Heliyon 10 (2024) e37096

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Clinical data and quantitative CT parameters combined with machine learning to predict short-term prognosis of severe COVID-19 in the elderly☆

Lifang Fan [a], Shujian Wu [b], Yimin Wu [c], Xiaoyan Xu [a], Zhengyuan Xu [a], Lei Huang [a], Guoxian Chen [d,*]

[a] School of Medical Imageology, Wannan Medical College, Wuhu, Anhui, China
[b] Department of Radiology, Yijishan Hospital of Wannan Medical College, Wuhu, Anhui, China
[c] Department of Ultrasound, WuHu Hospital, East China Normal University (The Second People's Hospital, WuHu), Wuhu, Anhui, China
[d] School of Clinical Medicine, Wannan Medical College, Wuhu, Anhui, China

ARTICLE INFO

Keywords:
Quantitative CT
Machine learning
COVID-19

ABSTRACT

Rationale and objectives: This study aims to evaluate the effectiveness of integrating clinical data and quantitative CT parameters with machine learning techniques in forecasting the short-term outcomes of severe COVID-19 in elderly patients.

Materials and methods: In this retrospective study, we analyzed the clinical profiles and chest quantitative CT parameters of 239 elderly patients with severe COVID-19 admitted for treatment. The cohort included 61 deceased patients (death group) and 178 who recovered and were discharged (survival group). The participants were randomly assigned into a training group (n = 167) and a validation group (n = 72). Quantitative CT parameters were measured using the 3D-Slicer software. Univariate and multivariate logistic regression analyses identified independent risk factors for mortality. Predictive models were developed employing four machine learning algorithms: Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM).

Results: Both univariate and multivariate logistic regression analyses revealed age, hypersensitive C-reactive protein (hs-CRP), and solid organ volume percentage (SOV%) as independent predictors of mortality. The Area Under the Curve (AUC) values for the LR, RF, DT, and SVM models in the training group were 0.795, 0.726, 0.854, and 0.589, respectively; for the validation group, they were 0.817, 0.634, 0.869, and 0.754, respectively. The DT algorithm outperformed other models in both the training and validation groups, emerging as the most effective predictive model in this study.

Conclusion: The combination of clinical data and quantitative CT parameters with machine learning approaches is highly valuable in predicting the short-term prognosis of severe COVID-19 in the elderly. Among the various models tested, the Decision Tree algorithm-based model proved to be the most accurate and reliable in this context.

☆ Lifang Fan, Shujian Wu and Yimin Wu have made an equal contribution to this article and share first authorship.

* Corresponding author. School of Clinical Medicine, Wannan Medical College, No.22 Wenchang West Road, Yijiang District, Wuhu, 241002, Anhui Province, China.
E-mail address: 20120012@wnmc.edu.cn (G. Chen).

https://doi.org/10.1016/j.heliyon.2024.e37096
Received 15 May 2024; Received in revised form 6 July 2024; Accepted 27 August 2024
Available online 7 September 2024

1. Introduction

Since its emergence, Coronavirus Disease 2019 (COVID-19) has profoundly disrupted global work and life patterns. The uncertain origins of the virus pose significant challenges in its complete eradication. While the probability of a global COVID-19 pandemic re-emerging is low, localized outbreaks continue to present a real risk. The virus affects individuals across all age groups; however, younger patients typically experience milder symptoms and favorable prognoses. In contrast, the elderly, often burdened with comorbidities, are predisposed to severe manifestations of the disease, resulting in poorer outcomes [1]. Recent studies report mortality rates as high as approximately 19 % in severe and critical cases [2]. Prior research has identified advanced age, elevated Sequential Organ Failure Assessment (SOFA) scores, increased D-dimer levels, and heightened inflammatory markers as key predictors of negative outcomes in COVID-19 patients. However, these studies lack specificity, as they encompass a broad age range of COVID-19 patients [3,4].

In recent years, machine learning has seen widespread application in medical research. Machine learning algorithms can extract valuable information from large, complex datasets, enabling pattern recognition and predictive analysis. This capability has shown tremendous potential in disease diagnosis, prognosis prediction, and treatment outcome evaluation. Commonly used machine learning algorithms, such as Logistic Regression(LR), Random Forest(RF), Decision Tree(DT), and Support Vector Machine(SVM), have demonstrated excellent performance in various medical applications. These algorithms are particularly effective at handling multi-dimensional data and identifying critical features, thereby assisting clinicians in making more accurate decisions.

This study, therefore, conducts a retrospective analysis of clinical data and quantitative CT parameters to pinpoint risk factors contributing to mortality in severely ill elderly COVID-19 patients. By leveraging various machine learning algorithms to develop predictive models, this research aims to facilitate the early identification of high-risk patients, enabling prompt, targeted clinical interventions. Such proactive measures are crucial in improving the prognosis of elderly patients with severe manifestations of COVID-19.

2. Materials and methods

2.1. Study population

This study entailed a retrospective examination of the clinical and imaging data of COVID-19 patients hospitalized at the First Affiliated Hospital of Wannan Medical College (Yijishan Hospital) between December 2022 and March 2023. The study's inclusion criteria were: (1) a confirmed positive result for the novel coronavirus via nucleic acid testing; (2) patients aged 60 years or older categorized as severe or critical based on the criteria set forth in the "Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 9)" [5]; and (3) comprehensive patient records, including chest High-Resolution Computed Tomography (HRCT) scans and complete clinical and laboratory data. The exclusion criterion was the presence of significant imaging artifacts that impaired measurement feasibility. Applying these criteria, 239 patients with severe COVID-19 were identified: 128 with severe cases and 111 with critical conditions. Of these, 61 patients succumbed during their hospital stay (death group), while 178 showed improvement and were discharged (survival group). The cohort was randomly divided into a training group (n = 167) and a validation group (n = 72). In the training group, there were 42 cases in the death group. In the validation group, there were 19 cases in the death group. This study is a retrospective study and informed consent exemption has been obtained.

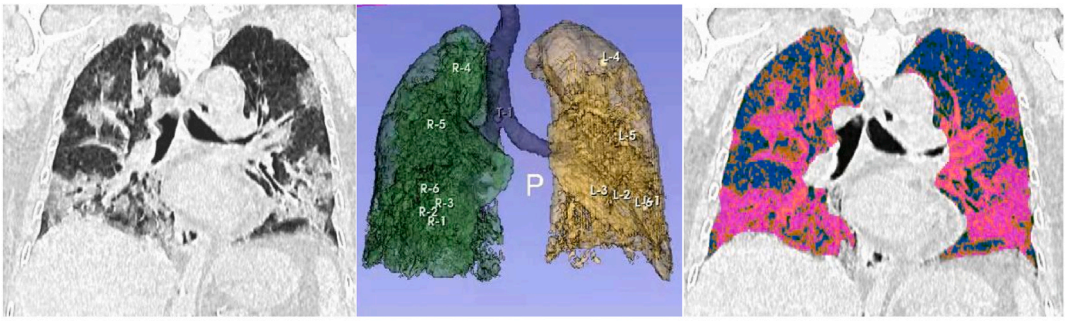

Fig. 1. Quantitative CT parameter assessment with 3D-Slicer software. This illustration pertains to an 88-year-old male patient with a critical case of COVID-19, characterized by extensive ground-glass and solid opacities across both lungs (1a). Advanced computer vision techniques are employed for a three-dimensional representation of the pulmonary abnormalities (1b). Lung segmentation is demonstrated in the image (1c), where the orange-yellow regions denote ground-glass density lesions, and the pink regions indicate solid density lesions. Abbreviations: GGOV: Ground-glass opacity volume, GGOV%: Ground-glass opacity volume percentage, SOV: Solid opacity volume, SOV%: Solid opacity volume percentage.

2.2. Clinical data

The study meticulously gathered baseline admission data of the patients, which included demographic information (gender, age), medical histories (hypertension, diabetes, coronary heart disease, chronic obstructive pulmonary disease), and comprehensive hematological profiles. Key biochemical markers were also evaluated, encompassing hypersensitive C-reactive protein (hs-CRP), creatine kinase (CK), creatine kinase isoenzyme-MB (CKMB), lactic dehydrogenase (LDH), α-hydroxybutyrate dehydrogenase (α-HBDH), D-dimer (D-D), and high sensitivity troponin (hs-TnI). Additionally, the investigation focused on short-term clinical outcomes, categorizing them into in-hospital mortality and patient discharge upon condition improvement. The analysis further extended to calculating critical hematological indices, namely the Neutrophil to Lymphocyte Ratio (NLR), Platelet to Lymphocyte Ratio (PLR), and Lymphocyte to Monocyte Ratio (LMR), providing a comprehensive insight into the patients' immunological status.

2.3. Measurement of quantitative CT parameters

The quantitative CT parameters were meticulously assessed using the "Chest Imaging Platform" module within the 3D-Slicer software, leveraging 1.5 mm thin-slice thoracic CT scans. This analysis was conducted independently by two seasoned radiologists, each with over a decade of experience (10 and 15 years, respectively, as shown in Fig. 1). The software was calibrated to automatically delineate lung areas affected by ground-glass and solid opacities through designated CT Hounsfield unit (HU) thresholds, specifically identifying ground-glass opacity regions (−700 HU to −300 HU) and solid opacity regions (−300 HU to 0 HU). This sophisticated tool enabled the precise quantification of several key parameters: the volume of ground-glass opacities (GGOV), the percentage of the lung volume represented by ground-glass opacities (GGOV%), the volume of solid opacities (SOV), the percentage of lung volume constituted by solid opacities (SOV%), the total volume of pulmonary lesions (LeV), and the percentage of total lung volume occupied by these lesions (LeV%).

Table 1
Comparison of clinical and quantitative CT data between the training and validation groups (n = 239).

Clinical and Quantitative CT Data	Training Group (n = 167)	Validation Group (n = 72)	t/Z/x^{2}Value	*P*-Value
Gender [n (%)]			0.008[a]	0.930
Male	110	47		
Female	57	25		
History of Hypertension [n (%)]			0.463[a]	0.496
Yes	106	49		
No	61	23		
History of Diabetes [n (%)]			0.825[a]	0.364
Yes	57	29		
No	110	43		
History of Coronary Heart Disease [n (%)]			0.768[a]	0.381
Yes	56	20		
No	111	52		
History of Chronic Obstructive Pulmonary Disease [n (%)]			0.017[a]	0.895
Yes	29	12		
No	138	60		
Age	79.2 ± 8.42	78.9 ± 8.57	0.292[b]	0.771
NLR	7.8(4.7 , 16.3)	7.3(4.1 , 18.1)	−0.527	0.598
PLR	241.4(150.0 , 375.0)	201.1(115.9 , 359.1)	−1.386	0.166
LMR	1.8(1.2 , 3.0)	1.9(1.2 , 3.0)	−0.284	0.777
hs-CRP (mg/L)	43.7(9.4 , 81.0)	57.1(16.7 , 81.7)	−0.642	0.521
CK (U/L)	87.0(49.0 , 168.0)	81.0(34.3 , 214.3)	−0.621	0.535
CKMB (U/L)	16.0(11.0 , 23.0)	17.0(11.3 , 25.0)	−0.527	0.598
LDH (U/L)	273.0(213.0 , 372.0)	269.0(205.3 , 357.8)	−0.908	0.364
HBDH (U/L)	192.0(151.0 , 255.0)	177.0(142.3 , 253.3)	−0.830	0.407
D-D (mg/L)	2.8(1.0 , 7.5)	1.9(0.9 , 4.6)	−1.781	0.075
hs-Tnl (ng/mL , $\times 10^{-2}$)	2.6(0.9 , 8.6)	2.4(0.8 , 9.4)	−0.328	0.743
HGB (g/L)	111.2 ± 24.28	117.3 ± 25.69	−1.754[b]	0.081
GGOV (mL)	524.6(395.7 , 673.7)	515.8(388.7 , 695.0)	−0.361	0.718
GGOV%	18.9(13.2 , 25.3)	18.1(13.1 , 27.2)	−0.201	0.841
SOV (mL)	220.6(161.6 , 341.1)	231.4(160.6 , 350.3)	−0.100	0.920
SOV%	7.8(5.3 , 12.2)	7.8(5.0 , 13.6)	−0.202	0.840
LeV (mL)	754.1(562.8 , 1003.7)	769.3(536.1 , 1025.8)	−0.075	0.940
LeV%	27.0(18.0 , 37.0)	26.0(18.3 , 40.0)	−0.189	0.850

The superscript 'a' next to the statistic indicates a χ2 value, 'b' indicates a t value, and the rest represent Z values.
Abbreviations:NLR:Neutrophil to Lymphocyte Ratio, PLR: Platelet to Lymphocyte Ratio, LMR: Lymphocyte to Monocyte Ratio, hs-CRP: Hypersensitive C-Reactive Protein,CK:Creatine Kinase, CKMB: Creatine Kinase Isoenzyme-MB, LDH: Lactic Dehydrogenase, HBDH: Hydroxybutyrate Dehydrogenase, D-D: D-Dimer, hs-TnI: High Sensitivity Troponin I, GGOV:Ground-glass opacity volume, GGOV%:Ground-glass opacity volume percentage, SOV: Solid opacity volume, SOV%:Solid opacity volume percentage, LeV:Lesion volume, LeV%:Lesion volume percentage.

2.4. CT scan parameters

CT scans were performed using a Toshiba Aquilion 16-slice spiral CT scanner. Patients were positioned supine and scanned during breath-hold after deep inspiration. The scanning range extended from the apex of the lungs to the lung base. The scanning parameters were as follows: tube voltage of 120 kV, automatic tube current modulation, pitch of 0.8, rotation time of 0.6 s per rotation, matrix size of 512 × 512, slice thickness of 5 mm, inter-slice gap of 5 mm, and field of view (FOV) of 40 cm × 40 cm. Axial reconstruction was performed with lung window settings (window width 1500 HU, window level −600 HU), and the reconstructed slice thickness was 1.5 mm.2.5 Statistical methods.

The analysis commenced with the Shapiro-Wilk test to determine the normality of continuous variables. Variables adhering to a normal distribution were evaluated using the *t*-test and are presented as $\bar{x} \pm s$. For those not conforming to normal distribution, the rank-sum test was employed, with results expressed as median M_{50} (25th percentile P_{25}, 75th percentile P_{75}). Categorical variables were examined using either the chi-square (χ^2) test or Fisher's exact test, as appropriate. In the case of variables demonstrating statistically significant differences (*P*-value< 0.05), a preliminary univariate logistic regression analysis was undertaken. Factors identified with a *P*-value <0.05 in this initial analysis were subsequently incorporated into a multivariate logistic regression to discern independent risk factors. The study utilized four distinct classification algorithms for model construction based on clinical data and quantitative CT parameters: LR, RF, DT, and SVM. The models' predictive performance was appraised using the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC). Additionally, the Intraclass Correlation Coefficient (ICC) was employed for assessing consistency. A *P*-value <0.05 was deemed indicative of statistical significance.

3. Results

3.1. Comparison of clinical and quantitative CT data between groups

In this study, a cohort of 239 patients with severe COVID-19 was enrolled. The training group included 167 patients categorized

Table 2
Comparison of clinical and quantitative CT data between the death and survival subgroups in the training group (n = 167).

Clinical and Quantitative CT Data	Death group(n = 42)	Survival group(n = 125)	t/Z/x^{2}Value	*P*-Value
Gender [n (%)]			0.016[a]	0.900
Male	28	82		
Female	14	43		
Age	81.0(78.0 , 87.0)	78.0(72.0 , 85.0)	−2.242	0.025
History of Hypertension [n (%)]			0.377[a]	0.539
Yes	25	81		
No	17	44		
History of Diabetes [n (%)]			0.772[a]	0.380
Yes	12	45		
No	30	80		
History of Coronary Heart Disease [n (%)]			0.168[a]	0.682
Yes	13	43		
No	29	82		
History of Chronic Obstructive Pulmonary Disease [n (%)]			13.165 a	< 0.001
Yes	15	14		
No	27	111		
NLR	13.7(7.8 , 23.8)	7.1(4.3 , 12.2)	−3.993	< 0.001
PLR	278.6(168.9 , 438.0)	226.6(147.7 , 332.5)	−1.265	0.206
LMR	1.5(0.9 , 2.5)	1.9(1.3 , 3.0)	−2.097	0.036
hs-CRP (mg/L)	78.4(30.1 , 127.9)	33.9(8.7 , 75.6)	−3.559	< 0.001
CK (U/L)	140.0(85.0 , 294.0)	77.0(44.8 , 142.0)	−3.305	0.001
CKMB (U/L)	21.0(13.0 , 29.0)	15.0(10.3 , 21.0)	−3.050	0.002
LDH (U/L)	381.0(280.0 , 595.0)	243.0(201.0 , 340.3)	−4.722	< 0.001
HBDH (U/L)	289.0(199.0 , 386.0)	177.5(144.8 , 230.0)	−5.051	< 0.001
D-D (mg/L)	6.0(2.6 , 9.8)	2.4(0.8 , 5.0)	−4.149	< 0.001
hs-Tnl (ng/mL , $\times 10^{-2}$)	10.3(2.7 , 28.1)	1.7(0.8 , 4.4)	−5.056	< 0.001
HGB (g/L)	108.9 ± 28.88	112.0 ± 22.55	−0.728b	0.468
GGOV (mL)	505.9(334.1 , 582.8)	525.0(408.3 , 682.2)	−1.490	0.136
GGOV%	19.8(12.5 , 29.1)	18.5(13.2 , 24.8)	−0.633	0.527
SOV (mL)	241.6(172.7 , 393.5)	213.4(161.5 , 309.1)	−0.904	0.366
SOV%	11.4(6.9 , 17.7)	7.7(5.3 , 11.1)	−3.186	0.001
LeV (mL)	747.1(512.2 , 1000.8)	755.6(576.3 , 1007.8)	−0.856	0.392
LeV%	32.0(16.0 , 43.0)	26.0(18.3 , 36.0)	−1.280	0.201

The superscript 'a' next to the statistic indicates a χ2 value, 'b' indicates a t value, and the rest represent Z values.
Abbreviations: NLR: Neutrophil to Lymphocyte Ratio, PLR: Platelet to Lymphocyte Ratio, LMR: Lymphocyte to Monocyte Ratio, hs-CRP: Hypersensitive C-Reactive Protein, CK: Creatine Kinase, CKMB: Creatine Kinase Isoenzyme-MB, LDH: Lactic Dehydrogenase, HBDH: Hydroxybutyrate Dehydrogenase, D-D: D-Dimer, hs-TnI: High Sensitivity Troponin I, GGOV: Ground-glass opacity volume, GGOV%: Ground-glass opacity volume percentage, SOV: Solid opacity volume, SOV%: Solid opacity volume percentage, LeV: Lesion volume, LeV%: Lesion volume percentage.

into 89 with severe and 78 with critical conditions. This group comprised 110 males and 57 females, ranging in age from 60 to 98 years, with a mean age of 79.2 ± 8.42 years. The validation group consisted of 72 patients, divided into 39 severe and 33 critical cases, with 47 males and 25 females, aged between 60 and 101 years, averaging 78.9 ± 8.57 years. Comparative analysis between the training and validation groups revealed no statistically significant differences ($P > 0.05$) as detailed in Table 1.

3.2. Intra-group comparison of clinical and quantitative CT data

The training group included 167 severe COVID-19 patients, categorized into 42 cases in the death group and 125 in the survival group. Comparative analysis between these subgroups revealed significant differences in age, COPD history, NLR, LMR, hs-CRP, CK, CKMB, LDH, HBDH, D-D, hs-TnI, and SOV% ($P < 0.05$). Other variables did not show significant differences ($P > 0.05$), as detailed in Table 2.

3.3. Risk factor analysis

Initially, age, history of chronic obstructive pulmonary disease, NLR, LMR, hs-CRP, CK, CKMB, LDH, HBDH, D-D, hs-TnI, and SOV% were included in a univariate logistic regression analysis. Subsequently, those indicators that exhibited statistically significant differences ($P < 0.05$) were incorporated into a multivariate logistic regression analysis to identify independent risk factors for predicting patient mortality. The analysis identified age, hs-CRP, and SOV% as key independent risk factors ($P < 0.05$) as presented in Table 3.

3.4. Construction and validation of machine learning models

Using age, hs-CRP, and SOV% as variables, predictive models for the short-term prognosis of patients with severe COVID-19 were developed employing four machine learning algorithms: LR, RF, DT, and SVM. The performance of these models in the training group was evaluated based on AUC, sensitivity, and specificity. The LR model demonstrated an AUC of 0.795 (0.726–0.865) with a sensitivity of 90.7 % and specificity of 62.9 %. The RF model showed an AUC of 0.726 (0.648–0.805) with 93.0 % sensitivity and 48.4 % specificity. The DT model achieved the highest AUC of 0.854 (0.794–0.915) coupled with 83.7 % sensitivity and 75.0 % specificity. In contrast, the SVM model exhibited the lowest AUC of 0.589 (0.485–0.694), a sensitivity of 72.1 % and specificity of 60.5 %. In the validation group, the LR model's AUC was 0.817 (0.701–0.934) with 78.9 % sensitivity and 83.0 % specificity. The RF model registered an AUC of 0.634 (0.462–0.806), 52.6 % sensitivity and 90.6 % specificity. The DT model outperformed others with an AUC of 0.869 (0.789–0.950), 84.2 % sensitivity and 71.7 % specificity, while the SVM model recorded an AUC of 0.754 (0.625–0.882), 73.7 % sensitivity and 75.5 % specificity as illustrated in Fig. 2a 2b. Notably, the SVM algorithm exhibited minimal discrimination ability in the training group. The DT algorithm emerged as the most effective, making its predictive model the most robust in this study. The DT algorithm was used as the output model for this study, and SHapley Additive exPlanations (SHAP) values were utilized for visualization, as shown in Fig. 2c.

4. Discussion

This study aimed to evaluate the effectiveness of combining clinical data and quantitative CT parameters with machine learning techniques to predict the short-term prognosis of elderly patients with severe COVID-19. The results showed that age, hs-CRP, and SOV % are independent predictors of mortality. Among the four machine learning algorithms, the Decision Tree algorithm performed the best in both the training and validation groups, with AUC values of 0.854 and 0.869, respectively, demonstrating high predictive

Table 3
Analysis of risk factors for predicting mortality during hospitalization in patients with severe COVID-19.

Clinical and Quantitative CT Data	Univariate	*P*-Value	Multivariate	*P*-Value
	OR(95%CI)		OR(95%CI)	
Age (years)	1.054(1.008–1.102)	0.022	1.121(1.049–1.198)	0.001
History of Chronic Obstructive Pulmonary Disease	4.209(1.820–9.733)	0.001	2.579(0.924–7.200)	0.070
NLR	1.024(1.005–1.043)	0.015	1.014(0.991–1.038)	0.221
LMR	0.784(0.588–1.046)	0.098	–	–
hs-CRP(mg/L)	1.012(1.005–1.018)	< 0.001	1.009(1.001–1.017)	0.020
CK(U/L)	1.001(1.000–1.002)	0.079	–	–
CKMB(U/L)	1.014(0.996–1.032)	0.120	–	–
LDH(U/L)	1.004(1.002–1.006)	< 0.001	0.998(0.993–1.003)	0.474
HBDH(U/L)	1.005(1.002–1.008)	< 0.001	1.006(0.998–1.013)	0.124
D-D(mg/L)	1.037(1.007–1.067)	0.015	1.029(0.995–1.065)	0.100
hs-Tnl(ng/mL)	1.001(1.000–1.002)	0.245	–	–
SOV%	1.105(1.045–1.168)	< 0.001	1.097(1.019–0.181)	0.014

Abbreviations:NLR: Neutrophil to Lymphocyte Ratio, LMR: Lymphocyte to Monocyte Ratio, hs-CRP: Hypersensitive C-Reactive Protein, CK: Creatine Kinase, CKMB: Creatine Kinase Isoenzyme-MB, LDH: Lactic Dehydrogenase, HBDH: Hydroxybutyrate Dehydrogenase, D-D: D-Dimer, hs-TnI: High Sensitivity Troponin I, SOV%: Solid opacity volume percentage.

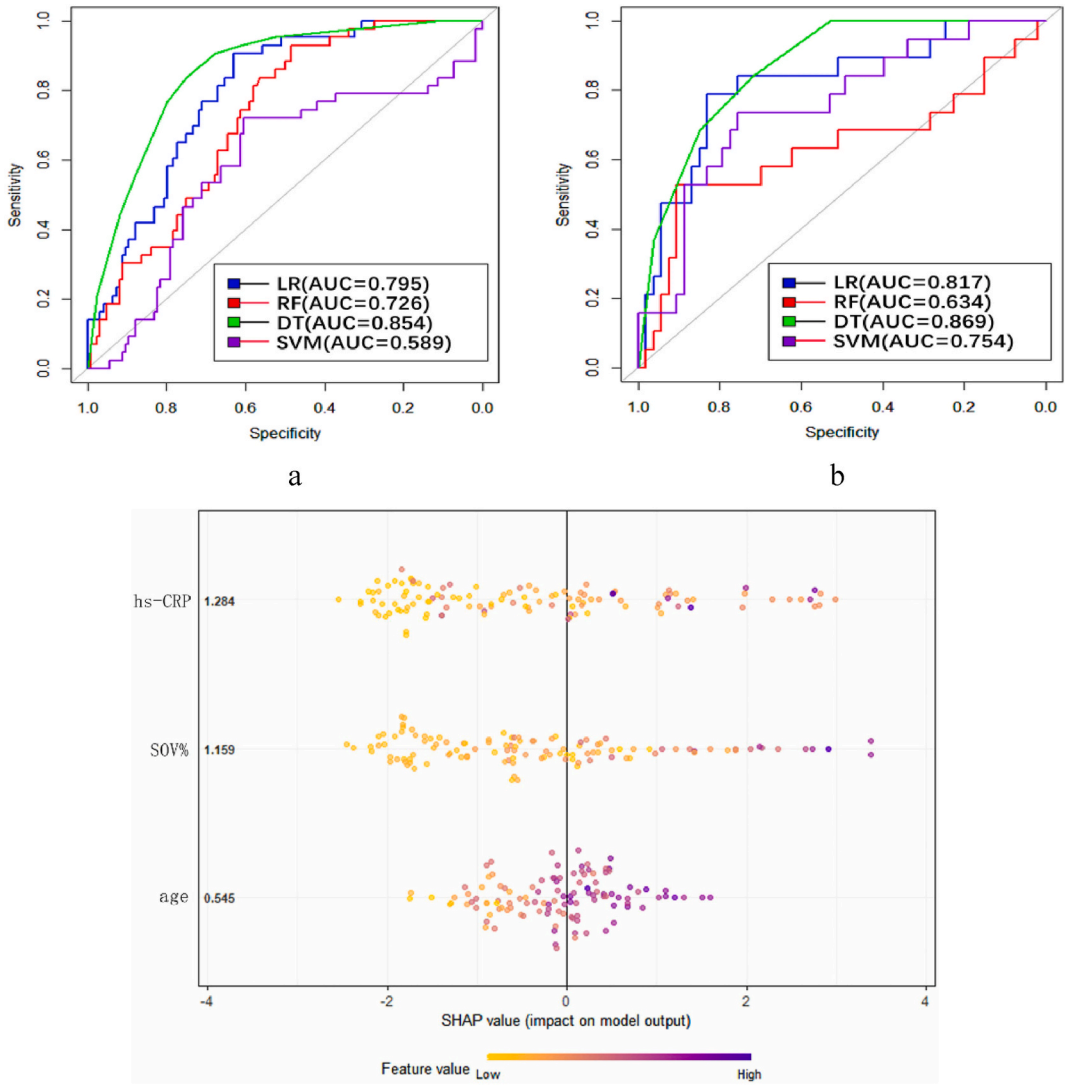

Fig. 2. a 2b Roc curves of predictive models constructed by four machine learning algorithms: a) training group, b) validation group Abbreviations: LR:Logistic Regression, RF: Random Forest, DT:Decision Tree, SVM:Support Vector Machine.
cA scatter plot of SHAP values illustrates the contribution of each feature to the model and its impact on the model's output. Y-axis: Each row represents a feature. In this plot, there are three features: hs-CRP, SOV%, and age. X-axis: SHAP values, which indicate the magnitude of each feature's contribution to the model's output. Positive values mean the feature pushes the prediction higher, while negative values mean the feature pushes the prediction lower. Color of the dots: Indicates the size of the feature values. Yellow represents higher feature values, and purple represents lower feature values. Abbreviations: SHAP: SHapley Additive exPlanations.

accuracy and reliability. These findings indicate that integrating clinical data and quantitative CT parameters with machine learning techniques can effectively predict the short-term prognosis of elderly patients with severe COVID-19.

COVID-19, known for its high transmissibility, poses significant threats to society. Although initial infection can trigger antibody production, neutralizing antibody levels typically reach their peak within one month post-infection and subsequently decline, thus maintaining a risk of reinfection over time [6]. The virus can affect multiple organs upon entry into the human body, particularly in severe cases, potentially leading to multi-organ failure [7]. Notably, in the early stages of the disease, individuals aged ≥60 exhibit significantly lower serum neutralizing antibody levels compared to those under 60, correlating strongly with increased severity and mortality rates [8]. Consequently, the elderly population faces heightened risks of severe outcomes and mortality. Our study specifically targets elderly severe COVID-19 patients aged ≥60, aiming to analyze risk factors associated with mortality. It was found that advanced age, elevated levels of several laboratory indicators (NLR, LMR, hs-CRP, CK, CKMB, LDH, HBDH, D-D, hs-TnI), and an increase in the quantitative CT parameter SOV% are all linked to an elevated mortality risk. Among these, advanced age, heightened hs-CRP, and increased SOV% are identified as independent risk factors for mortality during hospitalization.

Consistent with prior research, our study corroborates age as a pivotal independent risk factor for mortality in COVID-19 patients

[9]. Specifically, in our cohort of severe COVID-19 patients aged 60 and above, advanced age was significantly associated with increased mortality. The median age in the deceased group was 81 years, surpassing that of the surviving group, which was 78 years. The elderly, particularly vulnerable due to weakened immunity and a higher prevalence of comorbidities such as hypertension, diabetes, coronary heart disease, and chronic pulmonary diseases, are at an escalated risk of fatal outcomes. In the context of COVID-19, the infection triggers a cytokine storm, where NLR and LMR serve as critical markers of systemic inflammation. Existing literature substantiates the utility of NLR and LMR in predicting adverse outcomes across a spectrum of diseases [10,11]. During severe infections, an abnormal surge in neutrophil activation leads to unbalanced inflammatory responses, inflicting damage through two primary mechanisms: organ dysfunction caused by the inflammatory storm and impaired viral clearance due to reduced lymphocyte counts, culminating in increased mortality risk. Accordingly, our findings revealed higher median NLR and lower LMR values in the death group than in the survival group. Interestingly, NLR did not emerge as an independent predictor of mortality in this study, deviating from previous reports [12], potentially attributable to our focus on the elderly demographic. Research has shown [13] that early-stage infectious diseases provoke IL-6 production from Toll-like receptor-stimulated monocytes, tumor necrosis factor, and macrophages, elevating hs-CRP levels and reflecting the severity of the inflammatory storm [14]. Numerous studies have identified hs-CRP as a critical independent marker for foreseeing severe COVID-19 complications [15,16]. Furthermore, COVID-19's impact extends beyond the respiratory system, with a proportion of patients experiencing cardiac involvement, a factor closely linked to mortality rates [17]. In our study, the death group exhibited markedly higher levels of cardiac biomarkers (CK, CKMB, LDH, HBDH, and hs-TnI) than the survival group, aligning with previous findings. Additionally, the elevated D-Dimer levels observed in the death group suggest the formation of microthrombi secondary to an acute inflammatory cytokine storm, indicating progressive disease and an increased risk of respiratory failure [18].

CT scanning stands as the paramount and most efficacious imaging technique for the diagnosis and evaluation of COVID-19 severity, primarily identifying lung pathologies such as ground-glass opacities and consolidations [19]. Historically, CT evaluations of COVID-19 have predominantly concentrated on the qualitative and semi-quantitative aspects of these pulmonary abnormalities [20,21]. In this study, we employed advanced automatic segmentation software to derive quantitative CT parameters for COVID-19, facilitating a more objective and accurate assessment of the lung lesion severity. Our findings indicated that the death group exhibited a notably higher median SOV% compared to the survival group. Significantly, SOV% was identified as an independent risk factor for mortality in patients with severe COVID-19. The escalation of SOV% in more severe cases can be attributed to the accumulation of cellular fibromucinous exudates, inflammatory cells, and hemorrhage in the alveolar spaces and septa, leading to an increased solid component in the lesions, thus explaining the elevated SOV% observed in the death group [22].

This study faces several limitations. Primarily, the case selection demonstrates geographic limitations, resulting in a lack of external validity for our conclusions. Additionally, being a retrospective analysis, the study is potentially subject to selection biases. Moreover, the link between quantitative CT parameters in COVID-19 patients and their pathological mechanisms is yet to be corroborated by related pathological studies.

This study demonstrates that combining clinical data and quantitative CT parameters with machine learning techniques can predict, to some extent, the short-term prognosis of elderly patients with severe COVID-19. The results indicate that age, hs-CRP, and SOV% are independent predictors of mortality. The Decision Tree algorithm performed well, exhibiting high predictive accuracy and reliability. Although these findings provide valuable clinical insights, the study's limitations, including a small sample size and lack of external validation, necessitate further research to verify and expand upon these results.

Ethical statement

This retrospective study involving human participants met the ethical standards of the institutional and/or national research committee and with the Helsinki Declaration of 1975, as revised in 1983 or comparable ethical standards, was reviewed and approved by the Ethics Committee of Wannan Medical College(IRB No. 6, 2024) and obtained informed consent exemption.

Data availability

All of the data in this study are available from the corresponding authors.

Funding

This work was supported by Bengbu Medical University (Grant No. KFKT202307); the Education Department of Anhui Province (Grant No. 2023AH051743); Wannan Medical College (Grant No. WK202213) ;Wannan Medical College (Grant No. WK2023ZQNZ53) and Wannan Medical College (Grant No.WK2023ZZD09).

CRediT authorship contribution statement

Lifang Fan: Writing – original draft, Funding acquisition, Data curation, Conceptualization. **Shujian Wu:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yimin Wu:** Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Xiaoyan Xu:** Investigation, Conceptualization. **Zhengyuan Xu:** Writing – review & editing, Methodology. **Lei Huang:** Methodology. **Guoxian Chen:** Writing – review & editing, Visualization, Resources, Project administration, Funding acquisition, Data curation.

Declaration of competing interest

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

Acknowledgements

Thank you to all the friends who provided help and support during the data collection process of the thesis.

References

[1] F. Wu, M. Liu, A. Wang, et al., Evaluating the association of clinical characteristics with neutralizing antibody levels in patients who have recovered from mild COVID-19 in Shanghai, China, JAMA Intern. Med. 180 (10) (2020) 1356–1362.
[2] Z. Wu, J. Mcgoogan, 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 Intern. Med. 323 (13) (2020) 1239–1242.
[3] L.I. Nikolaeva, M.D. Stuchinskaya, A.V. Dedova, et al., Association of polymorphic variants of hemostatic system genes with the course of COVID-19, Vopr. Virusol. 68 (5) (2023) 445–453.
[4] G. Grasselli, M. Greco, A. Zanella, et al., Risk factors associated with mortality among patients with COVID-19 in intensive care units in lombardy, Italy, JAMA Intern. Med. 180 (10) (2020) 1345–1355.
[5] Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 9), Journal of Cardiovascular and Pulmonary Diseases 41 (5) (2022) 449–457.
[6] Š. Miroševič, J. Prins, N. Bešić, et al., Symptoms of anxiety and depression among slovenian breast cancer survivors post-treatment during the COVID-19 pandemic: a cross-sectional study, Zdr Varst 63 (1) (2024) 55–62.
[7] M. Yamamoto, Y. Kato, J. Takeuchi, Hand function and quality of life in patients with diabetes mellitus before and during the COVID-19 pandemic, Nagoya J. Med. Sci. 85 (4) (2023) 659–667.
[8] B. Bošnjak, S. Stein, S. Willenzon, et al., Low serum neutralizing anti-SARS-CoV-2 S antibody levels in mildly affected COVID-19 convalescent patients revealed by two different detection methods, Cell. Mol. Immunol. 18 (4) (2021) 936–944.
[9] S. Mustafić, E. Jusufović, F. Hukić, et al., Early predictors of severity and mortality in COVID-19 hospitalized patients, Med. Glas. 18 (2) (2021) 384–393.
[10] S. Ha, S. Choi, S. Park, et al., Prognostic effect of preoperative neutrophil-lymphocyte ratio is related with tumor necrosis and tumor-infiltrating lymphocytes in hepatocellular carcinoma, Virchows Arch. 477 (6) (2020) 807–816.
[11] N. Adamstein, J. Macfadyen, L. Rose, et al., The neutrophil-lymphocyte ratio and incident atherosclerotic events: analyses from five contemporary randomized trials, Eur. Heart J. 42 (9) (2021) 896–903.
[12] K. Kerboua, NLR: a cost-effective nomogram to guide therapeutic interventions in COVID-19, Immunol. Invest. 50 (1) (2021) 92–100.
[13] W.J. Yang, J.Y. Yoon, Dental trauma trends in emergency care:a comparative analysis before, during, and after COVID-19, J Korean Assoc Oral Maxillofac Surg. 49 (6) (2023) 339–346.
[14] M. Del Giudice, S. Gangestad, et al., Rethinking IL-6 and CRP: why they are more than inflammatory biomarkers, and why it matters, Brain Behav. Immun. 70 (2018) 61–75.
[15] H. Gozdas, S. Kayis, T. Damarsoy, et al., Multi-inflammatory index as a novel mortality predictor in critically ill COVID-19 patients, J. Intensive Care Med. 37 (11) (2022) 1480–1485.
[16] S. Prasad, S. Patel, A. Behera, et al., Influence of laboratory biomarkers on inflammatory indices for assessing severity progression in COVID-19 cases, J Family Med Prim Care 11 (5) (2022) 2056–2072.
[17] A. Perks, P. Brendt, Pericardiocentesis for COVID-19 associated cardiac tamponade using a central venous catheter in rural Australia: a case report, Air Med. J. 43 (1) (2024) 63–65.
[18] A.V. Fetyukov, D.S. Pshennikov, S.E. Zaitseva, Olfactory, gustatory and auditory disorders associated with COVID-19. A systematic review of their distribution and pathogenesis, Vestn. Otorinolaringol. 88 (6) (2023) 61–66.
[19] D. Sun, X. Li, D. Guo, et al., CT quantitative analysis and its relationship with clinical features for assessing the severity of patients with COVID-19, Korean J. Radiol. 21 (7) (2020) 859–868.
[20] E. Dumea, E.C. Barbu, C.E. Chiţu, et al., Clinical, biochemical and pulmonary CT imaging features for hepatobiliary involvement in COVID-19, Germs 13 (2) (2023) 121–129.
[21] A. Bidari, E. Zarei, M. Hassanzadeh, et al., Development of a scoring method based on a chest CT scan to determine the outcomes of COVID-19 patients, Cureus 15 (10) (2023) e47354.
[22] F. Ufuk, M. Demirci, E. Uğurlu, et al., Evaluation of disease severity with quantitative chest CT in COVID-19 patients, Diagn Interv Radiol 27 (2) (2021) 164–171.