Predicting Severe Disease and Critical Illness on Initial Diagnosis of COVID-19: Simple Triage Tools

Lutfi Ali S. Kurban[1], Sharina AlDhaheri[2], Abdulbaset Elkkari[2], Ramzi Khashkhusha[2], Shaikha AlEissaee[2], Amna AlZaabi[2], Mohamed Ismail[2] and Omran Bakoush[3]*

[1] Department of Radiology, Tawam Hospital, Al Ain, United Arab Emirates, [2] Department of Internal Medicine, Tawam Hospital, Al Ain, United Arab Emirates, [3] Department of Internal Medicine, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

Rationale: This study was conducted to develop, validate, and compare prediction models for severe disease and critical illness among symptomatic patients with confirmed COVID-19.

Methods: For development cohort, 433 symptomatic patients diagnosed with COVID-19 between April 15th 2020 and June 30th, 2020 presented to Tawam Public Hospital, Abu Dhabi, United Arab Emirates were included in this study. Our cohort included both severe and non-severe patients as all cases were admitted for purpose of isolation as per hospital policy. We examined 19 potential predictors of severe disease and critical illness that were recorded at the time of initial assessment. Univariate and multivariate logistic regression analyses were used to construct predictive models. Discrimination was assessed by the area under the receiver operating characteristic curve (AUC). Calibration and goodness of fit of the models were assessed. A cohort of 213 patients assessed at another public hospital in the country during the same period was used to validate the models.

Results: One hundred and eighty-six patients were classified as severe while the remaining 247 were categorized as non-severe. For prediction of progression to severe disease, the three independent predictive factors were age, serum lactate dehydrogenase (LDH) and serum albumin (ALA model). For progression to critical illness, the four independent predictive factors were age, serum LDH, kidney function (eGFR), and serum albumin (ALKA model). The AUC for the ALA and ALKA models were 0.88 (95% CI, 0.86–0.89) and 0.85 (95% CI, 0.83–0.86), respectively. Calibration of the two models showed good fit and the validation cohort showed excellent discrimination, with an AUC of 0.91 (95% CI, 0.83–0.99) for the ALA model and 0.89 (95% CI, 0.80–0.99) for the ALKA model. A free web-based risk calculator was developed.

Conclusions: The ALA and ALKA predictive models were developed and validated based on simple, readily available clinical and laboratory tests assessed at presentation. These models may help frontline clinicians to triage patients for admission or discharge, as well as for early identification of patients at risk of developing critical illness.

Keywords: COVID-19, disease progression, critical illness, prediction model, prognosis

INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a new respiratory infectious disease that was first reported in Wuhan, China and has subsequently spread worldwide. It is estimated that 80% of patients with COVID-19 have mild flu-like symptoms or no symptoms (1). However, up to a quarter of adult patients develop a severe respiratory illness that may progress to respiratory failure with high risk of mortality (2–5). This is causing a significant strain on health services worldwide.

Developing a reliable clinical tool to predict patient outcome at an early stage of the disease would greatly improve the management of patients and also ensure optimal utilization of health care resources. Moreover, early identification of patients with potential critical illness is crucial for early provision of the required supportive care and also appropriate selection for therapeutic trials. Furthermore, recognition of patients with potential non-severe disease is helpful to prompt discharge and alleviate the burden on the healthcare system (2, 3, 5). An ideal triage system in a pandemic setting should be quick, readily available, feasible, and affordable with the latter particularly relevant in resource-constraint setting. Various predictive factors have been evaluated to model disease outcomes in patients with COVID-19 (6–8), however, there are several limitations to the current literature. Firstly, most of these studies are from Chinese, European or American populations, with few or no reports from the Middle East. As the population age structure, socioeconomic status, and prevalence of underlying health comorbidities vary in different population groups, the proposed predicative factors may not be applicable to other populations. This study was conducted in the United Arab Emirates (UAE), which hosts over 200 nationalities, and about 88.5% of the population are expatriates and immigrants from other Middle Eastern and South Asian countries (9). Secondly, a recent critical appraisal and systemic analysis revealed that most published predictive models have significant selection bias and poorly defined predictors and outcomes, as well as lacking calibration and validation, making their generalizability and implementation across different settings and populations questionable (10, 11). Our predictive models were developed in accordance with the TRIPOD statement's recommendations with complete data on disease progression and clearly defined final outcomes (12). Finally, a recent review of the current screening and triage tools for COVID 19 concluded that almost all published validated triage tools rely on resource-intensive laboratory and imaging investigations, limiting their generalizability and utility in low-resource settings (13).

The aim of this study is to predict the risk of progression to severe disease and critical illness in symptomatic adult patients with confirmed COVID-19 infection based exclusively on simple and readily available baseline clinical parameters on initial diagnosis. We developed and validated a three- variable predictive model for severe disease and a four-variable model for critical illness and we compared their diagnostic accuracy with previously published predictive models. To the best of our best knowledge, this is the first study in the Middle East to develop validated predictive models for patients with COVID-19

infection to facilitate early prediction of severe disease and critical illness by measuring clinical and laboratory biomarkers at the time of presentation.

MATERIALS AND METHODS

Participants and Study Design

From the early period of the epidemic, the governmental hospitals in Abu Dhabi adopted an admission policy for all cases of confirmed COVID-19 infection, including non-severe cases, to ensure isolation and also to limit potential viral spread in the community. To identify predictive factors for severe disease and critical illness, we started by reviewing the electronic health records (EHR) of all symptomatic adult patients ($\geq$18 years) with suspected COVID-19 infection at Tawam Hospital, Abu Dhabi Emirate, UAE between 15th of April 2020 and 30th of June 2020. Of the 470 patients with suspected COVID-19, 37 patients had a negative swab and were excluded. Four hundred and thirty-three patients were included in our cohort and all had positive results on real time reverse transcription polymerase chain reaction (RT-PCR) assay of nasopharangeal swab specimens (AllplexTM 2019-nCoV Assay, Seegene Inc., Seoul, Republic of Korea). The results were validated using another cohort of 213 patients presented to another governmental hospital, Al Ain Hospital, Abu Dhabi, UAE during the same period and selected with similar inclusion/exclusion criteria. Our target population and outcomes were clearly defined, and we developed our predictive model using the recommendations and checklist of the TRIPOD statement (12). This study was approved by the Institutional Review Board and the Department of Health Ethical Committee, and written informed consent was waived.

Data Collection

A team of respiratory clinicians reviewed and extracted the data from EHR using a standardized data collection tool. The generated data were checked independently by one clinician to ensure accuracy and consistency. Demographic data, comorbidities, symptoms and signs at presentation, laboratory biomarkers, and outcomes were collected and evaluated. We recorded information on respiratory support (high oxygen supplementation and mechanical ventilation), admission to intensive care unit (ICU), discharge, and death. Comorbidities were identified from the medical history section, previous visits, or at the time of diagnoses. Obesity was defined as a body mass index (BMI) $\geq$ 30 and was calculated as weight in kilograms divided by height in meters squared. The patients' identities were anonymized, and the data were password protected.

Outcomes

We assessed two primary outcomes: severe disease and critical illness. The outcomes were evaluated longitudinally over the entire study period, not just at the time of the initial testing event. Final outcome of either discharge or death was obtained for all patients in our cohort and there were no censored data.

Severe disease was defined according to the Chinese management guideline for COVID-19 (14), version 6.0 as development of any of the following parameters during

admission: respiratory rate $\geq$ 30/min, resting pulse oxygen saturation (SpO2) $\leq$ 93%, partial pressure of oxygen (PaO$_2$)/fraction of inspired oxygen (FiO$_2$) $\leq$ 300 mmHg (1 mmHg = 0.133 kPa).

Critical illness was defined as a composite event of death, septic shock or respiratory failure requiring high-level supplemental oxygen or mechanical ventilation (invasive or non-invasive). Almost all critically ill patients were admitted to the intensive care unit (ICU).

The criteria for discharge were absence of fever for at least 3 days, clinical remission of respiratory symptoms, and two negative RT-PCR swabs for COVID-19 obtained at least 24 h apart.

Potential Predictors

We assessed 19 variables as potential predicators in our models based on the patients' demographic data, comorbidities, and laboratory biomarkers. Demographic variables included age, sex, and ethnicity/race. Comorbidities included obesity (according to the most recent BMI), hypertension, diabetes, cardiovascular disease, pulmonary disease (defined by chronic obstructive pulmonary disease or asthma), cerebrovascular disease (CVD), and chronic kidney disease (CKD). We assessed the laboratory biomarkers that have been reported to be commonly altered according to the recent recommendations in a systematic review and critical appraisal (10).

Laboratory biomarkers included the lymphocyte–neutrophil ratio (LN ratio), red cell distribution width (RDW), serum C-reactive protein (CRP), lactate dehydrogenase (LDH), albumin, D dimer and troponin T. The estimated glomerular filtration rate (eGFR) was calculated using CKD-EPI 2009 (15). We included only the first measurement obtained within 3 days of presentation.

Statistical Analysis

The quantitative variables are reported as median and range. Categorical variables are presented as frequencies and proportions. To explore the association of risk factors with severe disease and critical illness, the Chi square test was used.

Univariate and multivariate logistic regression models were used to determine which predictive factors are associated with disease outcomes. All variables with $p < 0.10$ in the univariate analysis were included in the final multivariable models to determine if each variable is an independent predictor of outcomes. Multiple imputation procedure in IDM SPSS software was used to address the missing laboratory values for RDW (1.4% missing), CRP (2.1%), LDH (2.1%), D dimer (5.3%), eGFR (6.2%), albumin (14.1%), direct bilirubin (30.5%), and troponin T (36.3%). Data were assumed to be missing at random. Five imputations were performed and regression coefficients were pooled using Rubin's rules (16). All the other variables, including the outcome variables, were complete.

The regression analysis results are summarized as relative risk with 95% confidence interval (95% CI) and p-value. $P \leq 0.05$ were considered statistically significant. All analyses were carried out using IBM SPSS software (version 27, Chicago, IL, USA).

Statistical Prediction Models

The predictive models were constructed using predefined protocols according to the TRIPOD guidelines.

Selection of Predictive Factors

We built our prediction models using a logistic regression model for the dichotomous outcomes of interest: disease severity and critical illness. Variables that strongly associated with the outcome of interest during stepwise multivariate analysis were selected for building the prediction model. Covariates that improved the performance of the model were included in the final model. The statistical assumptions of linearity for the model were met.

Discrimination

Discrimination is the predictive accuracy of a model and is a measure of its ability to distinguish between those with and those without the outcome. This was assessed with concordance statistic (C statistic) by measuring the area under the receiver operating characteristic curve (AUC), which ranges from 0.5 to 1. AUC $\geq$ 0.7 was considered acceptable, whereas an AUC between 0.8 and 0.9 indicated excellent diagnostic accuracy.

Calibration

Calibration refers to the agreement between the observed and predicted number of outcomes, and it measures the goodness of fit. The calibration was assessed by plotting the predicted risk of outcome on the x-axis against the observed risk of outcome on the y-axis. Further assessment of goodness of fit was performed by applying the Hosmer–Lemeshow test, which determines if the difference between the observed and predicted risks is significant or not. A $p \leq 0.05$ was considered indicative of model lack of fit.

Validation

To examine the prediction model's reproducibility and generalizability, we validated our predictive models on a validation cohort from Al-Ain hospital, Abu Dhabi, UAE. The predictive variables required to calculate risk score were collected from the validation cohort and the AUC was calculated.

RESULTS

Characteristics of the Patients

Of 470 patients triaged for suspected COVID-19 infection, 433 patients with a positive nasopharyngeal swab RT-PCR were included in the study. The median age was 48 years (range 18–90 years), and 99 patients (22%) were females. The disease was severe in 186 patients (43%). All 247 patients with non-severe disease recovered and were discharged. Of the 186 patients with severe disease, 93 patients developed critical illness, with 75 patients required mechanical ventilation. The other 18 patients received non-invasive ventilation support. The median time from admission to invasive mechanical ventilation was 5 days (range 1–21). The median time from mechanical ventilation to death

FIGURE 1 | Flow diagram of patient selection, categorization, and outcome.
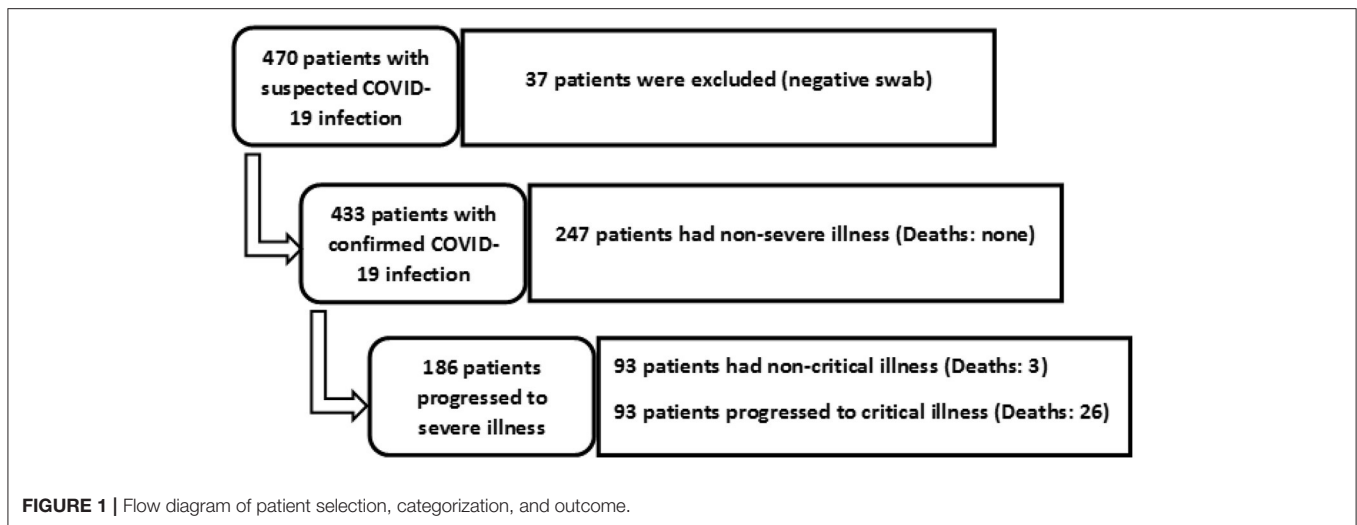
was 7 days (range 1–29 days). **Figure 1** illustrates the recruitment of participants and their final outcomes.

The most prevalent comorbidities were hypertension (34.6%), diabetes mellitus (33.7%), cardiovascular disease (11.3%) and CKD (8.8%). Non-Emirati patients represented 84.1% of the study sample, reflecting the national demographic composition. Asians were the most prevalent ethnic group (65.5%). Arabs were the second most prevalent ethnic group (32.1%) and Africans were the minority (2.4%). BMI was available for 363 patients, 129 (35%) of whom were obese with BMI ≥ 30. Patients with severe disease/ critical illness were more likely to have comorbidities than patients with non-severe illness. The most significant comorbidities were diabetes (40.3 vs. 28.7%), CKD (16.1 vs. 6.5%), and CVD (6.5 vs. 0.8%).

The most common clinical symptoms were fever (84.3%), cough (70.2%), shortness of breath (51.2%), fatigue (39.5%), sore throat (17.3%), and chest pain (17.1%). Median duration of symptoms was 4 days (range 1–15 days).

Predictors of Severe Disease

In univariate analysis, severe disease was significantly associated with increasing age, being male, diabetes, CKD, and CVD, as well as with low LN ratio, serum albumin, and eGFR. Severe disease was also associated with high RDW, CRP, LDH, D-dimer, troponin T, and direct bilirubin. Severe disease was not associated with ethnic origin, hypertension, cardiovascular disease, pulmonary disease, or BMI.

In stepwise multivariate logistic regression analysis, severe disease was significantly associated with increasing age, male gender, and high RDW, LDH, CRP and D dimer, and with low LN ratio and albumin (**Table 1**).

Predictors of Critical Illness

In univariate analysis, critical illness was significantly associated with older age, diabetes, CKD and CVD, with lower LN ratio, albumin and eGFR, and with high RDW, CRP, LDH, D-Dimer, and direct bilirubin. Critical illness was not associated with

gender, ethnic origin, hypertension, CVD, pulmonary disease, BMI, or serum troponin T.

In stepwise multivariate logistic regression analysis, critical illness was significantly associated with older age, being male, and high RDW, LDH, CRP, D dimer, and direct bilirubin, as well as with low LN ratio, albumin, and eGFR (**Table 1**).

Prediction Models of Outcomes
Prediction Model for Severe Disease
Three variables at presentation showed the best discriminating ability in predicting severe disease in patients with COVID-19 infection, namely, age, serum LDH, and serum albumin (ALA model). These three variables had an AUC of 0.881 (95% CI, 0.868–0.893), indicating excellent discrimination (**Figure 2**). The calibration plot indicated that the predicted probabilities matched the actual probabilities, suggesting very good calibration. The *p*-value for the Hosmer-Lemeshow test was 0.375, indicative of good fit.
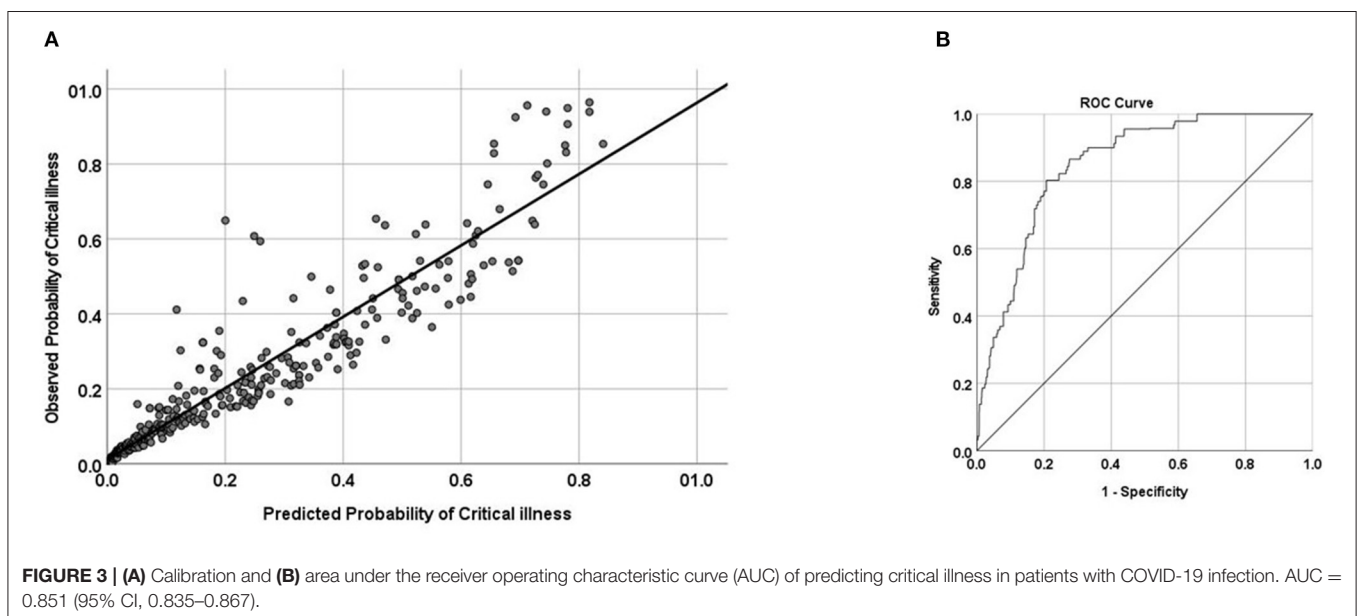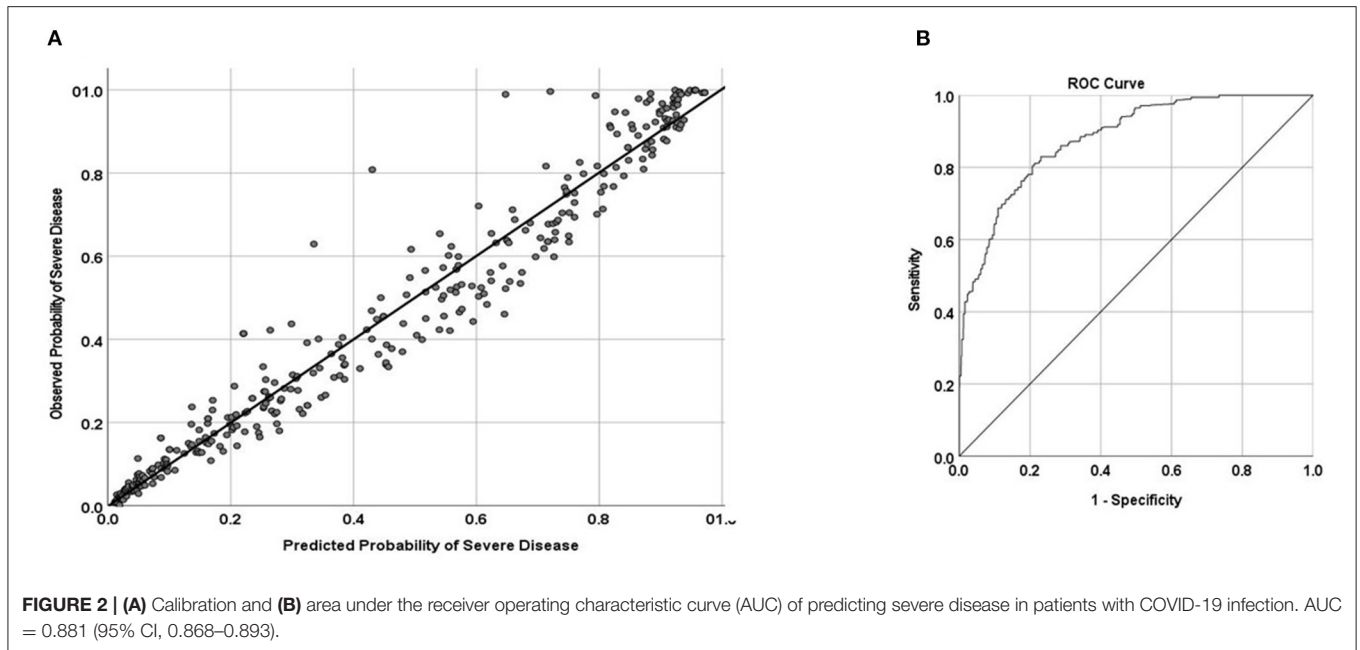
Prediction Model for Critical Illness
In this model, four variables were selected to predict critical COVID-19 illness at presentation. These factors, age, serum LDH, kidney function (eGFR) and serum albumin (ALKA predictive model), showed accurate discrimination, with an AUC of 0.851 (95% CI, 0.835–0.867, **Figure 3**). The calibration plot showed that predicted probabilities were very close to the observed probabilities, consistent with very good calibration. The *p*-value for the Hosmer-Lemeshow test was insignificant at 0.118, in keeping with good fit.

Construction of the Risk Score and Web-Based Calculator
Equations for predicting the risks of severe disease and critical illness in individual patients were constructed based on the fitted risk models. Web-based calculators are available online (https://covidriskscore.000webhostapp.com/).

TABLE 1 | Univariate and multivariate regression analysis of predictive factors of severe COVID-19, critical illness, and in-hospital mortality.

Demographics and comorbidities	Univariate analysis				Multivariate analysis			
	B	S.E.	Sig.	Exp(B) (95% CI)	B	S.E.	Sig.	Exp(B) (95% CI)
Disease severity								
Age	0.033	0.007	0.001	1.033 (1.019–1.048)	0.030	0.008	0.001	1.03 (1.013–1.04)
Sex, Female/Male	0.701	0.245	0.004	2.016 (1.248–3.256)	0.950	0.267	0.001	2.58 (1.530–4.36)
Arabs National	0.056	0.269	0.834	1.058 (0.624–1.793)				
BMI	0.016	0.016	0.330	1.016 (0.984–1.049)				
Hypertension	0.231	0.203	0.256	1.260 (0.845–1.877)				
DM	0.516	0.205	0.012	1.675 (1.120–2.504)	0.063	0.242	0.794	1.06 (0.663–1.71)
CKD	1.021	0.326	0.002	2.776 (1.464–5.265)	0.666	0.375	0.076	1.94 (0.934–4.06)
CVA	2.134	0.770	0.006	8.448 (1.867–38.224)	1.420	0.809	0.079	4.13 (0.848–20.2)
CVD	0.552	0.305	0.07	1.737 (0.955–3.160)	0.108	0.349	0.757	0.89 (0.453–1.77)
Asthma/COPD	−0.614	0.541	0.256	0.541 (0.187–1.563)				
Laboratory biomarkers								
LN ratio	−2.370	0.197	0.001	0.093 (0.063–0.138)	−0.607	0.194	0.002	0.545 (0.372–0.797)
RDW	0.091	0.019	0.001	1.095 (1.055–1.136)	0.108	0.026	0.001	1.114 (1.059–1.172)
CRP	0.013	0.001	0.001	1.013 (1.012–1.015)	0.003	0.001	0.002	1.003 (1.001–1.004)
LDH	0.012	0.001	0.001	1.012 (1.011–1.013)	0.011	0.001	0.001	1.011 (1.010–1.012)
D dimer	0.326	0.032	0.001	1.386 (1.301–1.476)	0.088	0.033	0.007	1.092 (1.025–1.164)
Albumin	−0.217	0.011	0.001	0.805 (0.788–0.822)	−0.129	0.014	0.001	0.879 (0.854–0.904)
eGFR	−0.007	0.001	0.001	0.993 (0.990–0.995)	−0.001	0.002	0.711	0.999 (0.996–1.003)
Direct bilirubin	0.058	0.008	0.001	1.060 (1.044–1.077)	0.013	0.010	0.192	1.013 (0.993–1.034)
Troponin	0.001	0.001	0.041	1.000 (1.000–1.000)	0.001	0.001	0.722	1.000 (1.000–1.000)
Critical Illness								
Age	0.049	0.009	0.001	1.050 (1.032–1.067)	0.039	0.010	0.001	1.039 (1.020–1.060)
Sex, Female/Male	0.316	0.291	0.278	1.371 (0.775–2.425)				
Arabs National	−0.211	0.308	0.493	0.810 (0.443–1.481)				
BMI	0.012	0.019	0.525	1.012 (0.976–1.049)				
Hypertension	0.388	0.237	0.102	1.474 (0.926–2.347)				
DM	0.948	0.237	0.001	2.580 (1.621–4.104)	0.353	0.271	0.193	1.423 (0.837–2.421)
CKD/ESKD	1.147	0.324	0.001	3.148 (1.669–5.938)	0.443	0.378	0.241	1.557 (0.743–3.263)
CVA	1.927	0.571	0.001	6.869 (2.245–21.019)	1.032	0.612	0.091	2.808 (0.847–9.310)
CVD	0.611	0.330	0.064	1.842 (0.966–3.515)	−0.214	0.383	0.577	0.808 (0.381–1.712)
Asthma/COPD	−0.295	0.647	0.648	0.744 (0.209–2.645)				
Laboratory biomarkers								
LN ratio	−3.017	0.294	0.001	0.049 (0.028–0.087)	−0.746	0.267	0.005	0.474 (0.281–0.800)
RDW	0.191	0.021	0.001	1.210 (1.163–1.260)	0.210	0.026	0.001	1.234 (1.173–1.297)
CRP	0.010	0.001	0.001	1.010 (1.009–1.012)	0.004	0.001	0.001	1.004 (1.003–1.006)
LDH	0.006	0.001	0.001	1.006 (1.005–1.007)	0.005	0.001	0.001	1.005 (1.004–1.005)
D dimer	0.293	0.028	0.001	1.340 (1.269–1.414)	0.107	0.026	0.001	1.112 (1.057–1.171)
Albumin	−0.203	0.011	0.001	0.817 (0.799–0.835)	−0.095	0.015	0.001	0.910 (0.884–0.936)
eGFR	−0.015	0.002	0.001	0.985 (0.982–0.988)	−0.008	0.002	0.001	0.992 (0.988–0.995)
Direct bilirubin	0.063	0.009	0.001	1.065 (1.048–1.084)	0.038	0.010	0.001	1.039 (1.018–1.060)
Troponin	0.001	0.001	0.169	1.000 (1.000–1.000)				

FIGURE 2 | (A) Calibration and (B) area under the receiver operating characteristic curve (AUC) of predicting severe disease in patients with COVID-19 infection. AUC = 0.881 (95% CI, 0.868–0.893).

FIGURE 3 | (A) Calibration and (B) area under the receiver operating characteristic curve (AUC) of predicting critical illness in patients with COVID-19 infection. AUC = 0.851 (95% CI, 0.835–0.867).

The prediction equation to assess an individual patient's severe disease risk score $= -0.96 + 0.03$ (age, years) $+ 0.011$ (LDH, U/L) $- 0.13$ (albumin, g/L).

The prediction equation to estimate the critical illness risk score $= -1.69 + 0.039$ (age, years) $+ 0.005$ (LDH, U/L) $- 0.095$ (albumin, g/L) $- 0.008$ (eGFR, ml/min/1.73 m^2).
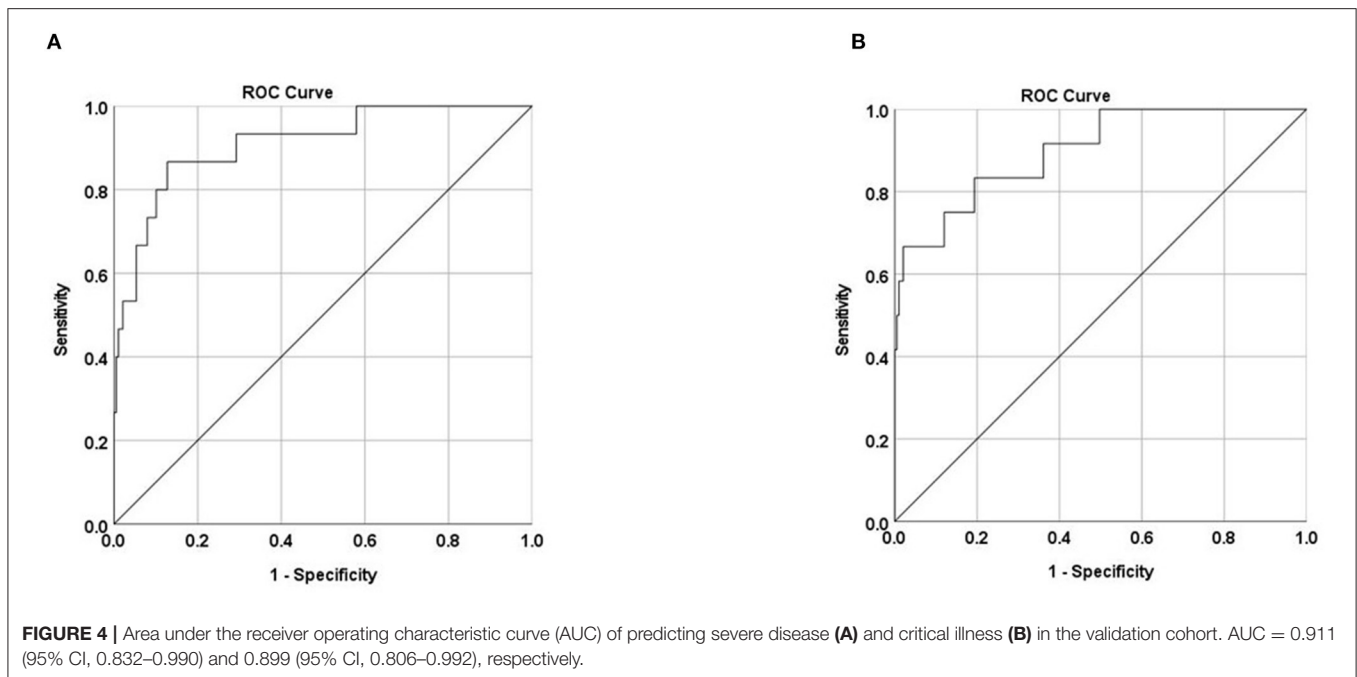
Validation

The validation cohort included 213 patients with median age of 42 years (range 18–82 years), 38 were females (17.8%). Sixteen patients developed severe disease and subsequently progressed to critical illness (7.5%). Of the 16 patients who developed critical illness, 3 patients died (18.7%).

The most prevalent comorbidities were hypertension (22.5%), diabetes mellitus (21.1%), cardiovascular disease (3.7%), and CKD (2%).

The accuracy of the models' prediction of disease severity and critical illness were comparable to the development cohort, with an AUC of 0.911 (95% CI, 0.832–0.990) and 0.899 (95% CI, 0.806–0.992), respectively, indicating excellent discrimination (**Figure 4**).

Comparison With Previously Published Models

Our ALA predictive model showed a high discriminatory performance (AUC, 0.88, 95% CI, 0.868–0.893) which

FIGURE 4 | Area under the receiver operating characteristic curve (AUC) of predicting severe disease **(A)** and critical illness **(B)** in the validation cohort. AUC = 0.911 (95% CI, 0.832–0.990) and 0.899 (95% CI, 0.806–0.992), respectively.

outperformed the exiting HNC-LL score model (AUC, 0.85, 95% CI, 0.82–0.89) and the NLR model (AUC, 0.71, 95% CI, 0.659 to −0.759).

Similarly, the ALKA predictive model demonstrated a higher diagnostic accuracy (AUC, 0.851, 95% CI, 0.835–0.867) compared to the existing predictive model of critical illness published by Zhang et al. (AUC, 0.74, 95% CI, 0.66 to −0.82).

DISCUSSION

We developed and validated models to predict disease outcomes in adult patients with confirmed COVID-19 infection. As the clinical course of the disease is unpredictable, COVID-19 infection continues to pose serious challenges to the healthcare systems globally. Identifying patients who are at risk of developing severe disease will enable informed decisions about admission and discharge as well as proper healthcare resource utilization. Furthermore, early recognition of patients at risk of critical illness and possibly death will improve patient monitoring and timely initiation of the required level of care and support. One of our models' key features is that they utilize simple, readily available clinical parameters that can be assessed at the time of presentation. The development of our predictive models in a unique cohort of patients with both severe and non-severe illness at the time of initial presentation is another major strength.

We found that increasing age is a strong predictor of severe disease and critical illness. During aging, major changes occur in the immune system, collectively described immunosenescence, a condition characterized by declining immune functions (17). These changes include T-cell and B-cell dysfunction as well as excess production of type 2 cytokines, which could lead to

uncontrolled viral replication, proinflammatory responses, and poor outcome (18).

The median age of our cohort was 48 years, which is lower than reported in other studies of American, European, and Chinese populations, reflecting demographic differences between these populations (19–22).

In our cohort, elevated serum LDH and low serum albumin at presentation were strong independent predictors of severe disease and critical illness. In patients with severe infection, elevated levels of LDH are attributed to the release of the intracellular content when the integrity of the cell membrane is compromised, such as in virally induced pulmonary damage. Additionally, LDH levels can be elevated in patients due to cytokine-mediated multiple organ injury (23, 24). In a pooled analysis of 1,532 patients with COVID-19 infection, elevated LDH levels were associated with a six-fold increase in the probability of developing severe disease and a 16-fold increase in the probability of death (24). Similarly, Muhammad et al. reported increased in-hospital mortality in COVID-19 patients with elevated LDH (25).

There are a few hypotheses regarding the cause of hypoalbuminemia in patients with severe COVID-19 infections. First, hypoalbuminemia is a recognized biomarker of acute and chronic inflammation. Moreover, essential amino acid consumption due to viral replication, transcriptional inhibition and albumin clearance might play a role in albumin depletion. Furthermore, reduced albumin synthesis by the liver and increased microvascular permeability and consequent redistribution of albumin into extravascular compartments have also been suggested to contribute to hypoalbuminemia (26). Albumin has strong anti-inflammatory, antioxidant and anticoagulation properties, so hypoalbuminemia

can have prothrombotic effects with an increased risk of arterial and venous thrombosis in different clinical settings (27, 28).

We showed that eGFR on presentation was an independent predictor of critical illness in patients with COVID-19. It has been reported that expression of angiotensin-converting enzyme 2 (ACE2), the major cell entry receptor for SARS-CoV-2, is almost 100 times higher in the kidneys than in the pulmonary parenchyma (29). Acute kidney disease in patients with COVID-19 is caused by several factors, including direct cytopathic effects on kidney tissue, endothelial damage, deposition of immune complexes, and virus-induced cytokines (30). Several studies have shown higher mortality in patients with CKD. Cheng et al. reported that during hospitalization of patients with COVID-19, there is a high prevalence of kidney disease on admission, and that the development of acute kidney injury is frequent and is associated with in-hospital mortality. Guan et al. found that mortality and ICU admission rates of patients with raised serum creatinine were higher than in patients with normal serum creatinine level (9.6 vs. 1%). Lim et al. reported that the development of acute kidney injury is more strongly associated with severe clinical outcomes and mortality in patients with chronic kidney disease (20, 30, 31).

Predictive Model of Severe Disease (ALA Model)

To predict severe disease at presentation in adult patients with COVID-19 infection, we developed a predictive model using three readily available clinical parameters (age, serum LDH, and albumin).

Our cohort included all symptomatic patients with both severe and non-severe COVID-19, as early on in the pandemic, the initial strategy in the country was to hospital isolate all confirmed positive cases to control the spread of the disease. All patients in our cohort who were predicted to have a course of non-severe illness at initial presentation were discharged without developing critical illness. Therefore, the ALA predictive model can be a useful tool for triaging patients with COVID-19 infection for admission or discharge.

The utility of the previously published models on predicting the risk of progression to severe COVID-19 and early identification of patients requiring hospital admission is limited (32–35). The COVID-19 vulnerability index model (CV19) has a serious limitation, as it was developed using proxy events and outcomes from non-COVID-19 pneumonia patients (32). The other predictive model developed by Meng et al. relies on subjective clinical symptoms such as breathlessness and serum biomarkers that are not readily available in the emergency department setting, e.g., serum interleukin 6 (35). We compared the accuracy of our ALA predictive model with other existing predictive models of severe disease. We applied the corresponding variables of the HNC-LL score model, hypertension, neutrophil count, serum CRP, lymphocyte count, and serum LDH, developed by Xiao et al. and the neutrophil-to-lymphocyte ratio (NLR) model developed by Liu et al. to our cohort (33, 34). We found that the discriminatory performance was lower for the HNC-LL model and the NLR model compared to our ALA predictive model. Noteworthy is that the HNC-LL and NLR models have been shown to outperform the CURB-65 and MuLBSTA scoring systems (33, 34).

Predictive Model of Critical Illness (ALKA Model)

In this model, we developed a scoring system to predict critical illness using four readily available parameters: age, serum LDH, kidney function (eGFR), and serum albumin. We would like to emphasize the importance of the timing of measurement in this model, as we focused exclusively on baseline assessment of the predictive variables at presentation. Therefore, this model could become an important tool for clinicians to identify patients, at or soon after admission, who are at risk of developing a critical illness. This is of paramount value in providing early high-level support and early intervention.

Several previous studies have reported risk prediction models of critical illness and in-hospital mortality based on demographic data, clinical findings and laboratory biomarkers assessed at the time of hospital admission or following transfer to ICU. Most of the studies have been from Chinese, European or American populations, with few reports from the Middle East (10, 36–43).

The previously reported predictive models have several limitations. Firstly, up to 10 predictive variables were employed, and some studies included subjective symptoms or laboratory biomarkers that are not readily available (10, 33–35, 39). In our model, we assessed the laboratory markers that have been reported to be commonly altered according to recent recommendations based on systematic review and critical appraisal (10). The predictive variables we selected are simple, readily available, and can be accurately measured.

Secondly, some studies measured their predictors at an inappropriate time, which may have influenced the outcomes (8, 19, 44, 45). Measuring variables not at presentation but later at their peak during ICU admission may lead to a look-ahead bias. As mentioned above, our predictive parameters are part of the baseline assessment at presentation, and the models can provide an early score of critical illness risk while avoiding look-ahead bias.

Thirdly, study participants were often excluded because they did not develop the outcome at the end of the study, which means that the final outcome was not determined. This generates a highly selected study sample (46). In our cohort, the end points and the outcome measures were clearly defined, and all our patients had a final outcome of either discharge or progression to critical illness.

Finally, some of the previously published predictive models lacked calibration and validation. Calibration, which measures the goodness of fit, refers to the agreement between the observed and predicted outcomes. Validation refers to the process of confirming that the model actually achieves its intended purpose, and it ensures the model's reproducibility. We strictly followed TRIPOD guidelines, including discrimination, calibration and validation of the predictive models. Our models showed good fit and were validated with a separate validation cohort from another governmental hospital to ensure generalizability to all

the UAE population, and probably to other nearby countries in the region.

We compared the discriminatory power of our ALKA model with the predictive model published by Zhang et al. which share two common predictive variables, age, and serum LDH with our ALKA model (41). Applying Zhang et al. model's predictive variables to our cohort showed a lower discriminatory power compared to our ALKA model. We were unable to evaluate the discriminatory performance of other previously published predictive models that utilized computed tomography findings, large numbers of parameters, or several uncommonly measured clinical and laboratory biomarkers.

Discussing prognostic triage tool for COVID 19 is not complete without referring to artificial intelligence (AI). There is no doubt that AI has the potential to transform how health care is delivered and it is becoming a more widely tested tool for emergency room triage including patients with COVID-19 infection. Many studies have investigated AI applications for diagnostic and prognostic triage of COVID 19 patients during the pandemic. However, the science behind the AI algorithm is deep and highly complex and it has the potential to suffer from a host of shortcomings, including bias and inapplicability outside of the training domain (47). A recent review of over 400 diagnostic and prognostic AI triage tools for COVID 19 infection highlighted serious methodological and reporting flaws that jeopardize reproducibility, generalizability, and usability for clinical practice (48). The utility of the best performing machine learning-based algorithm model was limited by using large number of unreadily available clinical parameters rendering it impractical in resource-constrained settings (49).

Our study has several limitations. Firstly, it was retrospective and all the data were collected from case records. Therefore, important information such as history of comorbidities might have been missed and some laboratory parameters were not available. To account for missing data, we used multiple imputations. Secondly, we limited our cohort to symptomatic patients with confirmed COVID-19 at presentation, potentially excluding patients with milder illness who did not seek medical advice. Thirdly, the number of patients included in the analysis was small, which may limit the interpretation of our model. Larger prospective studies are required to confirm our findings. Finally, predictive models need to be externally validated in other datasets to determine whether these models are generalizable to other cohorts outside the UAE. Given that our study was done early in the pandemic and the cohort was almost entirely comprised of unvaccinated patients, it would be interesting to test the ALA and ALKA models in vaccinated patients with breakthrough COVID-19 infection, to assess their validity.

CONCLUSIONS

Age, serum LDH and serum albumin are independent predictors of severe disease in patients with COVID-19 infection. In addition to these three predictor factors, eGFR is another strong independent predictor of critical illness. We developed and validated two simple, accurate predictive models (ALA and ALKA) to stratify patients at presentation into non-severe, severe or critical illness. The ALA scoring system may provide frontline clinicians with a useful tool to triage patients for admission and discharge. The ALKA scoring system may allow early identification of patients at risk of developing critical illness and therefore may enable very close monitoring and early active treatment. The predictive parameters in these two models are simple and readily available and can be easily incorporated into EHR as a clinical decision support tool. External validation of these two predictive models using a larger cohort is suggested.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board and the Department of Health Ethical Committee, Abu Dhabi, UAE. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

LK, SAlD, AE, RK, SAlE, AA, MI, and OB contributed to the design and implementation of the research. LK and OB contributed to the analysis of the results and to the writing of the manuscript. All authors discussed the results, commented on the manuscript, contributed to the article, and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

1. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72–314 cases from the Chinese Center for Disease Control and Prevention. *JAMA.* (2020) 323:1239–42. doi: 10.1001/jama.2020.2648

2. Fu L, Wang B, Yuan T, Chen X, Ao Y, Fitzpatrick T, et al. Clinical characteristics of coronavirus disease 2019 (COVID-19)

in China: a systematic review and meta-analysis. *J Infect.* (2020) 80:656–65. doi: 10.1016/j.jinf.2020.03.041

3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5

4. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, 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) 180:1–11. doi: 10.1001/jamainternmed.2020.0994

5. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3

6. Cummings MJ, Baldwin MR, Abrams D, Jacobson SD, Meyer BJ, Balough EM, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet.* (2020) 395:1763–70. doi: 10.1016/S0140-6736(20)31189-2

7. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ.* (2020) 369:m1966. doi: 10.1136/bmj.m1966

8. Shi Y, Yu X, Zhao H, Wang H, Zhao R, Sheng J. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. *Crit Care.* (2020) 24:108. doi: 10.1186/s13054-020-2833-7

9. World Bank. *International Migrant Stock, Total - United Arab Emirates United Nations Population Division, Trends in Total Migrant Stock: 2012 Revision.* (2021). Available online at: https://data.worldbank.org/indicator/SM.POP.TOTL?locations=AE (accessed October 1, 2021).

10. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ.* (2020) 369:m1328. doi: 10.1136/bmj.m1328

11. Hirner S, Pigoga JL, Naidoo AV, Calvello Hynes EJ, Omer YO, Wallis LA, et al. Potential solutions for screening, triage, and severity scoring of suspected COVID-19 positive patients in low-resource settings: a scoping review. *BMJ Open.* (2021) 11:e046130. doi: 10.1136/bmjopen-2020-046130

12. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* (2015) 350:g7594. doi: 10.1136/bmj.g7594

13. Barros LM, Pigoga JL, Chea S, Hansoti B, Hirner S, Papali A, et al. Pragmatic recommendations for identification and triage of patients with COVID-19 disease in low- and middle-income countries. *Am J Trop Med Hyg.* (2021) 104:3–11. doi: 10.4269/ajtmh.20-1064

14. Peng F, Tu L, Yang Y, Hu P, Wang R, Hu Q, et al. Management and treatment of COVID-19: the chinese experience. *Can J Cardiol.* (2020) 36:915–930. doi: 10.1016/j.cjca.2020.04.010

15. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF III, Feldman HI, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* (2009) 150:604–12. doi: 10.7326/0003-4819-150-9-200905050-00006

16. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology.* Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams and Wilkins (2008).

17. Goronzy JJ, Fang F, Cavanagh MM, Qi Q, Weyand CM. Naive T cell maintenance and function in human aging. *J Immunol.* (2015) 194:4073–80. doi: 10.4049/jimmunol.1500046

18. Opal SM, Girard TD, Ely EW. The immunopathogenesis of sepsis in elderly patients. *Clin Infect Dis Suppl.* (2005) 7:S504–12. doi: 10.1086/432007

19. Galloway JB, Norton S, Barker RD, Brookes A, Carey I, Clarke BD, et al. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. *J Infect.* (2020) 81:282–8. doi: 10.1016/j.jinf.2020.05.064

20. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* (2020) 382:1708–20. doi: 10.1056/NEJMoa2002032

21. Lian J, Jin X, Hao S, Cai H, Zhang S, Zheng L, et al. Analysis of epidemiological and clinical features in older patients with coronavirus disease 2019 (COVID-19) outside Wuhan. *Clin Infect Dis.* (2020) 71:740–7. doi: 10.1093/cid/ciaa242

22. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. *JAMA.* (2020) 323:2052–9. doi: 10.1001/jama.2020.6775

23. Zhang T, Chen H, Liang S, Chen D, Zheng C, Zeng C, et al. A non-invasive laboratory panel as a diagnostic and prognostic biomarker for thrombotic microangiopathy: development and application in a Chinese cohort study. *PLoS ONE.* (2014) 9:e111992. doi: 10.1371/journal.pone.0111992

24. Henry BM, Aggarwal G, Wong J, Benoit S, Vikse J, Plebani M, et al. Lactate dehydrogenase levels predict coronavirus disease 2019 (COVID-19) severity and mortality: a pooled analysis. *Am J Emerg Med.* (2020) 38:1722–6. doi: 10.1016/j.ajem.2020.05.073

25. Muhammad R, Ogunti R, Ahmed B, Munawar M, Donaldson S, Sumon M, et al. Clinical characteristics and predictors of mortality in minority patients hospitalized with COVID-19 infection. *J Racial Ethn Health Disparities.* (2021). doi: 10.1007/s40615-020-00961-x. [Epub ahead of print].

26. Ambade V. Biochemical rationale for hypoalbuminemia in COVID-19 patients. *J Med Virol.* (2021) 93:1207–9. doi: 10.1002/jmv.26542

27. Chi G, Gibson CM, Liu Y, Hernandez AF, Hull RD, Cohen AT, et al. Inverse relationship of serum albumin to the risk of venous thromboembolism among acutely ill hospitalized patients: analysis from the APEX trial. *Am J Hematol.* (2019) 94:21–8. doi: 10.1002/ajh.25296

28. Ronit A, Kirkegaard-Klitbo DM, Dohlmann TL, Lundgren J, Sabin CA, Phillips AN, et al. Plasma albumin and incident cardiovascular disease: results from the CGPS and an updated meta-analysis. *Arterioscler Thromb Vasc Biol.* (2020) 40:473–82. doi: 10.1161/ATVBAHA.119.313681

29. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7

30. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int.* (2020) 97:829–38. doi: 10.1016/j.kint.2020.03.005

31. Lim JH, Park SH, Jeon Y, Cho JH, Jung HY, Choi JY, et al. Fatal outcomes of COVID-19 in patients with severe acute kidney injury. *J Clin Med.* (2020) 9:1718. doi: 10.3390/jcm9061718

32. Decaprio D, Gartner J, Burgess T, Garcia K, Kothari S, Sayed S, et al. *Building a COVID-19 Vulnerability Index.* (2020). Available online at: https://ui.adsabs.harvard.edu/abs/2020arXiv200307347D

33. Liu J, Liu Y, Xiang P, Pu L, Xiong H, Li C, et al. Neutrophil-to-lymphocyte ratio predicts critical illness patients with 2019 coronavirus disease in the early stage. *J Transl Med.* (2020) 18:206. doi: 10.1186/s12967-020-02374-0

34. Xiao LS, Zhang WF, Gong MC, Zhang YP, Chen LY, Zhu HB, et al. Development and validation of the HNC-LL score for predicting the severity of coronavirus disease 2019. *EBioMedicine.* (2020) 57:102880. doi: 10.1016/j.ebiom.2020.102880

35. Meng Z, Wang M, Zhao Z, Zhou Y, Wu Y, Guo S, et al. Development and validation of a predictive model for severe COVID-19: a case-control study in China. *Front Med.* (2021) 8:663145. doi: 10.3389/fmed.2021.663145

36. Altschul DJ, Unda SR, Benton J, De La Garza Ramos R, Cezayirli P, Mehler M, et al. A novel severity score to predict inpatient mortality in COVID-19 patients. *Sci Rep.* (2020) 10:16726. doi: 10.1038/s41598-020-73962-9

37. Fumagalli C, Rozzini R, Vannini M, Coccia F, Cesaroni G, Mazzeo F, et al. Clinical risk score to predict in-hospital mortality in COVID-19 patients: a retrospective cohort study. *BMJ Open.* (2020) 10:e040729. doi: 10.1136/bmjopen-2020-040729

38. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C Mortality Score. *BMJ.* (2020) 370:m3339. doi: 10.1136/bmj.m3339

39. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med.* (2020) 180:1081–9. doi: 10.1001/jamainternmed.2020.2033

40. Pan P, Li Y, Xiao Y, Han B, Su L, Su M, et al. Prognostic assessment of COVID-19 in the intensive care unit by machine learning methods: model development and validation. *J Med Internet Res.* (2020) 22:e23128. doi: 10.2196/23128

41. Zhang S, Guo M, Duan L, Wu F, Hu G, Wang Z, et al. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. *Crit Care.* (2020) 24:438. doi: 10.1186/s13054-020-03123-x

42. Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. *PLoS ONE.* (2020) 15:e0236618. doi: 10.1371/journal.pone.0236618

43. Zhang P, He Z, Yu G, Peng D, Feng Y, Ling J, et al. The modified NUTRIC score can be used for nutritional risk assessment as well as prognosis prediction in critically ill COVID-19 patients. *Clin Nutr.* (2021) 40:534–41. doi: 10.1016/j.clnu.2020.05.051

44. Yan L, Zhang, H.-T., Xiao Y, Wang M, Guo Y, et al. Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. *medRxiv. [Preprint].* (2020). doi: 10.1101/2020.02.27.20028027

45. Alkaabi S, Alnuaimi A, Alharbi M, Amari MA, Ganapathy R, Iqbal I, et al. A clinical risk score to predict in-hospital mortality in critically ill patients with COVID-19: a retrospective cohort study. *BMJ Open.* (2021) 11:e048770. doi: 10.1136/bmjopen-2021-048770

46. Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. *medRxiv. [Preprint].* (2020). doi: 10.1101/2020.03.28.20045997

47. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* (2019) 17:195. doi: 10.1186/s12916-019-1426-2

48. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell.* (2021) 3:199–217. doi: 10.1038/s42256-021-00307-0

49. Cobre AF, Stremel DP, Noleto GR, Fachi MM, Surek M, Wiens A, et al. Diagnosis and prediction of COVID-19 severity: can biochemical tests and machine learning be used as prognostic indicators? *Comput Biol Med.* (2021) 134:104531. doi: 10.1016/j.compbiomed.2021.104531