

RESEARCH

Open Access



Prediction models based on machine learning algorithms for COVID-19 severity risk

Hansong Zhang¹, Ying Wang², Yan Xie³, Cuihan Wang^{4,5,6}, Yuqi Ma¹ and Xin Jin^{7,8*}

Abstract

Background The World Health Organization has highlighted the risk of Disease X, urging pandemic preparedness. Coronavirus disease 2019 (COVID-19) could be the first Disease X; therefore, understanding the epidemiological experiences of COVID-19 is crucial while preparing for future similar diseases.

Methods Prediction models for COVID-19 severity risk in hospitalized patients were constructed based on four machine learning algorithms, namely, logistic regression, Cox regression, support vector machine (SVM), and random forest. These models were evaluated for prediction accuracy, area under the curve (AUC), sensitivity, and specificity as well as were interpreted using SHapley Additive exPlanation.

Results Data were collected from 1,485 hospitalized patients across 6 centers, comprising 1,184 patients with severe or critical COVID-19 and 301 patients with nonsevere COVID-19. Among the four models, the SVM model achieved the highest prediction accuracy of 98.45%, with an AUC of 0.994, a sensitivity of 0.989, and a specificity of 0.969. Moreover, oxygenation index (OI), confusion, respiratory rate, and age were found to be predictors of COVID-19 severity risk.

Conclusions SVM could accurately predict COVID-19 severity risk; thus, it can be prioritized as a prediction model. OI is the most critical predictor of COVID-19 severity risk and can serve as the primary and independent evaluation indicator.

Keywords Prediction models, Machine learning algorithms, Severity risk, COVID-19

Introduction

The World Health Organization (WHO) recently warned of a possible outbreak of Disease X - an unpredictable infectious disease caused by unknown pathogens, and the next pandemic is a matter of when, not if [1], and coronavirus disease 2019 (COVID-19) could be the first Disease X. Although avoiding such an outbreak is difficult, the pandemic can indeed be significantly reduced and managed [2]. For this, the WHO has promoted the development of better early warning systems for detecting new diseases and improving the healthcare capacity globally. Among all efforts, machine learning is one of the widely used techniques for disease risk prediction, which can effectively identify early disease onset to reduce severity

*Correspondence:

Xin Jin

jx_123@tju.edu.cn

¹School of Mechanical Engineering, Tianjin University, Tianjin 300350, China

²Department of Nursing, Tianjin First Center Hospital, Tianjin 300196, China

³Department of Liver Transplantation, Tianjin First Center Hospital, Tianjin 300196, China

⁴Tianjin Nankai Hospital, Tianjin Medical University, Tianjin 300000, China

⁵Tianjin Key Laboratory of Acute Abdomen Disease Associated Organ Injury and ITCWM Repair, Tianjin 300000, China

⁶Institute of Integrative Medicine for Acute Abdominal Diseases, Tianjin 300000, China

⁷Medical School of Tianjin University, Tianjin 300072, China

⁸Tianjin Municipal Health Commission, Tianjin 300070, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

rate and subsequently mitigate the overcrowding of medical resources such as intensive care units (ICUs).

The existing prediction models [3–34] for COVID-19 utilized various data sources, focusing on risk factors and mortality or critical illness risks. Several studies construct predictive models based on basic clinical characteristics (e.g., age, gender, comorbidities, etc.). For instance, Abdulaal et al. [3] collected 22 primary clinical features, including demographics, lifestyle, and primary symptoms, to develop a mortality risk prediction model based on an artificial neural network (ANN) for COVID-19 patients upon admission, and identified confusion, dyspnea, and increasing age as the most significant predictors of mortality. Despite showing favorable prediction performance, incorporating laboratory measurements, such as blood biochemistry, may potentially improve the prediction accuracy. For example, Booth et al. [4] utilized 26 serum chemistry laboratory parameters and identified 5 critical clinical variables to construct a support vector machine (SVM) model for predicting mortality risk at least 48 h before death in COVID-19 patients, and identified C-reactive protein and calcium as the most influential laboratory indicators. Moreover, some studies emphasize medical imaging as a major clinical variable for prediction. Meng et al. [5] developed a 3D densely connected convolutional neural network (termed DeCOVID-19-Net) combining chest CT radiomic features and clinical information, enabling non-invasive prediction of short-term mortality based on initial CT scans. It is foreseeable that harmonizing data from different resources could further enhance the prediction performance, which is still to date insufficiently investigated.

Classic statistical methods prioritize clinical transparency: Liang et al. [6] identified 10 independent predictors via screening using LASSO regression and logistic regression (LR) and developed a risk score (COVID-GRAM) that predicted the development of critical illness. Ji et al. [7] proposed the CALL score by assessing the risk factors using Cox regression, which was applied to predict the risk of disease progression in patients with COVID-19. In the realm of deep learning, Shamout et al. [8] proposed a data-driven approach for the automatic prediction of deterioration risk using a deep neural network that learned from chest X-ray images and a gradient boosting model that learned from routine clinical variables. Fu et al. [9] adopted a multimodal AI strategy, integrating early-stage CT imaging and physiological biomarkers to assess severe COVID-19 risk, highlighting the potential of combining diverse data types for enhanced prediction. Mahajan et al. [10–11] developed an ensemble learning-based model to classify patients with infectious diseases like COVID-19 using electronic health records, demonstrating improvements in prediction accuracy compared to traditional models. Deep learning models

might achieve high accuracy but require large datasets with a lack of clinical interpretability. More importantly, model predictions often face challenges in data accessibility, such as the difficulty in acquiring high-quality and sufficient data within a short period. This delay in obtaining necessary data can impede rapid responses in critical situations, such as during the outbreak of infectious diseases.

The presented study aimed to construct prediction models for COVID-19 severity risk by considering the following. For this, large amounts of characteristic data containing primary clinical features and laboratory measurements on hospitalized patients from multiple centers were analyzed. Then, two prediction tasks are proposed based on the speed of clinical feature acquisition time: a rapid and effective prediction model (REPM) and an accurate and comprehensive prediction model (ACPM). REPM refers to the collection of available basic medical variables within a short period of a patient's visit, which is used to construct a model to provide a preliminary risk assessment for clinical decision-making in emergencies. ACPM involves the analysis of all available variables after fully collecting all the characteristics of a patient's examination, which is intended to be used as an auxiliary tool to provide a more accurate risk assessment under conditions of sufficient time. Moreover, Logistic regression (LR), Cox regression, SVM, and random forests (RF) were selected in combination with SHapley Additive explanation (SHAP) for their interpretability. Finally, the predictive performance and interpretable variables showing the highest contribution to risk prediction were examined and analyzed. Through comprehensive data analysis and predictive modeling, this study aims to make meaningful contributions to the field by: (1) systematically analyzing large-scale, multi-center datasets comprising primary clinical characteristics and laboratory indicators to identify key risk factors in COVID-19 patients, thereby offering evidence-based support for clinical decision-making; and (2) developing two risk prediction models tailored to the timing of clinical data availability: REPM for rapid response and ACPM for comprehensive evaluation. This modeling strategy addresses a critical but previously underexplored aspect of risk prediction in existing studies.

Methods

Figure 1 presents a comprehensive research framework to facilitate intuitive understanding of the study, which systematically outlines the analytical pipeline from data processing to model validation.

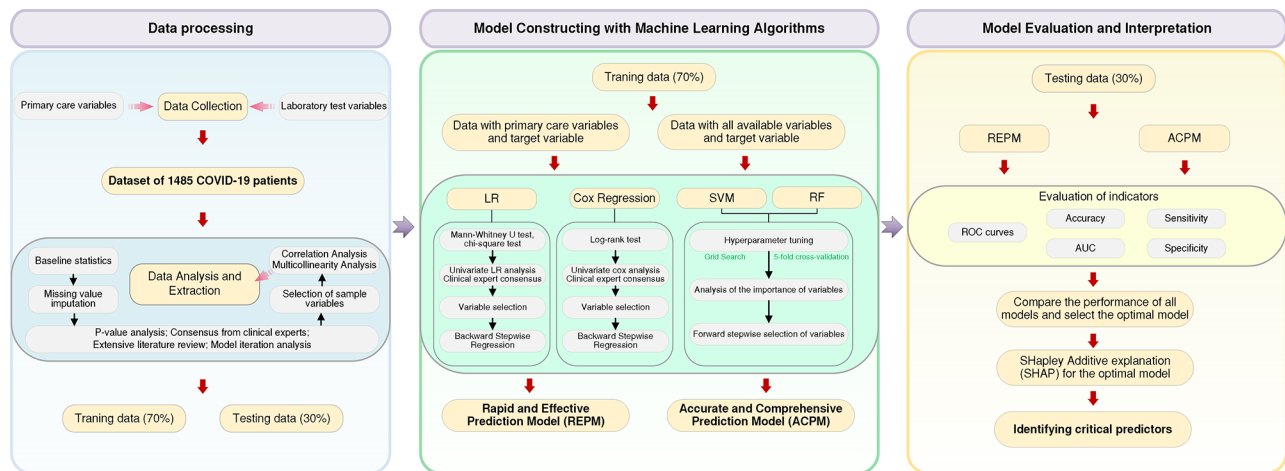


Fig. 1 Framework of the study

Data processing

Study population

Characteristic data were collected from 6 medical institutions in Tianjin City, China, where patients hospitalized from late 2022 to early 2023 were considered. The clinical characteristics included 10 primary care variables, including age, respiratory rate (RR), peripheral capillary oxygen saturation (SpO_2), oxygenation index (OI), systolic blood pressure, diastolic blood pressure (DBP), dehydration, confusion, time from onset of illness to presentation, number of comorbidities (hypertension, diabetes, coronary heart disease, chronic lung disease, oncosis, chronic kidney disease, chronic heart failure, hypohepatia, and renal failure), and 40 laboratory test variables, including CT images (unilateral pneumonia, double pneumonia, pleural effusion, and pleural thickening), infection markers (WBC, NEUT, CRP, ESR, and PCT), nutritional markers (HB, ALB, and five electrolytes), coagulation markers (D-dimer, pleural thickening [PT], and FIB), cardiac function markers (creatinine kinase [CK], CK-MB, α -HBDH, LDH, m-AST, MYO, cTn, and BNP), hepatic function indexes (ALT, TP, ALB, GLB, A/G, TBIL, DBIL, γ -GT, ALP, and AST), and renal function indexes (urea, Cr, and UA). This retrospective study was approved by the ethics committee of Tianjin University Hospital (2024-YLS-125) and was conducted in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was waived because of the retrospective nature of our study. Patients were grouped based on the severity of their condition [35]: (1) $\text{RR} \geq 30$ beats/min, (2) $\text{SpO}_2 \leq 93\%$, (3) $\text{OI} \leq 300$ mmHg, and (4) lung imaging showing notable progression of $> 50\%$ of the lesion within 24–48 h.

Minimum sample size

Determining the sample size for constructing a prediction model is a prerequisite for making the prediction results accurate and reliable. The minimum sample size required for constructing a prediction model can be obtained using the sample size formula of a clinical prediction model [36]. Herein, the model outcome was set as dichotomous, the total number of candidate variables was set at 50, and the proportion of severity was considered incidence of the outcome event, as determined using R software with the *pmsampsize* package.

Analysis of sample variables

Through SPSS, normally distributed continuous variables were compared using a t test, and non-normally distributed continuous variables were compared using a Mann–Whitney U test. Moreover, categorical variables were compared using a χ^2 test. P-values were obtained for differences between the two groups to form a clinical baseline table for patients. Because of the large number and repeatability of variables, it is necessary to screen the variables based on P-value analysis, consensus from clinical experts, extensive literature review, and model iteration analysis.

Machine learning

Different machine learning algorithms, such as LR, Cox regression, SVM, and random forest (RF), were employed to construct an REPM and an ACPM for COVID-19 severity risk. The dataset brought into the model is divided into training and test sets in a ratio of 7:3.

LR was utilized for its interpretability and adaptability to varied data distributions [37], with variables pre-screened by Mann–Whitney U or χ^2 tests (retained at $P < 0.1$). Model parameters optimized via maximum likelihood estimation. Cox regression analyzed time-to-event relationships using log-rank P-values (< 0.1) from

KM curves, adhering to proportional hazards assumptions [38]. As articulated by Muff et al. (2022) [39], conventional null-hypothesis significance testing employing arbitrary P-value thresholds (e.g., $P=0.05$) presents severe limitations. They indicated there was weak evidence that the input variable affects the target variable. Therefore, to retain more potentially clinically relevant variables, the criteria was expanded for multivariate analysis.

SVM handled nonlinear classification through kernel-based high-dimensional mapping, maximizing margin separation between classes [40]. RF generated ensemble predictions via feature-randomized decision trees, reducing overfitting while enabling importance ranking of variables [41]. For the SVM and RF, aiming to improve the prediction accuracy, hyperparameter tuning was performed on the training set by 5-fold cross-validation using the grid search method.

Model evaluation

The ROC curve and the area under the curve (AUC) are used to analyze the involved model performance in terms of predicting the severity in the training and test sets. The metrics selected for model performance evaluation were accuracy, sensitivity, specificity, and AUC.

ROC curves are one of the most commonly used tools for evaluating the performance of classification models, especially for binary classification problems. The relationship between True Positive Rate (TPR, True Positive Rate) and False Positive Rate (FPR, False Positive Rate) is plotted to show the performance of the model under different decision thresholds, which intuitively responds to the discriminative performance of the model. Specificity (denoted as TNR) indicates the proportion of samples that are actually in the negative category that the model

correctly recognizes as negative. The formulas for them are given below:

$$\begin{aligned} TPR &= \frac{TP}{TP + FN}, \\ FPR &= \frac{FP}{FP + TN}, \\ TNR &= \frac{TN}{TN + FP} \end{aligned} \quad (1)$$

Where, TP is the number of true positive samples, FN is the number of false negative samples, FP is the number of false positive samples, TN is the number of true negative samples.

The AUC is calculated by integrating the area under the ROC curve, and its value is between 0 and 1. In medical decision making, when the AUC value is between 0.90 and 1.00, the model has high accuracy and is suitable for triage or critical diagnosis [42, 43].

Predictive accuracy (Accuracy) is calculated by the formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Model interpretation

We used SHAP to interpret the predictions of the machine learning model. SHAP is a robust machine learning explanatory tool that is an extension of the Shapley value in game theory designed to characterize the impact of variables on model output [44]. SHAP can generate a prediction value for each prediction sample, based on which a unique value (SHAP value) is assigned to each influencing factor. This is expressed as a degree to which the influencing factor contributes to the prediction. By comparing the SHAP values of different influencing factors, those that contribute more to the prediction result can be determined. The relation between the influencing factors and the risk of critical illness was analyzed by observing the distribution of SHAP values to better understand the prediction logic of the model and the importance of variables.

Results

Data analysis

Patient characteristics

The characteristics of the 1485 patients at admission are presented in Table 1. Patients with various features indicating severity were categorized into the critical group (1184 patients, 79.7%), whereas the remaining were categorized into the non-critical group (301 patients, 20.3%). The minimum sample size was 1272 ($1485 > 1272$), assuming an expected C index (AUC) of 0.9. Thus, the

Table 1 Clinical characteristics of patients on admission

	Variables	Selection Methods
1	Age	P-values/ Clinical expert consensus
2	RR	Clinical expert consensus
3	OI	P-values/ Clinical expert consensus
4	Confusion	P-values
5	DBP	Iterative analysis/ Clinical expert consensus (blood pressure)
6	Time	Literature review
7	Number of Comorbidities	Clinical expert consensus
8	Pleural Thickening (CT)	Iterative analysis
9	CRP	Literature review/ Clinical expert consensus
10	ALB	Literature review/ Clinical expert consensus
11	D-dimer	Literature review
12	CK-MB	P-values
13	γ -GT	P-values
14	Cr	Iterative analysis

obtained sample size was adequate, thereby reducing model overfitting and enhancing its robustness.

The overall patient age distribution ranged from 9 to 98 years, with a median of 71 years, and 54.54% of the patients were aged >70 years; this is a reasonable distribution of data. OI in the severe group was much lower than that in the nonsevere group (221.9 vs. 344.8). Approximately 70% of all patients had hypertensive comorbidities. Compared with nonsevere patients, a larger proportion of severe patients had diabetic comorbidities (46.71% vs. 38.54%). The most common CT imaging manifestation was double pneumonia (80.94%), followed by PT (59.46%).

Selection of sample variables

In this study, a total of 50 clinical variables were collected and analyzed, in which some of them were measured for the same category of clinical parameters (e.g., SBP and DBP for blood pressure). Despite no correlation between them in the dataset, a rigorous variable selection process was implemented within each clinically coherent category to avoid potential information redundancy stemming from shared pathophysiological interpretations.

In the variable selection process for each category, we employed four methods: (1) P-value analysis, (2) consensus from clinical experts, (3) extensive literature review, and (4) model iteration analysis. Firstly, five variables were screened based on the inter-group difference P-values ($P < 0.05$). Secondly, based on experienced clinical experts' suggestions, two clinically relevant variables, e.g., RR and the number of comorbidities, were retained. Additionally, a thorough literature review supported the selection of the retention of the time from onset of illness to presentation [45–47], CRP [4, 14, 25] for infection status, ALB [4, 14] for nutrition, and D-dimer [22, 27, 28] for coagulation function. Finally, blood pressure, CT imaging, and renal function were examined by model iterative analysis, for controlling confounding factors. These variables were further filtered by a stepwise backward regression method, in which less-contributed variables were removed, retaining three variables: DBP (blood pressure indicator), PT (CT imaging manifestation), and Cr (renal function indicator). Overall, 14 representative variables (Table 1) were selected. The variable selection procedure combines both statistical and machine learning evaluation, prior knowledge from experts, and literature review, offering a robust set of variables for building the predicting model.

For these 14 variables, we conducted a correlation analysis to confirm that these variables are independent and diagnosed for multicollinearity. Nevertheless, owing to the complexity of laboratory tests, some biochemical indicators were missing. Thus, we need to populate this part of the data. These missing values for the variables

with <10% missing data were replaced by the median of the series, and missing values for the variables with >10% missing data were filled through multiple interpolations by chained equations. The missing variables shown in Table 2 are as follows: the data of 19 cases (1.3%) were missing for CRP, 85 cases (5.7%) were missing for ALB, 61 (4.1%) cases were missing for D-dimer, 90 (6%) cases were missing for CK-MB, and 18 cases (1.2%) were missing for Cr. The missing rate of data for the above five variables was less than 10%, so the median of the data for each variable was filled in the missing values. While γ -GT was missing 159 cases (10.7%) of data, the gaps were filled using multiple interpolation to reduce bias in the imputation (generating 5 complete datasets and aggregating the results by the MICE algorithm).

Then, 14 variables were analyzed for correlation; the heat map of correlation analysis is presented in Fig. 2. The highest correlation coefficient between the variables was 0.23 (time and RR), showing a weak correlation; the remaining variables were either weakly correlated or not correlated with each other. Thus, these 14 variables exhibited no strong correlation with each other and can therefore be substituted into the involved model as independent variables. Subsequently, 14 variables were analyzed for multiple covariance and included a VIF report (See Supplementary Table 1, Additional File 1). The tolerances were all close to 1, and the VIFs were all <5, indicating low multicollinearity of the variables.

Prediction model construction

The dataset of 1,485 patients was divided into 2 complementary subsets: training dataset containing 1,033 cases for model construction and test dataset containing 452 cases for validation analysis.

LR for prediction modeling

Table 3 identifies the differences between the two groups of variables ($P < 0.1$ indicates a significant difference). Based on P-values and the importance of variables determined by clinical expert experience, variables such as time, PT (CT), D-dimer, γ -GT, and Cr were excluded. Univariate LR was performed on the remaining variables (Table 2), and the P-values of DBP, number of comorbidities, CRP, and ALB were >0.1. However, because of the clinical importance of the above variables, they were still incorporated in the multivariate stepwise backward LR. The detection of covariance for the retained variables revealed that their VIF was <5, indicating that there is no multicollinearity problem. For the REPM, only primary care variables were used. Thus, the final model included variables such as OI ($P < 0.001$, $\beta = -0.096$, and OR (95% confidence interval [CI]) of 0.908 (0.892–0.926)), confusion ($P < 0.1$, $\beta = 4.064$, and OR (95% CI) of 58.217 (5.127–660.998)), and RR ($P < 0.1$, $\beta = 0.095$, and OR (95% CI)

Table 2 Clinical characteristics of patients on admission by primary outcome

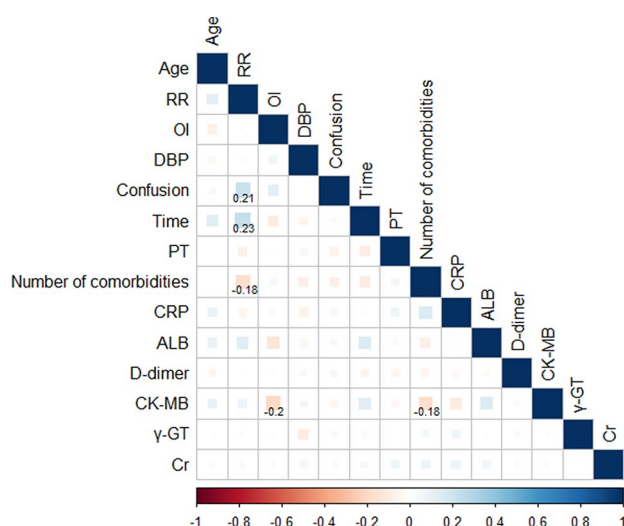
Variables		Missing	Overall (N = 1485)	Nonsevere group (n = 301)	Severe group (n = 1184)	P
Age ≥ 70 years			71 (60–81)	69 (59–79)	72 (61–81.00)	0.012
RR			21 (18–23)	20 (18–22)	21 (18–23)	0.171
SpO ₂ %			98 (97–99)	98 (97–99)	98 (97–99)	0.94
OI			238.97 (197.78–290.34)	344.83 (318.62–409.66)	221.86 (189.72–256.76)	< 0.001
SBP			141 (124–161)	137 (122–157)	142 (124–162)	0.073
DBP			80 (67–92)	77 (68–90)	80 (67–92)	0.273
Dehydration			19 (1.28)	3 (1.00)	16 (1.35)	0.84
Confusion			93 (6.26)	6 (1.99)	87 (7.34)	< 0.001
Time			7 (3–12)	7 (4–14)	7 (3–11)	0.127
Number of comorbidities			2 (2)	2 (2)	2 (2)	0.201
Hypertension			1065 (71.72)	215 (71.10)	851 (71.88)	0.789
Diabetes			669 (45.05)	116 (38.54)	553 (46.71)	0.011
Coronary heart disease			462 (31.11)	103 (34.22)	359 (30.32)	0.192
Chronic lung disease			130 (8.75)	28 (9.30)	102 (8.61)	0.706
Oncosis			152 (10.24)	22 (7.31)	130 (10.98)	0.06
Chronic kidney disease			340 (22.90)	63 (20.93)	277 (23.39)	0.363
Chronic heart failure			95 (6.40)	19 (6.31)	76 (6.42)	0.946
Hypohepatia			261 (17.58)	48 (15.95)	213 (17.99)	0.406
Renal failure			231 (15.56)	45 (14.95)	1816 (15.71)	0.746
CT	UP		248 (16.7)	52 (17.28)	196 (16.55)	0.764
	DP		1202 (80.94)	242 (80.40)	960 (81.08)	0.788
	PE		322 (21.68)	63 (20.93)	259 (21.86)	0.772
	PT		883 (59.46)	161 (53.49)	722 (60.998)	0.718
Indicators of infection	WBC	27	7.75 (5.78–9.27)	7.79 (5.65–9.37)	7.72 (5.86–9.26)	0.845
	NEUT	15	7.30 (4.70–9.95)	6.39 (4.17–8.89)	7.42 (4.74–12.80)	< 0.001
	CRP	19	11.44 (4.56–64.53)	12.03 (5.19–55.64)	10.90 (4.20–66.48)	0.985
	ESR	456	17.00 (9.00–24.00)	18.00 (10.20–25.00)	16.00 (8.00–24.00)	0.021
	PCT	69	0.06 (0.04–0.22)	0.06 (0.04–0.22)	0.06 (0.04–0.22)	0.96
Nutritional indicators	Hb	9	120.00 (98.00–134.00)	119.00 (98.00–133.00)	120.00 (98.00–134.75)	0.681
	ALB	85	35.20 (32.10–38.10)	35.1 (32.36–38.06)	35.30 (32.10–38.10)	0.85
	K	4	4.30 (4.00–4.70)	4.29 (3.90–4.59)	4.33 (4.00–4.70)	0.028
	Na	6	139.0 (135.40–142.00)	139.00 (135.95–141.90)	139.00 (135.40–142.00)	0.89
	Cl	8	105.84 (100.74–108.66)	105.55 (100.74–108.66)	106.15 (100.74–108.66)	0.686
	CO ₂ CP	213	25.90 (24.00–28.10)	25.55 (23.72–27.80)	26.00 (24.10–28.20)	0.097
	AG	63	9.60 (8.70–12.00)	9.70 (8.80–12.04)	9.60 (8.70–12.00)	0.79
Coagulation indicators	D-dimer	61	2.12 (1.16–3.77)	2.12 (1.07–3.89)	2.12 (1.18–3.77)	0.79
	PT	46	14.02 (12.30–16.67)	13.50 (12.00–15.72)	14.19 (12.40–16.88)	< 0.001
	FIB	47	3.82 (2.91–4.65)	3.82 (2.76–4.68)	3.83 (2.92–4.65)	0.879
Cardiac indicators	CK	36	50.40 (28.90–78.55)	52.4 (31.88–83.70)	49.0 (28.70–76.78)	0.104
	CK-MB	90	1.54 (0.82–3.54)	2.11 (1.00–8.69)	1.44 (0.80–3.02)	< 0.001
	α-HBDH	298	179.70 (144.00–244.00)	173.00 (140.00–234.45)	181.10 (145.30–247.90)	0.024
	LDH	41	246.50 (190.07–318.90)	246.00 (190.00–296.10)	246.50 (190.30–325.26)	0.175
	m-AST	344	6.60 (4.20–10.40)	6.40 (4.10–11.00)	6.60 (4.20–10.21)	0.899
	MYO	135	44.40 (25.71–79.26)	40.00 (24.00–79.24)	45.54 (26.21–79.23)	0.358
	cTn	135	9.41 (0.05–24.21)	8.79 (0.02–21.59)	9.56 (0.07–24.82)	0.06
	BNP	260	515.70 (133.38–1321.00)	423.50 (148.45–985.02)	563.60 (132.50–1398.25)	0.21

Table 2 (continued)

Variables		Missing	Overall (N = 1485)	Nonsevere group (n = 301)	Severe group (n = 1184)	P
Liver function indicators	ALT	13	24.70 (11.9–45.7)	23.55 (11.4–41.3)	25 (47.0–11.9)	0.168
	TP	20	62.30 (57.90–66.93)	61.95 (58.32–67.00)	62.30 (57.80–66.90)	0.883
	ALB	17	35.70 (32.60–38.80)	35.70 (33.20–38.63)	35.70 (32.50–38.80)	0.637
	GLB	25	27.40 (23.90–30.90)	27.10 (23.62–30.20)	27.44 (24.00–31.00)	0.216
	A/G	57	1.44 (1.21–1.68)	1.43 (1.21–1.65)	1.44 (1.22–1.68)	0.679
	TBIL	25	8.82 (7.12–10.75)	9.00 (7.20–12.32)	8.80 (7.10–10.20)	0.477
	DBIL	170	4.01 (3.07–5.22)	3.74 (2.99–4.97)	4.03 (3.07–5.22)	0.076
	γ-GT	159	33.11 (20.06–44.23)	32.00 (18.63–40.64)	34.85 (22.44–47.00)	0.005
	ALP	170	71.50 (59.49–86.12)	68.83 (57.25–84.44)	72.24 (60.04–86.12)	0.127
	AST	17	26.02 (16.90–41.00)	23.00 (16.20–38.20)	26.69 (17.20–42.29)	0.028
Renal function indicators	Urea	17	7.85 (5.69–9.51)	7.61 (5.32–9.44)	7.9 (5.71–9.55)	0.131
	Cr	18	77.55 (61.52–95.55)	76.00 (60.65–94.20)	78.00 (61.75–95.84)	0.212
	UA	20	265.42 (205.05–348.31)	268 (205.05–340.92)	265.42 (205.05–353.03)	0.637

Continuous variables (non-normal) are shown as medians (interquartile range, IQR), and categorical variables are presented as n (%)

Abbreviations: RR: respiratory rate; SpO₂: oxygen saturation; OI: oxygenation index; SBP: systolic blood pressure; DBP: diastolic blood pressure; UP: unilateral pneumonia; DP: double pneumonia; PE: pleural effusion; PT: pleural thickening; WBC: white blood cell; NEUT: neutrophil; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; PCT: procalcitonin; Hb: hemoglobin; ALB: albumin; CO₂CP: CO₂ combining power; AG: anion gap; PT: prothrombin time; FIB: fibrinogen; CK: creatine kinase; CK-MB: creatine kinase MB; α-HBDH: α-hydroxybutyric dehydrogenase; LDH, lactate dehydrogenase; m-AST: mitochondrial aspartate aminotransferase; MYO: myohemoglobin; cTn: cardiac troponin; BNP: brain natriuretic peptide; ALT: alanine aminotransferase; TP: total protein; GLB: globulin; A/G: ratio of albumin to globulin; TBIL: total bilirubin; DBIL: direct bilirubin; γ-GT: γ-glutamyl transpeptidase; ALP: alkaline phosphatase; AST: aspartate amino transferase; Cr: creatinine; UA: uric acid

**Fig. 2** Heatmap of correlation between 14 variables

of 1.100 (0.992–1.220)). The model was evaluated using the ROC curves, yielding an AUC of 0.994, a probability threshold of 0.5 (sensitivity=0.989; specificity=0.933), and a prediction accuracy of 97.77% for the training set and an AUC of 0.994 (sensitivity=0.986; specificity=0.967) and a prediction accuracy of 98.23% for the test set. For the ACPM, all variables were considered, yielding the same results as the REPM.

Cox regression for prediction modeling

According to Table 3, variables with P-values of <0.1 were retained. However, DBP ($P>0.1$), the number of comorbidities ($P>0.1$) were retained considering their

clinical importance. Univariate Cox regression was performed using these variables (Table 4); the P-values of age, the number of comorbidities and CRP were all >0.1. However, considering their importance, they were incorporated in the multivariate stepwise backward Cox regression analysis. The detection of covariance for the retained variables yielded a VIF of <5, without multicollinearity. For the REPM, only the primary care variables were incorporated. Thus, the final model contained OI ($P<0.001$, $\beta = -0.006$, and OR (95% CI) of 0.994 (0.993–0.995)), confusion ($P<0.001$, $\beta = 0.588$, and OR (95% CI) of 1.801 (1.374–2.360)), DBP ($P<0.1$, $\beta = 0.005$, and OR (95% CI) 1.005 (1.001–1.010)), and number of comorbidities ($P<0.1$, $\beta = -0.051$, and OR (95% CI) 0.951 (0.907–0.996)). Model evaluation using ROC curves yielded an AUC of 0.707, a probability threshold of 0.13 (sensitivity=0.903; specificity=0.267), and a prediction accuracy of 76.48% for the training set and an AUC of 0.742 (sensitivity=0.903; specificity=0.297) and a prediction accuracy of 80.31% for the test set.

For the ACPM, all variables were incorporated. The final model contained OI ($P<0.001$, $\beta = -0.006$, OR (95% CI) of 0.994 (0.993–0.995)); DBP ($P<0.1$, $\beta = 0.005$, and OR (95% CI) of 1.005 (1.001–1.010)); confusion ($P<0.1$, $\beta = 0.667$, and OR (95% CI) 1.948 (1.482–2.560)); number of comorbidities ($P<0.1$, $\beta = -0.007$, and OR (95% CI) 0.993 (0.992–0.994)); ALB ($P<0.1$, $\beta = 0.012$, and OR (95% CI) 1.012 (1.002–1.022)); and Cr ($P<0.1$, $\beta = -0.001$, and OR (95% CI) of 1.012 (1.002–1.022)). The model was evaluated using the ROC curves, yielding an AUC of 0.710, a probability threshold of 0.12 (sensitivity=0.909;

Table 3 P-values for statistical analysis and univariate LR analysis

	Predictors	P	Univariate P	OR (95% CI)
Primary care variables	Age	0.059	0.069	1.009 (0.999–1.018)
	RR	0.376	0.027	1.046 (1.005–1.088)
	OI	< 0.001	< 0.001	0.915 (0.900–0.930)
	DBP	0.435	0.581	1.003 (0.993–1.013)
	Confusion	0.001	0.004	7.885 (1.910–32.553)
	Time	0.115		
	Number of comorbidities	0.286	0.157	1.080 (0.971–1.202)
Laboratory test variables	PT (CT)	0.216		
	CRP	0.497	0.492	1.001 (0.998–1.004)
	ALB	0.661	0.812	0.999 (0.987–1.010)
	D-dimer	0.970		
	CK-MB	< 0.001	0.003	0.969 (0.949–0.990)
	γ-GT	0.105		
	Cr	0.296		

Table 4 P-values for log-rank test and univariate Cox regression analysis

	Predictors	P	Univariate P	OR (95% CI)
Primary care variables	Age	< 0.001	0.385	1.002 (0.998–1.007)
	RR	< 0.001	0.003	1.022 (1.008–1.037)
	OI	< 0.001	< 0.001	0.994 (0.993–0.995)
	DBP	0.102	0.045	1.005 (1.000–1.009)
	Confusion	< 0.001	< 0.001	2.005 (1.530–2.626)
	Number of comorbidities	0.152	0.222	0.971 (0.926–1.018)
	PT (CT)	0.357		
Laboratory test variables	CRP	< 0.001	0.301	1.001 (0.999–1.002)
	ALB	< 0.001	0.093	1.004 (0.999–1.008)
	D-dimer	< 0.001	0.623	1.000 (1.000–1.000)
	CK-MB	< 0.001	0.966	1.000 (0.989–1.011)
	γ-GT	0.001	0.345	1.000 (0.999–1.001)
	Cr	< 0.001	0.307	1.000 (0.999–1.001)

specificity=0.281) and a prediction accuracy of 78.21% for the training set and an AUC of 0.742 (sensitivity=0.936; specificity=0.275) and a prediction accuracy of 81.19% for the test set.

SVM for prediction modeling

The hyperparameters for SVM were selected from: ‘C’: [0.1, 1, 10, 15, 100, 1000], ‘gamma’: [0.001, 0.01, 0.1], ‘kernel’: [‘sigmoid’, ‘poly’, ‘linear’, ‘rbf’]. The optimal configuration of hyperparameters for the SVM has been identified as follows (Fig. 3): the kernel function is RBF; gamma=0.1; C=15.

The REPM was based on SVM, into which primary care variables were incorporated in their order of importance (Fig. 4a; Table 5). After validating the test set, four variables were input into the model, namely OI, confusion, age, and RR. The model achieved the highest accuracy of 98.45%, with an AUC of 0.994 and a probability threshold of 0.5 (sensitivity=0.989; specificity=0.967). Similarly, the variables were sequentially incorporated into the ACPM in their order of importance (Fig. 4b; Table 6). When variables included OI, confusion, and

the number of comorbidities, the model achieved the highest accuracy of 98.01%, with an AUC value of 0.994 and a probability threshold of 0.5 (sensitivity=0.983; specificity=0.967).

RF for prediction modeling

The number of trees is 200, specified by the parameter n_estimators=200. The hyperparameters were selected from: ‘max leaf nodes’: [50, 100, 200, 1000], ‘max depth’: [10, 20, 30, 50], ‘min samples split’: [5, 10, 20]. The optimal configuration of hyperparameters for the RF has been identified as follows (Fig. 5): the max leaf nodes is 100; the max depth is 10; and min samples split is 5.

The REPM was based on RF, and the primary care variables were incorporated into this model in their order of importance (Fig. 6a; Table 7). After validating the test set, when only one variable, OI, was input into the model, the model achieved the highest accuracy of 98.01%, with an AUC value of 0.991 and a probability threshold of 0.5 (sensitivity=0.986; specificity=0.956). Similarly, all variables were incorporated into the ACPM in the order of their importance (Fig. 6b; Table 8). This model achieved

[illegible]

Table 6 Predictions of different clinical variables selected by the SVM-based ACPM

Number of variables	OI	Confusion	Number of comorbidities	Age	CRP	ALB	γ-GT	DBP	CK-MB	Cr	Time	RR	D-dimer	PT (CT)	AUC	Accuracy
1	✓														0.994	97.35%
2	✓	✓													0.994	97.79%
3	✓	✓	✓												0.994	98.01%
4	✓	✓	✓	✓											0.994	97.79%
5	✓	✓	✓	✓	✓										0.994	97.57%
6	✓	✓	✓	✓	✓	✓									0.994	97.35%
7	✓	✓	✓	✓	✓	✓	✓								0.994	97.79%
8	✓	✓	✓	✓	✓	✓	✓	✓							0.994	97.57%
9	✓	✓	✓	✓	✓	✓	✓	✓	✓						0.994	96.9%
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					0.994	97.12%
11	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				0.994	97.79%
12	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			0.994	97.79%
13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		0.994	97.79%
14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0.994	98.01%

the highest accuracy of 98.01%, consistent with that of the REPM.

Comparison of model results

Accuracy and differentiation

Table 9 presents a comparison of results of the four machine learning methods. For the REPM, the prediction accuracy of SVM is the highest at 98.45%, followed by those of LR and RF at >98%. In contrast, Cox regression yielded a relatively low accuracy of 80.13%. The AUC values of SVM, LR, and RF are >0.99; these three models' differentiation is relatively good, and the AUC of Cox regression was relatively lower. SVM, LR and RF are all excellent with sensitivities and specificities above 0.95. Cox regression had poor specificity and weak ability to differentiate between nonsevere cases. For the ACPM, the prediction accuracy (98.23%) of LR is the highest, followed by those of SVM, RF, and Cox regression. SVM has the same AUC as LR at 0.994, followed by RF (0.991) and Cox regression (0.742). Both sensitivities and specificities of the SVM, LR, and RF models exceed the 0.95 threshold. In contrast, the Cox regression model exhibited low specificity and was less effective in distinguishing between nonsevere cases. Overall, the REPM based on SVM has the best prediction accuracy and differentiation, whereas those based on LR and RF also show better performance. By comparison, Cox regression has lower accuracy and AUC. As for the worse performance of Cox regression, the Schoenfeld residuals were employed to test if the proportional risk assumption is satisfied, yielding a p-value of <0.05 for the residuals of each covariate with time correlation (See Supplementary Table 2, Additional File 1). The proportional risk assumption was not fully met by the model and because cox regression is mainly used for survival analysis, which may be ineffective in comparison to classifiers such as the SVM.

To assess potential bias due to sample imbalance, the confusion matrix (See Supplementary Fig. 1, Additional File 1) of the SVM model was used, showing a high accuracy of 92.38% in predicting non-critical patients during training and 96.7% in testing, indicating strong performance across both classes without significant impact from class imbalance.

REPM and ACPM

The ROC curves of the four machine learning methods used for the REPM and ACPM were compared (Fig. 7a and b, respectively). The ROC curves were slightly different; however, the AUC values were identical. Therefore, the accuracy rates of these models were compared. The models based on LR contained the same variables and therefore exhibited the same accuracy. Based on Cox regression analysis, the accuracy of the ACPM (81.19%) was higher than that of the REPM (80.31%). However,

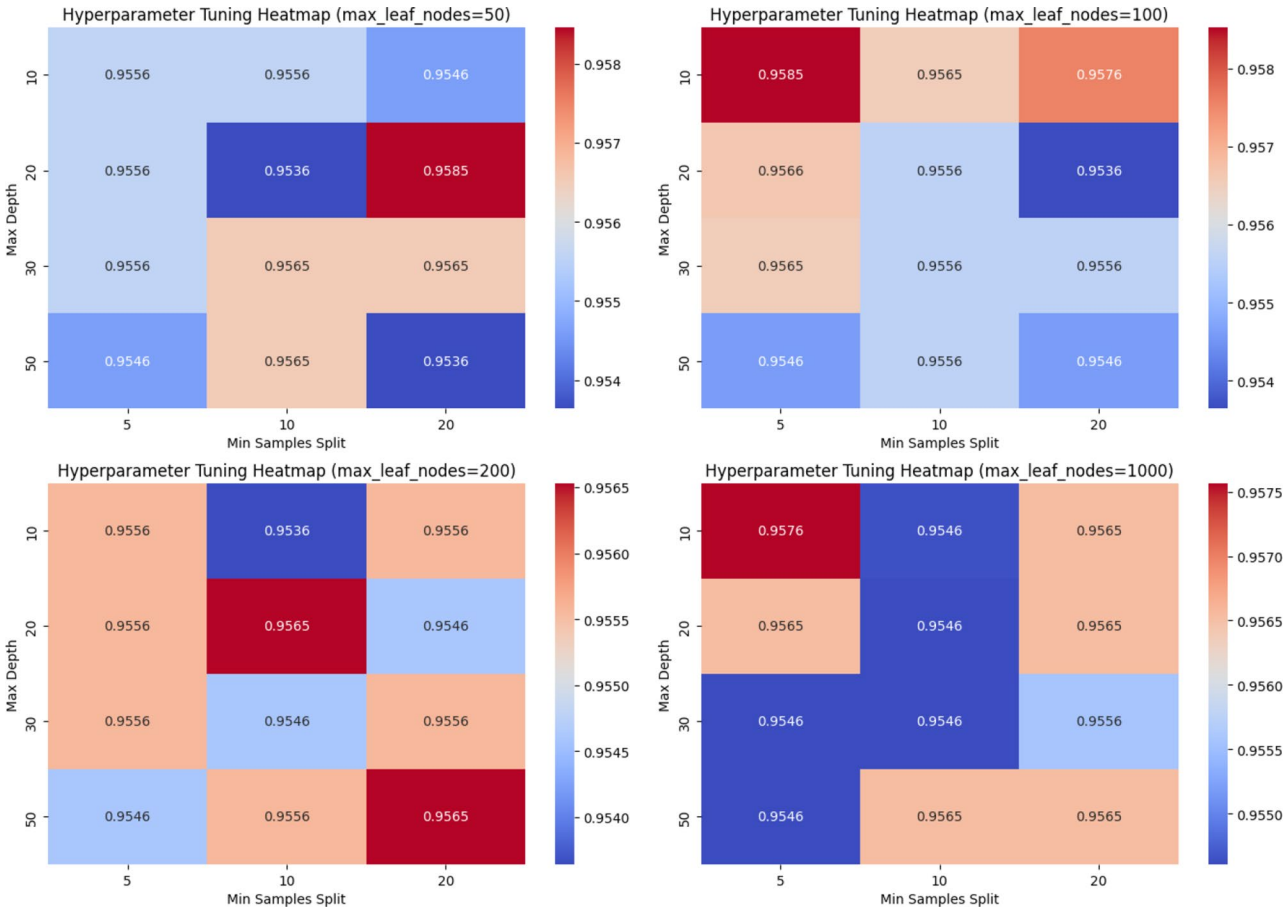


Fig. 5 Hyperparameter tuning for RF

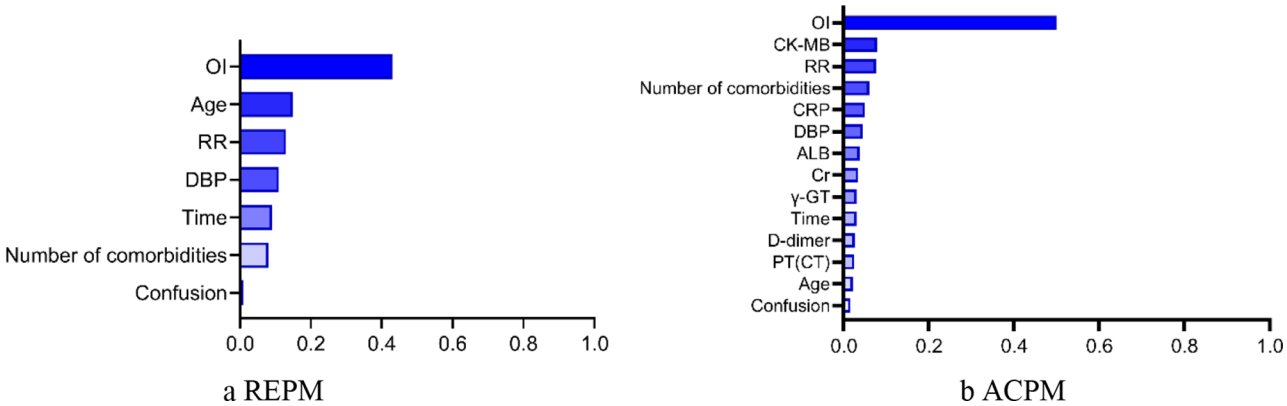


Fig. 6 Importance ranking of the RF model

the SVM-based REPM showed higher accuracy (98.45%) than the SVM-based ACPM (98.01%). Both RF-based models yielded the highest accuracy when only one variable, i.e., OI, was included. Therefore, when LR, SVM, and RF were used, the REPM could accurately predict the COVID-19 severity risk. However, when Cox regression was used, the ACPM could accurately predict the COVID-19 severity risk. In summary, the REPM has been

able to predict the risk of serious illness well based on the primary care variables, whereas the ACPM, despite not being significantly higher, also provides favorable prediction performance, which is also suitable as an auxiliary model for risk prediction and diagnosis under conditions of sufficient time.

Table 7 Predictions of different clinical variables selected by the RF-based REPM

Number of variables	OI	Age	RR	DBP	Time	Number of comorbidities	Confusion	AUC	Accuracy
1	✓							0.991	98.01%
2	✓	✓						0.968	93.81%
3	✓	✓	✓					0.989	95.58%
4	✓	✓	✓	✓				0.990	94.25%
5	✓	✓	✓	✓	✓			0.985	96.02%
6	✓	✓	✓	✓	✓	✓		0.990	97.12%
7	✓	✓	✓	✓	✓	✓	✓	0.990	97.35%

Predictors

As shown in Table 9, the LR-based models contained OI, confusion, and RR as variables. Cox regression-based REPM contained OI, confusion, DBP, and number of comorbidities as variables, whereas the ACPM contained the laboratory variables—ALB and Cr. The SVM-based REPM contained OI, confusion, age, and RR as variables, whereas the SVM-based ACPM contained OI, confusion, and number of comorbidities as variables. Both RF-based models contained only OI. Among the variables included in the two prediction models for the COVID-19 severity risk, OI appeared most frequently, followed by confusion. This indicates that these two characteristics, particularly OI, need to be focused on during the clinical observation of patients with COVID-19. Thus, OI can be considered an independent predictor.

Analysis of predictors via Shapley additive explanation (SHAP)

A comparison of model accuracy and differentiation above indicates that an REPM for determining COVID-19 severity risk is best constructed by SVM. Then the local and global interpretation of the SVM model by SHAP is realized using Python.

Local interpretation

The local interpretation is presented in a condensed force diagram, wherein the key predictors of the individual sample are indicated. Figure 8 shows a positive contribution of age to the predicted outcome, i.e., the predicted severity risk is elevated at the age of 50 years. However, it has limited contribution, as indicated by the short length of the red bar. OI has a negative contribution, i.e., an OI of 360.73 (> 300 mmHg); therefore, the predicted severity risk is reduced. The blue bar is the longest, indicating OI has the largest contribution to the predicted outcome. Confusion and RR, which are not shown in the diagram, contribute less to the predicted outcome in this case. Figure 9 shows that age and OI contribute positively to the predicted outcome. In other words, at the age of 89 years and an OI of 264.05 (< 300 mmHg), the predicted severity risk is high. Moreover, OI with the longest red bar contributes the most to the results. RR contributes negatively to the predicted outcome, i.e., at an RR of 15

(< 30 breaths/min), a lower severity risk is predicted. The contribution of OI is the largest in both cases, indicating its relatively large impact on the predicted risk of severe COVID-19 in patients. This result is consistent with the order of importance of the OI of variable obtained from the SVM.

Multivariate analysis

Global interpretation is presented as a bar and hive plot to rank the importance of the mean absolute SHAP values in the model. As shown in Fig. 10, OI had a considerable influence, confusion and RR had comparable influences, and age had a low influence on the predicted outcomes. SHAP values were visualized to analyze the positive and negative effects of each predictor on the predicted outcomes, and a honeycomb plot was constructed (Fig. 11). Each row represents a feature, with each point representing a sample. The redder and bluer the row color, the larger and smaller the variable value, respectively. Moreover, most red points of the OI were present in the region with SHAP values < 0, whereas the blue points were concentrated in the region with SHAP values > 0. Evidently, the risk of a serious illness can be high with a very low OI. Patients with confusion had a higher risk of developing a severe disease. The localized red points for RR were distributed in the region with SHAP values > 0, and the blue points were distributed in the region with SHAP values < 0. In other words, the possibility of severe risk increases with very high RR. The sample points for age were concentrated on the axis where the SHAP value was 0, and those for higher and lower ages were distributed in the positive and negative regions of the SHAP value, respectively. This showed a low correlation between age and severe risk, may be because the age structure is primarily middle-aged and older.

Univariate analysis

To further understand the relation between the influencing factors and predictors of the model, a dependency plot of the predictors analyzed using SHAP was constructed (Fig. 12). Age had a weak positive correlation with the SHAP value among patients aged > 70 years, i.e., severe risk may be high at an age of > 70 years. RR was positively correlated with SHAP value when RR > 25, i.e.,

Table 8 Predictions of different clinical variables selected by the RF-based ACPM

Number of variables	OI	CK-MB	RR	Number of comorbidities	CRP	DBP	ALB	Cr	γ-GT	Time	D-dimer	PT (CT)	Age	Confusion	AUC	Accuracy
1	✓														0.991	98.01%
2	✓	✓													0.990	96.46%
3	✓	✓	✓												0.985	94.47%
4	✓	✓	✓	✓											0.979	96.02%
5	✓	✓	✓	✓	✓										0.990	94.25%
6	✓	✓	✓	✓	✓	✓									0.990	95.89%
7	✓	✓	✓	✓	✓	✓	✓								0.979	97.12%
8	✓	✓	✓	✓	✓	✓	✓	✓							0.988	95.89%
9	✓	✓	✓	✓	✓	✓	✓	✓	✓						0.989	96.02%
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					0.989	96.9%
11	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				0.992	95.35%
12	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			0.990	96.46%
13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		0.986	94.69%
14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0.987	96.24%

the severe risk gradually increased with increasing RR. OI sharply decreased to 300 mmHg, consistent with the criterion for judging critical illness ($OI \leq 300$ mmHg). When the patient had confusion, the SHAP value was > 0 , with increasing severe risk.

The effect of the specific values of each predictor in a single sample and all samples on the predicted results is separately explained using SHAP. The SVM-based models contain important predictors, including OI that coincides with the categorization of patients according to severity of disease ($OI \leq 300$ mmHg as severe disease [35]), confusion that coincides with the CURB-65 score [20] and NEWS2 score [48], risk factors included in the COVID-GRAM [5] score, RR that coincides with the criterion for classifying patients with clinically severe COVID-19 ($RR \geq 30$ beats/min), and age that coincides with the risk factors included in the CURB-65 score [20], MuLBSTA score [17], CALL score [7], and the laboratory score developed by the University Hospital Center of Blida [14]. SHAP analysis verified that the OI is a crucial influencing factor, in agreement with the order of importance of the variables obtained by the model. Moreover, the influencing factors could be easily obtained, rendering model application simpler and more convenient.

Discussion

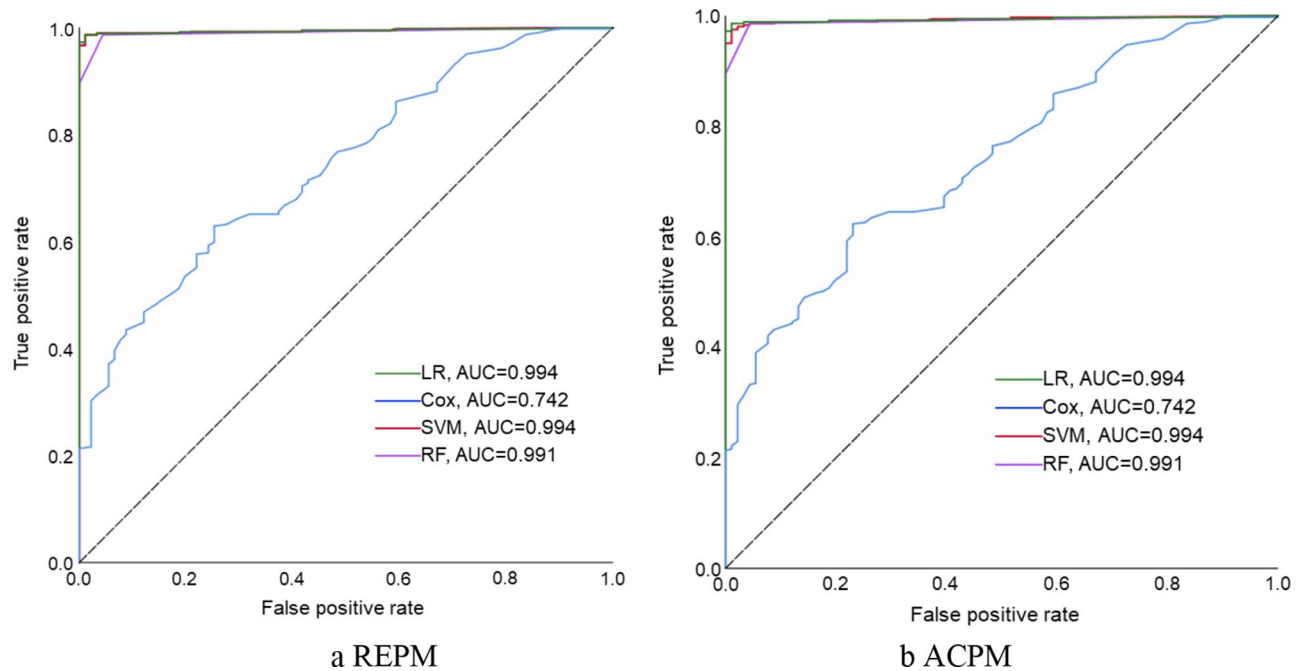
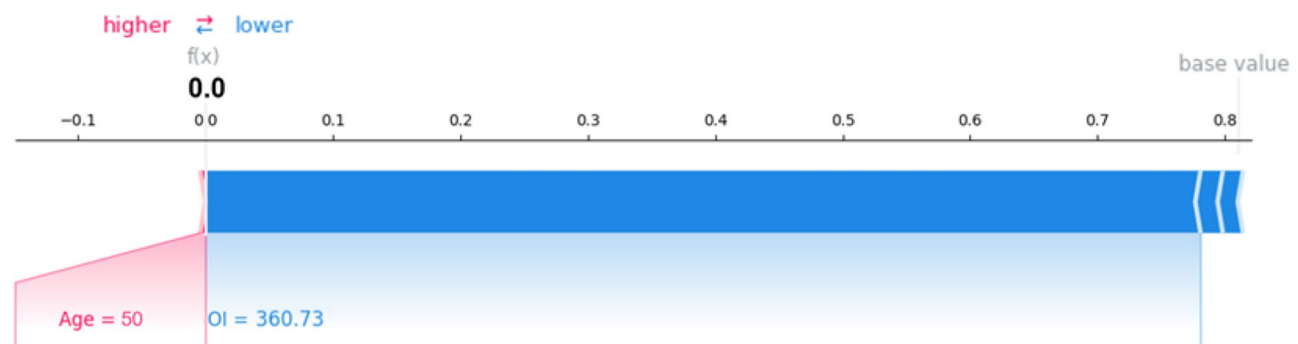
SVM model superiority

Herein, SVM in the REPM prediction task achieved the highest prediction accuracy of 98.45%, with an AUC of 0.994, sensitivity of 0.989, and specificity of 0.969. LR is the most widely used machine learning algorithm for predicting COVID-19 risk [6, 17, 29, 49–55], followed by XGBoost [8, 56–59], SVM [4, 60–62], and RF [63–65]. Contrarily, Cox regression [7, 14, 26, 66, 67] was more commonly accepted for screening and assessing the importance of variables. Table 10 lists some studies that used different machine learning algorithms to construct prediction models for severity or mortality risks in patients with COVID-19. Gong et al. [52] constructed a prediction model based on LR for severity risk in patients with COVID-19, which achieved an AUC of 0.853. Booth et al. [4] utilized an SVM model with an accuracy of 94.5%. Abdulaal et al. [3] employed an ANN model with an accuracy of 86.25%. Further, Xiong et al. [65] employed an RF model with an accuracy of 84.5%. In contrast, the SVM model achieved a predictive accuracy of 98.45%, thereby outperforming the aforementioned models. In addition, the sensitivity of 0.989, specificity of 0.967 and AUC values of the optimal model in this paper are the highest among all comparative models.

Compared to the data employed in the existing models, the multicenter data employed in our study encompassed a larger and more diverse patient population. The substantial sample size facilitated a greater depth of learning,

Table 9 Comparison of prediction modeling results for different methods

Models	Methods	Variables	AUC	Sensitivity	Specificity	Accuracy
REPM	LR	OI, confusion, and RR	0.994	0.986	0.967	98.23%
	Cox regression	OI, confusion, DBP, and number of comorbidities	0.742	0.903	0.297	80.31%
	SVM	OI, confusion, age, and RR	0.994	0.989	0.967	98.45%
	RF	OI	0.991	0.986	0.956	98.01%
ACPM	LR	OI, confusion, and RR	0.994	0.986	0.967	98.23%
	Cox regression	OI, confusion, DBP, number of comorbidities, ALB, and Cr	0.742	0.936	0.275	81.19%
	SVM	OI, confusion, and number of comorbidities	0.994	0.983	0.967	98.01%
	RF	OI	0.991	0.986	0.956	98.01%

**Fig. 7** ROC curves for different models**Fig. 8** Distribution of SHAP values for a patient in the nonsevere group

thereby enabling our model to discern subtle patterns and correlations that smaller datasets might not capture. Consequently, the model trained on these extensive data exhibited enhanced predictive performance. Among the various machine learning methods evaluated herein, the SVM-based model exhibited superior prediction accuracy. The SVM model was particularly well suited to this

task owing to its ability to efficiently handle high-dimensional data and robustness to overfitting. Its performance surpassed that of other algorithms, including LR, RF, and Cox regression, which are commonly employed in predictive modeling. In summary, the advantages of the SVM model suggest its strong candidacy for prioritization as a predictive model.

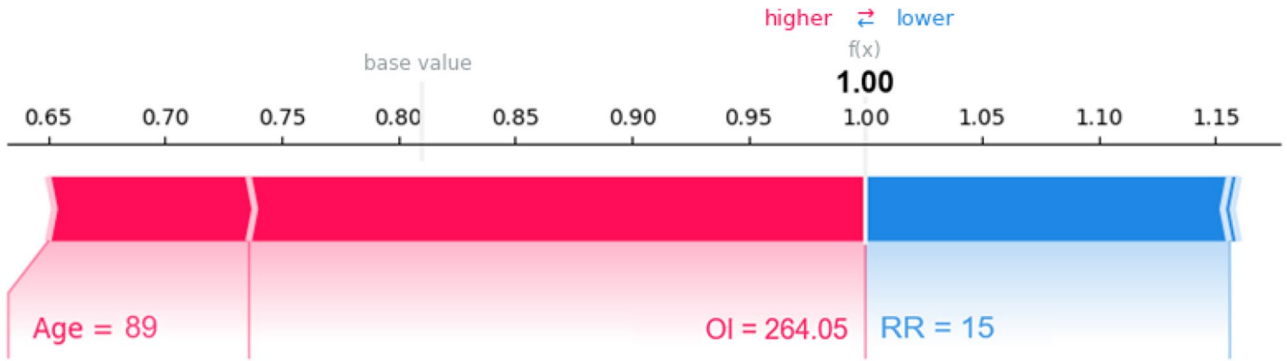


Fig. 9 Distribution of SHAP values for a patient in the severe group

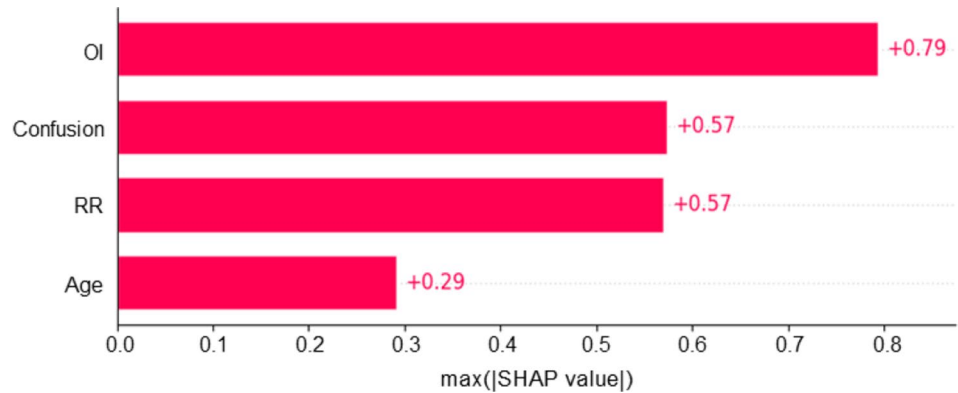


Fig. 10 Importance ranking of mean absolute SHAP values

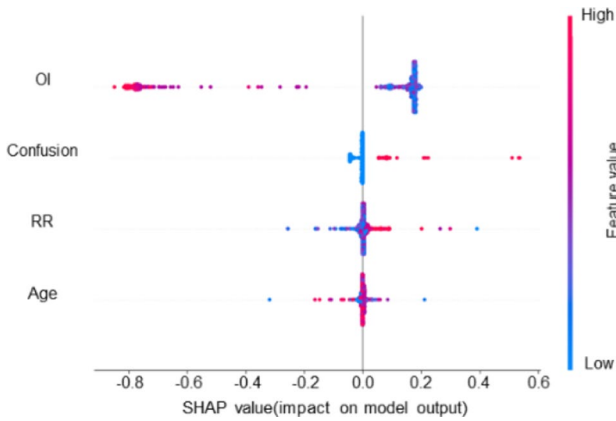


Fig. 11 Hive chart for SHAP value visualization

OI as a critical predictor

The most influential variables in the SVM model are OI, confusion, RR, and age, in order. OI is a crucial predictor of pulmonary gas exchange function, with values below 300 signaling impaired oxygenation. In patients with COVID-19, lower OI indicates reduced oxygen diffusion due to lung inflammation, guiding the need for oxygen therapy or mechanical ventilation. Confusion often reflects disease severity and can indicate neurological complications, drug reactions, or organ failure, impacting

prognosis and treatment decisions. RR provides real-time insight into the patient’s respiratory function. A persistently high RR (>25) suggests worsening lung function, requiring close monitoring and potential escalation of care. Age is a critical predictor of COVID-19 outcomes, with those over 70 at higher risk for severe disease and complications, aiding in risk stratification and treatment planning.

OI ranks first in importance in each model and is a critical predictor in SHAP analysis. OI is determined based on the oxygen flow rate and arterial oxygen partial pressure, which is a blood oxygen indicator besides SpO2. Oxygen flow rate and SpO2 are important reference variables for determining the Quick COVID-19 Severity Index score [18] developed by the Yale University School of Medicine in the United States, which has been verified by multiple sources [18, 68–70]. This score can be effectively used for the clinical triage of patients with COVID-19 in some Asian and European countries. The 4 C Mortality score [19] developed in a prospective cohort study jointly conducted by 260 hospitals in the UK also considered SpO2 in its scoring index. This approach exhibited excellent predictive performance [71–75], enabling informed decision-making by clinicians. SHAP analysis results showed that the lower the OI, the higher the severity risk, confirming the finding that hypoxemia

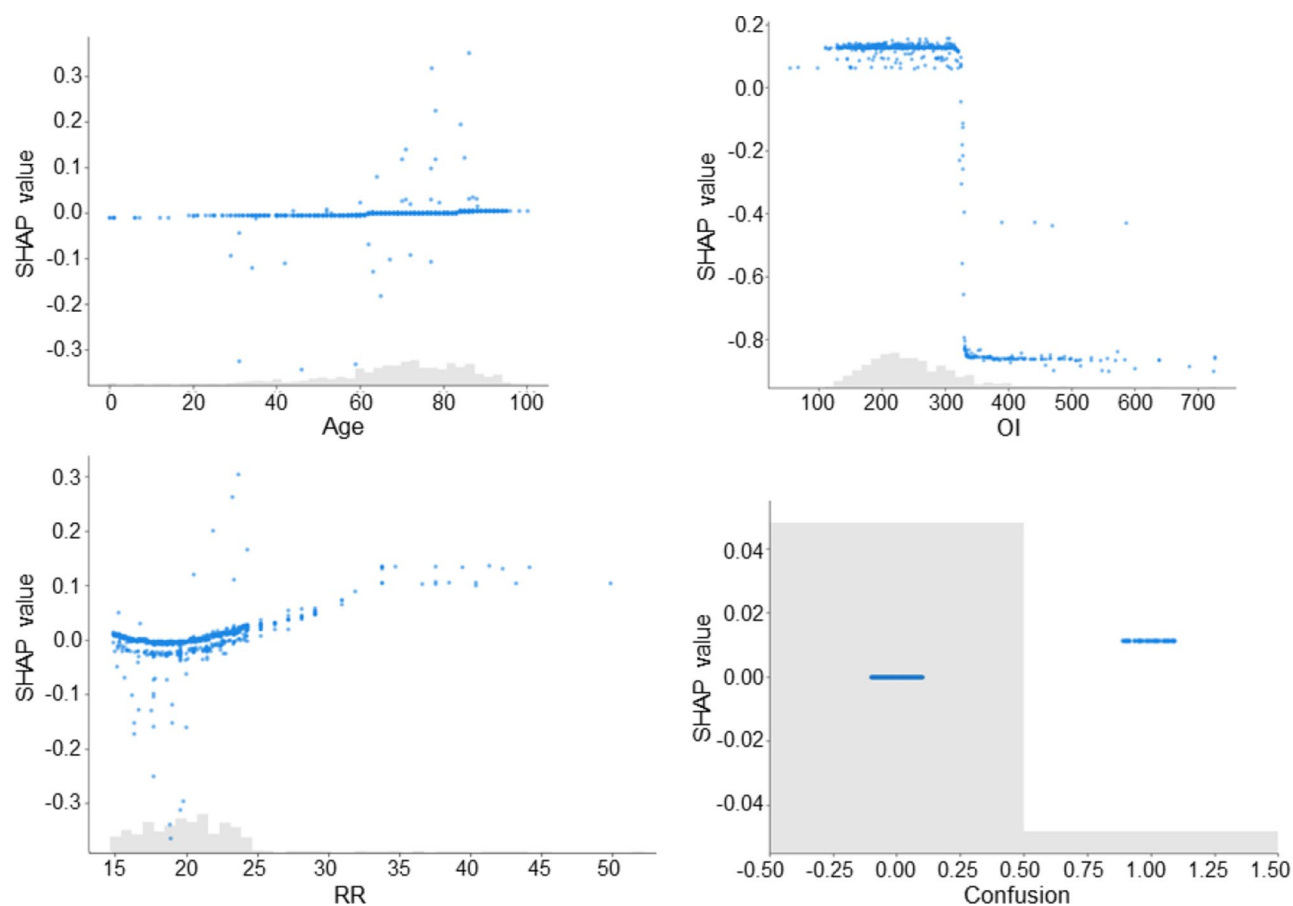


Fig. 12 Dependency plot of the predictors

Table 10 Comparison of machine learning algorithms in predicting severity and mortality risks in patients with COVID-19

	Machine learning	Accuracy	Sensitivity	Specificity	AUC
Proposed model	SVM	98.45%	0.989	0.967	0.994
Gong et al. [52]	LR		0.775	0.784	0.853
Booth et al. [4]	SVM	94.5%	0.91	0.91	0.93
Abdulaal et al. [3]	ANN	86.25%	0.875	0.859	
Xiong et al. [65]	RF	84.5%	0.967	0.695	0.970

was related to severe COVID-19 and mortality [76]. OI is also an indicator of respiratory distress and provides information on disease progression in a patient, thereby preventing deaths due to respiratory failure in patients with COVID-19 [77]. Therefore, blood oxygen indicators, such as OI, provide important reference values for predicting COVID-19 risk. Herein, OI was the critical predictor of COVID-19 severity risk; it can also serve as a primary and independent evaluation indicator.

Limitations and prospects

This study had certain limitations from the perspective of data structure and clinical medicine. First, the data were collected from 6 hospitals in a geographic region (Tianjin, China). While our test set was institutionally independent and we have put our best efforts to collect

data and avoid potential bias, the generalizability of the model may be constrained by region-specific factors. Future multi-center validations across diverse health-care systems are necessary to evaluate robustness against these biases. Second, as the model evaluation primarily focuses on predictive accuracy, the evaluation indicators (accuracy, sensitivity, specificity, and AUC) remain inadequate. In the future, when such infectious diseases over COVID-19 require MRI imaging or any other unstructured data for diagnosis, deep learning models such as LSTM, GRU and transformer could be adopted to construct disease risk prediction. Despite its limitations, the prediction model exhibited high accuracy and interpretability. Moreover, the variables required for the model were easily available. The severity risk of a disease can be self-tested using basic medical instruments, which can

reduce the burden on the hospital staff and assist clinicians in making informed treatment decisions.

Conclusion

In this study, prediction models for severity risk in patients with COVID-19 were constructed using four machine learning algorithms including LR, Cox regression, SVM, and RF. The SVM-based model demonstrated the highest prediction accuracy, rendering it a promising candidate for predicting. Furthermore, OI was deemed a critical predictor of COVID-19 severity risk, which can serve as a primary and independent evaluation indicator. The identification of OI as a key predictor of COVID-19 severity risk underpins its potential as a critical evaluation indicator. Although the WHO declared the end of the COVID-19 pandemic on May 5, 2023, COVID-19 cases recorded globally are still on the rise [78]. Our study assists physicians in more accurately predicting the disease progression in patients and provides a significant method as well as technical guidance for constructing risk prediction models to prevent increasing critical illness rates and mortality rates.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-025-22976-x>.

Supplementary Material 1

Acknowledgements

We are grateful for the technical assistance offered by Prof. Weizhong Chen from Chengdu Medical College University, China, and Prof. Jun Ma from Tianjin Medical University.

Author contributions

Xin Jin, Hansong Zhang, Ying Wang, Yan Xie, Cuihan Wang and Yuqi Ma all contributed to the study conception and design. Material preparation, data collection and analysis were performed by Hansong Zhang and Xin Jin. The first draft of the manuscript was written by Hansong Zhang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The dataset used in this study is not publicly stored due to the sensitive information involved. To access the data, please contact the authors at [jx_123@tju.edu.cn] and state the purpose of the study, and the data will be available upon authorization.

Declarations

Ethics approval and consent to participate

This retrospective study was approved by the ethics committee of Tianjin University Hospital (2024-YLS-125) and was conducted in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The ethics committee of Tianjin University Hospital waived informed consent due to the retrospective nature of our study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 13 October 2024 / Accepted: 29 April 2025

Published online: 13 May 2025

References

1. WHO Director-General's speech at the World Governments Summit-- 12 February. <https://www.who.int/director-general/speeches/detail/who-director-general-s-speech-at-the-world-governments-summit--12-february-2024>. Accessed 27 May 2024.
2. Wang Hesheng answered reporters' questions. at the press conference on the theme of people's livelihood at the second session of the 14th National People's Congress. https://www.ndcpa.gov.cn/jbkzzx/c100009/common/content/content_1766465153672060928.html. Accessed 5 May 2024.
3. Abdulaal A, Patel A, Charani E, Denny S, Mughal N, Moore L. Prognostic modeling of COVID-19 using artificial intelligence in the united Kingdom: model development and validation. *J Med Internet Res*. 2020;22(8):e20259.
4. Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Mod Pathol*. 2021;34(3):522–31.
5. Meng L, Dong D, Li L, Niu M, Bai Y, Wang M, Qiu X, Zha Y, Tian J. A deep learning prognosis model help alert for COVID-19 patients at High-Risk of death: A Multi-Center study. *IEEE J Biomed Health*. 2020;24(12):3576–84.
6. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, Li Y, Guan W, Sang L, Lu J, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *Jama Intern Med*. 2020;180(8):1081–9.
7. Ji D, Zhang D, Xu J, Chen Z, Yang T, Zhao P, Chen G, Cheng G, Wang Y, Bi J, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clin Infect Dis*. 2020;71(6):1393–9.
8. Shamout FE, Shen Y, Wu N, Kaku A, Park J, Makino T, Jastrzebski S, Witowski J, Wang D, Zhang B, et al. An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department. *NPJ Digit Med*. 2021;4(1):80.
9. Fu Y, Zeng L, Huang P, Liao M, Li J, Zhang M, Shi Q, Xia Z, Ning X, Mo J, et al. Severity-onset prediction of COVID-19 via artificial-intelligence analysis of multivariate factors. *Heliyon*. 2023;9(8):e18764.
10. Mahajan A, Toshniwal D. A Novel Ensemble-Based Framework for Feature Identification and Classification of COVID-19 Electronic Health Record Data. In 2023 IEEE International Conference on Big Data (BigData). 2023:3711–3720. <https://doi.org/10.1109/BigData59044.2023.10386343>
11. Mahajan A, Sharma N, Aparicio-Obregon S, Alyami H, Alharbi A, Anand D, Sharma M, Goyal N. A Novel Stacking-Based Deterministic Ensemble Model for Infectious Disease Prediction. In *Mathematics*. 2022. <https://doi.org/10.3390/math10101714>
12. Alakus TB, Turkoglu I. Comparison of deep learning approaches to predict COVID-19 infection. *Chaos Solitons Fractals*. 2020;140:110120.
13. Asteris PG, Gandomi AH, Armaghani DJ, Kokoris S, Papandreadi AT, Roumelioti A, Papanikolaou S, Tsoukalas MZ, Triantafyllidis L, Koutras EI, et al. Prognosis of COVID-19 severity using DERGA, a novel machine learning algorithm. *Eur J Intern Med*. 2024. <https://doi.org/10.1016/j.ejim.2024.02.037>.
14. Bennouar S, Bachir CA, Kessira A, Bennouar DE, Abdi S. Development and validation of a laboratory risk score for the early prediction of COVID-19 severity and in-hospital mortality. *Intens Crit Care Nur*. 2021;64:103012.
15. de Terwangne C, Laouni J, Jouffe L, Lechien JR, Bouillon V, Place S, Capulzini L, Machayekhi S, Ceccarelli A, Saussez S et al. Predictive Accuracy of COVID-19 World Health Organization (WHO) Severity Classification and Comparison with a Bayesian-Method-Based Severity Score (EPI-SCORE). In *Pathogens*. 2020. <https://doi.org/10.3390/pathogens9110880>
16. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN. A prediction rule to identify low-risk patients with community-acquired pneumonia. *New Engl J Med*. 1997;336(4):243–50.
17. Guo L, Wei D, Zhang X, Wu Y, Li Q, Zhou M, Qu J. Clinical features predicting mortality risk in patients with viral pneumonia: the mulbsta score. *Front Microbiol*. 2019;10:2752.

18. Haimovich AD, Ravindra NG, Stoytchev S, Young HP, Wilson FP, van Dijk D, Schulz WL, Taylor RA. Development and validation of the quick COVID-19 severity index: A prognostic tool for early clinical decompensation. *Ann Emerg Med*. 2020;76(4):442–53.
19. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, Dunning J, Fairfield CJ, Gamble C, Green CA, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4 C mortality score. *BMJ-Brit Med J*. 2020;370:m3339.
20. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, Lewis SA, Macfarlane JT. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003;58(5):377–82.
21. Mahanty C, Kumar R, Asteris PG, Gandomi AH. COVID-19 Patient Detection Based on Fusion of Transfer Learning and Fuzzy Ensemble Models Using CXR Images. In *Appl. Sci*. 2021. <https://doi.org/10.3390/app112311423>
22. Terpos E, Ntanasis-Stathopoulos I, Elalamy I, Kastritis E, Sergentanis TN, Politou M, Psaltopoulou T, Gerotziakas G, Dimopoulos MA. Hematological findings and complications of COVID-19. *Am J Hematol*. 2020;95(7):834–47.
23. Vincent JL, Moreno R, Takala J, Willatts S, Mendona AD, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related organ failure Assessment) score to describe organ dysfunction/failure. On behalf of the working group on Sepsis-Related problems of the European society of intensive care medicine. *Intens Care Med*; 1996. p. 22.
24. Willette AA, Willette SA, Wang Q, Pappas C, Klinedinst BS, Le S, Larsen B, Pollpeter A, Li T, Mochel JP, et al. Using machine learning to predict COVID-19 infection and severity risk among 4510 aged adults: a UK biobank cohort study. *Sci Rep-UK*. 2022;12(1):7736.
25. Xiao LS, Zhang WF, Gong MC, Zhang YP, Chen LY, Zhu HB, Hu CY, Kang P, Liu L, Zhu H. Development and validation of the HNC-LL score for predicting the severity of coronavirus disease 2019. *Ebiomedicine*. 2020;57:102880.
26. Xu J, Zhang W, Cai Y, Lin J, Yan C, Bai M, Cao Y, Ke S, Liu Y. Nomogram-based prediction model for survival of COVID-19 patients: A clinical study. *Heliyon*. 2023;9(9):e20137.
27. Zheng Z, Peng F, Xu B, Zhao J, Tang W. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect*. 2020;81(2).
28. Zhou MY, Xie XL, Peng YG, Wu MJ, Deng XZ, Wu Y, Xiong LJ, Shang LH. From SARS to COVID-19: what we have learned about children infected with COVID-19. *Int J Infect Dis*. 2020;96:710–4.
29. Zhu Z, Hu G, Ying Z, Wang J, Han W, Pan Z, Tian X, Song W, Sui X, Song L et al. Time-dependent CT score-based model for identifying severe/critical COVID-19 at a fever clinic after the emergence of Omicron variant. *Heliyon*. 2024:e27963.
30. Mahanty C, Kumar R, Patro SGK. Internet of medical Things-Based COVID-19 detection in CT images fused with fuzzy ensemble and transfer learning models. *New Generat Comput*. 2022;40(4):1125–41.
31. Dansana D, Kumar R, Bhattacharjee A, Mahanty C. Int J Reliab Qual E-Healthc. 2022;11(1):1–13. COVID-19 Outbreak Prediction and Analysis of E-Healthcare Data Using Random Forest Algorithms.
32. Mahanty C, Patro SGK, Rathor S, Rachapudi V, Muzammil K, Islam S, Razak A, Khan WA. Forecasting of coronavirus active cases by utilizing logistic growth model and fuzzy time series techniques. *Sci Rep-UK*. 2024;14(1):18039.
33. Mahanty C, Kumar R, Mishra BK, Hemanth DJ, Gupta D, Khanna A. Prediction of COVID-19 active cases using exponential and non-linear growth models. *Expert Syst*. 2022;39(3):e12648.
34. LC C, K CMR. B. KM: Brain Tumor Detection and Classification Using Convolutional Neural Network and Deep Neural Network. In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). 2020;1–4.
35. The guidelines for diagnosis and treatment of novel Coronavirus Infection Diagnosis. and Treatment Program (Trial 10th Edition) issued by the National Health Commission of China. <http://www.nhc.gov.cn/ylyjs/pqt/202301/32de5b2ff9bf4eaa88e75bdf7223a65a.shtml>. Accessed 27 May 2024.
36. Riley RD, Ensor J, Snell K, Harrell FJ, Martin GP, Reitsma JB, Moons K, Collins G, van Smeden M. Calculating the sample size required for developing a clinical prediction model. *BMJ-Brit Med J*. 2020;368:m441.
37. Schober P, Vetter TR. Logistic regression in medical research. *Anesth Analg*. 2021;132(2):365–6.
38. Jullum M, Hjort NL. What price semiparametric Cox regression? *Lifetime Data Anal*. 2019;25(3):406–38.
39. Muff S, Nilsen EB, O Hara RB, Nater CR. Rewriting results sections in the Language of evidence. *Trends Ecol Evol*. 2022;37(3):203–10.
40. Akram-Ali-Hammouri Z, Fernandez-Delgado M, Cernadas E, Barro S. Fast support vector classification for Large-Scale problems. *IEEE T Pattern Anal*. 2022;44(10):6184–95.
41. Scornet E, Biau, Gerard, Vert, Jean-Philippe: CONSISTENCY OF RANDOM FORESTS. *The Annals of Statistics: An Official Journal of the Institute of Mathematical Statistics* 2015.
42. Bradley P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn*. 1997;30(7):1145–59.
43. Swets J. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285–93.
44. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. In *NIPS*; 2017; 2017.
45. Dananche C, Elias C, Henaff LSEM. Baseline clinical features of COVID-19 patients, delay of hospital admission and clinical outcome: A complex relationship. *PLoS One*. 2022;17(1).
46. De La Calle M, García Reyne G, Lora-Tamayo A, Muñio Miguez J, Arnalich-Fernandez A, Beato Pérez F, Vargas Núñez JL, Caudevilla Martínez JA, Alcalá Rivera MA, Orviz García N. Impact of days elapsed from the onset of symptoms to hospitalization in COVID-19 in-hospital mortality: time matters. *Revista Clínica Española (English Edition)*. 2023;223(5):281–97.
47. Yadav KN, Hemmons J, Snider CK, Patel A, Childs M, Delgado MK. Association between patient-reported onset-to-door time and mortality in patients hospitalized with COVID-19 disease. *The American journal of emergency medicine*. 2024;77.
48. National Early Warning Score (NEWS) 2. <https://www.rcp.ac.uk/improving-care/resources/national-early-warning-score-news-2/>. Accessed 28 May 2024.
49. Alballa N, Al-Turaiki I. Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. *Inf Med Unlocked*. 2021;24:100564.
50. Aloisio E, Chibireva M, Serafini L, Pasqualetti S, Falvella FS, Dolci A, Panteghini M. A comprehensive appraisal of laboratory biochemistry tests as major predictors of COVID-19 severity. *Arch Pathol Lab Med*. 2020;144(12):1457–64.
51. Dehghani P, Schmidt CW, García S, Okeson B, Grines CL, Singh A, Patel R, Wiley J, Htun WW, Nayak KR, et al. North American COVID-19 myocardial infarction (NACMI) risk score for prediction of In-Hospital mortality. *J Soc Cardiovasc Angiogr Interv*. 2022;1(5):100404.
52. Gong J, Ou J, Qiu X, Jie Y, Chen Y, Yuan L, Cao J, Tan M, Xu W, Zheng F, et al. A tool for early prediction of severe coronavirus disease 2019 (COVID-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China. *Clin Infect Dis*. 2020;71(15):833–40.
53. Kamran F, Tang S, Otlés E, McEvoy DS, Saleh SN, Gong J, Li BY, Dutta S, Liu X, Medford RJ, et al. Early identification of patients admitted to hospital for covid-19 at risk of clinical deterioration: model development and multisite external validation study. *BMJ-Brit Med J*. 2022;376:e068576.
54. Kwekha-Rashid AS, Abduljabbar HN, Alhayani B. Coronavirus disease (COVID-19) cases analysis using machine-learning applications. *Appl Nanosci*. 2023;13(3):2013–25.
55. Luo M, Liu J, Jiang W, Yue S, Liu H, Wei S. IL-6 and CD8 + T cell counts combined are an early predictor of in-hospital mortality of patients with COVID-19. *JCI Insight*. 2020;5(13).
56. Bertsimas D, Lukin G, Mingardi L, Nohadani O, Orfanoudaki A, Stellato B, Wiberg H, Gonzalez-Garcia S, Parra-Calderon CL, Robinson K, et al. COVID-19 mortality risk assessment: an international multi-center study. *PLoS One*. 2020;15(12):e0243262.
57. Kim HJ, Han D, Kim JH, Kim D, Ha B, Seog W, Lee YK, Lim D, Hong SO, Park MJ, et al. An Easy-to-Use machine learning model to predict the prognosis of patients with COVID-19: retrospective cohort study. *J Med Internet Res*. 2020;22(11):e24225.
58. Pan P, Li Y, Xiao Y, Han B, Su L, Su M, Li Y, Zhang S, Jiang D, Chen X, 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(11):e23128.
59. Rechtman E, Curtin P, Navarro E, Nirenberg S, Horton MK. Vital signs assessed in initial clinical encounters predict COVID-19 mortality in an NYC hospital system. *Sci Rep-UK*. 2020;10(1).
60. Sun L, Song F, Shi N, Liu F, Li S, Li P, Zhang W, Jiang X, Zhang Y, Sun L, et al. Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19. *J Clin Virol*. 2020;128:104431.
61. Yao H, Zhang N, Zhang R, Duan M, Xie T, Pan J, Peng E, Huang J, Zhang Y, Xu X, et al. Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests. Volume 8. *Front Cell Dev Biol*; 2020. p. 683.

62. Zhao C, Bai Y, Wang C, Zhong Y, Jin R. Risk factors related to the severity of COVID-19 in Wuhan. *Int J Med Sci.* 2020;18(1).
63. Greco M, Angelotti G, Caruso PF, Zanella A, Stomeo N, Costantini E, Protti A, Pesenti A, Grasselli G, Cecconi M. Outcome prediction during an ICU surge using a purely data-driven approach: A supervised machine learning case-study in critically ill patients from COVID-19 Lombardy outbreak. *Int J Med Inf.* 2022;164:104807.
64. Jimenez-Solem E, Petersen TS, Hansen C, Hansen C, Lioma C, Igel C, Boomsma W, Krause O, Lorenzen S, Selvan R, et al. Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients. *Sci Rep-UK.* 2021;11(1):3246.
65. Xiong Y, Ma Y, Ruan L, Li D, Lu C, Huang L. Comparing different machine learning techniques for predicting COVID-19 severity. *Infect Dis Poverty.* 2022;11(1):19.
66. Ebrahimi V, Sharifi M, Mousavi-Roknabadi RS, Sadegh R, Khademian MH, Moghadami M, Dehbozorgi A. Predictive determinants of overall survival among re-infected COVID-19 patients using the elastic-net regularized Cox proportional hazards model: a machine-learning algorithm. *BMC Public Health.* 2022;22(1):10.
67. Roimi M, Gutman R, Somer J, Ben AA, Calman I, Bar-Lavie Y, Gelbshtein U, Liverant-Taub S, Ziv A, Eytan D, et al. Development and validation of a machine learning model predicting illness trajectory and hospital utilization of COVID-19 patients: A nationwide study. *J Am Med Inf Assn.* 2021;28(6):1188–96.
68. Ak R, Kurt E, Bahadirli S. Comparison of 2 risk prediction models specific for COVID-19: the Brescia-COVID respiratory severity scale versus the quick COVID-19 severity index. *Disaster Med Public.* 2021;15(4):e46–50.
69. Ngai K, Maher PP, Leibner E, Loo G, Legome E. 14 External validation of the quick COVID-19 severity index: A prognostic tool for early clinical decompensation. *Ann Emerg Med.* 2021;78(2):57–8.
70. Rodriguez-Nava G, Yanez-Bello MA, Trelles-Garcia DP, Chung CW, Hines DW. Performance of the Quick COVID-19 Severity Index and the Brescia-COVID Respiratory Severity Scale in hospitalized patients with COVID-19 in a community hospital setting. *Int J Infect Dis.* 2020;102(18).
71. Cárdenas-Fuentes G, Bosch De Basea M, Cobo I, Subirana I, Ceresa M, Famada E, Gimeno-Santos E, Delgado-Ortiz L, Faner R, Molina-Molina M, et al. Validity of prognostic models of critical COVID-19 is variable. A systematic review with external validation. *J Clin Epidemiol.* 2023;159:274–88.
72. Gordon AJ, Govindarajan P, Bennett CL, Matheson L, Kohn MA, Camargo C, Kline J. External validation of the 4 C mortality score for hospitalised patients with COVID-19 in the RECOVER network. *BMJ Open.* 2022;12(4):e054700.
73. Riley JM, Moeller PJ, Crawford AG, Schaefer JW, Cheney-Peters DR, Venkataraman CM, Li CJ, Smaltz CM, Bradley CG, Lee CY, et al. External validation of the COVID-19 4 C mortality score in an urban United States cohort. *Am J Med Sci.* 2022;364(4):409–13.
74. Vedovati MC, Barbieri G, Urbini C, D'Agostini E, Vanni S, Papalini C, Pucci G, Cimini LA, Valentino A, Ghiadoni L, et al. Clinical prediction models in hospitalized patients with COVID-19: A multicenter cohort study. *Resp Med.* 2022;202:106954.
75. Zahra A, van Smeden M, Abbink EJ, van den Berg JM, Blom MT, van den Dries CJ, Gussekloo J, Wouters F, Joling KJ, Melis R, et al. External validation of six COVID-19 prognostic models for predicting mortality risk in older populations in a hospital, primary care, and nursing home setting. *J Clin Epidemiol.* 2024;168:111270.
76. Wang Y, Wang Y, Chen Y, Qin Q. Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *J Med Virol.* 2020;92(6):568–76.
77. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intens Care Med.* 2020;46(5):846–8.
78. COVID - Coronavirus Statistics - Worldometer. <https://www.worldometers.info/coronavirus/>. Accessed 12 Sep 2024.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.