medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

# Predicting COVID-19 malignant progression with AI techniques

Xiang Bai[1*], Cong Fang[1*], Yu Zhou[1*], Song Bai[1*], Zaiyi Liu[5*], Qianlan Chen[2*], Yongchao Xu[1*], Tian Xia[1*], Shi Gong[1], Xudong Xie[1], Dejia Song[1], Ronghui Du[4], Chunhua Zhou[3], Chengyang Chen[2], Dianer Nie[2], Dandan Tu[6], Changzheng Zhang[6], Xiaowu Liu[6], Lixin Qin[3†], Weiwei Chen[2†]

[1]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China;

[2]Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, People's Republic of China;

[3]Department of Radiology, Wuhan Pulmonary Hospital, Wuhan 430030, People's Republic of China;

[4]Department of Pulmonary and Critical care Medicine, Wuhan Pulmonary Hospital, Wuhan 430030, People's Republic of China;

[5]Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, People's Republic of China;

[6]HUST-HW Joint Innovation Lab, Wuhan 430074, People's Republic of China.

*Contributed equally

Correspondence to: Weiwei Chen. Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, People's Republic of China. Email: chenweiwei_tjh@163.com.

Lixin Qin. Department of Radiology, Wuhan Pulmonary Hospital, Wuhan 430030, People's Republic of China. Email: qlxzch2020@163.com.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.


# Abstract

Background and purpose: The worldwide pandemic of coronavirus disease 2019 (COVID-19) greatly challenges public medical systems. With limited medical resources, the treatment priority is determined by the severity of patients. However, many mild outpatients quickly deteriorate into severe/critical stage. It is crucial to early identify them and give timely treatment for optimizing treatment strategy and reducing mortality. This study aims to establish an AI model to predict mild patients with potential malignant progression.

Methods: A total of 133 consecutively mild COVID-19 patients at admission who was hospitalized in Wuhan Pulmonary Hospital from January 3 to February 13, 2020, were selected in this retrospective IRB-approved study. All mild patients were categorized into groups with or without malignant progression. The clinical and laboratory data at admission, the first CT, and the follow-up CT at the severe/critical stage of the two groups were compared. Both multivariate logistic regression and deep learning-based methods were used to build the prediction models, with their area under ROC curves (AUC) compared.

Results: Multivariate logistic regression depicted 6 risk factors for malignant progression: age >55years (OR 5.334, 95%CI 1.8-15.803), comorbid with hypertension (OR 5.093, 95%CI 1.236-20.986), a decrease of albumin (OR 4.01, 95%CI 1.216-13.223), a decrease of lymphocyte (OR 3.459, 95%CI 1.067-11.209), the progressive consolidation from CT1 to CTsevere (OR 1.235, 95%CI 1.018-1.498), and elevated HCRP (OR 1.015, 95%CI 1.002-1.029); and one protective factor: the presence of fibrosis at CT1 (OR 0.656, 95%CI 0.473-0.91). By combining the clinical data and the temporal information of the CT data, our deep learning-based models achieved the best AUC of 0.954, which outperformed logistic regression (AUC: 0.893),

Conclusions: Our deep learning-based methods can identify the mild patients who are easy to deteriorate into severe/critical cases efficiently and accurately, which undoubtedly helps to optimize the treatment strategy, reduce mortality, and relieve the medical pressure.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

# Introduction

In mid-December 2019, the ongoing coronavirus disease 2019 (COVID-19) broke out in Wuhan and spread rapidly in the mainland of China (80813 cases, updated through March 12, 2020). So far, the infection had burst in countries outside China, evolving into a pandemic[4,5]. According to the Chinese epidemic data, the mild, severe, and critical types of COVID-19 were 81%, 14%, and 5% separately[6]. More seriously, as of 12 March, the mortality of COVID-19 was 3.93% (3176/80813) in the mainland of China, even reached 4.87% (2436/49991) in Wuhan City, which was much higher than that of other influenza[7,8]. In addition, the clinical course of COVID-19 varied individually. In order to prevent malignant progression and reduce the mortality of COVID-19, it is vital to identify mild patients who are easy to deteriorate into severe/critical cases and give them active treatment earlier.

However, most studies focused on cross-sectional description and comparison of clinical, laboratory and CT imaging findings[9-12]. Some studies focused on seeking risk factors for death outcome[3,13]. None of them used AI-based methods for progression prediction of mild COVID-19 patients up to date. To solve this problem, we aimed to apply AI techniques to study multivariate heterogeneous data (clinical data and serial chest CT imaging) and to further develop an accurate and effective prediction model. Specifically, we employed a deep learning-based model to effectively mine the complementary information in static clinical data and serial quantitative chest CT sequence. Since deep learning-based methods had been widely adopted and had achieved great performance in cancer outcome prediction[14], head CT scans detection[15] and antibiotic discovery[16]. Moreover, compared with the traditional multi-stage methods, the deep learning-based model could significantly improve the efficiency of patient stratification (Figure 1), which is very important when dealing with tremendous patients.

Therefore, the purposes of our study are: 1) to develop models to identify the mild patients who are easy to deteriorate into severe/critical cases using AI-based methods and logistic regression respectively; 2) to seek clinical and imaging factors for prediction of patients with potential malignant progression.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

# Methods

## Patients and clinical data

From January 3, 2020 to February 13, 2020, a total of 199 consecutive COVID-19 patients, who were hospitalized at Wuhan Pulmonary Hospital, one of the first two designated hospitals for COVID-19 battle, were enrolled in this retrospective study. This study was approved by the Institutional Review Board of Wuhan Pulmonary Hospital on February 6, 2020. The informed consent forms were waived by the IRB. All patients were confirmed based on either positive viral nucleic acid test result on throat swab samples (n=80) or clinical parameters (n=53) according to the guidelines for the diagnosis and treatment of novel coronavirus (COVID-19) infection by the Chinese National Health Commission (Version 5)[17]. As showed in the flowchart (Figure 2), a subset of 133 patients (COVID-19 dataset) assigned as the mild type of COVID-19 infection in admission assessments was selected for further investigation. The inclusion criteria were: 1) respiratory rate < 30 breaths per min; 2) resting blood oxygen saturation > 93%; 3) the ratio of arterial oxygen partial pressure to fraction of inspiration oxygen > 300mmHg; 4) non-ICU patients without shock, respiratory failure, mechanical ventilation, and failure of other organs. The clinical and laboratory data at the time point of admission, together with serial chest CT images of all patients were retrospectively analyzed. Based on the presence or absence of the severe/critical progression during the hospitalization, all patients were categorized into two groups. The diagnostic criteria for severe/critical progression were: 1) respiratory rate ≥ 30 breaths per min; 2) resting blood oxygen saturation ≤ 93%; 3) the ratio of arterial oxygen partial pressure to fraction of inspiration oxygen < 300mmHg; 4) ICU patients with one or multiple organ failure, shock, or mechanical ventilation. The time point of reaching the criteria of severe/critical progression was recorded.

## Chest CT imaging, analysis, and post-procession

All patients underwent serial chest CT exams on a dedicated CT scanner (SOMATOM go.NOW, Siemens Healthineers, Germany) with the following parameters: slice thickness 3mm, slice gap 0mm, 130kV, 50mAs. All serial CT images were reviewed by three radiologists (QC, CC, DN, 2 years' experience in

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

radiology) independently blinded to the clinical information, and the discrepancy was resolved by consulting another radiologist (WC, 15 years' experience in radiology). Lesions and imaging features were assessed in each lung segment of each patient. The number of involved segments was counted not only for each patient or each lobe but also for each imaging feature. If more than one type of imaging features present in a segment, the segment was counted for every involved feature. The imaging features assessed in this study included 1) ground glass opacity (GGO); 2) consolidation; 3) air bronchogram; 4) paving stone sign; 5) fibrosis; 6) nodule; and 7) halo sign. The first available CT after symptoms onset, the follow-up CT, the first available CT of the severe stage were assigned as $CT_1$, $CT_2 \sim CT_n$, and $CT_{severe}$ separately. In order to compare the longitudinal variation of CT features during the period of $CT_1$ and $CT_{severe}$ between the two patient groups, we chose $CT_2$ instead of $CT_{severe}$ for those patients without severe/critical progression.

Our raw COVID-19 dataset contained all the clinical data and the quantitative chest CT data. After excluding invalid and duplicate information, each sample contained 75 clinical data characteristics and a quantitative CT sequence obtained at different times. Since the sequence length of each sample varied from zero to seven, we adjusted the data structure of each sample to the same shape by zero-filling the uncollected or missing chest CT data. The original quantitative chest CT data contained twelve infection distribution features, eight infection sign type features, the thickness of thoracic diaphragm, and CT course. The lung was medically divided into 18 segments, and the infection sign characteristics at each checkpoint can be formatted as a matrix. This matrix composed of infection distribution features and sign type features was flattened into a vector and then concatenated with the original quantitative chest CT data.

**The pipeline of the AI model**

The pipeline of the prediction model is shown in Figure 3. The input data includes the static data and the dynamic data, where the static data is a 75-dimensional vector, containing the clinical data and personal information of patients. Dynamic data is a series of quantitative chest CT data collected at different times. Each CT data at different checkpoint consists of a $3 \times 6$ matrix and a 22-dimensional vector. In order to merge these two parts, we directly flattened the

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

matrix into an 18-dimensional vector and concatenated it with the 22-dimensional vector to form a 40-dimensional CT feature vector. According to the checkpoints, the CT data sequence with a length of seven and a dimension of 40 was formed. For the sake of combining static and dynamic data as the input of long short term memory (LSTM), a multi-layer perceptron (MLP) was applied to the static data to obtain a 40-dimensional feature vector, which is used as the input data of the first timestamp of the LSTM, followed by the other seven CT feature vectors[18]. The LSTM model employed in this study is a single-layer network with the embedding dimension of 40 and the hidden dimension of 32. The output of the LSTM, a 32 × 8 feature sequence, was then fed into fully connected layers. A Softmax layer was added at the top of the network to output the probability of the patient conversion to the severe/critical stage. A total of 133 samples were included in the COVID-19 dataset. The robustness of the model was evaluated by five-fold cross-validation repeating five times, and each fold was obtained by category-wise sampling.

**Statistical Analysis**

All the statistical analysis was performed using SPSS (Version 26) with statistical significance set at 0.05. Statistical optimization of the deep learning model was done through iterative training using Python (Version 3.6 with scipy, scikit-learn, and pytorch packages). The differences of clinical and laboratory data and imaging features between the patient with and without severe/critical progression were compared using Chi-square test, Fisher's exact test, independent t test and paired t test. AUC, accuracy, specificity, and sensitivity were compared among different AI methods and multivariable logistic regression. Two-sided 95% CIs were used to summarize the sample variability in the estimates. Specifically, the normal approximation CIs was used for accuracy, sensitivity, and specificity. The CI for the AUC was estimated using the bootstrap method with 2000 replications.

## Results

133 patients with mild COVID-19 pneumonia at admission included 66 male and 67 female, age ranged from 18 to 82 (52.82 ± 12.59) years, the interval from symptoms onset to admission ranged from 1 to 20 (8.76 ± 4.05) days. 54 patients (54/133, 40.6%) malignantly progressed to severe/critical periods during the

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

hospitalization, while the remaining 79 patients (79/133, 59.4%) did not (Figure 4). The whole clinical course of all patients, including assessment at admission, the severe or critical progression, and the outcome, was plotted in Figure 4.

The age, sex, exposure, comorbidity, signs and symptoms, laboratory results measured at admission, and serial CT imaging features of patients with and without severe/critical progression were separately summarized in Tables 1, 2, and 3. In brief, comparing to the patients without severe/critical progression, the patients with severe/critical progression showed older age, more comorbidities, higher respiratory rate, inflammatory cell factors, lower albumin and fewer counts of lymphocyte, T cell, and its subsets. The patients with severe/critical progression were more likely to involve organs other than the lung. On the first available CT, no difference was found in either the distribution of involved lung or the other CT imaging features, except paving stone sign and the presence of fibrosis. However, the patients with severe/critical progression showed significantly more lesions in all lobes, more lesions of consolidation, paving stone sign and halo sign than patients without severe/critical progression when they progressed to the period of severe/critical stage.

The logistic regression results depicted that age >55years (OR 5.334, 95%CI 1.800-15.803), comorbid with hypertension (OR 5.093, 95%CI 1.236-20.986), a decrease of albumin (OR 4.010, 95%CI 1.216-13.223), a decrease of lymphocyte (OR 3.459, 95%CI 1.067-11.209), the progressive consolidation from $CT_1$ to $CT_{severe}$ (OR 1.235, 95%CI 1.018-1.498), and elevated HCRP (OR 1.015, 95%CI 1.002-1.029) were the risk factors for severe/critical progression. However, the presence of fibrosis at $CT_1$ (OR 0.656, 95%CI 0.473-0.910) was the protective factor for severe/critical progression. The accuracy of the prediction is 79.2%.

We conducted comprehensive experiments to validate our hypotheses and compared the performance of various models. Table 4 summarized the performance of traditional multi-stage and deep learning-based methods. Static clinical data including personal information, dynamic quantitative chest CT data or both of them were used for predictive experiments.

For traditional multi-stage methods, PCA was used for data dimensionality reduction, and SVM or LDA was used for classification. The results indicate that quantitative chest CT data without time series modeling is also beneficial for

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

traditional multi-stage methods, which brought a considerable improvement of 0.069 (SVM) and 0.115 (LDA) respectively in AUC.

For deep learning-based methods, compared with the static data, the time series modeling of dynamic data through LSTM alone brought a dramatic improvement of 0.104 in AUC. On the other hand, feeding both static data and dynamic data without time-series information to the classifier of MLP, we got an improvement of 0.141 in AUC compared to using static data alone. Furthermore, when two complementary data and time-series data were added, our hybrid model achieved the AUC of 0.920 (95% CI 0.910-0.930), the accuracy of 0.891 (95% CI 0.879-0.902), the specificity of 0.857 (95% CI 0.836-0.878) and the sensitivity of 0.925 (95% CI 0.893-0.957). All these evaluation results were substantially superior to the traditional multi-stage methods.

The above results clearly supported the significance of complementary information from different medical data and time-series information from the chest CT sequence. Finally, our proposed method had a high probability of stabilizing at a high confidence interval, which is very important for clinical applications.

## Discussion

With the worldwide outbreak of COVID-19, early prediction and early aggressive treatment of mild patients at high risk of malignant progression to severe/critical stage are important ways to reduce mortality. In this work, we found that the complementarity of clinical data and quantitative chest CT sequence is important for predicting patients with malignant progression. In particular, the rich series information of the chest CT sequence, which has not been considered by other studies so far, is critical for this specific task. We also demonstrated that our method can effectively fuse these two complementary data and handle time-series information in the quantitative chest CT sequence, which achieved an AUC of 0.954 (95% CI 0.942-0.967) under five-fold cross-validation repeating five times. This compares favorably with other traditional multi-stage methods such as SVM (AUC 0.857 (0.823-0.891)), LDA (AUC 0.868 (0.860-0.877)) and MLP (AUC 0.797 (0.790-0.802)).

Although lots of clinical, laboratory, and imaging parameters varied significantly between patients with and without severe/critical progression, seven predictive

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

parameters were selected by logistic regression analysis in this study. Elder age and comorbidity with hypertension, reflecting the basic physical condition of susceptible people, are the most important risk factors for malignant progression[19]. Pathological studies about COVID-19 reported extensive cellular fibromyxoid exudates and interstitial lymphocytes inflammatory infiltrates[20,21]. Thus, the decrease of albumin and lymphocytes, the following two risk factors, probably suggest they migrate from the circulation into affected alveoli and also resulting in an increase of consolidation on follow-up CT imaging. Elevated HCRP is an indicator of anti-virus inflammation, also served as a risk factor for malignant progression. The risk factors mentioned above are all in accordance with the results from observation studies[22-26]. Fibrosis on CT imaging is reported to be closely associated with the late outcome of COVID-19 pneumonia[27,28]. However, our study depicted the early presence of fibrosis on CT imaging is a protective factor for malignant progression, which might reflect early absorption of inflammation and the occurrence of the repair stage.

Our work is novel because it is the first study in which complementary data of quantitative CT sequence and clinical data is used to analyze the problem of COVID-19 malignant progression prediction. Experimental results show that both of them have a significant reference value for this problem and can obtain more accurate prediction results. Furthermore, there is very little literature to date modeling the spatial information of the quantitative CT data and considering the time-series information of patients. This information has important reference value for the prediction of patients with potential malignant progression. Specifically, the quantified spatial CT data is converted into a two-dimensional matrix format according to the real distribution of lung areas. And then, the spatial CT data was combined with the disease course and other infection sign type features. Finally, the hybrid CT data was folded into sequences following the sequence of clinical and patient-specific information.

Unlike the traditional predictive model using a hand-crafted feature extractor and shallow classifiers, our deep learning-based method using a multilayer perceptron combined with an LSTM to this predictive task, which attempts to learn high-level hierarchical features from mass data, and expands the search space of the features for specific tasks. Moreover, this method jointly optimizes the feature extraction network and classifier through an end-to-end manner.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Our study has several limitations. First, samples available for malignant progression prediction were limited. The diverse data in the large scale dataset will allow deep learning-based methods to gain a more comprehensive understanding of what causes the malignant progression of mild patients. Second, the quantitative information of CT data is not detailed enough. Using the richer original features included in pixel-wise segmentation results of the CT scans, the predictive model may perform better.

In conclusion, the deep learning-based method using clinical and quantitative CT data to predict malignant progression to severe/critical stage. We modeled the spatial information in the quantitative CT data and organized the static clinical data and dynamic chest CT data into a time series form. We validated the significance of complementary data and its special formatting form for this particular prediction task. Compared with traditional multi-stage methods, we demonstrate that our deep learning-based method can extract spatial and temporal information efficiently and improve the prediction performance significantly. The ability to identify patients with potentially severe and critical COVID-19 outcomes using an inexpensive, widely available, the point-of-care test has important practical implications for preventing mild patients from becoming severe, effectively improving cure rate, and reducing mortality. Our future work will focus on mining richer spatial information from the CT scan sequence and using AI technologies to screen the risk factors of potential severe/critical patients.

**Contributors**

1. Organization of the research team and management of the development of the overall study: XB and WC
2. Study design: XB, WC, ZL, YZ, LQ, and CF.
3. Literature search: CF, WC, YZ, QC, DN, and CC.
4. Data collection: LQ, RD, CZ, QC and WC.
5. Imaging data analysis: WC, LQ, QC, DN, CC, DT, CZ, and XL.
6. Development of the machine learning algorithms: CF, YZ, SB, YX, XX, and SG.
7. Data interpretation: CF, WC, ZL, XB, SB, YX, TX, and XX.
8. Made the tables and figures, and contribute to the writing of the manuscript: CF, WC, YZ, SG, XX, DS, QC, and DN
9. Paper writing and revision: XB, YZ, SB, CF, WC, YX, ZL and TX.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Declaration of interests**

We declare no competing interests.

**Data availability statement**

According to the relevant scientific research policy of China, our research data cannot be shared online.

**Acknowledgments**

This work was supported by National Key R&D Program of China (No. 2018YFB1004600), HUST COVID-19 Rapid Response Call (No. 2020kfyXGYJ093, No. 2020kfyXGYJ094), National Key R&D Program of China (No.2017YFC1309100), National Science Fund for Distinguished Young Scholars (No.81925023), National Natural Science Foundation of China (No.61703049, No. 81771912, No. 81401390).

## Reference

1. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet (London, England) 2020; 395: 507-13.
2. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet (London, England) 2020; 395: 497-506.
3. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet (London, England) 2020.
4. National Health Commission of the People's Republic of China. The latest situation of COVID-19 in China http://www.nhc.gov.cn/xcs/yqtb/202003/948a03ad76f54d3583a018785efd7be9.shtml (accessed Mar 10, 2020).
5. WHO declares COVID-19 pandemic. http://www.chinadaily.com.cn/a/202003/12/WS5e6914dda31012821727e4a5.html (accesed March 12, 2020).
6. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. Jama 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

7. Iuliano AD, Roguski KM, Chang HH, et al. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. Lancet (London, England) 2018; 391(10127): 1285-300.
8. Li L, Liu Y, Wu P, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. The Lancet Public health 2019; 4: e473-e81.
9. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. Jama 2020.
10. Yang W, Cao Q, Qin L, et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19):A multi-center study in Wenzhou city, Zhejiang, China. The Journal of infection 2020.
11. Kim H. Outbreak of novel coronavirus (COVID-19): What is the role of radiologists? European radiology 2020.
12. Shi H, Han X, Jiang N, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. The Lancet Infectious diseases 2020.
13. 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. Intensive care medicine 2020.
14. Skrede OJ, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. Lancet (London, England) 2020; 395: 350-60.
15. Chilamkurthy S, Ghosh R, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet (London, England) 2018; 392: 2388-96.
16. Stokes JM, Yang K, Swanson K, et al. A Deep Learning Approach to Antibiotic Discovery. Cell 2020; 180: 688-702.e13.
17. National Health Commission of the People's Republic of China. Diagnosis and treatment guidelines of COVID-19 (Version 5.0) http://www.nhc.gov.cn/yzygj/s7653p/202002/d4b895337e19445f8d728fcaf1e3e13a/files/ab6bec7f93e64e7f998d802991203cd6.pdf (accessed Feb 8, 2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

18. Shi B, Bai X, Yao C. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. IEEE transactions on pattern analysis and machine intelligence 2017; 39: 2298-304.
19. Xie J, Tong Z, Guan X, Du B, Qiu H, Slutsky AS. Critical care crisis and some recommendations during the COVID-19 epidemic in China. Intensive care medicine 2020.
20. Xu Z, Shi L, Wang Y, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. The Lancet Respiratory medicine 2020.
21. Tian S, Hu W, Niu L, Liu H, Xu H, Xiao SY. Pulmonary pathology of early phase 2019 novel coronavirus (COVID-19) pneumonia in two patients with lung cancer. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer 2020.
22. Liu W, Tao ZW, Lei W, et al. Analysis of factors associated with disease outcomes in hospitalized patients with 2019 novel coronavirus disease. Chinese medical journal 2020.
23. Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. Clinical chemistry and laboratory medicine 2020.
24. Li K, Wu J, Wu F, et al. The Clinical and Chest CT Features Associated with Severe and Critical COVID-19 Pneumonia. Investigative radiology 2020.
25. Kanne JP, Little BP, Chung JH, Elicker BM, Ketai LH. RadiologyEssentials for Radiologists on COVID-19: An Update- Scientific Expert Panel. Radiology 2020: 200527.
26. Hosseiny M, Kooraki S, Gholamrezanezhad A, Reddy S, Myers L. Radiology Perspective of Coronavirus Disease 2019 (COVID-19): Lessons From Severe Acute Respiratory Syndrome and Middle East Respiratory Syndrome. AJR American journal of roentgenology 2020: 1-5.
27. Sun P, Qie S, Liu Z, Ren J, Li K, Xi J. Clinical characteristics of 50466 hospitalized patients with 2019-nCoV infection. Journal of medical virology 2020.
28. Pan Y, Guan H, Zhou S, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China. European radiology 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 1: Age, sex, exposure, comorbidity, signs, symptoms and death of patients with and without severe/critical progress**

| | Patients without severe/critical progress | Patients with severe/critical progress | Total | p (2-sided) |
|---|---|---|---|---|
| Number of patients | 79 | 54 | 133 | NA |
| Sex | 40M/39F | 26M/28F | 66M/67F | 0.860 |
| Age (years) | 48.51±12.04 | 59.13±10.66 | 52.82±12.59 | 0.000 |
| Days from symptom onset to admission | 8.51±4.08 | 9.13±4.01 | 8.76±4.05 | 0.385 |
| Exposure | | | | |
| Huanan seafood market exposure | 14 | 1 | 15 | 0.004 |
| Family cluster history | 14 | 14 | 28 | 0.282 |
| Comorbidity | | | | |
| Hypertension | 9 | 19 | 28 | 0.001 |
| Diabetes | 8 | 10 | 18 | 0.199 |
| Pulmonary tuberculosis | 4 | 4 | 8 | 0.714 |
| Malignant tumor | 1 | 2 | 3 | 0.566 |
| Cardiocerebrovascular diseases | 6 | 22 | 28 | 0.001 |
| Chronic digestive disorders | 11 | 9 | 20 | 0.805 |
| Chronic obstructive pulmonary disease | 1 | 4 | 5 | 0.157 |
| Signs and symptoms | | | | |
| Fever | 75 | 46 | 121 | 0.038 |
| Respiratory rate (breaths per minute) | 20.01±1.24 | 20.77±1.87 | 20.34±1.63 | 0.013 |
| Heart rate (beats per minute) | 88.00±13.23 | 88.15±13.21 | 88.17±12.97 | 0.949 |
| Cough | 60 | 43 | 103 | 0.656 |
| Gasp | 24 | 19 | 43 | 0.573 |
| Days from symptom onset to dyspnoea | 0.63±2.16 | 2.16±3.89 | 1.25±3.05 | 0.013 |
| Dyspnoea | 10 | 20 | 30 | 0.001 |
| Expectoration | 30 | 24 | 54 | 0.584 |
| Fatigue | 35 | 29 | 64 | 0.368 |
| Myalgia | 18 | 17 | 35 | 0.314 |
| Haemoptysis | 1 | 5 | 6 | 0.040 |

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 1: Age, sex, exposure, comorbidity, signs, symptoms and death of patients with and without severe/critical progress, continued**

| | Patients without severe/critical progress | Patients with severe/critical progress | Total | p (2-sided) |
|---|---|---|---|---|
| Gastrointestinal symptoms | 16 | 17 | 33 | 0.154 |
| Headache | 10 | 8 | 18 | 0.798 |
| Death | 0 | 3 | 3 | 0.065 |

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 2: Laboratory results of blood cells, blood oxygen, inflammatory factor, function of organs, coagulation function of patients with and without severe/critical progress**

| | Patients without severe/critical progress | Patients with severe/critical progress | Total | *p* (2-sided) |
|---|---|---|---|---|
| Blood oxygen content | | | | |
| Resting blood oxygen saturation % | 96.70±1.56 | 96.83±1.49 | 96.68±1.55 | 0.639 |
| PO2, mmHg | 94.08±33.63 | 88.69±38.03 | 91.63±35.28 | 0.428 |
| PCO2, mmHg | 39.80±4.76 | 37.48±5.03 | 38.81±4.99 | 0.014 |
| Blood routine | | | | |
| White cell count, $\times 10^9$/L | 5.03±1.96 | 5.56±2.30 | 5.47±3.32 | 0.167 |
| Neutrophil count, $\times 10^9$/L | 3.54±1.90 | 4.15±2.29 | 3.79±2.06 | 0.107 |
| Neuttrophil % | 66.94±12.16 | 74.60±13.42 | 70.12±13.05 | 0.001 |
| Lymphocyte count, $\times 10^9$/L | 1.12±0.36 | 0.90±0.43 | 1.03±0.41 | 0.002 |
| Lymphocyte % | 24.50±10.07 | 20.22±13.79 | 22.71±11.80 | 0.046 |
| Total T lymphocyte % | 66.00±13.27 | 60.34±12.11 | 63.22±13.27 | 0.026 |
| Total T lymphocyte count, cell/μl | 632.60±373.70 | 391.52±279.24 | 527.49±353.63 | 0.000 |
| Absolute count of CD3+CD4+ T cells, cell/μl | 355.59±238.57 | 242.79±180.40 | 307.05±221.53 | 0.009 |
| Absolute count of CD3+CD8+ T cells, cell/μl | 262.50±154.66 | 141.06±109.70 | 207.52±148.36 | 0.000 |
| Absolute count of lymphocyte CD45, cell/μl | 935.93±514.18 | 624.31±360.11 | 807.07±487.69 | 0.000 |
| Inflammatory factor | | | | |
| Hypersensitive C-reactive protein (HCRP), mg/L | 31.88±30.28 | 59.37±49.96 | 42.75±41.68 | 0.001 |
| Procalcitonin (PCT), ng/ml | 0.05±0.03 | 0.41±1.40 | 0.22±0.96 | 0.082 |
| Lactate dehydrogenase (LDH), u/l | 251.65±60.02 | 331.24±138.03 | 283.06±107.79 | 0.000 |
| Cardiac function | | | | |
| Myoglobin, ng/ml | 21.81±14.85 | 46.62±53.43 | 45.45±100.87 | 0.013 |
| Troponin, ng/ml | 0.03±0.00 | 0.04±0.03 | 0.06±0.22 | 0.165 |
| Brain natriuretic peptide (BNP), pg/ml | 141.05±200.42 | 467.24±773.60 | 338.79±727.39 | 0.020 |
| Liver function | | | | |
| Alanine aminotransferase (ALT), U/L | 24.80±21.13 | 35.25±31.48 | 29.09±26.14 | 0.040 |
| Aspartate aminotransferase (AST), U/L | 26.70±12.81 | 39.37±30.77 | 32.01±22.74 | 0.007 |
| Total bilirubin, μmol/L | 8.58±5.17 | 10.10±4.17 | 9.58±6.73 | 0.081 |

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 2: Laboratory results of blood cells, blood oxygen, inflammatory factor, function of organs, coagulation function of patients with and without severe/critical progress, continued**

| | Patients without severe/critical progress | Patients with severe/critical progress | Total | *p* (2-sided) |
|---|---|---|---|---|
| Direct bilirubin, μmol/L | 2.59±3.68 | 2.96±1.65 | 2.90±3.58 | 0.504 |
| Alkaline phosphatase, U/L | 60.70±24.85 | 65.35±27.41 | 62.98±25.61 | 0.324 |
| γ-Glutamyl transpeptidase (γ-GT), U/L | 42.82±58.82 | 45.94±62.97 | 43.20±59.36 | 0.776 |
| Albumin, g/L | 39.73±4.21 | 36.21±5.34 | 38.38±5.09 | 0.000 |
| Renal function | | | | |
| Urea, mmol/L | 4.13±1.38 | 5.20±2.24 | 4.62±1.94 | 0.003 |
| Creatinine, μmol/L | 65.44±15.24 | 75.19±27.78 | 69.83±21.78 | 0.025 |
| Coagulation function | | | | |
| D-dimers, mg/L | 0.61±1.95 | 1.40±3.65 | 1.23±4.09 | 0.183 |
| Prothrombin time (PT), s | 13.18±1.14 | 13.88±2.88 | 13.50±2.06 | 0.105 |
| Activated partial thromboplastin time (APTT), s | 34.40±4.45 | 34.66±5.83 | 34.52±4.98 | 0.783 |

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 3: Longitudinal variation of CT features from admission ($CT_1$) to the severe stage ($CT_{severe}$)**

| | Patients without severe/critical progress* | Patients with severe/critical progress | Total | *p* (2-sided) |
|---|---|---|---|---|
| Days from symptom onset to the $CT_1$ (at admission) | 9.87±4.79 | 9.23±5.45 | 9.61±5.06 | 0.486 |
| Days from the $CT_1$ to the $CT_{severe}$ | 3.8±2.39 | 4.48±5.17 | 3.76±4.98 | 0.387 |
| $CT_1$ (at admission) | | | | |
| $CT_1$_total involved lung segments | 11.75±4.58 | 12.31±5.14 | 11.53±5.24 | 0.521 |
| $CT_1$_involvd segments in right upper lobe | 1.76±1.21 | 1.88±1.10 | 1.74±1.19 | 0.556 |
| $CT_1$_involvd segments in left upper lobe | 2.38±1.50 | 2.69±1.45 | 2.51±1.48 | 0.245 |
| $CT_1$_involved segments in right middle lobe | 0.89±0.89 | 1.10±0.80 | 0.94±0.86 | 0.192 |
| $CT_1$_involved segments in right lower lobe | 3.67±1.31 | 3.48±1.46 | 3.46±1.51 | 0.443 |
| $CT_1$_involved segments in left lower lobe | 3.04±1.31 | 3.15±1.24 | 2.97±1.39 | 0.621 |
| $CT_1$_ground glass opacity | 7.95±5.16 | 8.13±5.06 | 7.72±5.23 | 0.839 |
| $CT_1$_consolidation | 5.63±4.10 | 5.35±4.59 | 5.31±4.34 | 0.713 |
| $CT_1$_air bronchogram | 1.29±1.97 | 1.54±2.22 | 1.34±2.05 | 0.506 |
| $CT_1$_paving stone sign | 0.84±2.07 | 1.85±3.09 | 1.20±2.53 | 0.044 |
| $CT_1$_fibrosis (fibrous stripes) | 2.00±2.36 | 0.62±2.06 | 1.38±2.31 | 0.001 |
| $CT_1$_nodule | 0.47±1.57 | 0.21±0.57 | 0.35±1.24 | 0.252 |
| $CT_1$_halo sign | 2.82±3.44 | 4.23±4.31 | 3.26±3.85 | 0.051 |
| Changes from $CT_1$ to $CT_{severe}$ | | | | |
| $CT_{1\text{-}severe}$_total involved lung segments | -0.95±4.77 | 1.65±3.60 | 0.11±4.50 | 0.001 |
| $CT_{1\text{-}severe}$_involvd segments in right upper lobe | -0.05±0.92 | 0.29±0.80 | 0.09±0.89 | 0.032 |
| $CT_{1\text{-}severe}$_involved segments in left upper lobe | -0.08±1.03 | 0.67±1.47 | 0.14±1.22 | 0.015 |
| $CT_{1\text{-}severe}$_involved segments in right middle lobe | -0.08±0.63 | 0.23±0.51 | 0.05±0.60 | 0.004 |
| $CT_{1\text{-}severe}$_involved segments in right lower lobe | -0.33±1.54 | 0.46±1.00 | -0.01±1.39 | 0.001 |
| $CT_{1\text{-}severe}$_involved segments in left lower lobe | -0.25±1.31 | 0.29±0.89 | -0.03±1.18 | 0.011 |
| $CT_{1\text{-}severe}$_ground glass opacity | -0.32±4.52 | 0.62±2.83 | 0.06±3.93 | 0.154 |
| $CT_{1\text{-}severe}$_consolidation | -1.08±3.81 | 0.87±2.39 | -0.29±3.43 | 0.001 |
| $CT_{1\text{-}severe}$_air bronchogram | -0.11±1.38 | 0.50±1.97 | 0.14±1.66 | 0.043 |
| $CT_{1\text{-}severe}$_paving stone sign | 0.14±1.51 | 0.73±2.24 | 0.38±1.86 | 0.103 |
| $CT_{1\text{-}severe}$_fibrosis(fibrous stripes) | 0.25±2.44 | 0.42±1.07 | 0.32±2.00 | 0.586 |

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 3: Longitudinal variation of CT features from admission ($CT_1$) to the severe stage ($CT_{severe}$), continued**

| | Patients without severe/critical progress* | Patients with severe/critical progress | Total | *p* (2-sided) |
|---|---|---|---|---|
| $CT_{1\text{-}severe}$_nodule | -0.20±1.51 | -0.02±0.42 | -0.13±1.19 | 0.331 |
| $CT_{1\text{-}severe}$_halo sign | -0.95±3.39 | 0.48±2.10 | -0.37±3.01 | 0.004 |

* Because this group of patients did not have a period of severe/critical progress, we chose the 2nd available follow-up CT exam instead for comparison.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.20.20037325; this version posted March 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 4: Performance comparison of different methods**

| | AUC 95% CI | SPE 95% CI | SEN 95% CI | ACC 95% CI | Static Data | Dynamic Data |
|---|---|---|---|---|---|---|
| PCA+LDA | 0.753(0.716-0.774) | 0.696(0.676-0.772) | 0.772(0.727-0.777) | 0.730(0.688-0.756) | √ | - |
| PCA+LDA | 0.868(0.860-0.877) | 0.729(0.711-0.747) | 0.842(0.830-0.854) | 0.789(0.783-0.795) | √ | √ |
| PCA+SVM | 0.751(0.745-0.758) | 0.716(0.702-0.731) | 0.745(0.732-0.761) | 0.728(0.721-0.736) | √ | - |
| PCA+SVM | 0.857(0.823-0.891) | 0.718(0.671-0.766) | 0.847(0.805-0.888) | 0.781(0.746-0.815) | √ | √ |
| MLP | 0.797(0.790-0.802) | 0.761(0.738-0.782) | 0.823(0.812-0.836) | 0.796(0.789-0.811) | √ | - |
| MLP | 0.938(0.925-0.949) | 0.863(0.843-0.884) | 0.898(0.869-0.921) | 0.878(0.865-0.892) | √ | √ |
| MLP+LSTM | 0.901(0.891,0.912) | 0.841(0.810-0.871) | 0.889(0.875-0.914) | 0.864(0.853-0.876) | - | √ |
| MLP+LSTM | 0.954(0.942-0.967) | 0.857(0.836-0.878) | 0.925(0.893-0.957) | 0.891(0.879-0.902) | √ | √ |

Static Data refers to clinical data and personal information. Dynamic data refers to quantitative chest CT sequence.

√=Use this data during training and testing. -=Not use this data during training and testing. ACC=accuracy. AUC=area under ROC curve. SPE=specificity. SEN=sensitivity. PCA=principal component analysis. LDA=linear discriminant analysis. SVM=support vector machine. MLP=multilayer perceptron. LSTM= long short term memory.

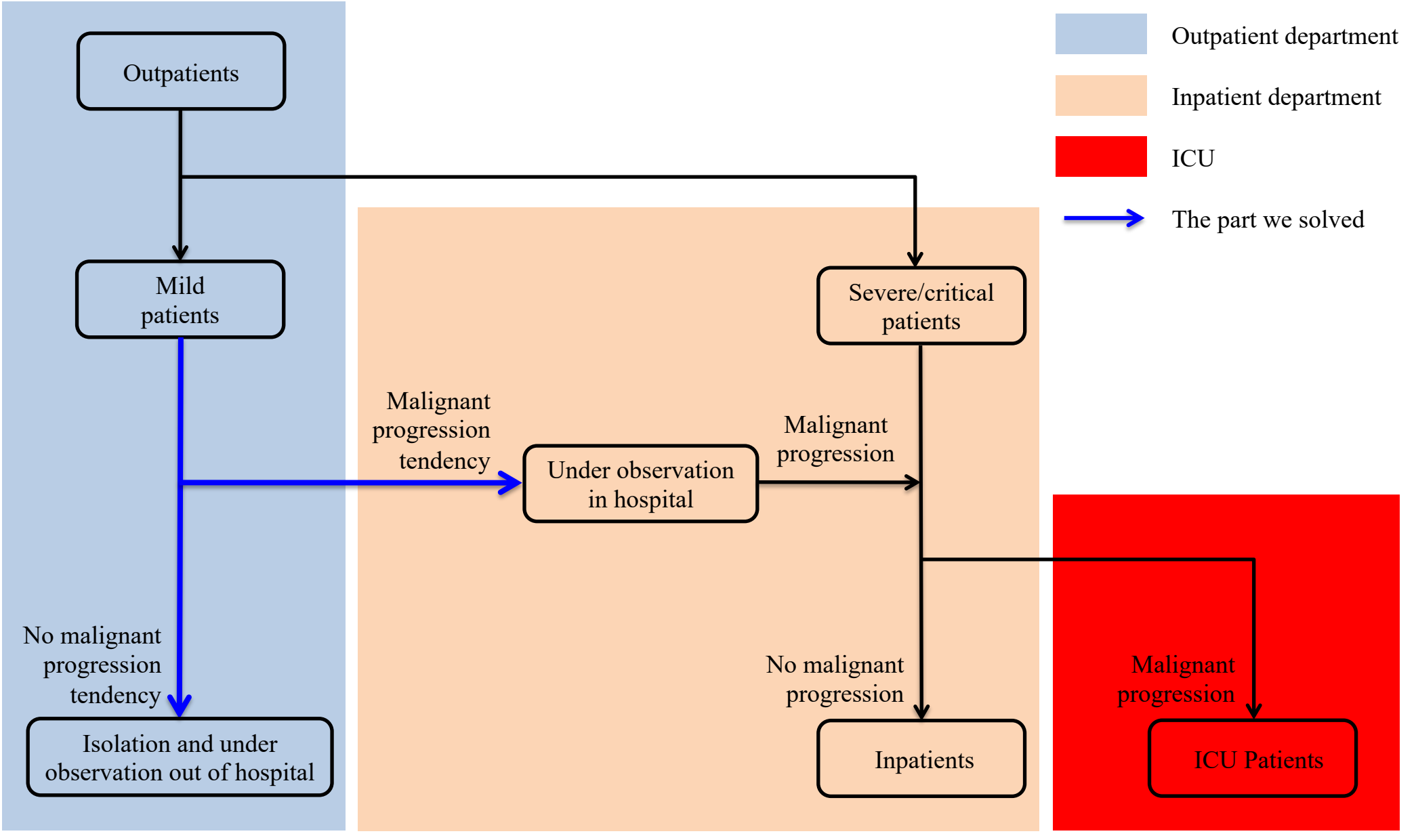


**Figure 1: Patient stratification from outpatients to ICU**

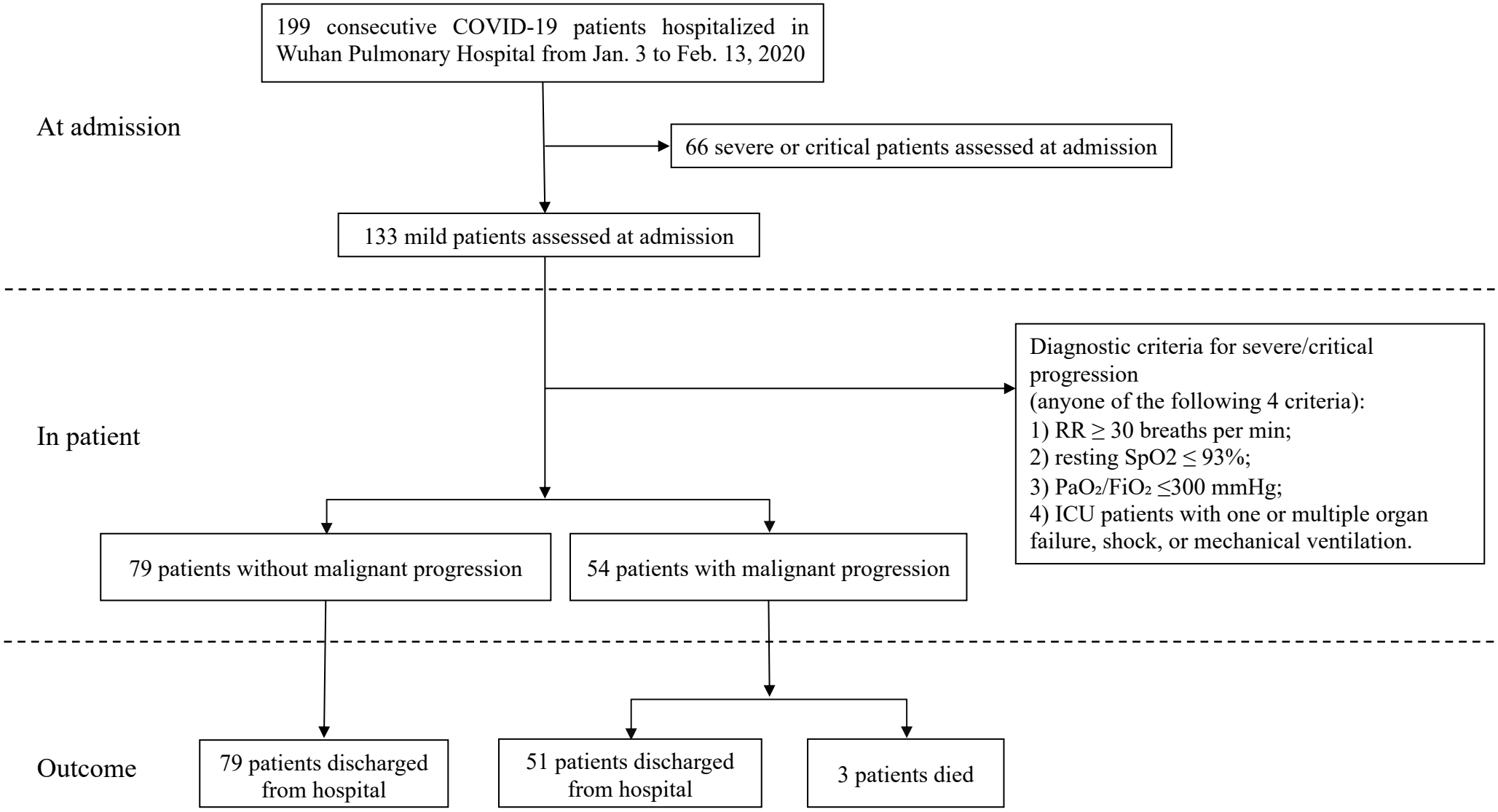


**Figure 2: Flowchart of patient selection**

RR=respiratory rate. SpO2=blood oxygen saturation. $PaO_2$=arterial oxygen partial pressure. $FiO_2$=fraction of inspiration oxygen.

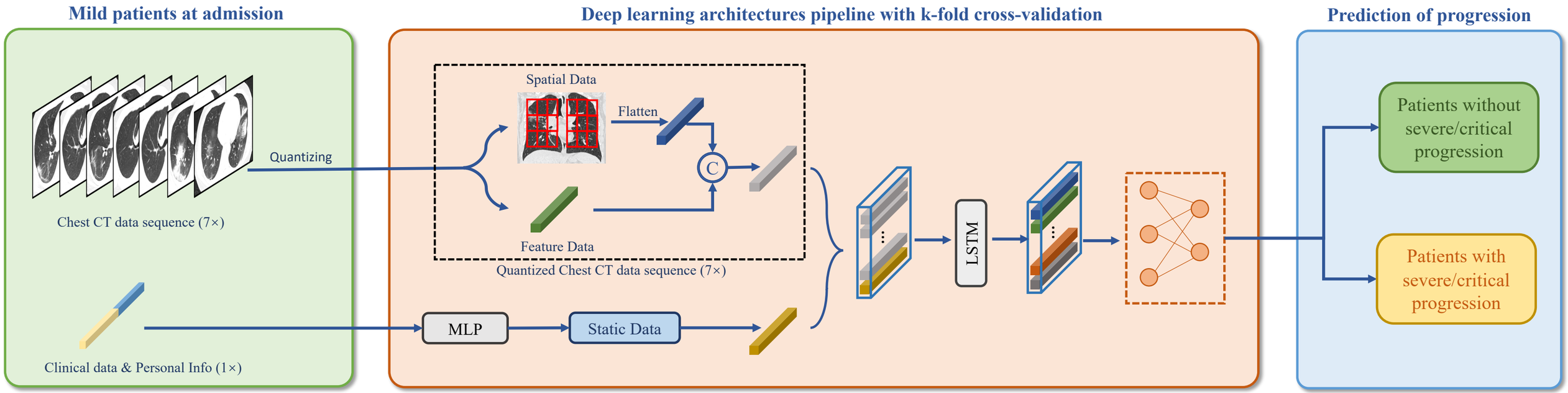


**Figure 3: The pipeline of COVID-19 severe/critical progression prediction model**

CT=computed tomography. MLP=multilayer perceptron. LSTM= long short term memory.

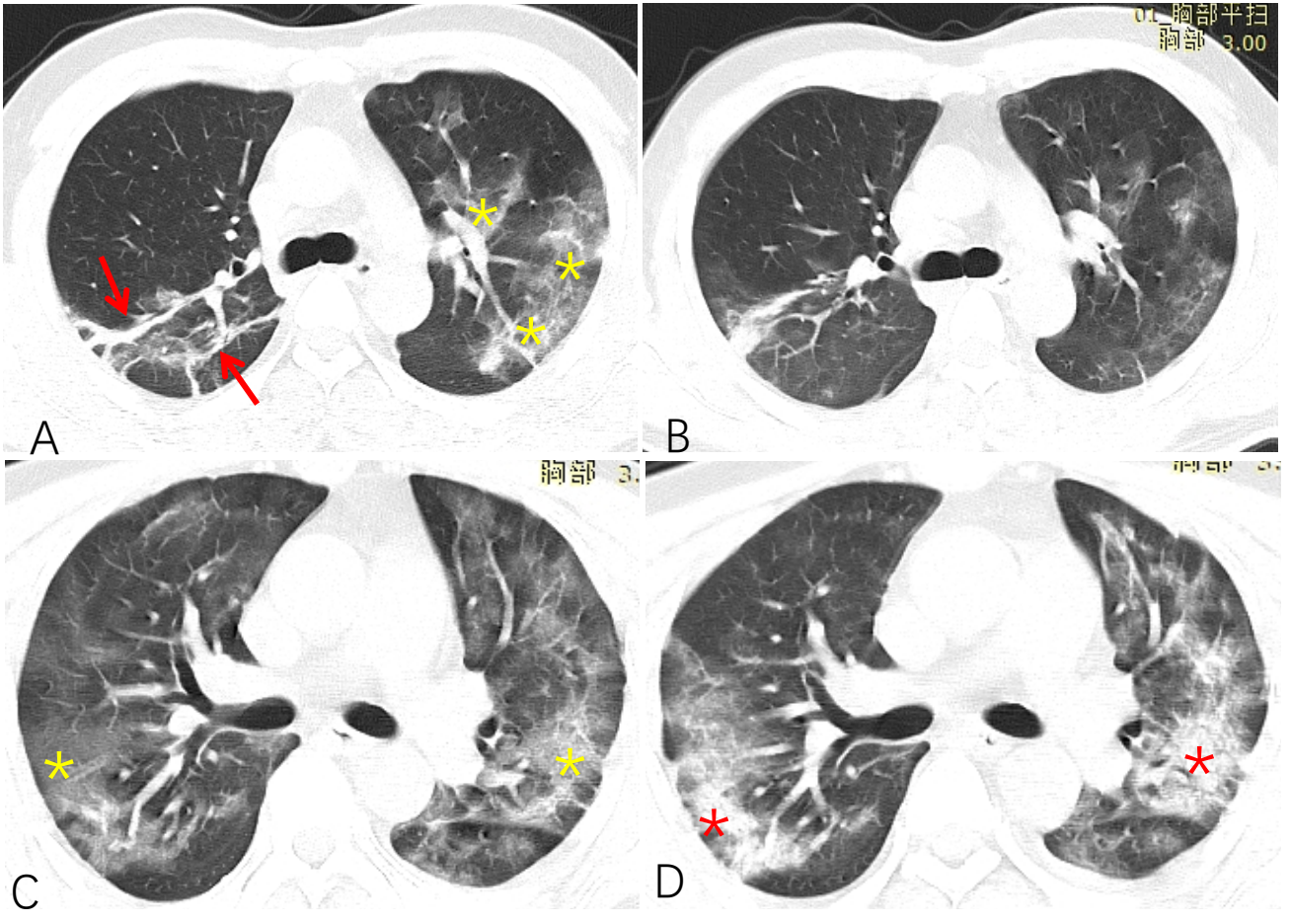


**Figure 4: Representative cases without malignant progression (A, B) and with malignant progression(C, D)**

A 32-year-old male with symptoms of fever, cough, and dyspnoea. $CT_1$ (A) shows ground glass opacity lesion (yellow stars) in the left upper lobe and some fibrosis (red arrows) in the right upper lobe. The CT image of 3 days follow-up (B) shows the prior lesions shrunk in size and decrease in density. A 64-year-old male with symptoms of fever, dyspnoea. $CT_1$ (C) shows ground glass opacity lesion (yellow stars) in bilateral upper lobes. The CT image of 3 days follow-up (D) shows the prior lesions progress to consolidation (red stars).

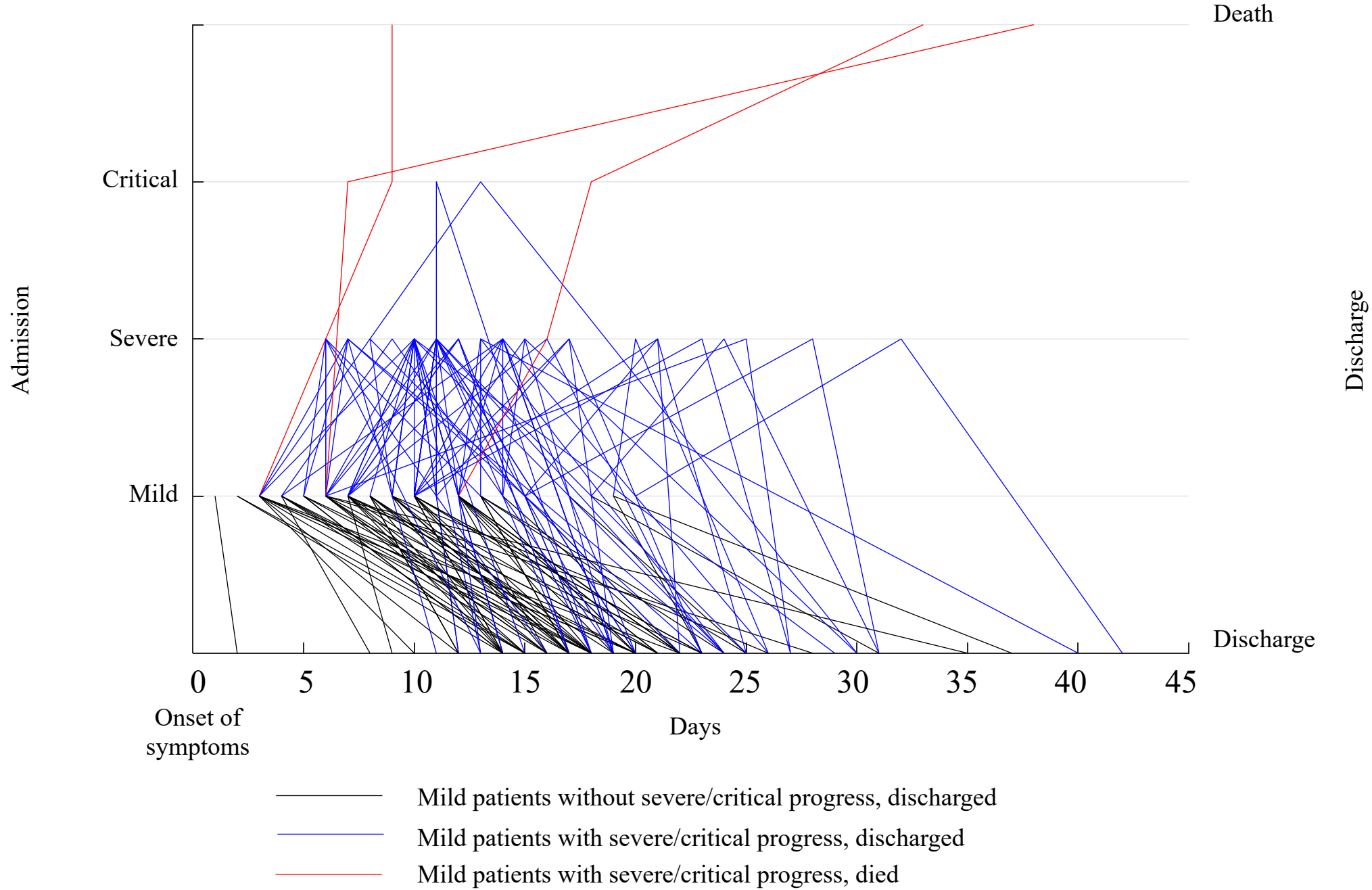


**Figure 5: The plots of clinical courses for all included patients**

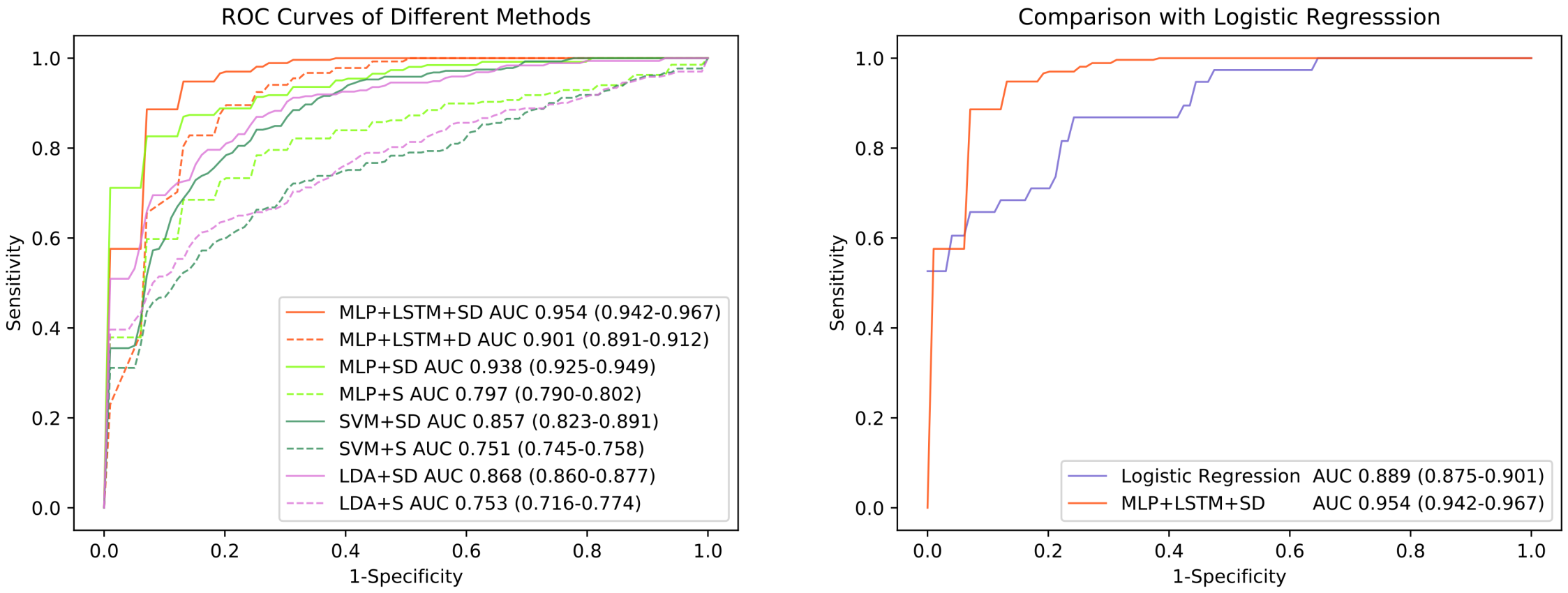


**Figure 6: ROC curves of different methods on COVID-19 dataset**

ROC=receiver operating characteristic. AUC=area under ROC curve. LDA=linear discriminant analysis. SVM=support vector machine. MLP=multilayer perceptron. LSTM= long short term memory. S=Static Data. D=Dynamic data.