

# Prognostic Predictors For COVID-19 in Daily Clinical Practice in Japan: a Propensity Score-Matched Case-Control Study

**Masahiro Shinoda**

Tokyo Shinagawa Hospital

**Shinichiro Ota**

Tokyo Shinagawa Hospital

**Yuto Yoshida**

Tokyo Shinagawa Hospital

**Takatomo Hirouchi**

Tokyo Shinagawa Hospital

**Kanako Shinada**

Tokyo Shinagawa Hospital

**Takashi Sato**

Tokyo Shinagawa Hospital

**Miwa Morikawa**

Tokyo Shinagawa Hospital

**Naoki Ishii**

Tokyo Shinagawa Hospital

**Masaharu Shinkai** (✉ [shinkai050169@gmail.com](mailto:shinkai050169@gmail.com))

Tokyo Shinagawa Hospital

---

## Research Article

**Keywords:** COVID-19, prognosis, computed tomography, ground-glass opacity, consolidation, fever

**Posted Date:** October 6th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-917302/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Prognostic predictors for COVID-19 in daily clinical practice in Japan: a propensity score-matched case-control study**

Masahiro Shinoda MD<sup>1)</sup>, Shinichiro Ota MD<sup>1)</sup>, Yuto Yoshida MD<sup>1)</sup>, Takatomo Hirouchi MD<sup>1)</sup>, Kanako Shinada MD<sup>1)</sup>, Takashi Sato MD, PhD<sup>1)</sup>, Miwa Morikawa MD, PhD<sup>1)</sup>, Naoki Ishii MD, PhD, MPH<sup>2)</sup>, Masaharu Shinkai MD, PhD<sup>1)</sup>

1) Department of Respiratory Medicine, Tokyo Shinagawa Hospital, Tokyo, Japan

2) Department of Gastroenterology, Tokyo Shinagawa Hospital, Tokyo, Japan

**Corresponding author:**

Masaharu Shinkai MD, PhD

Department of Respiratory Medicine, Tokyo Shinagawa Hospital

6-3-22 Higashi-Oi, Shinagawa, Tokyo 140-8522, Japan

Tel: +81-3-3764-0511

Fax: +81-3-3764-3415

E-mail: [shinkai050169@gmail.com](mailto:shinkai050169@gmail.com)

**Running title:** Prognostic predictors for COVID-19

## **Abstract**

**Introduction:** Blood tests and computed tomography (CT) findings at diagnosis are widely used in daily clinical practice and can offer useful prognostic factors for coronavirus disease 2019.

**Methods:** We retrospectively evaluated 66 patients who underwent a blood test and CT between January 1 and May 31, 2020, and performed a propensity score-matched case-control study. Cases and controls were a severe respiratory failure group (non-rebreather mask, nasal high-flow, positive-pressure ventilation) and a non-severe respiratory failure group, matched at a ratio of 1:3 by propensity scores constructed by age, sex, and medical history. We compared groups for maximum body temperature up to diagnosis, laboratory findings, and CT findings in the matched cohort. Two-tailed P-values <0.05 were considered statistically significant.

**Results:** Nine cases and 27 controls were included in the matched cohort. Significant differences were seen in maximum body temperature up to diagnosis ( $p=0.0043$ ), the number of shaded lobes ( $p=0.0434$ ), amount of ground-glass opacity (GGO) in the total lung field ( $p=0.0071$ ), amounts of GGO ( $p=0.0001$ ), and consolidation ( $p=0.0036$ ) in the upper lung field, and pleural effusion ( $p=0.0117$ ).

**Conclusions:** Fever and CT findings (such as GGO and consolidation) may be prognostic indicators that can be easily measured at diagnosis.

**Keywords:** COVID-19, prognosis, computed tomography, ground-glass opacity, consolidation, fever

## Introduction

Since the first death from coronavirus disease 2019 (COVID-19) pneumonia was reported in Wuhan, China, in December 2019<sup>1)</sup>, the epidemic has spread to become a global pandemic, with no end yet in sight. As of May 31, 2021, a total of 169,597,415 confirmed cases and 3,530,582 deaths related to COVID-19 had been reported to the World Health Organization, with an apparent mortality rate of around 2.1%<sup>2)</sup>. As of the same date in Japan, a total of 745,392 people had been infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in Japan, with 12,926 deaths<sup>3)</sup>.

Pneumonia has been a common complication in COVID-19<sup>4)</sup>. Although many patients show mild disease, some patients develop acute respiratory distress syndrome (ARDS) and progress to a severe within 8–14 days after onset<sup>5)</sup>. Various radiological patterns of COVID-19 pneumonia have been reported at different times throughout the disease course<sup>6)7)8)</sup>, which may make the progression of the disease difficult to predict.

Aggravating factors reported to date include age<sup>9)</sup>, obesity<sup>10)</sup>, type 2 diabetes<sup>11)</sup>, hypertension<sup>12)13)</sup>, chronic kidney disease<sup>14)</sup>, and dyslipidemia<sup>15)</sup>, but among patients of the same age, physique, and underlying illness, illness can still range from mild to severe.

In Japan, patients at risk of aggravation are supposed to be hospitalized in principle, but as the number of patients with COVID-19 increases, tools for early prediction of patients likely to experience severe illness are expected to be developed to more appropriately allocate limited medical resources. Several prognostic scores have been developed in other countries for inpatients, such as COVID-GRAM<sup>16)</sup> and the ISARIC WHO 4C Mortality Score<sup>17)</sup>. Studies have also shown the utility of community-acquired pneumonia severity scores such

as A-DROP<sup>18</sup>). However, these are not yet sufficient and do not include the spread of opacities or radiological findings of pneumonia, which are speculated to be important factors in aggravated COVID-19.

We, therefore, performed a matched case-control study using propensity scores to analyze prognostic factors from among the results of physical examinations, blood tests, and CT at diagnosis in daily clinical practice.

## **Results**

Regarding the background characteristics of the 66 patients (38 men, 28 women), the mean age was 51.5 years (standard deviation, 17.2 years). Mean BMI was 24.7 kg/m<sup>2</sup> (standard deviation, 4.34 kg/m<sup>2</sup>). Hypertension was the most common comorbidity (n=14, 21.2%), followed by diabetes (n=10, 15.2%), and dyslipidemia (n=2, 3%). Fifty-nine patients were hospitalized, and the mean duration of hospitalization was 18.7 days (standard deviation 14.5 days). Fifty-five patients were confirmed negative by RT-PCR after an average of 20.7 days from diagnosis (standard deviation, 9.19 days).

Twenty-seven patients had received oxygen therapy or ventilation management. Two patients received nasal high-flow therapy, three had NIPPV, seven had mechanical ventilation, and two had mechanical ventilation and extracorporeal membrane oxygenation. Thirty-six patients received antiviral treatments such as favipiravir and/or remdesivir and 50 patients were treated with corticosteroids. Four patients (6.1%) died.

Distributions of baseline characteristics in patients from the severe and non-severe respiratory groups are shown in Table 1 in both unmatched and matched samples. The area under the ROC of propensity scores for severe respiratory failure was 0.83 (95% confidence interval, 0.72–0.94).

In the unmatched sample, the severe respiratory failure group was older, with higher BMI and a higher prevalence of diabetes. There was no significant difference in the period from onset to diagnosis (median [IQR] 4.00 [3.00–6.75] days in non-severe respiratory failure group vs 5.00 [4.00–7.00] days in severe respiratory failure group,  $p=0.448$ ).

Although baseline characteristics seemed well balanced between severe and non-severe respiratory failure groups after propensity score matching, absolute standardized differences were not less than 0.1 in sex and history of diabetes mellitus (Table 1).

Table 2 shows the results of maximum fever up to diagnosis, blood findings, and CT findings at the first visit. CRP, LDH, uric acid, and fibrinogen were more elevated in the severe respiratory failure group. There was no significant difference in the period from the onset to the day CT was performed (median [IQR] 4.50 [3.00–6.75] days in non-severe respiratory failure group vs 5.00 [4.00–6.25] days in severe respiratory failure group,  $p=0.671$ ). GGO was more extensive than consolidation in both non-severe and severe respiratory failure groups.

Distributions of both GGO and consolidation were increased toward the bottom of the lungs. Opacities were almost always evident bilaterally. In the severe respiratory failure group, shadows were more frequently multilobed. Mediastinal lymphadenopathy was observed in 34 cases (51.5%) and tended to be more common in the severe respiratory failure group. Two patients classed as serious cases showed pleural effusion. Pleural thickening was observed in 11 cases (16.7%) and tended to be more common in the non-severe respiratory failure group. Twenty-one cases (31.8%) showed subpleural curvilinear lines/stripe shadows. Vascular enlargement was observed in two cases (3.0%). Traction bronchiectasis was observed in one severe case.

Among matched patients, significant differences in maximum body temperature up to diagnosis, number of shaded lung lobes, amount of GGO in the total lung field, amount of GGO in the upper lung field, amount of consolidation in the upper lung field, and presence of pleural effusion differed significantly. The property of severe respiratory failure was significantly associated with increases in these factors except pleural effusion on multivariate linear regression after adjusting the factors used for the construction of propensity scores in the matched cohort (Table 3). The optimal cutoff for GGO in the total lung field on logistic regression probabilities was 14.6% (Figure 2).

## **Discussion**

In our study using propensity score matching to reduce bias due to confounding factors by matching patients for baseline variables using multivariable logistic regression modeling, significant differences in maximum body temperature were seen up to diagnosis, the number of shaded lung lobes, amount of GGO in the total lung field, amount of GGO in the upper lung field, amount of consolidation in the upper lung field, and presence of pleural effusion. Blood findings did not differ significantly, and fever and radiological findings were more prognostic at diagnosis.

COVID-19 is considered a disease for which the prognosis changes significantly depending on age and underlying disease, and matching patient background characteristics as confounding factors and extracting prognosis-related factors from physical and laboratory findings is very important.

Regarding factors that may be confounders for prognosis in the patient background, Wu et al.<sup>9)</sup> found that older patients and those with more comorbidities were at greater risk of developing ARDS among patients

infected with COVID-19. Patients with COVID 19 with diabetes are known to progress to severe illness<sup>11</sup>.

Liang et al.<sup>19</sup> found that patients with COVID-19 and cancer showed a higher risk of severe events compared with patients without cancer. A dose-response meta-analysis by Pranata et al.<sup>10</sup> showed that increased BMI was associated with an increased frequency of poor outcomes among patients with COVID-19. Other prognostic factors included hypertension<sup>12)13</sup>, dyslipidemia<sup>14</sup>, chronic kidney disease<sup>15</sup>, and chronic obstructive pulmonary disease<sup>20</sup>. In this study, significant differences were seen in age, BMI, and diabetes, as in previous reports, but not in hypertension, dyslipidemia, or chronic renal failure. Regarding the prevalence of cancer, none of the patients in this study had cancer and could not be evaluated.

In terms of radiographic findings, the number of shaded lung lobes, amount of GGO in the total lung field, amount of GGO in the upper lung field, amount of consolidation in the upper lung field, and presence of pleural effusion differed significantly after propensity score matching.

In the early stages of COVID-19, lesions usually manifest as localized GGO, subpleural bands, vascular enlargement, and peripheral distribution on CT. This is because the virus invades and replicates in the alveolar epithelium and exudates leak mainly into the alveolar space with distributions mainly under the pleura or around the peribronchovascular regions<sup>21)22</sup>. With disease progresses, the range of involved alveoli and mucosa increases, and bronchial walls swell, contributing to patterns of air bronchograms with consolidation and bronchial wall thickening. Patterns of crazy-paving, interlobular septal thickening, and reticulation basically reflect the involvement of pulmonary interstitium, such as interlobular interstitial edema<sup>23</sup>.

In most patients, the lower lobes are involved more frequently than the upper and middle lobes. A peripheral distribution, multilobar involvement, and posterior involvement are other important characteristics of lesions distribution<sup>24</sup>).

According to a report by Ye et al.<sup>4</sup>), pleural effusion was observed in about 5% and pleural thickening in about 30%, subpleural curvilinear lines in about 20% and lymphadenopathy in 4–8%. Subpleural curvilinear lines are reportedly related to pulmonary edema or fibrosis<sup>4</sup>). Stripe shadows were thought to be related to the healing of pulmonary chronic inflammation or proliferative disease<sup>4</sup>), with gradual replacement of cellular components by scar tissue. Vascular enlargement is described as the dilation of pulmonary vessels around and within the lesions, which might be attributed to damage to and swelling of the capillary wall caused by pro-inflammatory factors<sup>4</sup>).

Song et al.<sup>25</sup>) found more consolidative lesions in patients with a longer interval between symptom onset or with age over 50 years and suggested that this manifestation may warrant greater attention in the management of patients. In a meta-analysis of chest features on CT by Zheng et al.<sup>24</sup>), compared to findings in common patients, some CT manifestations were more frequent in severe patients, such as traction bronchiectasis, interlobular septal thickening, consolidation, crazy-paving pattern, reticulation, pleural effusion, and lymphadenopathy. However, these findings are not seen in the early stages of the disease and are not considered predictors of prognosis in the early stage. In addition, relatively few reports have provided details of radiological findings in the early stages of the disease.

In our study, significant differences were observed in the amount of GGO, the amount of GGO or consolidation at each site, or the presence of pleural effusion. Traction bronchiectasis, as a finding in the advanced phase, was rarely observed. No significant differences in lymphadenopathy were identified. This was probably because cases in this study were rarely swollen, with a short axis diameter of 1 cm or more.

Based on the present results, early predictors of prognosis could be widespread GGO, widespread GGO extending to the upper lung field, early presence of consolidation in atypical sites that is more frequently seen in the late phase, and pleural effusion that is usually seen in the late phase.

In this study, maximum body temperature up to diagnosis was a predictor of prognosis. Fever was considered a predictor of mortality by Iftime et al.<sup>26)</sup> and the same result was obtained in this study.

Zhou et al.<sup>27)</sup> reported that elevated white blood cell count, decreased lymphocyte count, thrombocytopenia, elevated lactate dehydrogenase, elevated creatinine, elevated D-dimer, elevated ferritin, elevated high-sensitivity cardiac troponin I, and elevated interleukin (IL)-6 were seen in patients who died from COVID-19.

Ayanian et al.<sup>28)</sup> reported that monitoring IL-6, D-dimer, CRP, LDH, and ferritin was clinically useful in this respect, particularly when these markers were above the mentioned cutoff values. d'Alessandro et al.<sup>29)</sup> reported KL-6 measurement as potentially useful for evaluating the prognosis of COVID-19 patients. Sugiyama et al.<sup>30)</sup> reported that elevated IFN- $\lambda$ 3, elevated IP-10, elevated C-X-C motif Chemokine ligand (CXCL9), and low Thymus and activation-regulated chemokine (TARC) could be promising prognostic markers to distinguish between mild/moderate and severe/critical patients.

In this study, no significant difference in blood test findings were confirmed at the first visit in daily clinical practice, indicating the difficulty of predicting disease severity at the first visit and also revealing fever and radiological findings as more important factors than blood test.

Currently, severity prediction scores are being prepared. COVID-GRAM was developed in China and includes age, underlying disease, history of malignancy, chest X-ray, symptoms (dyspnea, bloody sputum, impaired consciousness), and blood findings (neutrophils/lymphocyte ratio, LDH, and direct bilirubin)<sup>16</sup>. The accuracy of the COVID risk score in the validation cohort was similar to that in the development cohort, with an AUC in the validation cohort of 0.88 (95%CI, 0.84–0.93)

The 4C mortality score developed in the United Kingdom consists of age, sex, number of underlying illnesses, respiratory rate, saturation of percutaneous oxygen (SpO<sub>2</sub>), Glasgow Coma Scale, and blood tests (blood urea nitrogen (BUN), CRP)<sup>17</sup>. This score showed high discrimination for mortality (derivation cohort: area under the receiver operating characteristic curve, 0.79; 95% confidence interval, 0.78–0.79; validation cohort, 0.77, 0.76–0.77) with excellent calibration (validation: calibration-in-the-large=0, slope=1.0).

Fan et al.<sup>18</sup>) compared the accuracy of several severity scores for community-acquired pneumonia including Pneumonia Severity Index (PSI), CURB-65 and CRB-65, A-DROP, SMART-COP, NEWS2 and qSOFA were also compared. Among all seven scores determined using patient information on admission, A-DROP presented the highest discrimination (AUC 0.87; 95%CI 0.84–0.90), following by CURB-65 (AUC 0.85; 95%CI 0.81–0.89), PSI (AUC 0.85; 95%CI 0.81–0.88), SMART-COP (AUC 0.84; 95%CI 0.80–0.88), NEWS2 (AUC 0.81;

95%CI 0.77–0.85), CRB-65 (AUC 0.80; 95%CI 0.76–0.84), and qSOFA (AUC 0.73; 95%CI 0.69–0.78) for predicting in-hospital death.

The prognosis prediction score for COVID-19, which shows few subjective symptoms in the early stage and potentially progresses, seems to depend on severe symptoms and hypoxemia, and may be unsuitable for predicting prognosis at an earlier stage. In addition, that score does not include CT findings, and addition of factors for the distribution of GGO and consolidation at an early stage would be better, based on the results of this study. The quantification of GGO and consolidation can generally be measured by installing the software. In addition, referring to the fact that the cutoff for the ratio of GGO to the entire lung field in all cases for severe and non-severe cases was 14.6%, for convenience, GGO that occupies one-sixth of the total lung field on chest X-ray, the presence of GGO in the upper lung field and the presence of consolidation in the upper lung field may be used as indexes.

Due to the urgency of the current global situation, trials that provide real-world evidence are more important than randomized controlled trials, which usually take months to complete and are too long in the midst of a pandemic. Quite key is real-world evidence to answer the question of drug effectiveness. Various medicines have been tried against COVID-19, and we plan to examine the efficacy, optimal dose, and optimal duration of administering drugs such as steroids that are currently in wide use by performing propensity score matching using the predictors in this study.

As limitations to this study, first, our hospital mainly treated patients with mild to moderate illness, and selection biases may thus have impact the enrollment of patients. This was also related to the slightly lower

frequency of comorbidities. Second, although we matched patient characteristics using propensity scores and by balancing variables, unmeasured confounders were not balanced between cases and controls in the present study. Third, the present study was a single-center experience, and the generalizability of the results obtained in the study was not sufficient. A multicenter study is required to confirm these results.

## **Conclusion**

Maximum body temperature up to diagnosis and radiological findings (number of shaded lung lobes, amount of GGO in the total lung field, amount of GGO and consolidation of the upper lung field, and presence of pleural effusion) could offer useful prognostic predictors in radiographic and blood test findings at the diagnosis of COVID-19 in daily clinical practice, rather than blood test findings.

## **Materials and Methods**

This study was approved by the ethics committee of Tokyo Shinagawa Hospital (approval number 20-A-06) and was performed in accordance with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The requirement to obtain informed consent for participation was waived due to the retrospective nature of this study. We retrospectively enrolled 66 patients with a final diagnosis of COVID-19 pneumonia who had undergone blood testing and chest CT from the early to middle phase (from onset to 10 days; if CT was performed multiple times during this period, only the CT performed earliest within the target period was analyzed) in our hospital between January 1 and May 31, 2020. COVID-19 was confirmed by reverse transcription-polymerase chain reaction (RT-PCR) or loop-mediated isothermal amplification. Based on the previous reports<sup>6,7)</sup>, we quantified ground-glass opacity (GGO) and consolidation, as the main components

of COVID-19 pneumonia, and analyzed shadow spread (uni- or bilateral, how many lobes showing shadows), mediastinal lymphadenopathy, traction bronchiectasis, stripe shadows/subpleural curvilinear lines, pleural thickening, vascular enlargement, pleural effusion, and pericardial effusion.

In terms of laboratory findings, we analyzed white blood cell counts, lymphocyte counts, platelet counts, lactate dehydrogenase (LDH) isozymes, uric acid, C-reactive protein (CRP), fibrinogen, and D-dimer measured at diagnosis.

For physical examination, fever that could be investigated from past medical records was analyzed.

### **Protocols for chest CT**

All images were obtained on one of two CT systems (Revolution Maxima; GE Healthcare, Chicago, IL, USA or Aquilion ONE; Canon, Tokyo, Japan) with the patient supine. The main scanning parameters were as follows: tube voltage, 120 kVp; automatic tube current modulation, 40–140 mAs; pitch, 0.828–1.375 mm; matrix,  $512 \times 512$ ; slice thickness, 2.5–3.0 mm; and field of view,  $400 \times 400$  mm. All images were then reconstructed with a slice thickness of 0.5–1.25 mm and 1-mm increments.

### **Image interpretation**

Images were analyzed by three senior fellows of the Japanese Respiratory Society. Evaluators independently and freely assessed CT features using axial CT, multiplanar reconstruction images, and high-resolution CT.

After separate evaluations, any disagreements were resolved by discussion as a consensus decision of all three senior fellows of the Japanese Respiratory Society.

### **Calculation of volumes of GGOs and consolidation**

We used a tool newly developed by Fujitsu Limited (Tokyo, Japan) to annotate shadows manually and automatically calculated the volume based on annotation results (Fig. 1). Specifically, we selected 30 images including lung fields from one series of CT images for each case at as close to equal intervals as possible to calculate volumes of GGO and consolidation. All CT images were monochrome, 512 pixels × 512 pixels on the axial view and including both lungs. The following processing was performed on each selected image:

1. Division of image into 1024 (32 × 32) patches, each as a 16 pixel × 16 pixel square.
2. Confirmation of CT images by senior fellows of the Japanese Respiratory Society and identification of each patch as "Consolidation due to COVID-19 pneumonia", "GGO due to COVID-19 pneumonia", "Other lung", or "Extrapulmonary field". CT images are rendered into 3 dimensions by linear interpolation between the selected 30 images, and patches were also rendered into 3 dimensions. Volumes of patches classified as "Consolidation due to COVID-19 pneumonia", "GGO due to COVID-19 pneumonia" or "Other lung" were obtained.

### **Statistical analysis**

We selected a propensity score-matched case-control study design to evaluate prognostic predictors in radiographic and blood test findings at diagnosis in COVID-19. We defined a severe respiratory failure group (severe respiratory failure group: non-rebreather mask, nasal high-flow, noninvasive positive-pressure ventilation (NIPPV) or invasive positive-pressure ventilation) as cases and a non-severe respiratory failure group (non-severe respiratory failure group: nasal cannula or mask or no oxygen administration) as controls. A multivariable logistic regression model was used to construct case propensity scores, and confounding factors of patient characteristics that impacted on patient prognosis and were examined based on findings from the

previous studies<sup>9-15</sup>) were included for the construction of propensity scores: 1) sex; 2) age; 3) body mass index (BMI); 4) presence of hypertension; 5) presence of diabetes; 6) presence of dyslipidemia, and 7) presence of chronic renal failure. Age and BMI were used as the binary data (age over 60 or not; BMI over 25 or not) for the construction of propensity scores.

Propensity score-matched groups were created at a ratio of 1:3 based on the nearest neighbor matching algorithm with a 0.2-caliper distance of the standard deviation of the logit of the propensity score with replacements. Discrimination (i.e., the capability to classify individuals with and without events) was evaluated by the C-statistic or the area under the receiving operating characteristic curve. We compared the severe and non-severe respiratory failure groups for maximum body temperature up to diagnosis, laboratory findings, CT findings (uni- or bilateral, number of lobes with shadows, the volume ratio of GGO to whole lung field, the volume ratio of consolidation to whole lung field, volume ratios of GGO to each of the upper, middle and lower lung fields, volume ratios of consolidation to each of the upper, middle and lower lung fields, mediastinal lymphadenopathy (including those with short-axis diameter <1 cm), traction bronchiectasis, fibrous stripes/subpleural curvilinear lines, pleural thickening, vascular enlargement, pleural effusion, and pericardial effusion.

Continuous and categorical variables of patient characteristics were compared using Student's t-test and the  $\chi^2$  test within each set. Ratios of GGO consolidation to total/upper/middle/lower lung field were compared using the Mann-Whitney U test. Multivariate linear regression was also used to evaluate whether significantly different factors were associated with severe/non-severe respiratory failure due to insufficient absolute

standardized differences in the patient characteristics used for constructing propensity scores between propensity score-matched cases and controls and to confirm double robustness by the propensity-score matched study and multivariate linear regression analyses.

Moreover, we used likelihoods to plot the receiver operating characteristic (ROC) curve and identified optimal cut-offs for predicted probability by maximizing the Youden index. Two-tailed P-values  $<0.05$  were considered significant. Data were analyzed using STATA<sup>®</sup> version 16 (StataCorp, College Station, TX, USA). All files of extracted data were encoded to anonymize patient data and prevent personal identification.

## References

- 1) Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395:497-506.
- 2) WHO Coronavirus Disease (COVID-19) Dashboard. <https://covid19.who.int> (accessed 31 may 2021)
- 3) Ministry of Health, Labour and Welfare open data. <https://www.mhlw.go.jp/stf/covid-19/open-data.html> (accessed 31 May 2021).
- 4) Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. *JAMA*. 2020;324:782-93.
- 5) Zhou M, Zhang X, Qu J. Coronavirus disease 2019 (COVID-19): a clinical update. *Front Med*. 2020;14:126-35.
- 6) Ye Z, Zhang Y, Wang Y, Huang Z, Song B. Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *Eur Radiol*. 2020;30:4381-9.
- 7) Li M, Lei P, Zeng B, Li Z, Yu P, Fan B, et al. Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease. *Acad Radiol*. 2020;27:603-8.
- 8) Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology*. 2020;295:715-21.
- 9) Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. 2020;180:934-43.

- 10) Pranata R, Lim M, Yonas E, Vania R, Lukito AA, Siswanto BB, et al. Body mass index and outcome in patients with COVID-19: A dose-response meta-analysis. *Diabetes Metab.* 2021;47:101178.
- 11) Zhou Y, Chi J, Lv W, Wang Y. Obesity and diabetes as high-risk factors for severe coronavirus disease 2019 (Covid-19). *Diabetes Metab Res Rev.* 2021;37:e3377.
- 12) 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):e16-25.
- 13) Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis* 2020;94:91-95.
- 14) Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int.* 2020;97:829-38.
- 15) Hariyanto TI, Kurniawan A. Dyslipidemia is associated with severe coronavirus disease 2019 (COVID-19) infection. *Diabetes Metab Syndr.* 2020;14:1463-5.
- 16) 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: 1081-9.
- 17) Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ.* 2020; 370: m3339.

- 18) Fan G, Tu C, Zhou F, Liu Z, Wang Y, Song B, et al. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J.* 2020; 56: 2002113.
- 19) Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol.* 2020;21:335-7.
- 20) Lippi G, Henry BM. Chronic obstructive pulmonary disease is associated with severe coronavirus disease 2019 (COVID-19). *Respir Med.* 2020;167:105941.
- 21) Dai WC, Zhang HW, Yu J, Xu HJ, Chen H, Luo SP, et al. CT imaging and differential diagnosis of COVID-19. *Can Assoc Radiol J.* 2020;71:195-200.
- 22) Chen H, Ai L, Lu H, Li H. Clinical and imaging features of COVID-19. *Radiol Infect Dis.* 2020;7:43-50.
- 23) Yang W, Sirajuddin A, Zhang X, Liu G, Teng Z, Zhao S et al. The role of imaging in 2019 novel coronavirus pneumonia (COVID-19). *Eur Radiol.* 2020;30:4874-4882.
- 24) Zheng Y, Wang L, Ben S. Meta-analysis of chest CT features of patients with COVID-19 pneumonia. *J Med Virol.* 2021;93:241-9.
- 25) Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, et al. Emerging 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology.* 2020;295:210-7.
- 26) Iftimie S, López-Azcona AF, Vicente-Miralles M, Descarrega-Reina R, Hernández-Aguilera A, Riu F. Risk factors associated with mortality in hospitalized patients with SARS-CoV-2 infection. A prospective, longitudinal, unicenter study in Reus, Spain. *PLoS One.* 2020;15:e0234452.

- 27) Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020; 395: 1054–62.
- 28) Ayanian S, Reyes J, Lynn L, , Teufel K. The association between biomarkers and clinical outcomes in novel coronavirus pneumonia in a US cohort. *Biomark Med*. 2020;14(12):1091-7.
- 29) d'Alessandro M, Cameli P, Refini RM, Bergantini L, Alonzi V, Lanzarone N, et al. Serum KL-6 concentrations as a novel biomarker of severe COVID-19. *J Med Virol*. 2020;92:2216-20.
- 30) Sugiyama M, Kinoshita N, Ide S, Nomoto H, Nakamoto T, Saito S, et al. Serum CCL17 level becomes a predictive marker to distinguish between mild/moderate and severe/critical disease in patients with COVID-19. *Gene*. 2021; 766: 145145.

**Acknowledgments**

We are grateful to all patients who participated in this study and their families.

**Author Contributions:**

M. Shinoda and M. Shinkai conceptualized the article; S.O., M. Shinoda, N.I., T.S., and M. Shinkai analyzed the data, with input from Y.Y., T.H., K.S., and M.M.; M. Shinoda wrote the initial draft, with all authors providing critical feedback and edits to subsequent revisions. All authors approved the final draft of the manuscript. The corresponding author attests that all listed authors meet the authorship criteria and that no others meeting the criteria have been omitted.

**Funding sources:** This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

**Disclosure:** The authors have stated that they have no conflicts of interest.

## **Figure legends**

### **Figure 1**

The annotation tool developed by Fujitsu Limited to annotate shadows manually for automatic calculation of volume.

### **Figure 2**

ROC curve for GGO/total lung field and sensitivity/specificity as a function of probability cut-off. The optimal cutoff on logistic regression probabilities was the intersection point reflecting the optimal balance between sensitivity and specificity (GGOs/total lung field 14.6%).

**Table 1** Baseline characteristics in unmatched and matched patients

	Unmatched patients			Matched patients			
	Non-severe respiratory failure group	Severe respiratory failure group	p-value	Non-severe respiratory failure group	Severe respiratory failure group	ASD	p-value
	n=50	n=16		n=27	n=9		
<b>Age, years</b>							
mean (SD)	47.9 (16.1)	62.7 (16.0)	0.0021	51.1 (18.0)	59.7 (18.0)		0.2263
median [IQR]	48.5 [37–59]	61.5 [49–75.5]	0.0057	50 [37–62]	50 [49–70]		0.3218
<b>Over 60, n (%) of patients</b>	10 (20.0)	8 (50.0)	0.019	12 (44.4)	4 (44.4)	0	1.000
<b>Sex, n (%) of patients</b>							
male	27 (54.0)	11 (68.7)	0.2988	21 (77.8)	8 (88.9)	0.291	0.4658
female	23 (46.0)	5 (31.3)		6 (22.2)	1 (11.1)		
<b>Body mass index, kg/m<sup>2</sup></b>							
mean (SD)	23.9 (3.78)	27.3 (5.00)	0.0063	27.1 (2.35)	29.2 (5.31)		0.1070
median [IQR]	23 [21.4–25.9]	28.1 [23.8–29.6]	0.0135	26.7 [25.8–27.6]	28.7 [28.4–30.0]		0.1112
<b>Over 25, n (%) of patients</b>	14 (28.0)	10 (62.5)	0.013	21 (77.8)	7 (77.8)	0	1.000
<b>Comorbidity, n (%) of patients</b>							
hypertension	11 (22.0)	3 (18.8)	0.7819	6 (22.2)	2 (22.2)	0	1.0000
dyslipidemia	1 (2.0)	1 (6.3)	0.3880	0	0	-	
diabetes mellitus	4 (8.0)	6 (37.5)	0.0042	6 (22.2)	1 (11.1)	0.291	0.4658
chronic renal failure	0	1 (6.3)	0.0749	0	0	-	

**Note.** A two-tailed p-value < 0.05 was considered to indicate statistical significance.

**Abbreviations:** ASD, absolute standardized difference; IQR, interquartile range; SD, standard deviation

**Table 2** Outcomes of comparison between non-severe and severe respiratory failure groups in unmatched and matched patients

	Unmatched patients			Matched patients		
	Non-severe respiratory failure group n=50	Severe respiratory failure group n=16	p-value	Non-severe respiratory failure group n=27	Severe respiratory failure group n=9	p-value
<b>Maximum body temperature until diagnosis, °C</b>						
mean (SD)	38.5 (0.76)	38.9 (0.91)		38.4 (0.39)	39.3 (0.87)	
median [IQR]	38.5 [38–39]	39 [38.2–39.7]	0.0819	38.5 [38.2–38.5]	39.5 [39–39.9]	0.0043
<b>CRP, mg/dL</b>						
mean (SD)	2.12 (2.00)	8.19 (7.08)		3.28 (2.39)	8.69 (9.03)	
median [IQR]	1.57 [0.44–3.34]	6.29 [2.96–10.0]	<0.0001	2.16 [1.74–4.29]	5.84 [3.14–7.67]	0.0886
<b>LDH, U/L</b>						
mean (SD)	218 (60.4)	348 (171)		234 (46.6)	335 (158)	
median [IQR]	201 [179–254]	275 [218–460]	0.0009	224 [203–266]	269 [211–420]	0.0615
<b>White blood cell count, /<math>\mu</math>L</b>						
mean (SD)	4652 (1558)	5356 (2230)		5367 (1698)	5933 (2090)	
median [IQR]	4200 [3800–5200]	5100 [4000–6500]	0.1593	4400 [4000–6400]	5500 [4400–7000]	0.4395
<b>Lymphocyte count, /<math>\mu</math>L</b>						
mean (SD)	1172 (551.5)	940.8 (484.4)		1182 (458.7)	974.7 (484.7)	
median [IQR]	1067 [823.2–1310]	952.5 [594.4–1291]	0.2143	1170 [1100–1216]	959 [616–1270]	0.4752
<b>Platelet count, <math>\times 10^4/\mu</math>L</b>						

mean (SD)	18.8 (5.18)	19.8 (7.26)		19.2 (4.47)	18.6 (5.78)	
median [IQR]	18.4 [14.9–20.9]	17.0 [14.8–24.3]	0.9582	19.7 [17.8–20.8]	16.7 [14.6–23.1]	0.6205
<b>Uric acid, mg/dL</b>						
mean (SD)	4.35 (1.55)	5.69 (2.03)		4.40 (1.64)	5.73 (1.29)	
median [IQR]	4.30 [3.1–5.2]	5.70 [3.8–6.6]	0.0297	4.07 [3.80–5.60]	6.30 [4.65–6.50]	0.0585
<b>Fibrinogen, mg/dL</b>						
mean (SD)	429 (133)	516 (132)		458 (163)	556 (161)	
median [IQR]	409 [327–488]	493 [443–570]	0.0217	443 [289–641]	570 [443–742]	0.1501
<b>D-dimer, µg/mL</b>						
mean (SD)	0.49 (0.59)	3.76 (12.1)		0.41 (0.41)	0.66 (0.71)	
median [IQR]	0.1 [0.1–0.8]	0.4 [0.1–1.6]	0.1386	0.10 [0.10–0.90]	0.4 [0.10–1.15]	0.4297
<b>CT findings</b>						
unilateral	10 (20.0)	1 (6.2)	0.1990	6 (22.2)	0	0.1213
bilateral	40 (80.0)	15 (93.8)		21 (77.8)	9 (100)	
Lobe, n						
	3.68 (1.89)	5.19 (1.38)		4.11 (1.83)	5.33 (1.41)	
	3.5 [2–6]	6 [4.5–6]	0.0038	5 [3–6]	6 [6–6]	0.0434
Ground glass opacity						
to total lung field ratio, %						
mean (SD)	3.27 (3.39)	21.3 (23.7)	<0.0001	3.13 (2.34)	20.3 (21.1)	

median [IQR]	1.98 [0.81–4.54]	11.1 [3.38–30.3]	0.0002	2.89 [2.08–3.37]	20.7 [3.51–26.5]	0.0071
to upper lung field ratio, %						
mean (SD)	8.38 (11.0)	31.0 (18.1)		9.30 (11.3)	33.2 (12.0)	
median [IQR]	3.09 [0–16.23]	27.2 [16.5–41.9]	<0.0001	3.12 [2.55–8.94]	31.54 [27.1–36.1]	0.0001
to middle lung field ratio, %						
mean (SD)	36.8 (25.4)	33.0 (13.0)		44.3 (30.0)	40.5 (8.55)	
median [IQR]	35.7 [17.9–51.1]	35.9 [24.5–39.5]	0.7307	47.0 [17.3–55.1]	39.2 [36.3–43.5]	0.4100
to lower lung field ratio, %						
mean (SD)	54.8 (28.2)	36.0 (22.2)		46.4 (32.6)	26.3 (16.7)	
median [IQR]	57.3 [32.8–74.0]	36.9 [13.9–48.6]	0.0181	42.4 [21.7–73.8]	22.1 [12.4–37.2]	0.1112
Consolidation/total lung field, %						
to total lung field ratio, %						
mean (SD)	0.65 (1.21)	1.38 (2.86)	0.1464	0.28 (0.37)	1.74 (3.63)	
median [IQR]	0.21 [0.05–0.89]	0.21 [0.07–1.46]	0.4906	0.15 [0.11–0.26]	0.19 [0.08–1.59]	0.2484
to upper lung field ratio, %						
mean (SD)	7.56 (18.4)	29.8 (33.7)		9.24 (15.3)	36.6 (32.8)	
median [IQR]	0 [0–3.17]	11.7 [7.18–52.6]	0.0001	0 [0–16.8]	27.2 [11.5–53.6]	0.0036
to middle lung field ratio, %						
mean (SD)	32.1 (31.4)	27.3 (22.8)		39.6 (31.1)	31.7 (24.8)	
median [IQR]	23.6 [0–56.0]	29.2 [4.9–36.7]	0.6286	38.2 [19.6–59.7]	29.8 [25.8–36.5]	0.5437
to lower lung field ratio, %						

mean (SD)	46.3 (35.7)	30.5 (30.5)		40.0 (30.7)	20.6 (26.5)	
median [IQR]	44.2 [11.4–81.0]	28.0 [0–59.8]	0.1312	40.3 [22.6–59.6]	5.28 [0–37.4]	0.1341
Subpleural curvilinear lines/fibrosis stripes	17 (34.0)	4 (25.0)	0.5011	9(33.3)	2 (22.2)	0.5309
Traction bronchiectasis	0	1 (6.25)	0.0749	0	0	
Vascular enlargement	1 (2.00)	1 (6.25)	0.3880	3 (11.1)	1 (11.1)	1.0000
Lymphadenopathy	22 (44.0)	12 (75.0)	0.0308	9 (33.3)	6 (66.7)	0.0790
Pleural thickening	10 (20.0)	1 (6.25)	0.1990	3 (11.1)	1 (11.1)	1.0000
Pleural effusion	0	2 (12.5)	0.0111	0	2 (22.2)	0.0117
Pericardial effusion	0	0		0	0	

**Note.** A two-tailed p-value < 0.05 was considered to indicate statistical significance.

**Abbreviations:** LDH, lactate dehydrogenase isozyme; CRP, C-reactive protein; IQR, interquartile range; SD, standard deviation

**Table 3.** Association with maximum body temperature until diagnosis, CT findings (lobe, ground glass opacity-to-total lung field ratio, ground glass opacity-to-upper lung field ratio, and consolidation-to-upper lung field ratio) on multivariate linear regressions in the matched cohort.

	<b>Matched patients (n=36)</b>	
	<b>Maximum body temperature until diagnosis</b>	
	$\beta$ coefficient, 95%CI	p-value
<b>Non-severe respiratory failure</b>	0	
<b>Severe respiratory failure</b>	0.93, 0.50–1.36	<0.0001
	<b>CT finding</b>	
	<b>Lobe</b>	
	$\beta$ coefficient, 95%CI	p-value
<b>Non-severe respiratory failure</b>	0	
<b>Severe respiratory failure</b>	1.38, 0.42–2.34	0.006
	<b>Ground glass opacity to total lung field ratio</b>	
	$\beta$ coefficient, 95%CI	p-value
<b>Non-severe respiratory failure</b>	0	
<b>Severe respiratory failure</b>	16.7, 9.12–24.4	<0.0001
	<b>Ground glass opacity to upper lung field ratio</b>	
	$\beta$ coefficient, 95%CI	p-value
<b>Non-severe respiratory failure</b>	0	
<b>Severe respiratory failure</b>	24.9, 17.7–32.2	<0.0001
	<b>Consolidation to upper lung field ratio</b>	
	$\beta$ coefficient, 95%CI	p-value
<b>Non-severe respiratory failure</b>	0	
<b>Severe respiratory failure</b>	30.0, 16.0–43.9	<0.0001

**Note.** Adjusted for patient characteristics used for the construction of propensity scores (age, sex, BMI, diabetes, hypertension, dyslipidemia, and chronic renal failure). A two-tailed P-value < 0.05 was considered to indicate statistical significance.

**Abbreviations:** CI, confidence interval.

Figure1

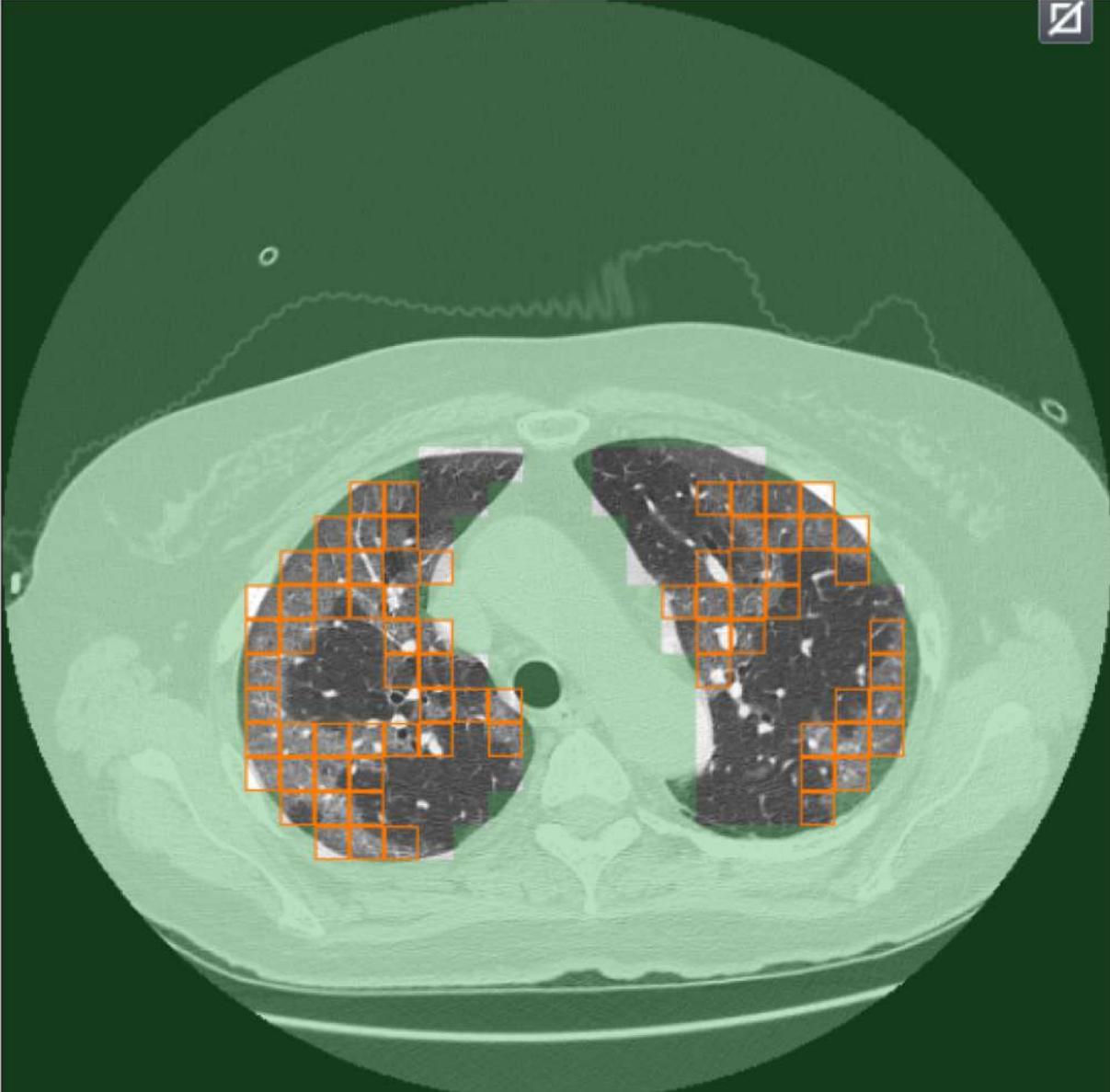


Figure2

