CPS Score 1 to 4) | 112 (62.2) | 112 (100.0) | 112 (68.3) |
| Yes, without mechanic ventilation (WHO CPS Score 5-6) | 52 (28.9) | - | 52 (31.7) |
| Yes, with invasive mechanic ventilation (WHO-CPS Score 7-9) | 16 (8.9) | - | - |

Categorical variables were expressed as numbers and percentages, and continuous variables were presented as medians and interquartile ranges. IQR: Interquartile range, COPD: chronic obstructive pulmonary disease, WHO-CPS: World Health Organization Clinical Progression Scale.

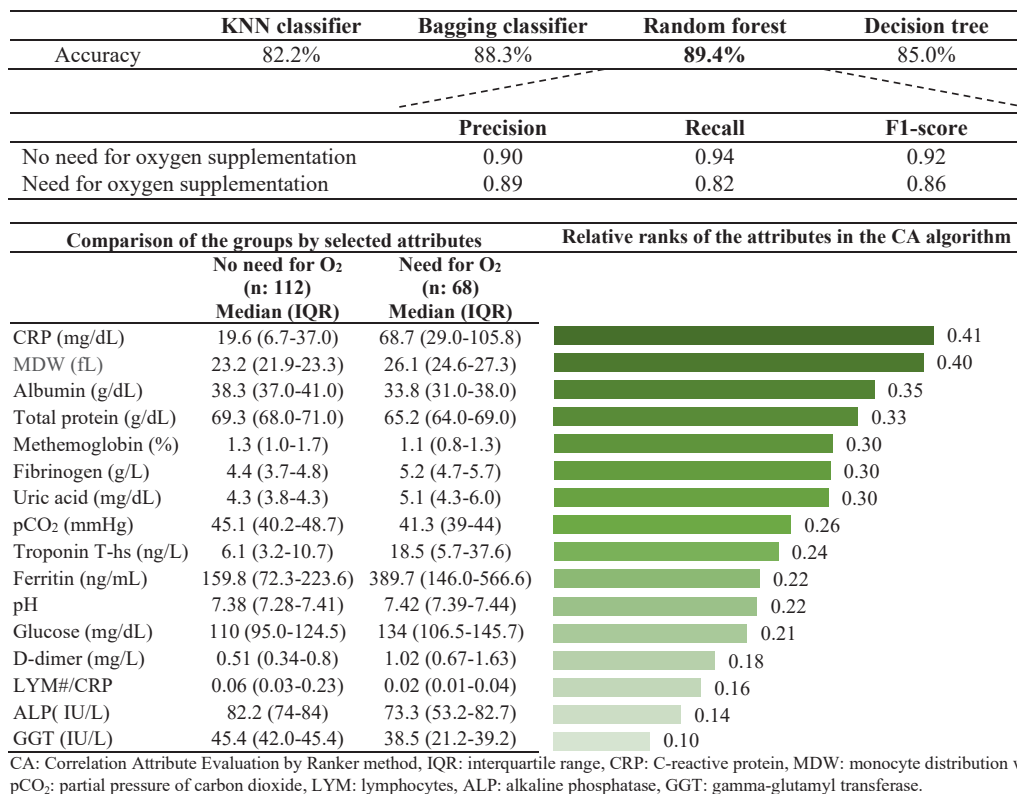


Figure 2. Classification accuracy and F1 scores of machine learning algorithms for oxygen requirement on admission (*upper panel*) with the relative rankings (*right lower*) and statistical comparisons (*left lower*) of the selected attributes.

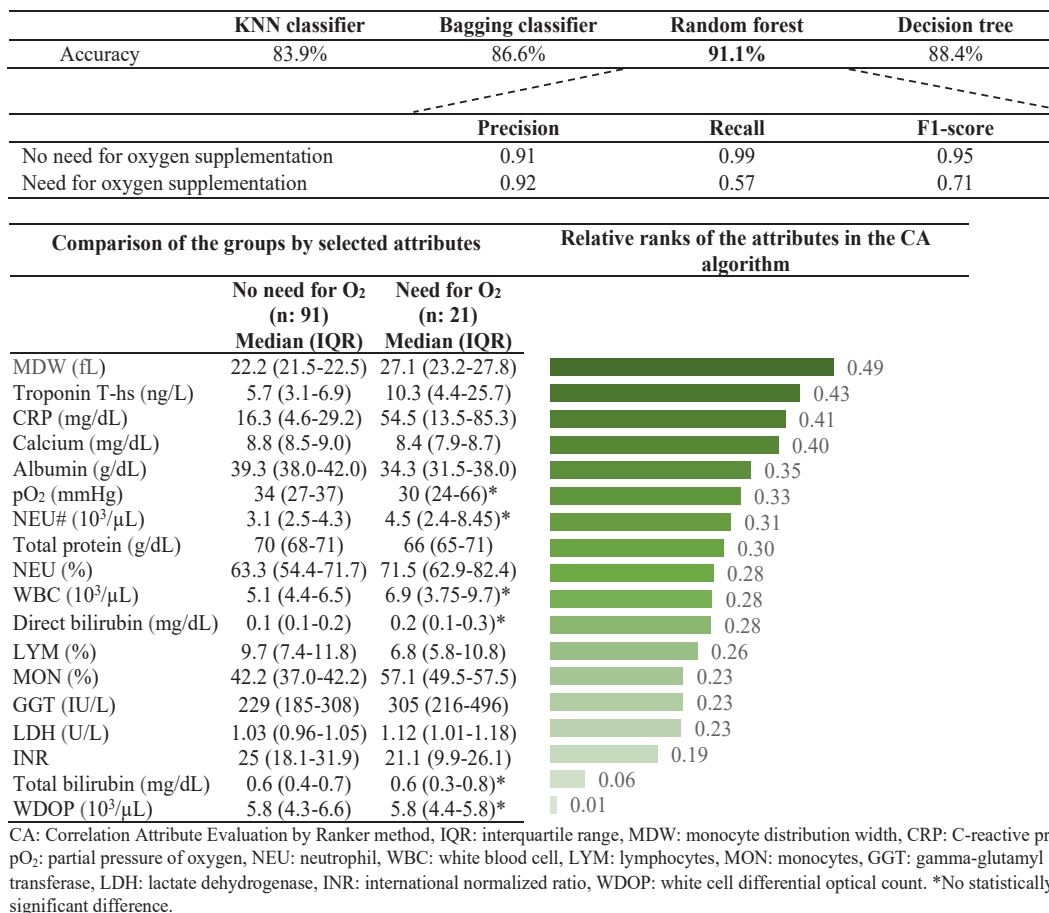


Figure 3. Classification accuracy and F1 scores of machine learning algorithms for oxygen requirement during hospitalization (*upper panel*) with the relative rankings (*right lower*) and statistical comparisons (*left lower*) of the selected attributes.

| | KNN classifier | Bagging classifier | Random forest | Decision tree |
|----------|----------------|--------------------|---------------|---------------|
| Accuracy | 87.4% | 91.0% | 92.2% | 90.4% |

| | Precision | Recall | F1-score |
|--|-----------|--------|----------|
| No invasive mechanical ventilation | 0.95 | 0.93 | 0.94 |
| Need for invasive mechanical ventilation | 0.71 | 0.76 | 0.73 |

| Comparison of the groups by selected attributes | | Relative ranks of the attributes in the CA algorithm | |
|---|---|--|------|
| | No mechanical ventilation (n=135) Median (IQR) | Need for mechanical ventilation (n=29) Median (IQR) | |
| Initial WHO-CPS category | 4 (4-4) | 5 (5-6) | 0.56 |
| Troponin T-hs (ng/L) | 6.6 (3.4-10.8) | 36.3 (8.1-41.3) | 0.46 |
| CK-MB (IU/L) | 1.3 (0.7-1.5) | 2.7 (1.7-3.14) | 0.42 |
| Hematocrit (%) | 38.1 (35.5-40.9) | 33.1 (29.7-37.7) | 0.33 |
| Hemoglobin (g/dL) | 13.3 (12.2-14.3) | 11.5 (9.9-13.2) | 0.32 |
| RBC (10 ³ /μL) | 4.5 (4.2-4.9) | 4.1 (3.7-4.5) | 0.30 |
| MDW (fL) | 23.6 (22.8-23.7) | 26.2 (24.4-27.4) | 0.29 |
| pO ₂ (mmHg) | 36 (26.0-37.5) | 38 (26.5-49.7) | 0.19 |
| MON (%) | 9.6 (7.2-11.6) | 7.2 (4.9-8.4) | 0.11 |
| Methemoglobin (%) | 1.26 (0.95-1.6) | 1.15 (0.85-1.45)* | 0.10 |
| MON (10 ³ /μL) | 0.5 (0.4-0.6) | 0.4 (0.3-0.5) | 0.02 |
| Male (%) | 50.3 | 48.2* | 0.01 |

CA: Correlation Attribute Evaluation by Ranker method, IQR: interquartile range, WHO-CPS: World Health Organization Clinical Progression Scale, Troponin T-hs: Troponin T-high sensitivity, CK-MB: creatine kinase myocardial band, RBC: red blood cell, MDW: monocyte distribution width, pO₂: partial pressure of oxygen, MON: monocytes. *No statistically significant difference.

Figure 4. Classification accuracy and F1 scores of machine learning algorithms for mechanic ventilation requirement during hospitalization (*upper panel*) with the relative rankings (*right lower*) and statistical comparisons (*left lower*) of the selected attributes.

DISCUSSION

Healthcare systems face many difficulties managing resources and healthcare personnel during a pandemic. Although there have been studies on many parameters that predict disease severity or mortality risk of COVID-19, such as laboratory features (e.g., CRP, ferritin, D-dimer, lymphocyte count), using these parameters in traditional statistical methods are complex, heterogeneous, and not cost-effective.^{27,28} Accurately predicting severity allows managing COVID-19-infected patients on admission, which will help decrease hospital burden and pressure on healthcare workers.²⁹ In this single-center retrospective study focusing on testing ML models to predict the need for oxygen supplementation or mechanical ventilation in hospitalized COVID-19, using baseline laboratory biomarkers on admission, we have demonstrated that our models might help discriminate patients who would need oxygen supplementation or mechanical ventilation during their COVID-19 infection and allocate health services for them. These findings would be clinically significant in a resource-limited setting, where ML algorithms could aid clinicians in decision-making.

There has been a plethora of evidence regarding conventional scoring systems, such as NEWS, NEWS2, MEWS, and other scores, to predict severe COVID-19 and COVID-19-related mortality; all scores show moderate-to-high discriminatory power based on the clinical scenario they have been used for.^{10,30,31} However, due to the complex nature of the COVID-19 pandemic and multi-faceted causes of severe infection and mortality, there have been efforts to develop different ML applications to predict COVID-19 prognosis better: purposes: Kamran et al.¹² developed an ML model that can define patients at risk for clinical deterioration in patients with COVID-19 infection with external validation; Yu et al.¹⁹ demonstrated that different ML methods using blood inflammatory cytokines could help predict COVID-19 death; Elhazmi et al.³² developed a successful decision tree ML algorithm to predict mortality in critically ill adult patients, Liu et al.³³ used different ML algorithms to successfully predict mild, regular, severe and critical cases using clinical and radiological data, and Kocadagli et al.³⁴ investigated hybrid ML models to predict disease severity based on clinical and laboratory parameters. Similar to our study endpoint, in a multicenter retrospective study,

Yamanaka et al.³⁵ successfully predicted oxygen therapy needs in COVID-19 patients using a modern XGboost model with the eight clinical and laboratory variables, with a high negative predictive value of 0.93, and the authors underlined that compared to conventional scoring approaches, the ML model had better results. Although most of the data on ML use in COVID-19 are heterogeneous in terms of included data and endpoints, our results are consistent with the literature on the usefulness of ML in predicting COVID-19 cases that require oxygen supplementation and mechanical ventilation using objective and easily obtainable laboratory data. Since most traditional scoring systems and ML studies incorporate subjective clinical data like age, comorbidities, physical examination findings, and radiological data, we believe that this study could make a significant contribution to the literature, as our results suggest that easily accessible, quickly reported, and objective biochemical tests have comparable prognostic effectiveness, which could be beneficial for clinicians as an easily accessible, rapid tool in the future.

The International Federation of Clinical Chemistry (IFCC) stated that no single biochemical or hematological marker is sensitive enough or specific to predict the outcome of COVID-19 infection.³⁶ In particular, the IFCC recommends interpreting laboratory abnormalities based on groups of parameters. In our study, twelve to eighteen results identified patients needing oxygen supplementation and/or respiratory support. Nevertheless, two laboratory measurements were found within all three analyses in our study: MDW and high-sensitive troponin T. Monocytes undergo significant morphological changes, alterations in surface markers, and cytokine production during sepsis, both as they become activated and during the immunosuppressive phase. The volumetric changes can be detected as variations in MDW. The magnitude of MDW elevation correlates with organ dysfunction and sepsis severity, suggesting that MDW can be used as “a red flag,” a marker for the intensity of the inflammatory response.³⁷⁻³⁹ MDW, as a prognostic marker, can even outperform other early sepsis detection markers such as CRP and procalcitonin (PCT).⁴⁰ Apart from being a useful diagnostic and prognostic tool in sepsis, previous data have also demonstrated that MDW levels were elevated in COVID-19, correlated with disease activity, which is explained by the presence of hyperinflammatory state during COVID-19,

similar to sepsis, where pro-inflammatory cytokines are overexpressed, leads to morphological changes in monocytes, including increased cell size and variability.⁴¹ Moreover, in a retrospective study, MDW was higher in patients who needed respiratory support, and an MDW ≥ 25 had an area under the curve of 0.7 to identify oxygen requirement.⁴² Similar to MDW, higher troponin levels were associated with poor prognosis in COVID-19 infection and have been shown to identify a need for oxygen supplementation and mechanical ventilation.^{43,44} Our results on three different analyses are similar to the previously published literature on MDW and troponin, with similar differences between groups. ML approach provides a different perspective to these results, which define patients not based on conventional statistics but using stratification algorithms, which help discriminate beneficial patterns in extensive dimensional data to define subgroups of patients more accurately.⁴⁵

This study has several limitations. First, it is a retrospective study performed in a single hospital. Secondly, although we assessed models to determine if COVID-19 patients would need supplemental oxygen or mechanical ventilation, we did not consider comorbidities, treatments, radiological findings, or viral load while building ML models, which may have impacted disease severity. Third, the small sample size may restrict the precision of the identity of severity status. This may have affected our results because ML models involving multiple parameters require large datasets to train effectively and avoid overfitting. Nevertheless, we believe that the homogenous nature of the patient population still provides some insight regarding using ML in predicting clinical deterioration in patients with COVID-19. Additional studies focused on different waves and variants of COVID-19 spread are needed to validate the predictive accuracy of the evaluated scores, considering vaccination status as well.

CONCLUSIONS

The COVID-19 pandemic led to overwhelming complex clinical cases, with a significant percentage of patients rapidly deteriorating. Our data on a single-center retrospective cohort of hospitalized COVID-19 patients highlights the potential of integrating machine

learning algorithms into routine clinical practice as a valuable tool for analyzing complex clinical scenarios comprehensively. This emphasizes the importance of leveraging ML methods to predict clinical deterioration in COVID-19 patients, particularly in predicting the need for oxygen supplementation and mechanical ventilation.

Conflict of Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding Sources

This manuscript received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

Ethical Approval

The protocol of the study was approved by the Medical Ethics Committee of Marmara University Clinical Studies Ethics Committee (Approval date: 27.04.2020, Approval number: 09.2020.487).

Authors' Contribution

Study Conception: All authors contribute; Study Design: All authors contribute; Literature Review: All authors contribute; Critical Review: All authors contribute; Data Collection and/or Processing: All authors contribute; Analysis and/or Data Interpretation: All authors contribute; Manuscript preparing: Tazegul G, Aydın V.

REFERENCES

1. WHO. WHO Coronavirus (COVID-19) Dashboard 2020, 2022. Available at: www.covid19.who.int Accessed April 1, 2024.
2. Fontanarosa PB, Bauchner H. COVID-19-looking beyond tomorrow for health care and society. *JAMA*. 2020 May 19;323(19):1907-8. doi: 10.1001/jama.2020.6582.
3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020 Feb 15;395(10223):497-506. doi: 10.1016/S0140-6736(20)30183-5.
4. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA*. 2020 Mar 17;323(11):1061-9. doi: 10.1001/jama.2020.1585.
5. Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: Systematic review. *J Med Internet Res*. 2020 Feb 20;22(2):e16866. doi: 10.2196/16866.
6. WHO working group on the clinical characterisation and management of COVID-19 infection. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis*. 2020 Aug;20(8):e192-e197. doi: 10.1016/S1473-3099(20)30483-7.
7. Su Y, Ju MJ, Xie RC, Yu SJ, Zheng JL, Ma GG, Liu K, Ma JF, Yu KH, Tu GW, Luo Z. Prognostic accuracy of early warning scores for clinical deterioration in patients with COVID-19. *Front Med (Lausanne)*. 2021 Feb 1;7:624255. doi: 10.3389/fmed.2020.624255.
8. Myrstad M, Ihle-Hansen H, Tveita AA, Andersen EL, Nygård S, Tveit A, Berge T. National Early Warning Score 2 (NEWS2) on admission predicts severe disease and in-hospital mortality from Covid-19 - a prospective cohort study. *Scand J Trauma Resusc Emerg Med*. 2020 Jul 13;28(1):66. doi: 10.1186/s13049-020-00764-3.
9. Gidari A, De Socio GV, Sabbatini S, Francisci D. Predictive value of National Early Warning Score 2 (NEWS2) for intensive care unit admission in patients with SARS-CoV-2 infection. *Infect Dis (Lond)*. 2020 Oct;52(10):698-704. doi: 10.1080/23744235.2020.1784457.
10. Brajkovic M, Vukcevic M, Nikolic S, Dukic M, Brankovic M, Sekulic A, Popadic V, Stjepanovic M, Radojevic A, Markovic-Denic L, Rajovic N, Milic N, Tanasilovic S, Todorovic Z, Zdravkovic M. The predictive value of risk factors and prognostic scores in hospitalized COVID-19 patients. *Diagnostics (Basel)*. 2023 Aug 11;13(16):2653. doi: 10.3390/diagnostics13162653.
11. Veldhuis L, Ridderikhof ML, Schinkel M, van den Bergh J, Beudel M, Dormans T, Douma R, Gritters van den Oever N, de Haan L, Koopman K, de Kruif MD, Noordzij P, Reidinga A, de Ruijter W, Simsek S, Wyers C, Nanayakkara PW, Hollmann M. Early warning scores to assess the probability of critical illness in patients with COVID-19. *Emerg Med J*. 2021 Dec;38(12):901-5. doi: 10.1136/emermed-2020-211054.
12. Kamran F, Tang S, Otles E, McEvoy DS, Saleh SN, Gong J, Li BY, Dutta S, Liu X, Medford RJ, Valley TS, West LR, Singh K, Blumberg S, Donnelly JP, Shenoy ES, Ayanian JZ, Nallamotheu BK, Sjoding MW, Wiens J. Early identification of patients admitted to hospital for covid-19 at risk of clinical deterioration: model development and multisite external validation study. *BMJ*. 2022 Feb 17;376:e068576. doi: 10.1136/bmj-2021-068576.

13. Rauseo M, Perrini M, Gallo C, Mirabella L, Mariano K, Ferrara G, Santoro F, Tullo L, La Bella D, Vetuschi P, Cinnella G. Machine learning and predictive models: 2 years of Sars-CoV-2 pandemic in a single-center retrospective analysis. *J Anesth Analg Crit Care*. 2022 Oct 14;2(1):42. doi: 10.1186/s44158-022-00071-6.
14. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2014 Nov 15;13:8-17. doi: 10.1016/j.csbj.2014.11.005.
15. Pan L, Liu G, Lin F, Zhong S, Xia H, Sun X, Liang H. Machine learning applications for prediction of relapse in childhood acute lymphoblastic leukemia. *Sci Rep*. 2017 Aug 7;7(1):7402. doi: 10.1038/s41598-017-07408-0.
16. Karthikeyan A, Garg A, Vinod PK, Priyakumar UD. machine learning based clinical decision support system for early COVID-19 mortality prediction. *Front Public Health*. 2021 May 12;9:626697. doi: 10.3389/fpubh.2021.626697.
17. Bian Y, Han Q, Zheng Y, Yao Y, Fan X, Lv R, Pang J, Xu F, Chen Y. SUPER score contributes to warning and management in early-stage COVID-19. *Infect Med (Beijing)*. 2023 Oct 19;2(4):308-14. doi: 10.1016/j.imj.2023.09.003.
18. Datta D, George Dalmida S, Martinez L, Newman D, Hashemi J, Khoshgoftaar TM, Shorten C, Sareli C, Eckardt P. Using machine learning to identify patient characteristics to predict mortality of inpatients with COVID-19 in south Florida. *Front Digit Health*. 2023 Jul 28;5:1193467. doi: 10.3389/fdgh.2023.1193467.
19. Yu Z, Li X, Zhao J, Sun S. Identification of hospitalized mortality of patients with COVID-19 by machine learning models based on blood inflammatory cytokines. *Front Public Health*. 2022 Nov 17;10:1001340. doi: 10.3389/fpubh.2022.1001340.
20. Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. *Comput Biol Med*. 2019 Sep;112:103375. doi: 10.1016/j.combiomed.2019.103375.
21. Pereira RB, Plastino A, Zadrozny B, Merschmann LH. Categorizing feature selection methods for multi-label classification. *Artif Intell Rev*. 2018;49(1):57-78. doi: 10.1007/s10462-016-9516-4.
22. Bhargava N, Sharma S, Purohit R, Rathore PS. Prediction of recurrence cancer using J48 algorithm. Presented at: 2nd International Conference on Communication and Electronics Systems (ICCES); 19–20 October, 2017; Coimbatore, India. doi: 10.1109/CESYS.2017.8321306.
23. Yadav S, Shukla S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. Presented at: IEEE 6th International conference on advanced computing (IACC); 27-28 February, 2016; Bhimavaram, India. doi: 10.1109/IACC.2016.25.
24. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123-40. doi: 10.1007/BF00058655.
25. Sun Y, Zhang H, Zhao T, Zou Z, Shen B, Yang L. A new convolutional neural network with random forest method for hydrogen sensor fault diagnosis. *IEEE Access*. 2020;8:85421-30. doi: 10.1109/ACCESS.2020.2992231.
26. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. 2015 Apr 25;27(2):130-5. doi: 10.11919/j.issn.1002-0829.215044.
27. Gao Y, Cai GY, Fang W, Li HY, Wang SY, Chen L, Yu Y, Liu D, Xu S, Cui PF, Zeng SQ, Feng XX, Yu RD, Wang Y, Yuan Y, Jiao XF, Chi JH, Liu JH, Li RY, Zheng X, Song CY, Jin N, Gong WJ, Liu XY, Huang L, Tian X, Li L, Xing H, Ma D, Li CR, Ye F, Gao QL. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun*. 2020 Oct 6;11(1):5033. doi: 10.1038/s41467-020-18684-2.
28. Huang I, Pranata R, Lim MA, Oehadian A, Alisjahbana B. C-reactive protein, procalcitonin, D-dimer, and ferritin in severe coronavirus disease-2019: a meta-analysis. *Ther Adv Respir Dis*. 2020 Jan-Dec;14:1753466620937175. doi: 10.1177/1753466620937175.
29. Sun Q, Qiu H, Huang M, Yang Y. Lower mortality of COVID-19 by early recognition and intervention: experience from Jiangsu Province. *Ann Intensive Care*. 2020 Mar 18;10(1):33. doi: 10.1186/s13613-020-00650-2.
30. Hu H, Yao N, Qiu Y. Predictive value of 5 early warning scores for critical COVID-19 patients. *Disaster Med Public Health Prep*. 2022 Feb;16(1):232-239. doi: 10.1017/dmp.2020.324.
31. Fan G, Tu C, Zhou F, Liu Z, Wang Y, Song B, Gu X, Wang Y, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Wu W, Cao B. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J*. 2020 Sep 10;56(3):2002113. doi: 10.1183/13993003.02113-2020.
32. Elhazmi A, Al-Omari A, Sallam H, Mufti HN, Rabie AA, Alshahrani M, Mady A, Alghamdi A, Altalaq A, Azzam MH, Sindi A, Kharaba A, Al-Aseri ZA, Almekhlafi GA, Tashkandi W, Alajmi SA, Faqihi F, Alharthy A, Al-Tawfiq JA, Melibari RG, Al-Hazzani W, Arabi YM. Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. *J Infect Public Health*. 2022 Jul;15(7):826-34. doi: 10.1016/j.jiph.2022.06.008.
33. Liu H, Wang J, Geng Y, Li K, Wu H, Chen J, Chai X, Li S, Zheng D. Fine-grained assessment of COVID-19 severity based on clinico-radiological data using machine learning. *Int J Environ Res Public Health*. 2022 Aug 26;19(17):10665. doi: 10.3390/ijerph191710665.
34. Kocadagli O, Baygul A, Gokmen N, Incir S, Aktan C. Clinical prognosis evaluation of

- COVID-19 patients: An interpretable hybrid machine learning approach. *Curr Res Transl Med.* 2022 Jan;70(1):103319. doi: 10.1016/j.retram.2021.103319.
35. Yamanaka S, Morikawa K, Azuma H, Yamanaka M, Shimada Y, Wada T, Matano H, Yamada N, Yamamura O, Hayashi H. Machine-learning approaches for predicting the need of oxygen therapy in early-stage COVID-19 in Japan: Multicenter retrospective observational study. *Front Med (Lausanne).* 2022 Feb 23;9:846525. doi: 10.3389/fmed.2022.846525.
36. Bohn MK, Lippi G, Horvath A, Sethi S, Koch D, Ferrari M, Wang CB, Mancini N, Steele S, Adeli K. Molecular, serological, and biochemical diagnosis and monitoring of COVID-19: IFCC taskforce evaluation of the latest evidence. *Clin Chem Lab Med.* 2020 Jun 25;58(7):1037-52. doi: 10.1515/cclm-2020-0722.
37. Polilli E, Sozio F, Frattari A, Persichitti L, Sensi M, Posata R, Di Gregorio M, Sciacca A, Flacco ME, Manzoli L, Di Iorio G, Parruti G. Comparison of monocyte distribution width (MDW) and procalcitonin for early recognition of sepsis. *PLoS One.* 2020 Jan 10;15(1):e0227300. doi: 10.1371/journal.pone.0227300.
38. Agnello L, Vidali M, Lo Sasso B, Giglio RV, Gambino CM, Scazzone C, Ciaccio AM, Bivona G, Ciaccio M. Monocyte distribution width (MDW) as a screening tool for early detecting sepsis: a systematic review and meta-analysis. *Clin Chem Lab Med.* 2022 Feb 15;60(5):786-792. doi: 10.1515/cclm-2021-1331.
39. Piva E, Zuin J, Pelloso M, Tosato F, Fogar P, Plebani M. Monocyte distribution width (MDW) parameter as a sepsis indicator in intensive care units. *Clin Chem Lab Med.* 2021 Mar 5;59(7):1307-14. doi: 10.1515/cclm-2021-0192.
40. Li CH, Seak CJ, Chaou CH, Su TH, Gao SY, Chien CY, Ng CJ. Comparison of the diagnostic accuracy of monocyte distribution width and procalcitonin in sepsis cases in the emergency department: a prospective cohort study. *BMC Infect Dis.* 2022 Jan 4;22(1):26. doi: 10.1186/s12879-021-06999-4.
41. Kim SW, Lee H, Lee SH, Jo SJ, Lee J, Lim J. Usefulness of monocyte distribution width and presepsin for early assessment of disease severity in COVID-19 patients. *Medicine (Baltimore).* 2022 Jul 8;101(27):e29592. doi: 10.1097/MD.00000000000029592.
42. Daorattanachai K, Hirunrut C, Pirompanich P, Weschawalit S, Srivilaithon W. Association of monocyte distribution width with the need for respiratory support in hospitalized COVID-19 patients. *Indian J Crit Care Med.* 2023 May;27(5):352-7. doi: 10.5005/jp-journals-10071-24447.
43. Cordeanu EM, Duthil N, Severac F, Lambach H, Tusch J, Jambert L, Mirea C, Delatte A, Younes W, Frantz AS, Merdji H, Schini-Kerth V, Bilbault P, Ohlmann P, Andres E, Stephan D. Prognostic value of troponin elevation in COVID-19 hospitalized patients. *J Clin Med.* 2020 Dec 17;9(12):4078. doi: 10.3390/jcm9124078.
44. Alhindi T, Awad H, Alfaraj D, Elabdein Salih S, Abdelmoaty M, Muammar A. Troponin levels and the severity of COVID-19 pneumonia. *Cureus.* 2022 Mar 15;14(3):e23193. doi: 10.7759/cureus.23193.
45. Zimmerman A, Kalra D. Usefulness of machine learning in COVID-19 for the detection and prognosis of cardiovascular complications. *Rev Cardiovasc Med.* 2020 Sep 30;21(3):345-52. doi: 10.31083/j.rcm.2020.03.120.



This is an open access article distributed under the terms of [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).