

Comparative Characterization and Risk Stratification of Asymptomatic and Presymptomatic Patients with COVID-19

Lei Shi¹, Rong Ding², Tingting Zhang², Wei Wu², Ziyu Wang², Xiuzhi Jia³, Kening Li², Yuan Liang², Jie Li², Mengyan Zhu², Bin Huang², Lingxiang Wu², Min Wu², Jing Chen⁴, Chaochen Wang⁴, Bin Huang⁵, Caidong Liu², Hongbing Shen², Qianghu Wang², Xinyi Xia⁶, Pengping Li², Sali Lyu², and Ying Xiao⁴

¹Jiangsu University

²Nanjing Medical University

³Zhejiang University School of Medicine Sir Run Run Shaw Hospital

⁴Zhejiang University

⁵Nanjing University

⁶Jinling Hospital, Nanjing University School of Medicine, Nanjing Clinical College of Southern Medical University

January 30, 2024

Abstract

The identification of asymptomatic, non-severe presymptomatic, and severe presymptomatic coronavirus disease (COVID-19) in patients may help optimize risk-stratified clinical management and improve prognosis. This single-center case series from Wuhan Huoshenshan Hospital, China, included 2,980 patients with COVID-19 who were hospitalized between February 4, 2020 and April 10, 2020. Patients were diagnosed as asymptomatic (n=39), presymptomatic (n=34), and symptomatic (n=2,907) upon admission. This study provided an overview of asymptomatic, presymptomatic, and symptomatic COVID-19 patients, including detection, demographics, clinical characteristics, and outcomes. Upon admission, there was no significant difference in clinical symptoms and CT image between asymptomatic and presymptomatic patients for diagnosis reference. The mean area under the receiver operating characteristic curve (AUC) of the differential diagnosis model to discriminate presymptomatic patients from asymptomatic patients was 0.89 (95% CI, 0.81-0.98). The severe and non-severe presymptomatic patients can be further stratified (AUC = 0.82). The two-step risk-stratification model based on 10 laboratory indicators on admission can facilitate early identification of asymptomatic patients with COVID-19 and predict illness severity for appropriate clinical management. Moreover, single-cell data analyses revealed that the CD8+T cell exhaustion contributed to the progression of COVID-19.

Introduction

Coronavirus disease (COVID-19) is a self-limiting disease in more than 80% of patients, and severe pneumonia occurs in approximately 15% of patients. However, rapid disease transmission turned the COVID-19 outbreak into a pandemic (1-3). The COVID-19 outbreak began and spread worldwide in December 2019 (4, 5), and on January 4, 2021, the World Health Organization reported a total of 83,910,386 COVID-19 cases globally.

Epidemiological estimates suggest that 50% of patients testing positive for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) on nucleic acid testing have no typical symptoms (i.e., fever, cough, or shortness of breath) (6-9). Asymptomatic and presymptomatic viral shedding not only poses a big challenge to infection control but also adds complexity to appropriate clinical decision-making and resource allocation(10, 11). It

is noteworthy that the mortality rate varies from country to country, possibly reflecting the differences in how rapidly local health authorities respond to isolate and initiate effective stratification and management strategies for the infected population (12). Hospitalization of all infected persons places too much burden on the healthcare system, but management based only on symptoms may not be appropriate(13). The prediction of asymptomatic, non-severe presymptomatic, and severe presymptomatic COVID-19 in patients could facilitate clinical resource allocation by health authorities, and improve the prognosis of patients. However, at present, there is no method for the risk stratification of asymptomatic and presymptomatic COVID-19 patients.

This study was conducted to comparatively evaluate the clinical characteristics of COVID-19 patients and construct a two-step risk-stratification model based on clinical indicators to distinguish among asymptomatic, severe presymptomatic, and non-severe presymptomatic COVID-19 patients on admission.

Methods

Data collection

From February 4, 2020 to April 10, 2020, a total of 3,046 COVID-19 patients' data was collected from Wuhan Huoshenshan Hospital in China. The extracted information from electronic medical records included clinical diagnosis, viral RNA load, chest CT imaging, and laboratory indicators. This study was approved by the hospital ethics committee, and the written informed consent was provided by every patient enrolled.

Patients classification

Patients were classified into asymptomatic, presymptomatic, and symptomatic groups based on the retrospective review of the medical records. Asymptomatic COVID-19 patients were those with SARS-CoV-2 nucleic acid test positive but without any clinical symptom or only with stable pre-exist chronic symptoms (e.g., chronic cough without worsening). Presymptomatic patients were those asymptomatic inpatients on admission but then developed typical (fever, cough, shortness of breath) or atypical symptom (14, 15). Symptomatic patients were those who had typical clinical symptoms on admission.

The diagnosis and severity degree of each patient was determined according to the clinical classification criterion in Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia released by the National Health Commission (6th or 7th edition). In clinical, patients were classified into mild or moderate, severe, and critical cases, depending on the above criterion. In our analysis, the clinically defined severe or critical COVID-19 patients were grouped into severe cases group, and the clinically defined mild or moderate COVID-19 patients were defined as non-severe group cases.

Baseline correction and statistical analysis

The propensity score matching (PSM) method was used to screen symptomatic patients based on the clinical profiles of asymptomatic and presymptomatic COVID-19 patients upon admission to hospital (16). Confounding factors include age, gender, hypertension, diabetes, cardiovascular disease, cerebrovascular disease, cancer, chronic obstructive pulmonary disease, chronic kidney disease, chronic liver disease, and immunodeficiency. For continuous variables, the median and interquartile range (IQR) were used to show the distribution, and the Wilcoxon rank-sum test with two-sided was used to compare the differences between groups. Number and percentage were used to show discrete variables, and Fisher's exact test with two-sided was used to compare the differences. All statistical analyses were performed using R software.

Single-cell transcriptional profiles of COVID-19 patients

Single-cell sequencing data of eight healthy lung tissues, three mild, and six severe lung lavage fluid from COVID-19 patients were downloaded from the GEO database with the accession number GSE145926 (17). The quality control, standardization, integration, dimensionality reduction, and unsupervised cell clustering of the samples were carried out using R package Seurat (v3.0.2). For each patient, the same standard for quality control was used: 1) each cell expressed 200-6,000 genes, and the number of expressed genes was

greater than 1,000; 2) each gene was expressed in at least three cells; 3) The proportion of mitochondrial genes was less than 10%.

Two-step risk stratification model

In order to improve prognosis of patients without clinical symptoms on admission, we designed a two-step risk-stratification model. (1) Initially, we constructed a random forest classifier based on AdaBoost with 39 asymptomatic and 34 presymptomatic patients. The diagnostic model could accurately classify asymptomatic and presymptomatic patients in 5-fold cross-validation. (2) Next, to predict the progression of presymptomatic COVID-19 patients, we constructed a random forest classifier based on AdaBoost with the training set of 1,751 symptomatic patients. The model was then validated in 9 non-severe and 25 severe presymptomatic patients.

The r-adabag package was used to perform the ensemble learning. Candidate features used to construct the classifier were 63 laboratory indicators indicated in Table S2, Table S3. Patients whose laboratory tests with more than 80% missing value were excluded. To apply the model for early application and avoid the interference of therapy, we only selected the laboratory test data within one day after admission. If the patient had multiple tests, the mean value was used for analysis. Missing values were supplied by the method of Conditional Mean Completer (CMC).

Results

Demographic and clinical characteristics of asymptomatic, presymptomatic, and symptomatic COVID-19 patients

Of the initial 3,046 COVID-19 patients who were screened for study inclusion, 66 were excluded, including those who were readmitted ($n=11$), had negative nucleic acid test results ($n=19$), or had missing data on the degree of illness severity ($n=36$). Thus, 2,980 patients were included in the study: 39, 34, and 2,907 asymptomatic, presymptomatic (non-severe, $n=9$, 26%; severe, $n=25$, 74%), and symptomatic (non-severe, $n=1,385$, 58%; severe, $n=1,522$, 52%) patients, respectively (Fig. 1). To ensure the even distribution of confounders among the study groups, baseline correction was performed (16), and we obtained a sample of 1,751 symptomatic COVID-19 patients, of whom 907 (51.8%) and 844 (48.2%) had non-severe and severe illness, respectively (Fig. 2A). The commonest symptoms in symptomatic patients from disease onset to hospital admission were fever ($n=2,178$, 75%) and cough ($n=2,097$, 72%). Shortness of breath, respiratory distress, and fatigue were observed in 1,320 (45%), 1,157 (40%), and 681 (23%) patients, respectively. Less common symptoms in symptomatic COVID-19 patients included muscle soreness ($n=118$, 4%), diarrhea ($n=66$, 2%), headache ($n=18$, 1%), dizziness ($n=26$, 1%), nausea ($n=21$, 1%), and vomiting ($n=21$, 1%). Of note, respiratory failure, a sign of severe illness, was observed in only 36 symptomatic COVID-19 patients with pneumonia (<1%; Table S1).

The median age of the patients was 60 years (IQR: 50–68 years), and there were no significant differences in age and sex among the three groups. Basic diseases (such as hypertension and diabetes) were detected in asymptomatic (15.38%), presymptomatic (20.58%), and symptomatic (16.44%) COVID-19 patients (Fig. 2B, Table S1). Surprisingly, those in the presymptomatic severe group had higher rates of admission to the intensive care unit (ICU; 8.3% vs 3.3%, P -value=0.2, Fig. 2C) and mortality (4% vs 1.8%, P -value=0.37, Fig. 2D) than patients in the symptomatic severe group. Additionally, there were no significant differences in viral RNA load among asymptomatic, presymptomatic, and symptomatic patients (Fig. 2E). It indicated that infected persons may still transmit the virus to others and should be quarantined, regardless of symptoms(6). The levels of specific immunoglobulin G (IgG) and immunoglobulin M (IgM) antibodies against SARS-CoV-2 were significantly higher in symptomatic patients than in asymptomatic and presymptomatic patients (Fig. 2F), indicating that patients might have variable antibody reactions despite similar viral RNA loads.

Laboratory indicators in asymptomatic and presymptomatic COVID-19 patients

On admission, the most common computed tomography (CT) findings were ground-glass opacity and lung-

texture increase in asymptomatic (83% and 50%, respectively), presymptomatic (83% and 61%, respectively), and symptomatic (92% and 62%, respectively) patients, with no significant intergroup differences (18). Other chest CT characteristics such as fuzzy boundary (range 8–17%), consolidation (0–12%), and pleural thickening (range 4–7%) were observed in asymptomatic, presymptomatic, and symptomatic patients, but without statistical significance. Interestingly, pleural effusion (7%), fine reticular opacity (2%), and lymph node enlargement (1%) were only observed in symptomatic patients (Fig. 3A). The presence of enlarged lymph nodes in symptomatic patients may indicate a stronger immune response.

To obtain more information on the affected organs, we compared laboratory indicators between the groups of asymptomatic, presymptomatic, and symptomatic COVID-19 patients. The rates of abnormal (one-fold away from the reference range) laboratory indicators associated with liver and kidney function were all above 60% in each group (Fig. 3B), and the rates of abnormal laboratory indicators associated with heart function were less than 40%. Specifically, when the threshold for abnormality was set to two-fold points from the reference range, asymptomatic patients showed lower rates of abnormal laboratory indicators of hepatic function (2.8% *vs* 20%, P -value < 0.05) and cardiac function (5.7% *vs* 10%, P -value > 0.05) and comparable rates in terms of kidney function (11.43% *vs* 10%, P -value > 0.05 , Fig. 3C) than presymptomatic patients. Abnormal levels of lactate dehydrogenase, albumin, uric acid, D-dimer (DD), C-reactive protein, and calcium were less frequently observed in asymptomatic and presymptomatic patients than in symptomatic patients (Fig. 3D, Table S2 and Table S3). These analyses indicate that the liver and kidney were the most common affected organs in all COVID-19 patients, whereas indicators for heart and liver function showed the most significant differences between asymptomatic and presymptomatic patients.

Diagnosis and prognosis stratification model for asymptomatic and presymptomatic patients

To ascertain the differences between severe and non-severe COVID-19 patients who were asymptomatic at the time of admission, data of the 63 indicators listed in Table S2 and Table S3 were retrieved and systematically compared. We found systematic differences in laboratory indicators between severe and non-severe COVID-19 patients who were asymptomatic on admission (Fig. S1 and Table S3). To optimize precise risk-stratification management and improve the prognosis of patients, we constructed a two-step risk-stratification model with the data obtained from 39 asymptomatic, 34 presymptomatic, and 1,751 symptomatic COVID-19 patients. The initial diagnostic model developed with the top 10 laboratory indicators could accurately classify asymptomatic and presymptomatic patients (area under the curve [AUC]=0.89, 5-fold cross-validation 95% confidence interval [CI] 0.81–0.98; Fig. 4A), and the degrees of contribution of the relevant 10 indicators are shown in Fig. 4B. It is noteworthy that the top five indicators (brain natriuretic peptide (BNP), IgG, IgM, glucose (GLU), and DD) might have similar stratification efficiencies (AUC=0.86, 5-fold cross-validation 95% CI, 0.73–1.00, Fig. S2A).

To stratify the non-severe and severe presymptomatic patients, we included the non-severe and severe symptomatic patients in a training dataset, and the constructed model with the top 10 laboratory indicators was then used to predict the severity of disease in presymptomatic patients with an AUC of 0.82 (Fig. 4C). The degrees of contribution of the 10 indicators included in this model are shown in Fig. 4D. The top five indicators were BNP, IgG, IgM, procalcitonin (PCT), and albumin/globulin ratio (A/G ratio), which could accurately predict illness severity (AUC=0.75, Fig. S2B), and their degrees of contribution were 26%, 16%, 11%, 7%, and 4%, respectively (Fig. 4D). Severe presymptomatic patients had significantly higher levels of BNP, PCT, magnesium, etc. than non-severe presymptomatic patients (Fig. 4E). Moreover, we obtained the validation data of 27 presymptomatic patients (non-severe, $n=25$; severe, $n=2$) from another study center (the First People's Hospital, Jiangxia District, Wuhan, China) and found that the levels of N-Terminal-proBNP (NT-proBNP), PCT, magnesium, etc., were high in severely ill COVID-19 patients (Fig. 4F). These results suggest that the application of laboratory indicators to stratify asymptomatic and presymptomatic COVID-19 patients is feasible, and such a risk-stratification model could be used to screen non-severe and severe presymptomatic patients on admission. The package-based stratification calculator was provided online (<https://github.com/liangyuan-njmu/SeverePredictModel>) for clinicians to assess whether presymptomatic patients will show disease progression.

CD8+ T-cell exhaustion contributes to the progression of COVID-19

Lymphocytopenia, defined as a lymphocyte count of $<1,500/\text{mm}^3$, was seen in symptomatic patients (35%) but was less frequently observed in asymptomatic (14%) and presymptomatic patients (17%). There were significant differences in lymphocyte percentages between asymptomatic and symptomatic patients (29.7% *vs.* 26.1%, P -value < 0.05 , Fig. 5A). Moreover, we analyzed the expression levels of cytokines in symptomatic COVID-19 patients and found that the expression levels of interleukin 6 (IL-6), IL-10, and IL-17 in severely ill patients were significantly higher than those in non-severe patients ($P < 0.001$, Fig. S3). In addition, those in the asymptomatic and presymptomatic groups had lower levels of IL-6 (median: 1.5 *vs.* 2 pg/ml, P -value = 0.012, Table S2) than patients in the symptomatic group. These findings were consistent with previous reports (19, 20) and indicated that asymptomatic patients have a weaker immune response to SARS-CoV-2 infection than symptomatic patients and play an immune protective role without cytokine storms.

To investigate the association of lymphocyte function with COVID-19 severity, we further analyzed single-cell sequencing data obtained from lung lavage fluids. Compared with that in severe COVID-19 patients, the proportion of lymphocytes (mainly T lymphocytes) increased significantly in mild COVID-19 patients (Fig. 5B). The T-distribution stochastic neighbor embedding (t-SNE) plot showed that 11 main cell clusters were identified from COVID-19 patients (Fig. 5C and 5D). CD8+ T-cells were subclustered by the known markers of CD3D and CD8A (Fig. 5E). Furthermore, the inhibitory membrane receptors expressed by CD8+ T-cells, such as PDCD1 (PD1), CTLA4, and HAVCR2 (TIM3), were significantly upregulated in severe COVID-19 patients when compared with those in healthy people and mild COVID-19 patients (Fig. 5F). Surprisingly, the expression of interferon gamma (IFN γ) in severely ill patients was significantly higher than that in non-severely ill patients. To identify the specific inhibitor phenotype related to IFN γ , the co-expression of cytokines and inhibitory membrane receptors was further analyzed. This revealed that the LAG3+IFN γ +CD8+ T-cells (10.12% CD8+T-cells in mildly symptomatic COVID-19 patients and 25.83% CD8+T-cells in severely symptomatic COVID-19 patients) and CTLA4+IL10RA+CD8+ T-cells (4.42% CD8+T in mildly symptomatic COVID-19 patients and 13.88% CD8+T in severely symptomatic COVID-19 patients) might contribute to the T-cell exhaustion associated with the severe phenotype of COVID-19 in patients (Fig. 5G–J).

Discussion

To control the worldwide pandemic of SARS-CoV-2 infection and efficiently allocate medical resources, the screening and management of asymptomatic or presymptomatic COVID-19 patients in the population are urgent challenges. This retrospective study systematically characterized the asymptomatic, presymptomatic, and symptomatic COVID-19 patients who were admitted to Wuhan Huoshenshan Hospital and developed a risk-stratification calculator that will help clinicians worldwide to make appropriate clinical decisions at the time of admission.

Asymptomatic and presymptomatic COVID-19 patients have similar CT findings, which indicates the urgent need for a risk-stratification model to differentiate these patients. Despite the absence of validation cohorts to confirm the performance accuracy of the differential diagnosis model, the performance of this risk-stratification model is satisfactory and accurate (based on AUCs > 0.82 in both the development and validation cohorts). The package-based prediction calculator can be utilized by clinicians to perform risk stratification of asymptomatic, non-severe presymptomatic, and severe presymptomatic COVID-19 patients based on 5–10 commonly evaluated indicators. This makes the package-based calculator easy to utilize, facilitating clinical resource reallocation, such as for ICU beds and ventilators.

Furthermore, the risk-stratification model can be utilized for symptomatic and presymptomatic COVID-19 patients, which implies that the indicators identified in the analysis possibly have intrinsic associations with the development of COVID-19. A high BNP level is related to acute respiratory distress syndrome, sepsis, and congestive heart failure and contributes to higher mortality in patients with pneumonia (21, 22). PCT, as an indicator for bacterial coinfection, is approved by the US Food and Drug Administration to guide antibiotic treatment for suspected lower respiratory tract infections (23). Previously published literature reviews and

meta-analyses have indicated that the results of laboratory tests for PCT could indicate progression of COVID-19 (8, 24, 25), which greatly supports our results. There is limited information about the frequency and bacterial spectrum of pulmonary coinfections and superinfections in COVID-19 patients (26); however, our analysis strengthens the vital role of BNP and PCT evaluation in the clinic to verify the possibility of coinfections associated with SARS-CoV-2 infection.

Functional CD8⁺ T-cell exhaustion contributes to the progression of SARS-CoV-2 infection (27, 28). As indicated in a previous report, loss of function occurs hierarchically in exhausted CD8⁺ T-cells (29). In the initial stage, CD8⁺ T-cells are unable to produce IL-2. During the intermediate stage, CD8⁺ T-cells may lose the ability to produce tumor necrosis factor (TNF). In the severe stage, CD8⁺ T-cells completely lose the ability to produce large amounts of IFN γ or there may be a physical deletion of virus-specific CD8⁺ T-cells. Based on the single-cell sequencing data, our re-analysis indicates that SARS-CoV-2 can induce intermediate-stage CD8⁺ T-cell exhaustion involving a lack of IL-2 and TNF α production but with limited IFN γ production (Fig. 5F), which may be mediated by LAG-3. Because LAG-3 cross-linking can inhibit IL-2, TNF α , and IFN γ production and T-cell proliferation (30). Therefore, LAG3+IFN γ +CD8⁺ T-cells may be the predominant phenotype showing exhaustion and may need to be further investigated to elucidate the pathogenic clinical severity of COVID-19.

As the retrieved and analyzed clinical data were based on all data from the clinical practice and were not derived from random selection, the data are deemed to be descriptive to some extent. For a comparable data analysis, baseline normalization by the propensity score matching method was performed, and this strategy helped reveal more information from these data. We were unable to verify the accuracy of the model due to lack of data from another research center, resulting in only partial validation of our risk-stratification model.

Conclusion

This risk-stratification model can facilitate the early prediction of disease progression in asymptomatic, severe presymptomatic, and non-severe presymptomatic patients based on 10 laboratory indicators and thereby help make appropriate clinical decisions and optimize the use of limited resources. This model and the 10 laboratory indicators can help improve the clinical management of COVID-19 patients who are at different stages of disease evolution.

Declarations

Ethical Approval

This study was approved by the Medical Ethical Committee of Wuhan Huoshenshan Hospital (HSSLL011), and the Ethical Committee of Nanjing Medical University (2020-511). Written informed consent was obtained from each patient.

Consent to Participate

Not applicable

Consent to Publish

Not applicable

Authors Contributions

YX, SL, PL, and XX had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. LS, RD, TZ, and WW contributed equally. Concept and design: YX, SL, PL, and XX. Experiments and data collection: XX. Data analysis and interpretation: SL, RD, TZ, WW, ZW, XJ, KL, YL, JL, MZ, BH, LW, MW, JC, CW, BH, HS, CL, QW.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 81972358, 91959113), Key Foundation of Wuhan Huoshenshan Hospital (Grant No. 2020[18]), Key Research & Devel-

opment Program of Jiangsu Province (Grant No. BE2017733, BE2018713), Medical Innovation Project of Logistics Service (Grant No. 18JS005), Basic Research Program of Jiangsu Province (Grant No. BK20180036). Novel coronavirus pneumonia technology project of Kunshan first people's Hospital (XGF202007).

Competing Interests

The authors declare that they have no conflict of interest.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung SM, Hayashi K, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis.* 2020;94:154-5.
2. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill.* 2020;25(10).
3. Sawano T, Ozaki A, Rodriguez-Morales AJ, Tanimoto T, Sah R. Limiting spread of COVID-19 from cruise ships: lessons to be learnt from Japan. *QJM.* 2020;113(5):309-10.
4. Lin Q, Zhao S, Gao D, Lou Y, Yang S, Musa SS, et al. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int J Infect Dis.* 2020;93:211-6.
5. Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM, et al. Coronavirus Disease 2019 (COVID-19): A Perspective from China. *Radiology.* 2020;296(2):E15-E25.
6. Lee S, Kim T, Lee E, Lee C, Kim H, Rhee H, et al. Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *JAMA Intern Med.* 2020.
7. Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al. Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(13):377-81.
8. 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;323(13):1239-42.
9. Acter T, Uddin N, Das J, Akhter A, Choudhury TR, Kim S. Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: A global health emergency. *Sci Total Environ.* 2020;730:138996.
10. Moghadas SM, Fitzpatrick MC, Sah P, Pandey A, Shoukat A, Singer BH, et al. The implications of silent transmission for the control of COVID-19 outbreaks. *Proc Natl Acad Sci U S A.* 2020;117(30):17513-5.
11. Qiu J. Covert coronavirus infections could be seeding new outbreaks. *Nature.* 2020.
12. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Intern Med.* 2020;180(8):1081-9.
13. Zheng X, Luo S, Sun Y, Han M, Liu J, Sun L, et al. Asymptomatic patients and asymptomatic phases of Coronavirus Disease 2019 (COVID-19): a population-based surveillance study. *National Science Review.* 2020;7(10):1527-39.

14. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19. *N Engl J Med.* 2020;382(22):2158-60.
15. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med.* 2020;382(22):2081-90.
16. Baek S, Park SH, Won E, Park YR, Kim HJ. Propensity score matching: a conceptual review for radiology researchers. *Korean J Radiol.* 2015;16(2):286-96.
17. Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med.* 2020;26(6):842-4.
18. Meng H, Xiong R, He R, Lin W, Hao B, Zhang L, et al. CT imaging and clinical course of asymptomatic cases with COVID-19 pneumonia at admission in Wuhan, China. *J Infect.* 2020;81(1):e33-e9.
19. Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang YQ, et al. Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther.* 2020;5(1):33.
20. Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med.* 2020;26(8):1200-4.
21. Mendez R, Aldas I, Menendez R. Biomarkers in Community-Acquired Pneumonia (Cardiac and Non-Cardiac). *J Clin Med.* 2020;9(2).
22. Karakioulaki M, Stolz D. Biomarkers in Pneumonia-Beyond Procalcitonin. *Int J Mol Sci.* 2019;20(8).
23. Kamat IS, Ramachandran V, Eswaran H, Guffey D, Musher DM. Procalcitonin to Distinguish Viral From Bacterial Pneumonia: A Systematic Review and Meta-analysis. *Clin Infect Dis.* 2020;70(3):538-42.
24. Lippi G, Plebani M. Procalcitonin in patients with severe coronavirus disease 2019 (COVID-19): A meta-analysis. *Clin Chim Acta.* 2020;505:190-1.
25. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect.* 2020;81(2):e16-e25.
26. Zhao D, Yao F, Wang L, Zheng L, Gao Y, Ye J, et al. A Comparative Study on the Clinical Features of Coronavirus 2019 (COVID-19) Pneumonia With Other Pneumonias. *Clin Infect Dis.* 2020;71(15):756-61.
27. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, et al. Dysregulation of Immune Response in Patients With Coronavirus 2019 (COVID-19) in Wuhan, China. *Clin Infect Dis.* 2020;71(15):762-8.
28. Zheng M, Gao Y, Wang G, Song G, Liu S, Sun D, et al. Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cell Mol Immunol.* 2020;17(5):533-5.
29. Wherry EJ. T cell exhaustion. *Nat Immunol.* 2011;12(6):492-9.
30. Macon-Lemaitre L, Triebel F. The negative regulatory function of the lymphocyte-activation gene-3 co-receptor (CD223) on human T cells. *Immunology.* 2005;115(2):170-8.

Figure legends

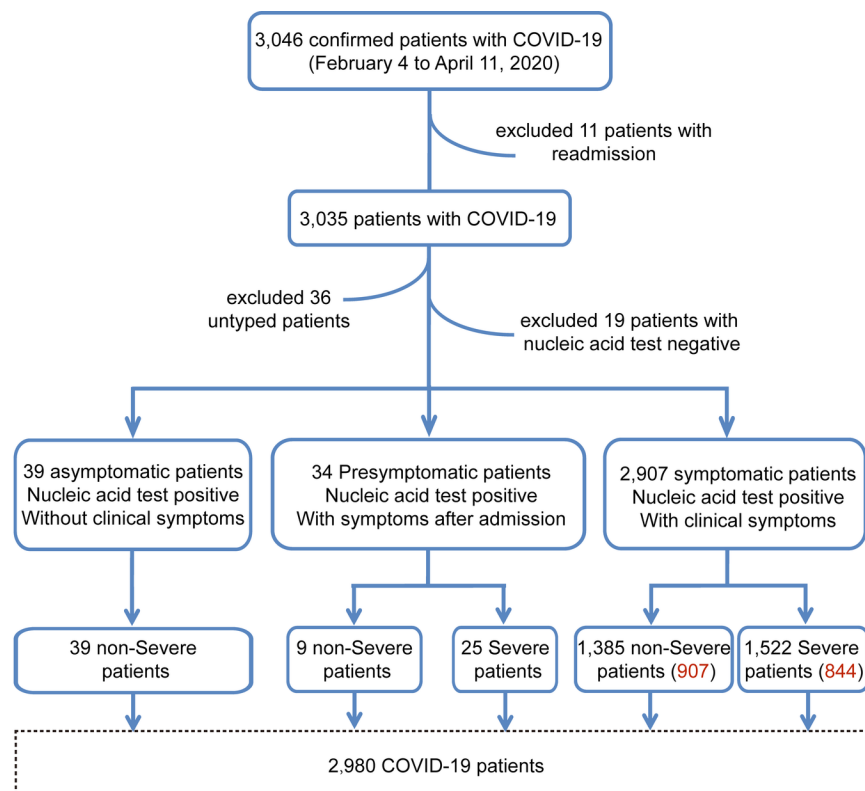


Fig. 1. The enrollment workflow of the COVID-19 cohort. The workflow chart shows the screening process of the cohort samples. The number of red words in parentheses represents the sample size after baseline correction.

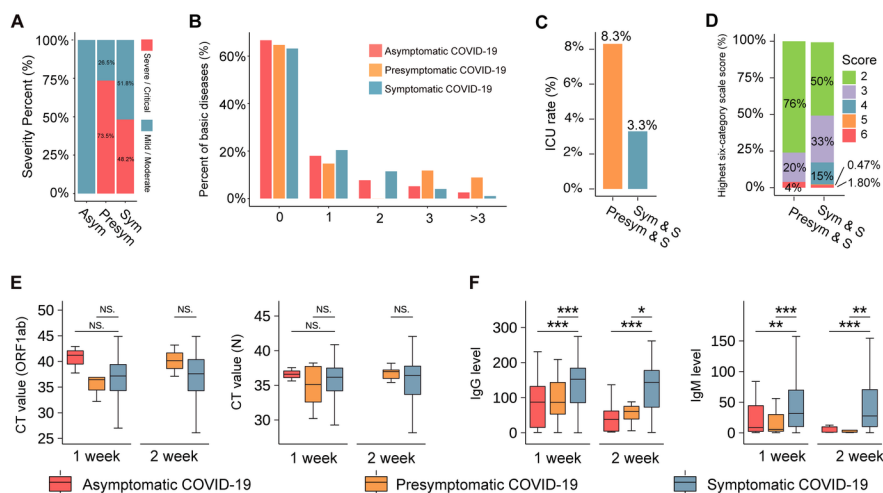


Fig. 2. The basic information of the COVID-19 cohort. (A) The severity percentage of asymptomatic, presymptomatic, and symptomatic patients with COVID-19. Blue represents non-severe (mild or moderate), and red represents severe or critical. (B) The basic disease rate of patients with COVID-19. The vertical axis represents the percentage of basic diseases (%). The horizontal axis represents the number of

basic diseases one patient has. (C) A comparison of the ICU rate between patients with presymptomatic and symptomatic severe COVID-19. The vertical axis indicates the rate of ICU. 'Pre-sym & S' and 'Sym & S' represent the severe presymptomatic and symptomatic patient group, respectively. (D) The highest six-category scale scores in severe COVID-19 patients. The vertical axis indicates the number of highest six-category scale scores (%). Scores represent their severity, including 1-discharged, 2-hospitalization without requiring oxygen therapy, 3-hospitalized with low-flow oxygen therapy, 4-hospitalized with non-invasive ventilation and/or high-flow oxygen therapy, 5-hospitalized for ECMO and/or mechanical ventilation, and 6-death. (E) and (F) The boxplot reflects the viral load (CT value) and SARS-COV-2 antibody IgG and IgM levels in three groups as a function of hospital stay. *** represents P -value < 0.001 . ** represents P -value < 0.01 . * represents P -value < 0.05 . NS. represents P -value > 0.05 . The three colors of red, orange and blue represent the asymptomatic, presymptomatic, and symptomatic group, respectively.

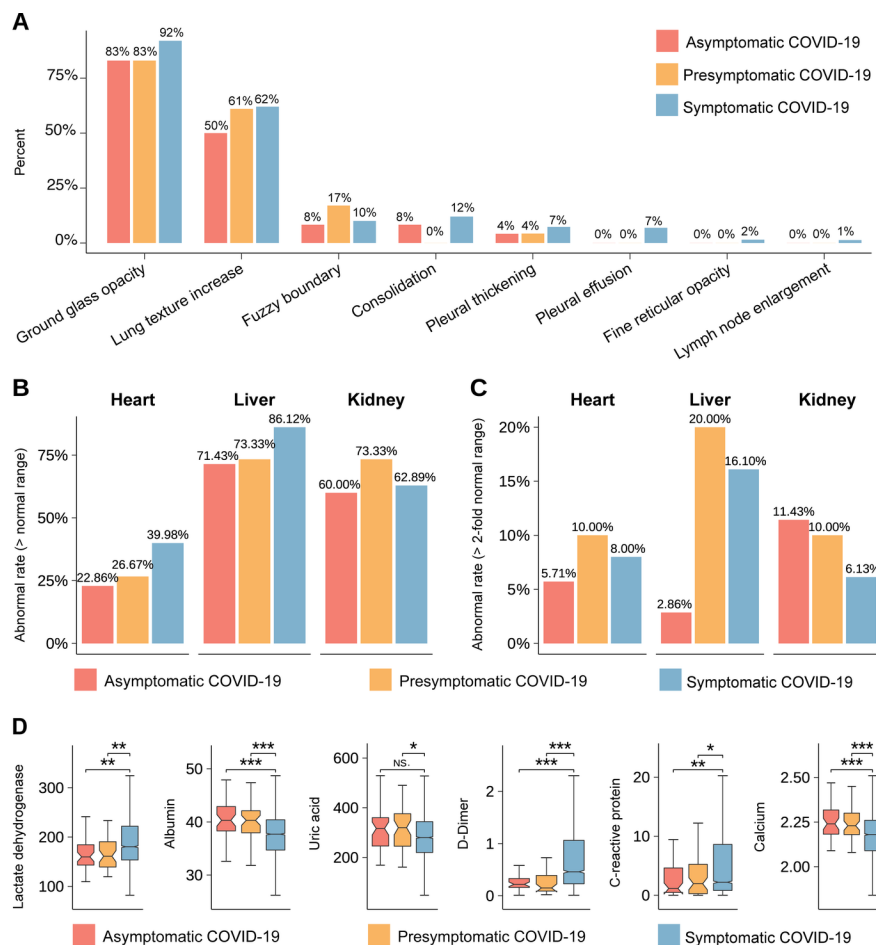


Fig. 3. Radiological characteristics and laboratory indicators detected in asymptomatic, presymptomatic, and symptomatic patients. (A) The percentages of typical characteristics in chest CT imaging. The vertical axis indicates the percentages of typical characteristics. The horizontal axis represents the typical characteristics of the lung. Red, orange and blue represent the asymptomatic, presymptomatic, and symptomatic group, respectively. (B) The abnormal rate of organs, including heart, liver, and kidney ($>$ normal range). The vertical axis indicates the rate of abnormal. (C) The abnormal rate of organs, including heart, liver, and kidney (>2 -fold normal range). (D) The values of laboratory indicators after admission. The vertical axis indicates the values of laboratory indicators. The P -value was calculated using the Wilcoxon's rank-sum test and shown at the top of each panel. Boxplots represent the 25th and 75th

percentiles, with midlines indicating the median values. *** represents P -value < 0.001 . ** represents P -value < 0.01 . * represents P -value < 0.05 . NS. represents P -value > 0.05 .

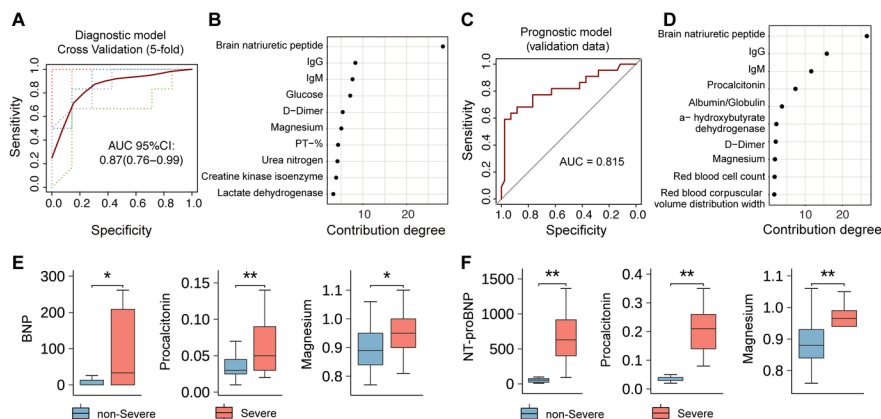


Fig. 4. A diagnosis and severity prediction model for COVID-19 patients. (A) The diagnostic model for asymptomatic and presymptomatic patients. The Receiver Operating Characteristic (ROC) curve of diagnostics of training (red line) and five-fold cross-validation (dotted lines). The dotted lines in different colors represent the results of 5-fold cross-validation. (B) The contribution degrees of markers (Top 10) in the diagnostic model. The horizontal axis represents the contribution degree of laboratory indicators. (C) The prognostic model for the severity of presymptomatic patients. (D) The contribution degrees of markers (Top 10) in the prognostic model. The horizontal axis represents the contribution degree of laboratory indicators. (E) and (F) A comparison of laboratory indicators between non-severe and severe patients without symptom on admission. The data set in (E) comes from Wuhan Huoshenshan Hospital in China. The validation data set in (F) comes from the First People's Hospital of Jiangxia District of Wuhan, Hubei, China. ** represents P -value < 0.01 . * represents P -value < 0.05 .

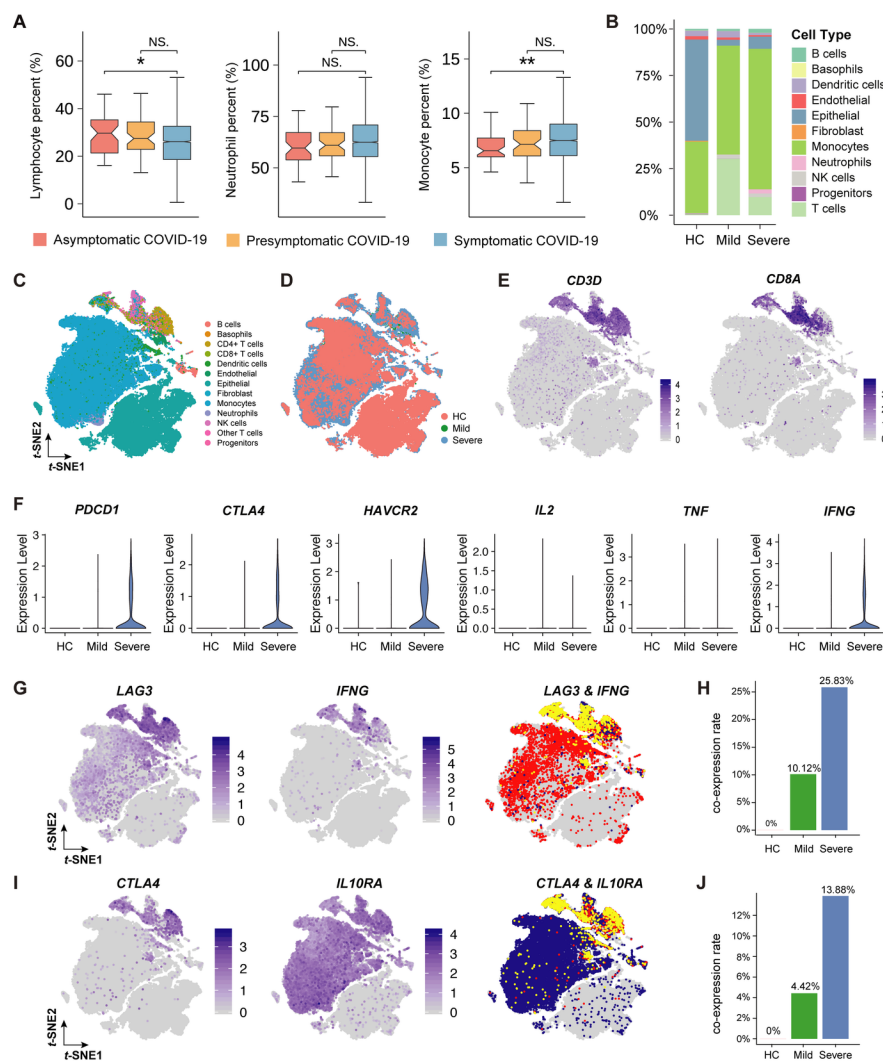


Fig. 5. Single-cell analysis of immune cells in COVID-19 patients. (A) The box plot describes the proportion of lymphocytes, neutrophil, and monocyte. The vertical axis indicates the values of laboratory indicators. The P -value was calculated using Wilcoxon's rank-sum test and shown at the top of each panel. Boxplots represent the 25th and 75th percentiles, with midlines indicating the median values. ** represents P -value < 0.01 . * represents P -value < 0.05 . NS. represents P -value > 0.05 . (B) The histogram shows the proportion of each cell type in healthy people (HC), mild patients, and severe patients. The vertical axis indicates the fraction of the cell type. Each color represents a cell type. (C) The t-distributed stochastic neighbor embedding (tSNE) plot of T cells for seventeen patients with COVID-19. (D) The tSNE plot of T cells colored by patient type. Red, green, and blue represent the HC, mild, and severe, respectively. (E) The tSNE plot of CD8+ T cells colored by cell-type-specific markers. The shade of the color represents the level of gene expression. (F) Violin plots of the inhibitory-receptors and cytokines' gene expression. The vertical axis indicates the level of gene expression. (G) and (I) The tSNE plot of inhibitory-receptor coexpression patterns. Yellow represents the cells expressing two genes at the same time. (H) and (J) The bar plot shows the proportion of inhibitory-receptor coexpression patterns. The vertical axis indicates the rate of coexpression patterns.