

# Association of Initial Symptoms or Comorbidities With Pneumonia Lesions in COVID-19 Patients: Based on Artificial Intelligence-Enabled CT Quantitation

**Fangzhengyuan Yuan**

Army Medical University

**Chuan Liu**

Army Medical University

**Jie Yang**

Army Medical University

**Hu Tan**

Army Medical University

**Shizhu Bian**

Army Medical University

**Xubin Gao**

Army Medical University

**Jihang Zhang**

Army Medical University

**Mingdong Hu**

Army Medical University

**Renzheng Chen**

Army Medical University

**Yang Shen**

Army Medical University

**Jingbin Ke**

Army Medical University

**Yuanqi Yang**

Army Medical University

**Chunyan He**

Army Medical University

**Ran Cheng**

Army Medical University

**Lan Huang** (✉ [huanglan260@126.com](mailto:huanglan260@126.com))

Army Medical University <https://orcid.org/0000-0001-6200-2309>

---

## Research

**Keywords:** Artificial intelligence, Coronavirus disease 2019, Comorbidities, Initial symptoms, Quantitative computed tomography, Lung lesion.

**DOI:** <https://doi.org/10.21203/rs.3.rs-78075/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** Coronavirus disease 2019 (COVID-19) patients with a larger ratio of pneumonia lesions are more likely to progress to acute respiratory distress syndrome and death. This study aimed to investigate the relationship of baseline parameters with pneumonia lesions on admission, as quantified by an artificial intelligence (AI) algorithm using computed tomography (CT) images.

**Methods:** This retrospective study quantitatively assessed lung lesions on CT using an AI algorithm in 1630 consecutive patients confirmed with COVID-19 on admission and classified the patients into none (0%), mild (>0–25%), intermediate (>25–50%), and severe (>50%) groups, according to the lesion ratio of the whole lung. A multivariate linear regression model was established to explore the relationship between the lesion ratio and laboratory parameters. The baseline parameters associated with lung lesions, including demographics, initial symptoms, and comorbidities, were determined using a multivariate ordinal regression model.

**Results:** The 1630 patients confirmed with COVID-19 had a median whole lung lesion ratio of 4.1%, and the right lower lung lobe had the most lesions among the five lung lobes based on the evaluation of CT using AI algorithm. The whole lung lesion ratio was associated with the levels of plasma fibrinogen ( $r=0.280$ ,  $p<0.001$ ), plasma D-dimer ( $r=0.248$ ,  $p<0.001$ ), serum  $\alpha$ -hydroxybutyrate dehydrogenase ( $r=0.363$ ,  $p<0.001$ ), serum albumin ( $r=-0.300$ ,  $p<0.001$ ), and peripheral blood leukocyte count ( $r=0.194$ ,  $p<0.001$ ). Among the four patients groups categorised by whole lung lesion ratio, the highest frequency of cough ( $p<0.001$ ) and shortness of breath ( $p<0.001$ ) were found in the severe group, and the highest frequency of hypertension ( $p<0.001$ ), diabetes ( $p<0.001$ ) and anemia ( $p=0.039$ ) were observed in the intermediate group. Based on baseline ordinal regression analysis, cough ( $p=0.009$ ), shortness of breath ( $p<0.001$ ), hypertension ( $p=0.002$ ), diabetes ( $p=0.005$ ), and anemia ( $p=0.006$ ) were independent risk factors for more severe lung lesions.

**Conclusions:** Based on AI-enabled CT quantitation, patients with initial symptoms of cough/shortness of breath, or with comorbidities of hypertension, diabetes, or anemia, had a higher risk for more severe lung lesions on admission in COVID-19 patients.

## Introduction

Since the outbreak of Coronavirus disease 2019 (COVID-19) in December 2019, more than 27 million confirmed cases and 894,241 deaths were reported until 9 September 2020. The trend of the world pandemic shows no sign of reversing and heavily burdens the health system all around the world. In contrast to mild symptoms caused by common cold coronaviruses, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection leads to lethal pneumonia, and acute respiratory distress syndrome (ARDS) is the primary cause of death in COVID-19 (1). As the lung was the main organ attacked by SARS-CoV-2, the COVID-19 patients with a larger proportion of lung lesions on computed tomography (CT) in early days had a higher likelihood of developing ARDS, admitted to the intensive care unit (ICU), and

progress to death (2–5). However, the baseline determinants of early COVID-19 pneumonia lesions have not been illustrated yet.

SARS-CoV-2 invades target cells by binding to angiotensin-converting enzyme 2, which is highly expressed in the lung and widely distributed in the heart, kidney, gastrointestinal tract, and even active in all organs (6, 7). Accordingly, the COVID-19 symptom profiles vary, including respiratory and gastrointestinal symptoms, such as cough, expectoration, shortness of breath, diarrhea, nausea, and other symptoms, namely, headache, chest distress, and myalgia (8, 9). With the exception of asymptomatic cases, COVID-19 patients suffered from diverse symptoms in the early stage of COVID-19, although efforts have been devoted to the management of COVID-19 symptoms, none of the research has focussed on the association of initial symptoms with COVID-19 pneumonia lesions (10, 11). Previous studies reported that in COVID-19 patients, comorbidities were associated with a higher rate of ICU admission, invasive ventilation, and death, however, the association of comorbidities with early COVID-19 pneumonia is not clear and worthy of investigation (12).

Emerging artificial intelligence (AI) technology has already been employed in the evaluation of CT images, with deep learning, the AI owned human-level performance in the classification of fibrotic lung disease (13). In the COVID-19 pandemic, AI assessment of lung CT images has been widely used in the diagnosis of COVID-19, evaluating drug efficacy and assessing the prognosis (2). AI performed with a similar accuracy in the diagnosis of COVID-19 compared to the experienced radiologist and provided an objective and instant way to quantitatively evaluate the pneumonia lesion on CT, however, the traditional way of evaluating pneumonia lesion on CT was based on semi-quantitative estimation by the radiologist (2, 14, 15). Thus, we take advantage of AI algorithm to quantify pneumonia lesions on CT to explore the potential baseline determinants associated with the severity of lung lesions in COVID-19 patients on admission.

## Methods

### Study participants

A total of 1630 adult patients ( $\geq 18$  years old) diagnosed with COVID-19 in Huo Shen Shan Hospital, Wuhan, China, from 4 February 2020 to 9 March 2020 and underwent chest CT scans with AI quantitative assessment of lung lesions on admission were included in the retrospective study. This study was approved by the Human Ethics Committee of Huo Shen Shan Hospital (No. HSSLL023), with a waiver of informed consent.

### Data collection

Standardised clinical electronic medical records were investigated by the research team, and all the extracted data related to patients' demographic, initial symptoms, comorbidities, laboratory tests, and chest CT images were crosschecked by two experienced doctors. Demographic data, such as age, sex, smoking status, and drinking status were collected. Initial symptoms and comorbidities were obtained

according to the patients' self-reports. The values of laboratory parameters and the proportion and volume of lung lesions based on chest CT images were recorded according to the examination results at admission.

### **Laboratory testing**

In order to confirm the SARS-CoV-2 infection, real-time reverse transcriptase–polymerase chain reaction was used to detect SARS-CoV-2 from throat swab specimens. Laboratory tests included complete blood count, coagulation profile, biochemical profile (including renal and liver function, lactate dehydrogenase), cardiac function test, procalcitonin, and high-sensitivity C-reactive protein (hsCRP). Laboratory test results at admission were selected.

### **CT acquisition and AI-based quantitation of lung lesion**

The COVID-19 confirmed patients holding their breath at the end of inspiration in the supine position underwent unenhanced CT scans using multi-detector CT (uCT 760, United Imaging, Shanghai, China) at admission. The parameters adopted in the CT scans were as follows: tube voltage, 120 kV; tube current-time product, 145 mAs; collimation, 0.625 mm; reconstructed slice thickness, 10 mm; matrix size, 512×512.

Based on CT scans, the lung lesions in COVID-19 patients were quantified using AI technology (DAMO Academy Medical AI, Alibaba, China). The reported parameters of the lung lesions included the lesion volume (mL) and the lesion ratio (%), calculated as the ratio of the lesion volume to the total volume of the corresponding lung or lobes. The severity of lung lesions was determined by the level of whole lung lesion ratio, and the patients were categorised into none (0%), mild (>0–25%), intermediate (>25–50%), and severe (>50%) groups (2).

### **Statistical analysis**

According to the severity of lung lesions, the patients were categorised into none, mild, intermediate, and severe groups. Categorical variables were presented as the number of patients and percentages and were compared using  $\chi^2$  tests or Fisher's exact tests, as appropriate. According to the results of the Kolmogorov-Smirnov tests, the continuous variables with skewed distribution are presented as median with interquartile range (IQR) and the differences between the four groups were compared using the Kruskal-Wallis H test. The association between the lesion ratio of the whole lung and laboratory findings was checked by univariate linear regression analysis, and the variables associated with the lesion ratio of the whole lung ( $p < 0.05$ ) were included in a stepwise multivariate linear regression model. The baseline determinants (including demographics, initial symptoms, and comorbidities) of lung lesion severity on admission were investigated using univariate ordinal logistic regression analysis. A test of parallel lines was performed to determine the feasibility of ordinary regression. Variables with  $p$ -values  $< 0.05$  and  $p$ -values of parallel line test  $> 0.05$  were included in a multivariate ordinal logistic model. A two-sided  $\alpha$ -

value less than 0.05 was considered statistically significant. Statistical analyses were performed using the SPSS software (Version 22.0, IBM Corp., Armonk, NY, USA).

## Results

### Quantitative assessment of lung lesions

The AI algorithm was employed to quantitatively evaluate the lung lesion on CT, and the distribution of lung lesions in COVID-19 patients is shown in Table 1. The median lesion ratio and median lesion volume of the whole lung were 4.1% (IQR 0.5–12.9) and 153.9 ml (IQR, 20.8–426.2), respectively. A larger proportion and volume of lung lesions were observed in the right lung, rather than in the left lung. Of the total five lung lobes, the highest proportion and volume of lung lesions were found in the right lower lobe. The median ratio and the median volume of lung lesions in the right lower lobe were 5.9% (IQR 0.3–22.8%) and 48.5 ml (IQR 2.7–149.6), respectively.

Table 1  
Distribution of the lung lesion on CT quantitatively assessed by AI algorithm

Lesion location	Lesion ratio (%), median (IQR) n = 1630	Lesion volume (ml), median (IQR) n = 1630
Whole lung	4.1 (0.5, 12.9)	153.9 (20.8, 426.2)
Left lung	2.4 (0.2, 9.8)	41.6 (3.8, 156.8)
Right lung	4.3 (0.4, 14.6)	85.9 (8.8, 264.0)
Left upper lobe	1.1 (0.1, 6.9)	10.7 (0.6, 61.9)
Left lower lobe	2.7 (0.2, 14.4)	19.8 (1.2, 87.6)
Right upper lobe	1.8 (0, 11.7)	13.6 (0.4, 85.3)
Right middle lobe	0.6 (0, 6.4)	2.2 (0, 26.3)
Right lower lobe	5.9 (0.3, 22.8)	48.5 (2.7, 149.6)
IQR, interquartile range		

### Characteristics of COVID-19 patients with different level of lung lesion ratio

A total of 1630 hospitalised COVID-19 patients were included in this study, and with respect to the whole lung lesion ratio, 128 (7.9%), 1336 (82.0%), 139 (8.5%), and 27 (1.7%) patients were categorised into none, mild, intermediate, and severe groups, respectively (2). The demographics, initial symptoms, and comorbidities of the patients are presented in Table 2. Of all 1630 cases, the median age was 62 years (IQR 52–69), and 839 patients (51.5%) were male. The proportions of smokers and drinkers were 5.6% and 3.2%, respectively. No differences were found in sex and smoking status among the four groups. The highest median age was observed in the intermediate group. With the gradually worsening of lung lesions

in the four groups, the frequency of fever, cough, shortness of breath, fatigue, and chest distress increased accordingly, and the highest proportion of myalgia was found in the intermediate group. Discrepancies were not observed in the other initial symptoms. The highest proportions of hypertension, diabetes, anemia, and stroke were found in the intermediate group. Other comorbidities among the four groups showed no significant differences.

Table 2  
Baseline characteristics of the patients

Baseline parameters	All patients (n = 1630)	None (n = 128)	Mild (n = 1336)	Intermediate (n = 139)	Severe (n = 27)	p
<b>Demographics</b>						
Age (y); median (IQR)	62 (52, 69)	48 (37, 58)	62 (54, 69)	66 (57, 72)	62 (50, 71)	< 0.001
Male, no (%)	839 (51.5)	55 (43.0)	696 (52.1)	73 (52.5)	15 (55.6)	0.245
Smoking, no (%)	92 (5.6)	4 (3.1)	81 (6.1)	5 (3.6)	2 (7.4)	0.359
Drinking, no (%)	52 (3.2)	2 (1.6)	48 (3.6)	0 (0)	2 (7.4)	0.047
<b>Initial symptoms</b>						
Fever, no (%)	1166 (71.5)	74 (57.8)	964 (72.2)	104 (74.8)	24 (88.9)	0.001
Cough, no (%)	1108 (68.0)	65 (50.8)	917 (68.6)	103 (74.1)	23 (85.2)	< 0.001
Expectoration, no (%)	232 (14.2)	9 (7.0)	201 (15.0)	17 (12.2)	5 (18.5)	0.071
Shortness of breath, no (%)	662 (40.6)	31 (24.2)	536 (40.1)	78 (56.1)	17 (63.0)	< 0.001
Myalgia, no (%)	437 (26.8)	19 (14.8)	363 (27.2)	47 (33.8)	8 (29.6)	0.005
Fatigue, no (%)	578 (35.5)	25 (19.5)	482 (36.1)	57 (41.0)	14 (51.9)	< 0.001
Diarrhea, no (%)	72 (4.4)	5 (3.9)	61 (4.6)	6 (4.3)	0 (0)	0.705
Nausea, no (%)	51 (3.1)	2 (1.6)	44 (3.3)	5 (3.6)	0 (0)	0.546
Headache, no (%)	46 (2.8)	5 (3.9)	36 (2.7)	4 (2.9)	1 (3.7)	0.872
Loss of appetite, no (%)	63 (3.9)	2 (1.6)	54 (4.0)	5 (3.6)	2 (7.4)	0.411
Chest distress, no (%)	135 (8.3)	8 (6.3)	106 (7.9)	15 (10.8)	6 (22.2)	0.030
<b>Comorbidities</b>						
Hypertension, no (%)	537 (32.9)	23 (18.0)	444 (33.2)	59 (42.5)	11 (40.7)	< 0.001

IQR, interquartile range; COPD, Chronic obstructive pulmonary disease



<b>Baseline parameters</b>	<b>All patients (n = 1630)</b>	<b>None (n = 128)</b>	<b>Mild (n = 1336)</b>	<b>Intermediate (n = 139)</b>	<b>Severe (n = 27)</b>	<b>p</b>
Arrhythmia, no (%)	45 (2.8)	0 (0)	38 (2.8)	6 (4.3)	1 (3.7)	0.171
Diabetes, no (%)	266 (16.3)	6 (4.7)	222 (16.6)	32 (23.0)	6 (22.2)	< 0.001
Coronary artery disease, no (%)	103 (6.3)	2 (1.6)	92 (6.9)	8 (5.8)	1 (3.7)	0.111
Hyperlipemia, no (%)	27 (1.7)	3 (2.3)	23 (1.7)	1 (0.7)	0 (0)	0.657
COPD, no (%)	67 (4.1)	2 (1.6)	58 (4.3)	6 (4.3)	1 (3.7)	0.510
Cancer, no (%)	35 (2.2)	2 (1.6)	30 (2.3)	2 (1.4)	1 (3.7)	0.822
Chronic liver disease, no (%)	46 (2.8)	5 (3.9)	37 (2.8)	3 (2.2)	1 (3.7)	0.835
Chronic kidney disease, no (%)	13 (0.8)	0 (0)	12 (0.9)	1 (0.7)	0 (0)	0.699
Anemia, no (%)	26 (1.6)	1 (0.8)	18 (1.4)	6 (4.3)	1 (3.7)	0.039
Stroke, no (%)	53 (3.3)	2 (1.6)	41 (3.1)	10 (7.2)	0 (0)	0.028
Gout, no (%)	22 (1.4)	0 (0)	18 (1.4)	4 (2.9)	0 (0)	0.207
Gastritis , no (%)	34 (2.1)	3 (2.3)	26 (2.0)	4 (2.9)	1 (3.7)	0.815
IQR, interquartile range; COPD, Chronic obstructive pulmonary disease						

Laboratory findings on admission in COVID-19 patients with different levels of lung lesion ratios were studied (Table 3). With the augment of lung lesion ratio in the four groups, the levels of plasma fibrinogen, plasma D-dimer, serum lactate dehydrogenase, serum  $\alpha$ -hydroxybutyrate dehydrogenase, serum high-sensitive cardiac troponin I, serum myoglobin, serum B-type natriuretic peptide, serum procalcitonin, serum hsCRP, peripheral blood leukocyte count, peripheral blood neutrophil percentage and prothrombin time increased gradually. Accordingly, the levels of peripheral blood hemoglobin, serum total protein, serum albumin, activated partial thromboplastin time, and peripheral blood lymphocyte percentage gradually decreased. The highest levels of serum alanine aminotransferase and serum creatinine kinase-myocardial band were observed in the severe group and the highest levels of serum aspartate transaminase, serum  $\gamma$ -glutamyltransferase, serum urea nitrogen, serum cystatin C and thrombin time were found in the intermediate group.

Table 3  
Laboratory features of Covid-19 patients with different levels of whole lung lesion ratio

Laboratory parameters	All patients (n = 1630)	None (n = 128)	Mild (n = 1336)	Intermediate (n = 139)	Severe (n = 27)	p
<b>Coagulation profile</b>						
Prothrombin time (s)	12.8 (12.2, 13.5)	12.8 (12.3, 13.5)	12.8 (12.2, 13.4)	13.4 (12.5, 14.4)	13.7 (12.5, 15.4)	< 0.001
Activated partial thromboplastin time (s)	28.0 (26.1, 30.1)	28.5 (26.9, 30.0)	28.0 (26.1, 30.1)	27.3 (25.5, 29.9)	27.1 (23.8, 29.1)	0.016
Fibrinogen (g/L)	3.0 (2.6, 3.4)	2.7 (2.4, 3.0)	3.0 (2.6, 3.4)	3.4 (3.0, 3.8)	3.4 (2.8, 4.1)	< 0.001
Thrombin time (s)	15.2 (14.5, 16.0)	14.9 (14.2, 15.5)	15.2 (14.5, 16.0)	15.7 (14.7, 16.5)	15.1 (14.5, 16.6)	< 0.001
D-dimer (mg/L)	0.4 (0.2, 0.9)	0.2 (0.2, 0.3)	0.4 (0.2, 0.8)	1.1 (0.6, 2.7)	1.1 (0.7, 3.9)	< 0.001
<b>Complete blood count</b>						
Leucocytes count (10 <sup>9</sup> /L)	5.7 (4.7, 7.1)	5.7 (4.8, 6.9)	5.7 (4.7, 7.0)	6.6 (5.2, 8.0)	8.3 (6.0, 10.5)	< 0.001
Neutrophil percentage (%)	62.0 (55.6, 69.1)	59.3 (54.2, 64.9)	61.7 (55.5, 68.0)	68.9 (60.5, 77.5)	75.5 (57.2, 82.7)	< 0.001
Lymphocyte percentage (%)	26.9 (20.5, 32.7)	30.2 (25.8, 35.6)	27.1 (21.3, 32.8)	19.7 (13.4, 27.0)	14.1 (9.3, 29.0)	< 0.001
Monocyte percentage (%)	7.7 (6.5, 9.0)	7.1 (6.1, 8.3)	7.8 (6.5, 9.1)	7.7 (6.3, 9.2)	7.1 (5.5, 8.2)	0.002
Eosinophil percentage (%)	2.0 (1.2, 3.3)	2.1 (1.5, 3.3)	2.1 (1.2, 3.3)	1.8 (1.0, 3.2)	2.0 (1.6, 3.1)	0.454
Basophil percentage (%)	0.4 (0.2, 0.5)	0.4 (0.3, 0.5)	0.4 (0.2, 0.6)	0.3 (0.2, 0.5)	0.3 (0.2, 0.5)	0.043
Platelet count (10 <sup>9</sup> /L)	224.0 (183.0, 270.5)	213.5 (184.0, 251.8)	223.5 (182.8, 271.0)	234.3 (184.0, 283.6)	228.5 (189.0, 292.0)	0.308
Hemoglobin (g/L)	123.0 (113.0, 134.0)	128.5 (118.8, 139.0)	123.0 (112.0, 134.0)	120.0 (109.2, 129.0)	117.0 (109.5, 128.0)	< 0.001

Laboratory parameters	All patients (n = 1630)	None (n = 128)	Mild (n = 1336)	Intermediate (n = 139)	Severe (n = 27)	p
<b>Biochemical parameters</b>						
Alanine aminotransferase (IU/L)	23.0 (14.7, 38.3)	23.5 (12.2, 38.2)	22.6 (14.7, 37.4)	27.4 (18.0, 48.1)	30.6 (15.4, 41.4)	0.010
Aspartate transaminase (IU/L)	19.6 (15.7, 26.4)	19.7 (15.7, 25.0)	19.3 (15.6, 26.1)	21.9 (17.3, 31.4)	20.5 (15.7, 29.8)	0.006
Total protein (g/L)	64.8 (60.7, 68.6)	67.7 (63.9, 70.3)	64.9 (60.8, 68.4)	61.8 (58.2, 66.9)	59.6 (57.3, 64.0)	< 0.001
Albumin (g/L)	37.6 (34.8, 40.2)	40.5 (38.6, 42.4)	37.7 (35.1, 40.1)	34.6 (31.6, 37.5)	34.2 (31.4, 37.4)	< 0.001
Total bilirubin (µmol/L)	9.3 (7.2, 12.1)	10.6 (8.5, 12.7)	9.1 (7.1, 12.1)	9.2 (7.1, 12.1)	10.5 (7.2, 12.3)	0.009
γ-glutamyltransferase (IU/L)	30.2 (20.4, 50.6)	23.3 (16.8, 37.2)	30.1 (20.3, 49.3)	39.2 (25.4, 64.9)	35.7 (26.7, 68.1)	< 0.001
Urea nitrogen (mmol/L)	4.4 (3.7, 5.5)	4.2 (3.4, 5.1)	4.4 (3.6, 5.4)	4.8 (3.9, 6.2)	4.5 (3.9, 6.1)	0.007
Creatinine (µmol/L)	64.5 (55.4, 75.4)	64.1 (54.2, 75.2)	64.8 (55.7, 75.5)	62.5 (54.6, 78.2)	59.3 (50.5, 66.5)	0.120
Cystatin C (mg/L)	0.94 (0.84, 1.08)	0.85 (0.77, 0.94)	0.94 (0.84, 1.08)	1.0 (0.86, 1.19)	0.94 (0.88, 1.12)	< 0.001
Lactate dehydrogenase (IU/L)	176.7 (152.3, 211.3)	156.4 (134.6, 178.1)	174.9 (151.9, 205.2)	222.7 (184.4, 268.6)	246.4 (209.2, 333.2)	< 0.001
<b>Myocardial injury markers</b>						
α-hydroxybutyrate dehydrogenase (IU/L)	143.9 (123.9, 172.8)	125.0 (106.8, 141.4)	142.6 (123.9, 167.7)	179.2 (148.5, 225.5)	196.8 (166.4, 271.1)	< 0.001
Creatinine kinase-myocardial band (IU/L)	8.5 (6.9, 10.8)	8.1 (6.8, 10.6)	8.5 (6.9, 10.6)	8.4 (7.0, 11.5)	10.0 (7.6, 13.0)	0.102

Laboratory parameters	All patients (n = 1630)	None (n = 128)	Mild (n = 1336)	Intermediate (n = 139)	Severe (n = 27)	p
High-sensitive cardiac troponin I (ng/mL)	0.01 (0.01, 0.01)	0.01 (0.01, 0.01)	0.01 (0.01, 0.01)	0.01 (0.01, 0.013)	0.011 (0.010, 0.025)	0.008
Myoglobin (ng/mL)	5.2 (2.2, 9.0)	5.0 (2.8, 8.7)	5.1 (1.9, 8.7)	5.8 (4.0, 10.2)	8.4 (4.8, 13.2)	0.009
B-type natriuretic peptide (pg/mL)	0.01 (0.01, 29.48)	0.01 (0.01, 0.01)	0.01 (0.01, 27.29)	22.16 (0.01, 74.87)	36.29 (9.99, 59.49)	< 0.001
<b>Inflammation markers</b>						
Procalcitonin (ng/mL)	0.04 (0.03, 0.08)	0.03 (0.03, 0.05)	0.04 (0.03, 0.07)	0.07 (0.04, 0.11)	0.08 (0.06, 0.37)	< 0.001
High-sensitive C-reactive protein (mg/L)	2.4 (0.8, 8.5)	0.8 (0.3, 2.4)	2.2 (0.8, 7.3)	10.9 (3.2, 35.4)	12.2 (3.6, 48.3)	< 0.001

### Association between laboratory findings and lung lesion ratio

The linear regression analysis between the whole lung lesion ratio and laboratory findings showed that the levels of plasma fibrinogen ( $r = 0.280$ ,  $p < 0.001$ ), plasma D-dimer ( $r = 0.248$ ,  $p < 0.001$ ), peripheral blood neutrophil percentage ( $r = 0.243$ ,  $p < 0.001$ ), serum hsCRP ( $r = 0.212$ ,  $p < 0.001$ ), serum lactate dehydrogenase ( $r = 0.375$ ,  $p < 0.001$ ), serum  $\alpha$ -hydroxybutyrate dehydrogenase ( $r = 0.363$ ,  $p < 0.001$ ), and peripheral blood leukocyte count ( $r = 0.194$ ,  $p < 0.001$ ) were positively correlated with the lesion ratio of the whole lung. The levels of serum albumin ( $r = -0.300$ ,  $p < 0.001$ ) and peripheral blood lymphocyte rate ( $r = 0.272$ ,  $p < 0.001$ ) were negatively correlated with the lesion ratio of the whole lung (Fig. 1 and **Supplementary Material online Table S1**). Stepwise multivariate regression analysis revealed that the plasma fibrinogen level ( $\beta = 0.183$ ,  $p < 0.001$ ), plasma D-dimer ( $\beta = 0.167$ ,  $p < 0.001$ ), serum albumin ( $\beta = -0.195$ ,  $p < 0.001$ ), serum  $\alpha$ -hydroxybutyrate dehydrogenase ( $\beta = 0.222$ ,  $p < 0.001$ ), and peripheral blood leukocyte count ( $\beta = 0.073$ ,  $p = 0.026$ ) were independently correlated with the lesion ratio of the whole lung (**Supplementary Material online Table S1**).

### Baseline determinants of lung lesion severity

Ordinal logistic regression was employed to evaluate the correlation of demographics, initial symptoms, and comorbidities (baseline variables) with the severity of lung lesions (an ordinal outcome variable). According to the results of the parallel line test ( $p$  value of parallel line test  $< 0.05$ ), age, drinking, and expectoration were not applicable to the ordinal logistic regression. In the univariate and multivariate analysis (Table 4), male, smoking and drinking were not found associated with the severity of lung lesion; however, fever (odds ratio [OR], 1.428; 95% confidence interval [CI], 1.064–1.916;  $p = 0.018$ ), cough (OR, 1.470; 95% CI, 1.099–1.966;  $p = 0.009$ ), shortness of breath (OR, 1.765; 95% CI, 1.328–2.344;  $p < 0.001$ ),

fatigue (OR, 1.363; 95% CI, 1.024–1.813;  $p = 0.033$ ), hypertension (OR, 1.570; 95% CI, 1.183–2.083;  $p = 0.002$ ), diabetes (OR, 1.650; 95% CI, 1.168–2.335;  $p = 0.005$ ), and anemia (OR, 3.497; 95% CI, 1.428–8.559;  $p = 0.006$ ) were confirmed to be independent risk factors associated with the severity of lung lesion (Table 4). The OR in patients with shortness of breath to develop into more severe lung lesions is 1.765 (1.328–2.344) over subjects without shortness of breath. The patients with anemia were 3.497 (95% CI 1.428–8.559) times more likely to develop into more severe lung lesions, compared to patients without anemia.

Table 4

Ordinal logistic regression analyses of baseline determinants with the severity of lung lesions

Variables	Univariate Analysis			Multivariate Analysis	
	p value of parallel line test	p value	OR (95% CI)	p value	OR (95% CI)
Age	< 0.001	< 0.001	0.950 (0.942, 0.960)	Not applicable	
Male	0.426	0.118	1.223 (0.950, 1.573)	Not selected	
Smoking	0.167	0.828	1.062 (0.617, 1.828)	Not selected	
Drinking	0.002	0.726	0.880 (0.430, 1.802)	Not selected	
Fever	0.114	< 0.001	1.662 (1.251, 2.208)	0.018	1.428 (1.064, 1.916)
Cough	0.190	< 0.001	1.868 (1.415, 2.467)	0.009	1.470 (1.099, 1.966)
Expectoration	0.039	0.213	1.255 (0.878, 1.791)	Not selected	
Shortness of breath	0.864	< 0.001	2.172 (1.667, 2.841)	< 0.001	1.765 (1.328, 2.344)
Myalgia	0.250	0.001	1.621 (1.221, 2.151)	---	
Fatigue	0.089	< 0.001	1.707 (1.309, 2.228)	0.033	1.363 (1.024, 1.813)
Diarrhea	0.281	0.821	0.931 (0.505, 1.719)	Not selected	
Nausea	0.717	0.365	1.380 (0.687, 2.773)	Not selected	
Headache	0.776	0.710	0.866 (0.405, 1.853)	Not selected	
Loss of appetite	0.358	0.283	1.416 (0.751, 2.672)	Not selected	
Chest distress	0.320	0.027	1.644 (1.059, 2.555)	---	
Coronary artery disease	0.031	0.306	1.303 (0.784, 2.166)	Not selected	
Hypertension	0.284	< 0.001	1.790 (1.366, 2.347)	0.002	1.570 (1.183, 2.083)
Arrhythmia	0.120	0.039	2.100 (1.037, 4.255)	---	
Diabetes	0.059	< 0.001	1.968 (1.412, 2.743)	0.005	1.650 (1.168, 2.335)
Hyperlipemia	0.681	0.228	0.553 (0.211, 1.449)	Not selected	
COPD	0.365	0.335	1.357 (0.731, 2.519)	Not selected	
Cancer	0.633	0.929	1.040 (0.438, 2.469)	Not selected	
Chronic liver disease	0.861	0.469	0.755 (0.353, 1.614)	Not selected	
Chronic kidney disease	0.289	0.684	1.331 (0.335, 5.296)	Not selected	
Anemia	0.876	0.009	3.136 (1.331, 7.389)	0.006	3.497 (1.428, 8.559)
Stroke	0.134	0.031	2.061 (1.069, 3.971)	---	
Gout	0.187	0.095	2.305 (0.865, 6.147)	Not selected	
Gastritis	0.753	0.598	1.262 (0.532, 2.995)	Not selected	

CI, confidence interval; COPD, Chronic obstructive pulmonary disease; "Not applicable" indicates the variable not applicable to the ordinal logistic regression for the p value of parallel line test < 0.05.

The lesion ratio of the right lung, left lung, and the whole lung in patients with or without cough, shortness of breath, hypertension, diabetes, or anemia are shown in Fig. 2. Patients with one of the aforementioned initial symptoms or comorbidities had higher lesion ratios in the whole lung, right lung, and right lung, compared to those in the patients without these initial symptoms or comorbidities.

## Discussion

To the best of our knowledge, this is the first study to investigate the association of demographics, initial symptoms, and comorbidities with COVID-19 pneumonia lesions quantitatively evaluated by AI algorithm

on CT when admitted in a relatively large population of 1630 patients in Huo Shen Shan Hospital. We found that in COVID-19 patients, the right lung had a larger percentage of lesions compared to that in the left lung, and the right lower lobe had the most lesions among the five lung lobes. In addition, the levels of plasma fibrinogen, plasma D-dimer, serum albumin, serum  $\alpha$ -hydroxybutyrate dehydrogenase, and peripheral blood leukocyte count were independently correlated with the lesion ratio of the whole lung. The COVID-19 patients with specific initial symptoms or comorbidities, such as cough, shortness of breath, fatigue, hypertension, diabetes, or anemia were more likely to develop into more severe lung lesions at admission.

Similar to Zhang et al.'s study, the right lower lobe was found to have the most lesions among the five lung lobes. Additionally, we found that the right lung had a larger percentage of lesions compared to that in the left lung. We reported a larger median lesion ratio and median lesion volume in COVID-19 patients on admission, compared to those observed in COVID-19 patients ready for discharge in the study by Zhang et al (16).

The highly correlated laboratory parameters with the lesion ratio of the whole lung in this study consisted of coagulation profile, serum enzymes, complete blood count, and hsCRP. These laboratory parameters were multiorgan related and confoundedly associated with the lesion ratio of the whole lung. Consistent with He et al.'s finding, the levels of serum hsCRP and serum lactate dehydrogenase were positively correlated with the lesion ratio of the whole lung, and the level of serum albumin was negatively correlated with the lesion ratio of the whole lung (2). However, adjusted by multivariate linear regression model, we found that only the plasma fibrinogen level, plasma D-dimer level, serum albumin level, serum  $\alpha$ -hydroxybutyrate dehydrogenase level, and peripheral blood leukocyte count were independently correlated with the lesion ratio of the whole lung. This observation shares the same perspective with previous researches that reported elevated plasma D-dimer or serum  $\alpha$ -hydroxybutyrate dehydrogenase were associated with poor clinical outcomes in COVID-19 patients (17, 18). Laboratory findings and post-mortem findings in COVID-19 patients revealed activation of coagulation pathways and inflammation responses in the COVID-19 (19–21). However, D-dimer and fibrinogen levels were independently correlated with the lesion ratio of the whole lung, highlighting the vital role of the coagulation system in lung injury. Procoagulant state overproduced proinflammatory cytokines via proteinase-activated receptors participating in the lung injury, or it was only the consequence of lung lesion initiated by endothelial cell injury requires further validation (22). The higher level of peripheral blood leukocyte count facilitated the overproduction of cytokines, which promoted the formation of cytokine storm and led to worse severe lung lesions. The deterioration of serum albumin and  $\alpha$ -hydroxybutyrate dehydrogenase concomitantly with the severity of the whole lung lesion ratio confirmed the involvement of multiorgan injury in SARS-CoV-2 infection. However, multiple organs that were injured by the invasion of SARS-CoV-2 synchronously or secondary to respiratory failure needs further investigation.

Initial symptoms of fever, cough, shortness of breath, and fatigue were risk factors of more severe COVID-19 pneumonia lesions on admission. It is believed that vascular endothelial cells play a central role in the pathophysiological process of COVID-19 (23). In severe COVID-19 infection, cytokine storms are always

induced by the early response of proinflammatory cytokines and injury of the vascular endothelial cells, especially in the lung. The formation of cytokine storms in early stage prompt fever. The invasion of SARS-CoV-2 stimulated cough as a defensive reflex. Patients with coughs are more likely to have infections in the respiratory tract and have more lesions. Shortness of breath as a sign of hypoxia, indicating the injury of the cardiopulmonary system and is often present in severe pneumonia. The higher lung-lesion ratio was shown to be associated with a higher respiratory rate in COVID-19 patients (2). We also found that the patients with initial symptoms of shortness of breath had a larger ratio of lung lesions on admission. In agreement with a previous study, we found that fatigue was associated with poor outcomes. However, a previous finding in the poor prognosis of expectoration was not found in our study (11, 24). Different from those studies, our research focussed on the association of expectoration with the severity of lung lesions and carried out a larger study population.

Compared to patients who died with H1N1, SARS-CoV-2 infection prompted endotheliitis and induced the apoptosis of vascular endothelial cells, and 9 times the number of alveolar capillary microthrombi that were found in COVID-19 victims (19). It is believed that the injury of vascular endothelial cells, particularly in the lung, promotes inflammation, induces tissue oedema, and leads to ARDS (25). Thus, hypertension and diabetes with pre-existing injured endothelial cells were associated with poor clinical outcomes in COVID-19 patients (12, 25, 26). In our study, we found that diabetes and hypertension, other than chronic obstructive pulmonary disease (COPD), were associated with the severity of lung lesions, which suggests that COVID-19 patients with hypertension and/or diabetes might have more severe pneumonia and thus predicts worse clinical outcomes. A recent study reported that initial anemia was correlated with longer hospital stay and higher mortality in COVID-19 patients, and patients with anemia of inflammation (68.8% of all kinds of anemia) had an eight-fold higher likelihood of developing mechanical ventilation (27). Anemia of inflammation was the most common anemia in hospitalized patients (28). A higher level of inflammation in patients with anemia of inflammation may participate in the injury of endothelial cells in the lung. In addition, COVID-19 patients with anemia were more likely to suffer from hypoxia, which might increase cytokines and lead to the apoptosis of endothelial cells, and further progress to more severe pneumonia (29).

Several limitations of this study should be noted. First, this was a retrospective study and was subject to selection bias. Second, findings in this single-centre study need a multicentre prospective study for validation. Finally, the documented initial symptoms were self-reported by patients on admission, which inevitably involves recall bias.

## Conclusions

Based on AI quantitative assessment on CT, the initial symptoms of fever, cough, shortness of breath, fatigue, and comorbidities, such as hypertension, diabetes, and anemia, independently predict a worse severity of lung lesions on admission in COVID-19 patients. This finding provides early-warning indicators of more severe COVID-19 pneumonia and might inform and improve the individualised treatments of COVID-19 patients.

## Abbreviations

AI: artificial intelligence; ARDS: acute respiratory distress syndrome; COVID-19: Coronavirus disease 2019; CI: confidence interval; COPD: Chronic obstructive pulmonary disease; CT: computed tomography; hsCRP: high-sensitivity C-reactive protein; ICU: intensive care unit; IQR: interquartile range; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

## Declarations

### Acknowledgements

We thank all the hospital staff in Huo Shen Shan Hospital, which gathered across the country in response to the outbreak of COVID-19.

### Ethics approval and consent to participate

This study was approved by the Human Ethics Committee of Huo Shen Shan Hospital (No. HSSLL023).

### Consent for publication

Not applicable

### Availability of data and materials

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

### Competing interests

The authors declare that they have no competing interests.

### Funding

This work was supported by Scientific Research Project of Huo Shen Shan Hospital.

### Authors' contributions

LH, FZYY and CL conceived or designed the work. HT, FZYY, RZC, YS, JBK, YQY, CYH and RC contributed to the data collection. FZYY and JY performed statistical analyses. SZB, XBG, JHZ and MDH interpreted the results of statistical analyses. FZYY and CL drafted the manuscript. LH, FZYY, CL and JY critically revised the manuscript. All authors read and approved the final manuscript.

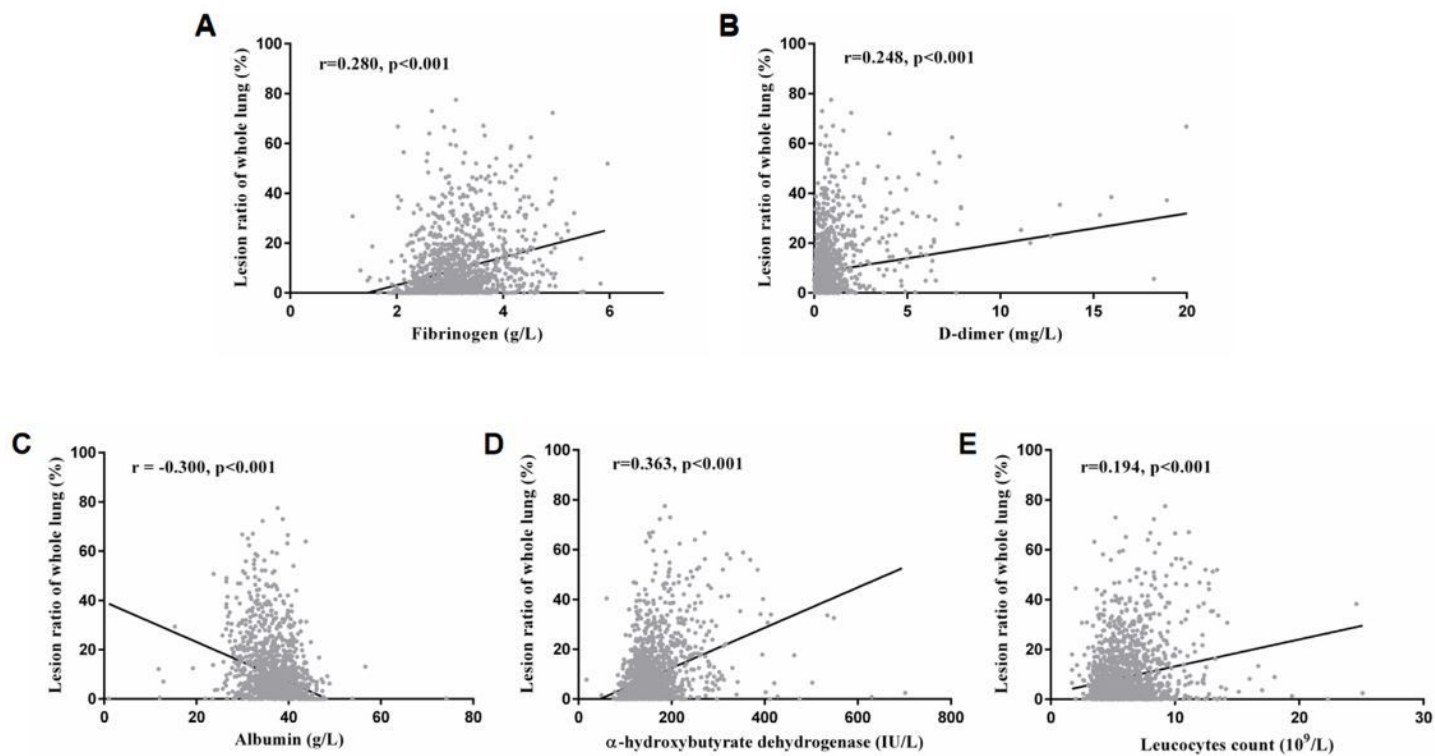
## References



1. 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.
2. Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, et al. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. *Cell*. 2020;181(6):1423-33.e11.
3. Liu F, Zhang Q, Huang C, Shi C, Wang L, Shi N, et al. CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients. *Theranostics*. 2020;10(12):5613-22.
4. Li Y, Yang Z, Ai T, Wu S, Xia L. Association of "initial CT" findings with mortality in older patients with coronavirus disease 2019 (COVID-19). *Eur Radiol*. 2020.
5. Colombi D, Bodini FC, Petrini M, Maffi G, Morelli N, Milanese G, et al. Well-aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia. *Radiology*. 2020;296(2):E86-E96.
6. Vaduganathan M, Vardeny O, Michel T, McMurray JJV, Pfeffer MA, Solomon SD. Renin-Angiotensin-Aldosterone System Inhibitors in Patients with Covid-19. *N Engl J Med*. 2020;382(17):1653-9.
7. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol*. 2004;203(2):631-7.
8. Ramachandran P, Onukogu I, Ghanta S, Gajendran M, Perisetti A, Goyal H, et al. Gastrointestinal Symptoms and outcomes in hospitalized COVID-19 patients. *Dig Dis*. 2020.
9. 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. *The Lancet*. 2020;395(10223):497-506.
10. National Institute for H, Care Excellence in collaboration with NHSE, Improvement NHS. Managing COVID-19 symptoms (including at the end of life) in the community: summary of NICE guidelines. *BMJ*. 2020;369:m1461.
11. Li J, Chen Z, Nie Y, Ma Y, Guo Q, Dai X. Identification of Symptoms Prognostic of COVID-19 Severity: Multivariate Data Analysis of a Case Series in Henan Province. *J Med Internet Res*. 2020;22(6):e19636.
12. Guan WJ, Liang WH, Zhao Y, Liang HR, Chen ZS, Li YM, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J*. 2020;55(5).
13. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *Lancet Respir Med*. 2018;6(11):837-45.
14. Shen C, Yu N, Cai S, Zhou J, Sheng J, Liu K, et al. Quantitative computed tomography analysis for stratifying the severity of Coronavirus Disease 2019. *J Pharm Anal*. 2020.
15. Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. AI Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Etiology on Chest CT. *Radiology*. 2020:201491.

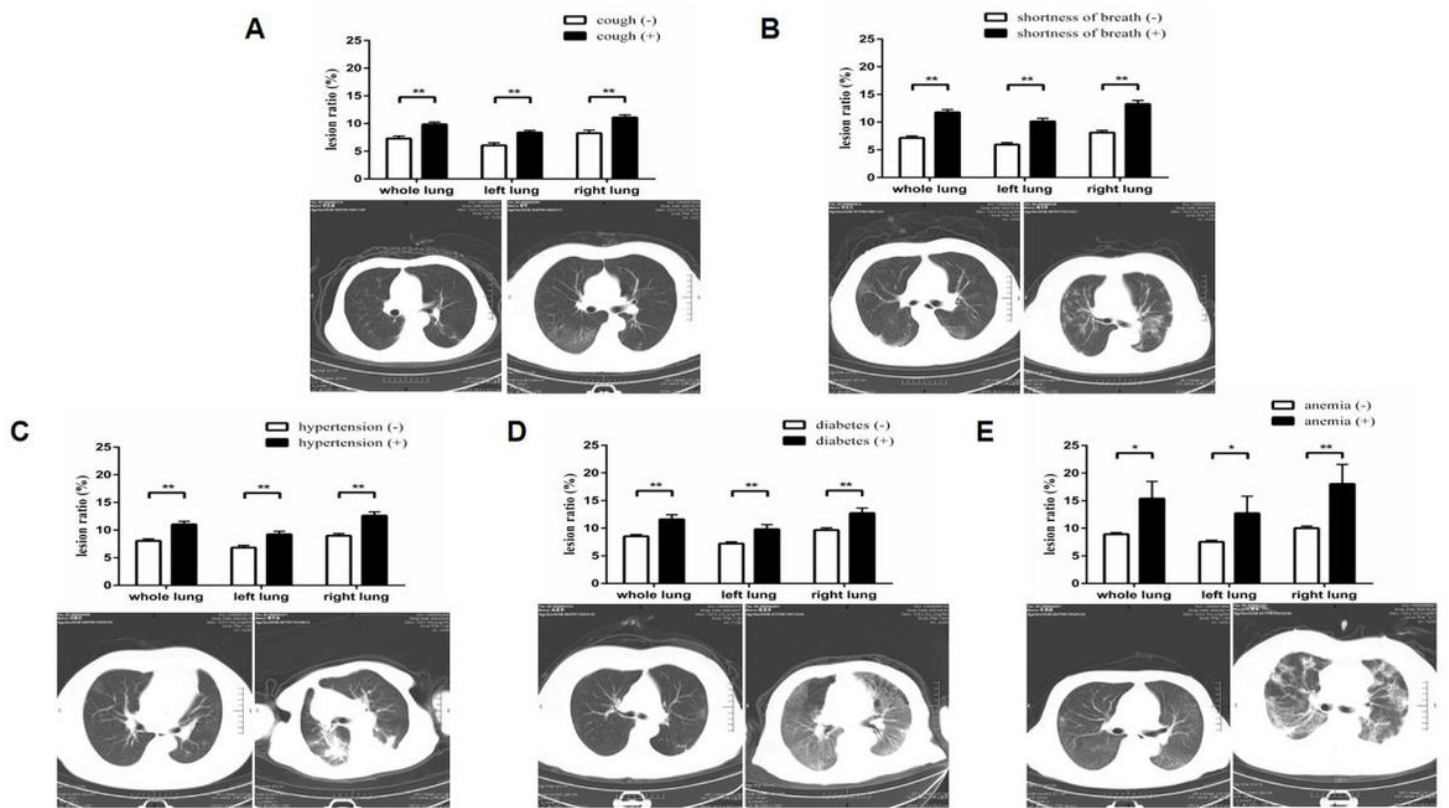
16. Du S, Gao S, Huang G, Li S, Chong W, Jia Z, et al. Chest lesion CT radiological features and quantitative analysis in RT-PCR turned negative and clinical symptoms resolved COVID-19 patients. *Quantitative Imaging in Medicine and Surgery*. 2020;10(6):1307-17.
17. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*. 2020;395(10229):1054-62.
18. Cen Y, Chen X, Shen Y, Zhang XH, Lei Y, Xu C, et al. Risk factors for disease progression in patients with mild to moderate coronavirus disease 2019-a multi-centre observational study. *Clin Microbiol Infect*. 2020.
19. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N Engl J Med*. 2020;383(2):120-8.
20. Tang N, Li D, Wang X, Sun Z. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost*. 2020;18(4):844-7.
21. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. *Lancet*. 2020;395(10234):1417-8.
22. Jose RJ, Manuel A. COVID-19 cytokine storm: the interplay between inflammation and coagulation. *Lancet Respir Med*. 2020;8(6):e46-e7.
23. Teuwen LA, Geldhof V, Pasut A, Carmeliet P. COVID-19: the vasculature unleashed. *Nat Rev Immunol*. 2020;20(7):389-91.
24. Zhang J, Wang X, Jia X, Li J, Hu K, Chen G, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. *Clin Microbiol Infect*. 2020;26(6):767-72.
25. Maticic C. Blood vessel injury may spur disease's fatal second phase. *Science*. 2020;368(6495):1039-40.
26. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev*. 2020:e3319.
27. Bellmann-Weiler R, Lanser L, Barket R, Rangger L, Schapfl A, Schaber M, et al. Prevalence and Predictive Value of Anemia and Dysregulated Iron Homeostasis in Patients with COVID-19 Infection. *J Clin Med*. 2020;9(8).
28. Weiss G, Ganz T, Goodnough LT. Anemia of inflammation. *Blood*. 2019;133(1):40-50.
29. Xiao F, Li X, Wang J, Cao J. Mechanisms of vascular endothelial cell injury in response to intermittent and/or continuous hypoxia exposure and protective effects of anti-inflammatory and anti-oxidant agents. *Sleep and Breathing*. 2019;23(2):515-22.

## Figures



**Figure 1**

The association between laboratory findings and the lesion ratio of whole lung Multivariate linear regression analysis revealed that laboratory parameters independently correlated with the lesion ratio of the whole lung. (A) The level of plasma fibrinogen ( $r=0.280, p<0.001$ ) positively correlated with the lesion ratio of the whole lung. (B) The level of plasma D-dimer ( $r=0.248, p<0.001$ ) positively correlated with the lesion ratio of the whole lung. (C) The level of serum albumin ( $r=-0.300, p<0.001$ ) negatively correlated with the lesion ratio of the whole lung. (D) The level of serum  $\alpha$ -hydroxybutyrate dehydrogenase ( $r=0.363, p<0.001$ ) positively correlated with the lesion ratio of the whole lung. (E) Leucocyte count ( $r=0.194, p<0.001$ ) positively correlated with the lesion ratio of the whole lung.



**Figure 2**

Distribution of lung lesion in patients with or without a specific initial symptom or comorbidity. The lung lesion ratio and images of patients with or without cough, shortness of breath, hypertension, diabetes, and anemia are shown from A to E. The left panel shows images of patients without a specific initial symptom or comorbidity. Right panel: images of patients with a specific initial symptom or comorbidity. (A–E) Patients with cough, shortness of breath, hypertension, diabetes, or anemia had a higher median lesion ratio of lung lesions in the left lung, right lung, and whole lung compared to patients without. \* $p < 0.05$ ; \*\* $p < 0.01$ .

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)