

81109_Auto_Edited.docx

Name of Journal: *World Journal of Clinical Cases*

Manuscript NO: 81109

Manuscript Type: ORIGINAL ARTICLE

Observational Study

Severe/critical COVID-19 early warning system based on machine learning algorithms using novel imaging scores

Li QY *et al.* COVID-19 warning system

Qiu-Yu Li, Zhuo-Yu An, Zi-Han Pan, Zi-Zhen Wang, Yi-Ren Wang, Xi-Gong Zhang, Ning Shen

Abstract

BACKGROUND

Early identification of severe/critical coronavirus disease 2019 (COVID-19) is crucial for timely treatment and intervention. Chest computed tomography (CT) score has been shown to be a significant factor in the diagnosis and treatment of pneumonia, however, there is currently a lack of effective early warning systems for severe/critical COVID-19 based on dynamic CT evolution.

AIM

To develop a severe/critical COVID-19 prediction model using a combination of imaging scores, clinical features, and biomarker levels.

METHODS

This study used an improved scoring system to extract and describe the chest CT characteristics of COVID-19 patients. The study also took into consideration the general clinical indicators such as dyspnea, oxygen saturation, alternative lengthening of

telomeres (ALT), and androgen suppression treatment (AST), which are commonly associated with severe/critical COVID-19 cases. The study employed lasso regression to evaluate and rank the significance of different disease characteristics.

RESULTS

The results showed that blood oxygen saturation, ALT, IL-6/IL-10, combined score, ground glass opacity score, age, crazy paving mode score, qsofa, AST, and overall lung involvement score were key factors in predicting severe/critical COVID-19 cases. The study established a COVID-19 severe/critical early warning system using various machine learning algorithms, including XGBClassifier, Logistic Regression, MLPClassifier, RandomForestClassifier, and AdaBoost Classifier. The study concluded that the prediction model based on the improved CT score and machine learning algorithms is a feasible method for early detection of severe/critical COVID-19 evolution.

CONCLUSION

The findings of this study suggest that a prediction model based on improved CT scores and machine learning algorithms is effective in detecting the early warning signals of severe/critical COVID-19.

Key Words: COVID-19; Clinical prediction model; Electron computed tomography; Machine learning.

Li Q, An ZY, Pan ZH, Wang ZZ, Wang YR, Zhang XG, Shen N. Severe/critical COVID-19 early warning system based on machine learning algorithms using novel imaging scores. *World J Clin Cases* 2023; In press

Core Tip: The computed tomography (CT) score is a relatively objective and clinically accessible semiquantitative assessment tool for patients with coronavirus disease 2019

(COVID-19). The CT scores of common, severe, and critically ill patients showed different trends, and there were differences between the groups of patients as the disease progressed. Patients who are recovering from the disease can be monitored *via* CT at reduced intervals to reduce their radiation exposure and financial burden. The 2 wk CT scores of the patients were important for predicting disease deterioration in hospitalized patients who have an average admission severity rating. The qSOFA score, aspartate aminotransferase, oxygenation, and dyspnea were important for the prediction of severe/critical COVID-19 disease.

INTRODUCTION

⁶ Since the outbreak of novel coronavirus pneumonia (COVID-19) in late December 2019, approximately 521 million more patients worldwide have been infected as of late May 2022, and over 6 million patients could have died from COVID-19. Some COVID-19 patients are not critical at the time of initial diagnosis, but their conditions may deteriorate, leading to severe illness. Currently, case reports in China show that most COVID-19 patients have mild disease, but 15% and 5% develop severe disease and become critically ill, respectively. The intensive care unit (ICU) mortality rate for critically ill patients is as high as 50%-60%. Hence, ¹ early identification of the warning signs of severe/critical COVID-19 cases and prompt intervention and treatment can help reduce mortality and improve cure rates^[1,2]. At present, CT chest examination is an important part of ⁸ the diagnosis and treatment of COVID-19 pneumonia and has been used to diagnose patients clinically. Its significance in indicating patient deterioration has been confirmed. However, there is still a lack of means to predict the early progression of severe/critical cases based on the dynamic evolution of CT, and how to determine the progression of CT lesions objectively and quantitatively has become an urgent clinical issue that needs to be addressed^[3,4]. During the progression of COVID-⁹ 19, the occurrence of an inflammatory storm plays a crucial role and the alteration of related inflammatory factors can directly damage the pulmonary capillary mucosa, promoting alveolar edema and inactivating surface active proteins, leading to diffuse

4
alveolar damage and ventilation dysfunction. Inflammatory storms are also a major cause of acute respiratory distress syndrome and multiple organ dysfunction syndrome^[5,6]. Numerous prognostic studies on COVID-19 have shown that the trends in related cytokines are crucial for early identification and treatment of critical cases^[7,8]. Therefore, exploring the trends in biomarkers, such as cytokines, is important for further understanding the mechanisms of disease progression and regression in COVID-19 patients^[9].

Identifying patients at the time of initial diagnosis and providing timely and aggressive interventions are currently the main challenges in treating COVID-19. There are several COVID-19 warning and scoring systems^[9], such as the MuLBSTA scoring system, a mortality risk prediction model for COVID-19 that integrates four ML algorithms^[10], 11 the National Early Warning Score and Rapid Emergency Medicine Score for ER COVID-19 patients, and NEWS+age^[11,12]. To optimize patient treatment and recovery using limited healthcare resources, early detection of prognostic biomarkers is crucial in distinguishing patients who may develop severe COVID-19 and assessing their associated mortality risk during a global pandemic. A model that combines multiple variables for early prediction of the prognosis in COVID-19 patients would help allocate healthcare resources effectively and reduce mortality. In this study, we aim to develop a severe/critical COVID-19 prediction model based on imaging scores, patient clinical features, and biomarker levels (Table 1).

MATERIALS AND METHODS

Study design and study population

A retrospective analysis was used to examine the medical records of adult COVID-19 patients who were admitted to the infectious disease wards of three medical centers in Beijing, Wuhan, and Nanchang. The retrospective cohort collected primary clinical data and CT imaging data, as well as collected the patients' clinical data, including the patients' demographic data, medical history, laboratory tests after admission, inflammatory markers and cytokine levels, and CURB65 scores. All data were collected

and obtained from the electronic medical record system, and if the electronic medical record system lacked relevant data records, these records were obtained by communicating with the attending physicians. This study was reviewed by the Ethics Committee of Peking University Third Hospital.

Case diagnostic criteria and clinical staging criteria

The case diagnostic criteria and clinical typing criteria included the following. All of the deceased patients were confirmed to have COVID-19 based on the pneumonia diagnosis and treatment protocol for novel coronavirus infection (trial version 7) that was issued by the National Health and Wellness Commission. 10 Confirmed cases were required to conform to the clinical manifestations in the diagnosis and treatment protocols, and pharyngeal swabs, sputum, and lower respiratory secretions were tested by real-time fluorescence RT-PCR using the Wuhan Centers for Disease Control and Prevention's 2019-nCoV nucleic acid test. The clinical typing criteria included the following: (1) Light: The clinical symptoms are mild, and no pneumonia manifestations are seen on imaging; (2) Common type: Fever and respiratory symptoms are present, and pneumonia is visible on imaging; (3) Severe disease: Any of the following are present: respiratory distress, a respiratory rate ≥ 30 breaths/min; a resting state oxygen saturation of $\leq 93\%$; and a $\text{PaO}_2/\text{FiO}_2 \leq 300$ mmHg (1 mmHg = 0.133 kPa); and (4) Critical type: One of the following conditions is present: Respiratory failure requiring mechanical ventilation; shock, which may combined with other organ failure that requires ICU monitoring and treatment.

Chest CT scoring

The CT images were independently interpreted by two emergency physicians with more than 10 years of experience, and the CT images were interpreted based on the Fleischner Society definition^[13]. The training of the artificial intelligence was by respiratory and radiology specialists. If the scores were inconsistent, they were reassessed, and if an agreement could not be reached, the closer score was used to

calculate the mean value. A new scoring system was developed based on the previous lung CT severity score 12, which is widely used in patients with interstitial lung disease, and this scoring system was modified by experts from the departments of respiratory medicine, radiology, critical care medicine, and emergency medicine: (1) The partitioning of the lung field was performed as follows. The lung field was divided into upper, middle, and lower parts, and a total of 6 regions were portioned on the left and right sides by the plane of the tracheal bulge and the plane of inferior pulmonary veins; (2) The target lesion types were defined as follows: (a) Ground glass opacity (GGO): Widespread, blurred increased density of the lung parenchyma with visible bronchial and vascular textures; (b) Pulmonary solidity: Uniformly increased density of the lung parenchyma, obscuring the bronchial and vascular shadows within it, with bronchial inflation signs; (c) Paving stone sign: Ground glass opacity combined with lobular septa; and (d) Pavement stone sign: The combination of lobular septum thickening with a ground glass shadow. The other types of lesions were not analyzed because they were relatively rare and were noncharacteristic^[6]; (3) Scoring: The overall extent of the involvement of the lesions, the extent of the ground glass shadows, and the extent of solid lesions in each region were scored separately: (a) 0: Normal lung tissue; (b) 1: The extent of the lesions is < 25%; (c) 2: The extent of lesions is 26%-50%; (d) 3: The extent of the lesions is 51%-75%; and (e) 4: The extent of the lesions is > 75%. The regional scores were accumulated (the total scores for the overall extent of the involvement, ground glass shadows, and solid lesions were 0-24 for each region). The total score of each was 0 to 24 points. The pavement stone sign was scored as 0 or 1 for the presence or absence of this sign, and the 6 area scores were also cumulative (total score 0 to 6)^[13].

To understand the changes in the CT scores over time (time series), the median interval time between the recheck lung CTs was used as a cutoff point to segment the disease course time axis and to compare the differences in the lung CT scores among patients with common, severe, and critical disease at different periods within the course of the disease. The mean values of the 2 CT scores were taken within the same time period. To understand the spatial distribution characteristics (spatial sequence) of the

intrapulmonary lesions, the lung fields were divided based on the aforementioned methods to compare the differences in the CT scores in the upper, middle and lower lung field regions (Table 2).

Statistical methods

Cronbach's alpha index was used to evaluate the reliability of the scores of the two reviewers. The measurement data are expressed as the median and interquartile range, and the count data are expressed as the frequency (percentage). The Mann-Whitney U test was used to evaluate the continuous variables, and the chi-square test or Fisher exact test was used to evaluate the rank variables. The Wilcoxon test or the Friedman test was used to compare two/multiple samples of interest. To evaluate the predictive validity of the CT scores, receiver operating characteristic (ROC) curves were traced, and the area under the curve (AUC) was calculated. The statistical analysis was performed using SPSS 22.0 software. MedCalc 19.1 software was used for the comparison of the AUCs of the ROC curves. All of the tests were two-sided; the significance level was set at $P < 0.05$.

RESULTS

General clinical characteristics

A cohort of 153 COVID-19 patients was included in the study, with 105 patients in severe or critical condition and 48 with common disease. The median age of the patients was 89 years and 50.98% were male. The results of the chi-square test, Mann-Whitney U test, and t test showed that there was no significant difference between the two groups in terms of basic information and vital signs, such as age, sex, heart rate, blood pressure, maximum body temperature, and body mass index. However, the respiratory rate and oxygenation index were significantly different. No significant difference was found between the two groups in terms of underlying diseases, such as coronary artery disease, pulmonary disease, and diabetes mellitus. Clinical symptoms were not significantly different between the two groups, except for dyspnea. Significant

differences were found between the two groups in terms of absolute neutrophil values, hemoglobin, liver function tests, albumin, urea, blood coagulation tests, CRP, ferritin, glucose, and T3. The qSOFA score and CURB-65 score were also significantly different between the groups. Significant differences were also found in the characteristics of lung CT scores for patients with different types of COVID-19. Of the 110 patients who underwent CT scans, 114 (94.6%) had GGOs, 68 (61.9%) had solid lesions, and 43 (39.1%) had paving stone signs. Bilateral involvement was seen in 103 (93.6%) of the patients. The CT scans were performed at intervals of 3 to 25 d, with a median interval of 7 d. Disease progression was segmented by week, with the CT scores of common and severe patients peaking during the 3rd to 4th week, while the CT scores of critically ill patients progressed more rapidly, peaking during the 2nd to 3rd week. During the first week, only the GGO score differed between common and critical patients. From the second week onward, the GGO score, solid lesion score, paving stone sign score, and overall extent of involvement score differed between the patients with different degrees of criticality. The spatial distribution of the lesions showed that the upper lung region involvement extent scores were lower for common and severe patients compared to the middle and lower lung regions, while for critically ill patients, there was no statistically significant difference in the CT scores of the upper, middle, and lower lung regions.

The machine learning-based risk factor mining for severe/critical COVID-19 disease

To explore the possible risk factors for severe/critical COVID-19 disease, this study used Lasso regression for the variable importance analysis based on machine learning, and the model parameters were as follows: cv (cross-validation fold): 10; max_iter (number of iterations): 1000; tol (convergence measure): 0.0001; alpha (L1 regularization factor): 0.01. By Lasso regression, the 10 variables with the highest importance (from highest to lowest) were found to be the blood oxygen saturation, ALT, IL-6/IL-10, consolidation score, GGO score, age, crazy-paving pattern score, qSOFA, AST, and overall lung involvement score (Figure 1).

The development and validation of a machine learning-based clinical prediction model for severe/critical COVID-19 disease

A combination of the results of variable importance analysis and other factors, such as clinical experience, age, IL-6/IL10 Levels, ALT levels, oxygen saturation, qSOFA score, and consolidation score, were selected to be included in the model. A classification task for the data sample was conducted using several machine learning models, including XGBClassifier, RandomForestClassifier, LogisticRegression, LGBMClassifier, and MLPClassifier. A forest plot displays the ROC results of each model for the prediction of severe/critical COVID-19 cases, with the error lines representing the mean and standard deviation of the ROC (Table 3). The means and standard deviations of the ROC were computed by repeating the sampling five times, with each resampled training set accounting for 20% of the overall sample and 80% of the training set. Among the models tested, the best performer in the validation set was the RandomForestClassifier (sorted by AUC), while the best performer in the test set was the LogisticRegression (sorted by AUC). The performance of the algorithms was inconsistent between the training and validation sets. The RandomForestClassifier was more prone to overfitting, while the LogisticRegression appeared to be relatively more stable and was ultimately chosen for the final modeling (Figure 2).

The machine learning models' performances for predicting severe/critical COVID-19 cases

The calibration plots also confirmed good consistency between the “LogisticRegression” algorithm that was predicted and the observed actual risk of severe/critical COVID-19 (Figure 3). A further decision curve analysis was performed, and this analysis showed that the present model provides an excellent net benefit when the clinical decision threshold is between 0% and 100% (Supplementary Figure 1).

DISCUSSION

Most of the warning scoring systems only use the clinical measurements, such as the level of consciousness and the patient's vital signs or laboratory testing and do not include the imaging features^[1,9,10,12,14]. Since SARS-CoV-2 mainly invades the lungs, a lung CT can often be the best indicator for the severity of the disease and can provide further guidance for the development and prognosis of COVID-19 disease. Several studies have used the CT findings as important indicators to assess the prognosis of COVID-19 patients^[15-21]. The prognosis of COVID-19 patients has been predicted by scoring factors, such as the presence of gross glassy shadows on CT images, bronchial and pleural involvement, the presence of a lobular septum, and the lesion morphology, distribution, and size^[18,19]. The present study differs from previous studies because the current study used a modified CT severity score that had been previously used in patients with interstitial lung disease, and this score has also been used in patients with COVID-19. This study was the first study that adjusted for complex anatomical localizations by selecting easily recognizable anatomical landmarks for the partitioning and simplification. In this study, the main manifestations of the lung CTs in patients with COVID-19 included GGOs in both lungs, thickening of the lobular septa in the form of the "pavement sign", solid changes, fibrous cords, *etc.* Multiple signs could exist simultaneously, and a "white lung" could be seen in severe cases. In contrast, pleural effusion, lymph node enlargement, cavitation, and nodular lesions were less common. Therefore, only the typical lesion types, such as GGOs, solid lesions and paving stone signs, were selected for the modified score, and this simplified the lesion types. Furthermore, the modified imaging score was not time-consuming, and the reliability of the two reviewers' scores for each type of lesion was high. Thus, the improved scoring increased the ability of clinicians to make rapid judgments. In addition, the overall extent of the involvement score downplays the identification of the lesion type, and the upper lung regional involvement score narrows the target area. Both of these scores have a statistically similar predictive validity, as compared to the GGO score, and both of the former scores reduce the complexity and sensitivity of the assessment. Therefore, these novel scores could be a relatively simple alternative to the GGO score. Regarding

the predictive value of the lung CT score, this study found that a GGO score of more than 5 on the lung CT at week 2 could be used as a significant indicator for severe COVID-19, even before the development of oxygen reserve depletion, clinical decompensation, and sudden deterioration. In conclusion, clinicians should perform a comprehensive evaluation in COVID-19 patients, which should be used in conjunction with the patient's lung CT score, in order to provide enough respiratory support to the patient as early as possible. The inclusion of the CT score in the clinical severity grading criteria may be considered in the future. In addition to the imaging findings, a variety of clinical features and laboratory indicators have been included in the models of different studies to develop appropriate predictive models. These factors mainly included age, sex, the patient's vital signs (temperature, arterial systolic blood pressure, respiration, heart rate), and the patient's laboratory indicators (neutrophil count, platelet count, CRP, arterial partial pressure of oxygen, blood creatinine value, eGFR, serum albumin value)^[16,17,20,22,23]; in addition, the patients' underlying diseases (hypertension, COPD, *etc.*) have also been included in some risk prediction models^[17,23]. In this study, qSOFA, aspartate aminotransferase, oxygenation, and dyspnea were found to have a good predictive value for the prediction of patients with severe/critical COVID-19.

The qSOFA score is an important diagnostic tool for organ failure. Current studies on the role of inflammatory factors in organ failure have suggested that, compared with COVID-19 patients who were not admitted to the ICU for treatment, critically ill COVID-19 patients in the ICU have higher levels of IP-10, macrophage inflammatory protein 1A (MIP-1A), serum granulocyte colony-stimulating factor (G-CSF), G-CSF and TNF- α expression levels, suggesting a positive correlation between the inflammatory storm and the disease severity^[5,24,25]. It is known that the levels of cytokines, including interleukin, play a crucial role in the progression of COVID-19. During an inflammatory storm, a sustained increase in the expression of proinflammatory factors, produced by the body's immune system, can exacerbate the disease progression, while anti-inflammatory factors can promote pathogen clearance and tissue repair. Monitoring the

levels of both pro- and anti-inflammatory cytokines early in the course of COVID-19 is important for determining the patient's condition, treatment plan, and prognosis. Multiple studies have shown that the trends of relevant cytokines are essential for early identification and treatment of critical COVID-19 cases^[5,6,24,26]. Above all, this study not only simplifies the complex anatomical location of the lesions that are seen on lung CT but also combines the findings on medical imaging with the patient's clinical features and laboratory indicators, and the inclusion of all of these factors will more accurately predict the prognosis of COVID-19 patients during the early course of the disease. In this study, there are several limitations to consider. Firstly, it is a retrospective study, which may introduce bias in the results and difficulties in the statistical analysis due to the absence of CT data from the first examination and from critical patients. Secondly, not all of the CT scans included in the analysis were high-resolution scans, which could affect the accuracy of the readings. Finally, the low proportion of deaths in the sample size limited the ability to analyze the predictive value of CT on in-hospital patient outcomes. To address these limitations, future studies could consider expanding the sample size and conducting prospective studies. In the future, it is necessary to evaluate data and perform statistics from multiple medical centers to further evaluate adult COVID-19 confirmed cases, to establish a "COVID-19 clinical-imaging database", and to systematically analyze the patients' clinical information, laboratory tests and imaging data of admitted and discharged patients with the help of imaging and histological analysis methods. These methods can also help to more accurately assess the lesion progression, establish a quantitative assessment criteria, determine the early warning signals for severe/critical COVID-19 disease, and establish a predictive model for early warning for the progression and development of severe/critical COVID-19 disease, based on the dynamic evolution of CT, and all of these factors can help in the early intervention and treatment of COVID-19 patients clinically^[3,27]. In addition, the relationship between the occurrence of inflammatory storms and CT information features is currently unclear. In recent years, big data analysis technology and artificial intelligence have become important tools for evaluating the findings of CT using clinical

precision judgment. By using imagingomics technology to extract high-dimensional quantitative features from the CT images of COVID-19 patients, we can conduct in-depth mining of the CT information features to evaluate the lesions, screen the CT information features with high sensitivity and specificity, and observe the dynamic evolution of the CT information features and related cytokines. It is scientifically important to further explore the mechanisms that are involved with COVID-19 disease progression and regression^[28-30]. The identification of the associations between the dynamic changes in the patient's imaging and histological findings and the trends of the patient's related cytokines, the mining of the CT information features that reflect the trends in related pro/anti-inflammatory factors, and the exploration of the potential inflammation based on the associated features in COVID-19 will all provide an important basis for an early and accurate clinical judgment. The development of these projects will provide an objective basis for the effective prevention and control of COVID-19. Validity studies that are based on chest CT should be performed, and these studies can provide strong support for the application of imaging and histological tests, especially when combined with artificial intelligence technology, in the diagnosis and treatment of COVID-19^[31].

CONCLUSION

In conclusion, CT scores provide a valuable and objective measure of the progression of COVID-19 in patients. The trends of CT scores differed between common, severe, and critical patients, and monitoring these scores over time can help reduce unnecessary exposure to radiation and cost. The 2-wk CT scores of patients can also be useful in predicting disease deterioration in hospital patients with an average admission severity. Factors such as qSOFA score, aspartate aminotransferase, oxygen saturation, and dyspnea were found to be significant predictors of severe or critical COVID-19.

4%

SIMILARITY INDEX

PRIMARY SOURCES

- 1

Xiaoyue Wang, Desheng Jiang, Huang Huang, Xiaofeng Chen, Chunlei Zhou, Dongsheng Jiao, Ping Fan, Qian Cui, Hui Liao, Binbin Shi. "An increased pretreatment neutrophil-to-lymphocyte ratio predicts severe novel coronavirus-infected pneumonia", Research Square, 2020
Crossref Posted Content

16 words — < 1%
- 2

Yi Zuo, Gang Cheng, Da-Kuan Gao, Xiang Zhang, Hai-Ning Zhen, Wei Zhang, San-Chao Xiao. "Gross-total hematoma removal of hypertensive basal ganglia hemorrhages: A long-term follow-up", Journal of the Neurological Sciences, 2009
Crossref

16 words — < 1%
- 3

Sauson Soldozy, Michelle Yeghyayan, Kaan Yağmurlu, Pedro Norat et al. "Endoscopic endonasal surgery outcomes for pediatric craniopharyngioma: a systematic review", Neurosurgical Focus, 2020
Crossref

13 words — < 1%
- 4

www.frontiersin.org
Internet

13 words — < 1%
- 5

ngmu.ru
Internet

12 words — < 1%
- 6

Zhiru Gao, Yinghui Xu, Chao Sun, Xu Wang, Ye Guo, Shi Qiu, Kewei Ma. "A Systematic Review of

11 words — < 1%

7	www.jstage.jst.go.jp Internet	11 words — < 1%
8	Li-Gang Song, Su-Rong Bai, Deng-Hua Hui, Li-Ping Ding, Lu Sun. "Association of COVID-19 patient's condition with fasting blood glucose and body mass index: A retrospective study", Technology and Health Care, 2022 Crossref	10 words — < 1%
9	academic-accelerator.com Internet	10 words — < 1%
10	archive.lstmed.ac.uk Internet	10 words — < 1%
11	pubmed.ncbi.nlm.nih.gov Internet	10 words — < 1%
12	repository.usmf.md:8080 Internet	10 words — < 1%
13	www.ncbi.nlm.nih.gov Internet	10 words — < 1%
14	www.science.gov Internet	10 words — < 1%