

## Original Paper

# Identification of Symptoms Prognostic of COVID-19 Severity: Multivariate Data Analysis of a Case Series in Henan Province

Jitian Li<sup>1\*</sup>, PhD; Zhe Chen<sup>2\*</sup>, PhD; Yifei Nie<sup>3</sup>, MPH; Yan Ma<sup>1</sup>, MPH; Qiaoyun Guo<sup>3</sup>, MPH; Xiaofeng Dai<sup>4\*</sup>, PhD

<sup>1</sup>Henan Provincial Orthopedic Institute, Zhengzhou, China

<sup>2</sup>Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

<sup>3</sup>Henan Center for Disease Control and Prevention/Henan Key Laboratory of Pathogenic Microorganisms, Zhengzhou, China

<sup>4</sup>Wuxi School of Medicine, Jiangnan University, Wuxi, China

\*these authors contributed equally

**Corresponding Author:**

Xiaofeng Dai, PhD

Wuxi School of Medicine

Jiangnan University

Lihu Avenue 1800

Wuxi, 214122

China

Phone: 86 18168870169

Email: [1281423490@qq.com](mailto:1281423490@qq.com)

## Abstract

**Background:** The outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease (COVID-19), has been declared a global pandemic. Identifying individuals whose infection can potentially become severe is critical to control the case fatality rate of COVID-19. However, knowledge of symptoms that are prognostic of COVID-19 severity is lacking.

**Objective:** The objective of our study was to identify symptoms prognostic of COVID-19 infection severity.

**Methods:** We analyzed documented symptoms, including fever, cough, fatigue, expectoration, sore throat, chest distress, headache, diarrhea, rhinorrhea, stuffed nose, nausea, vomiting, muscle or joint ache, shortness of breath, and their associations with disease severity using a case series, including 655 confirmed cases from January 23 to February 5, 2020 in Henan Province, China. We also analyzed the influence of individual characteristics, including age, gender, and comorbidities, on symptoms with prognostic value.

**Results:** Fatigue (95% CI 0.141 to 0.334,  $P < .001$ ), expectoration (95% CI 0.107 to 0.305,  $P < .001$ ) and stuffed nose (95% CI  $-0.499$  to  $-0.082$ ,  $P = .006$ ) were identified as the prognostic symptoms of COVID-19 patients from the multivariate analysis. Fever occurred in 603/655 (92.1%) of the patients but was not associated with disease severity. Fatigue accounted for 184/655 (28.1%) of the patients and was linearly associated with infection severity with statistical significance. Expectoration occurred in 169/655 (25.8%) patients in the cohort and was the sole prognostic factor for patients with cardiovascular complications, including hypertension. Shortness of breath, chest distress, muscle or joint ache, and dry cough, which occurred in 33 (5%), 83 (12.7%), 78 (11.9%), and 276 (42.1%) of the 655 patients, respectively, were significantly enriched among patients classified as severe. Stuffed nose and nausea were associated with favorable disease severity, especially among male patients. More female than male patients were documented as having muscle or joint ache. Headache was most enriched in patients aged 15 to 39 years, followed by those aged 40 to 64 years, with statistical significance.

**Conclusions:** Fatigue and expectoration are signs of severe COVID-19 infection. Shortness of breath, chest distress, muscle or joint ache, and dry cough are prevalent in severe patients. Expectoration is commonly present in older individuals and patients with cardiovascular disorders, including hypertension. Shortness of breath is prognostic of severe infection in male patients. Stuffed nose and nausea are favorable prognostic factors of severe infection, especially among male patients.

(*J Med Internet Res* 2020;22(6):e19636) doi: [10.2196/19636](https://doi.org/10.2196/19636)

**KEYWORDS**

prognostic symptoms; COVID-19; severity; CVD; Henan Province

## Introduction

In early December 2019, a pneumonia of unknown etiology emerged in Wuhan, a city in China with 11 million permanent residents and 5 million recurrent residents. On December 29, 2019, the first four cases of this pneumonia were reported, all of which were linked to the Huanan Seafood Wholesale Market in Wuhan. On January 7, 2020, a novel coronavirus was identified from the bronchoalveolar lavage fluid of a patient [1] and was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the World Health Organization (WHO) [2]. SARS-CoV-2 is the seventh enveloped RNA coronavirus to be identified [3]; it is transmittable via humans and has a 3-day median incubation time [4]. This virus has rapidly spread worldwide and has become a global health threat [5]. The high binding affinity of SARS-CoV-2 to angiotensin-converting enzyme 2 enables its rapid transmission [6]. Approximately 8 million individuals were infected and over 0.4 million deaths were reported worldwide as of June 2020 [7]. The death rate varies among countries and reached as high as 27.1% in Yemen [7].

Suspected cases were identified as having “fever or respiratory symptoms” and “traveling history or contact with confirmed infections within 2 weeks” [8]. Unlike the symptoms of severe acute respiratory syndrome (SARS), where fever was the symptom in approximately 100% of infected individuals [9], fever is only observed in 87.9% of patients with COVID-19 on admission [4]; other symptoms, such as cough (67.7%) and fatigue (38.1%), also frequently occur in SARS-CoV-2-infected individuals. Despite the many reports on symptoms associated with COVID-19, little effort has been devoted to the identification of symptoms associated with its severity. This study aims to identify symptoms with prognostic value on disease severity and their correlations with individual characteristics such as age, gender, and comorbidities to aid the prognosis of COVID-19 severity.

## Methods

### Data Source

This case series was collected by the Center for Disease Control and Prevention of Henan Province (Henan CDC) from 279 hospitals in the province. It includes 655 confirmed patients with COVID-19 who showed symptoms on admission and were admitted to hospitals from January 23, 2020, to February 5, 2020, with February 5 being the last follow-up date. Oral consent was obtained from the patients. All enrolled patients were diagnosed according to the WHO interim guidance [10].

Epidemic, clinical, and severity data were obtained with data collection forms from electronic medical records as part of standard care. The information recorded included demographic data, comorbidities, symptoms, and chest computed tomography (CT) scans. The date of disease onset was defined as the day when the symptom was noted. This study was approved by the ethics commissions of the Henan CDC with a waiver of informed consent. Cardiovascular disease (CVD) is a group of disorders of the heart and blood vessels, including

cerebrovascular disease, coronary heart disease, cerebrovascular disease, peripheral arterial disease, and rheumatic heart disease.

We stratified severity into three groups: light, normal, and severe. According to the Novel Coronavirus Diagnostic and Therapeutic Plan (Seventh Edition) [11], COVID-19 severity was initially divided into four types: light, normal, severe, and terminal. We merged severe and terminal patients into one group, named the severe group. The clinical symptoms of patients in the light group were mild, with no pneumonia found in imaging. Patients in the normal group had fever, respiratory tract infection, and other symptoms, with manifestation of pneumonia observable in imaging. In adults, if one of the following three conditions was satisfied, the patient was classified as severe: 1) shortness of breath, respiratory rate  $\geq 30$  times per minute; oxygen saturation  $\leq 93\%$  in the resting state; arterial blood sample partial pressure ( $\text{PaO}_2$ )/oxygen concentration ( $\text{FiO}_2$ )  $\leq 300$  millimeters of mercury. Among children, if any of the following criteria were met, the child was classified as severe: shortness of breath ( $\leq 2$  months of age, respiratory rate  $\geq 60$  times/min; 2 to 12 months of age, respiratory rate  $\geq 50$  times/min; 1-5 years of age, respiratory rate  $\geq 40$  times/min;  $>5$  years of age, respiratory rate  $\geq 30$  times/min), excluding the effects of fever and crying; oxygen saturation in the resting state is  $\leq 92\%$ ; assisted respiration (moaning, alar fluttering, three-concave sign); cyanosis; intermittent apnea; drowsiness; convulsion; refusing to eat or difficulty in feeding; and signs of dehydration. Patients with critically severe cases satisfied the following criteria: respiratory failure that requires mechanical ventilation; shock; other organ failure that requires intensive care unit monitoring and treatment.

### Laboratory Testing

Throat or nose swab samples were collected from patients suspected of having SARS-CoV-2 infection for total RNA extraction using the respiratory sample RNA isolation kit (Shanghai BioGerm Medical Technology Co Ltd, Catalog No. ZC-HX-201-2), followed by real time reverse transcription-polymerase chain reaction (RT-PCR) using a SARS-CoV-2 nucleic acid detection kit (Shanghai BioGerm Medical Technology Co Ltd) in the biosafety level 2 lab at Henan CDC. Targeting the open reading frame (ORF1a/b), the primers and sequences were forward primer CCCTGTGGGTTTTACTTAA, reverse primer ACGATTGTGCATCAGCTGA, and probe 5'-FAM-CCGTCTGCGGTATGTGAAAGGTTATGG-BHQ1-3'. Targeting nucleocapsid protein, the primers and sequences were forward primer GGGGAACCTTCTCCTGCTAGAAT, reverse primer CAGACATTTGCTCTCAAGCTG, and probe 5'-FAM-TTGCTGCTGCTTGACAGATT-TAMRA-3'. Conditions for the amplifications were 50 °C for 10 minutes and 95 °C for 5 minutes, followed by 40 cycles of 95 °C for 10 seconds and 55 °C for 40 seconds. Following recommendations by the Chinese National Institute for Viral Disease Control and Prevention [12], positive and negative tests were defined as cycle threshold ( $C_t$ ) values  $<37$  and  $\geq 40$ , respectively; samples with  $C_t$  values between these thresholds were subjected to retesting. A case was confirmed if two targets (ORF1a or 1b,

nucleocapsid protein) tested positive by real-time RT-PCR in the initial test or both tests (when a retest was needed).

### Statistical Analysis

Continuous variables were described using mean, SD, median, interquartile range (IQR), and range, and categorical variables were described by frequency and percentage. Means for continuous variables were compared using independent group tests when the data were normally distributed (Shapiro-Wilk test); otherwise, the Kruskal-Wallis H test was used (adjusted by Bonferroni correction). Proportions for categorical variables were compared using the chi-square test or Fisher exact test. The correlation of two variables was compared using Spearman rank correlation. The influencing variables for the severity of COVID-19 were analyzed using a linear regression model (forward method). All statistical analyses were performed using SPSS version 23.0 (IBM Corp). A 2-sided  $\alpha$  less than .05 was considered statistically significant.

### IT Infrastructure

We implemented the new Public Health Emergency Management Information System of Henan Province and extracted data from the system to analyze the epidemiological characteristics of COVID-19 patients in Henan Province and develop a COVID-19 cluster statistical information template. In addition, the hospital information system, dashboards, electronic prescription system, and cloud-based medical image sharing system were used to facilitate the analysis of the clinical data of COVID-19 patients.

## Results

In this case series including 655 COVID-19 patients, regarding disease severity, 163 (24.9%) patients were classified as light, 420 (64.1%) as normal, and 72 (11%) as severe (Table 1). Of the infected individuals, 12/655 (1.8%) were less than 15 years of age, 265 (40.5%) were between 15 and 39 years of age, 322 (49.2%) were between 40 and 64 years of age, and 71 (8.5%) were aged 65 years or older (Table 1). Most patients were male (367/655, 56.0%). Of the 150/655 patients (22.9%) who had coexisting medical disorders, 89 (59.3%) had CVD. Given the high percentage of CVD comorbidities in this dataset, we analyzed symptoms prognostic of COVID-19 severity separately among patients with CVD, patients without CVD, and patients without comorbidities. The average number of days from illness onset to diagnosis was 5.66 (SD 3.64, range 0-32). Among the 655 patients included in this case series, 634 (96.8%) had chest CT scans, and 535 (84.4%) showed typical pneumonia features (bilateral ground glass opacities, Figure 1).

Among all documented symptoms in this case series, fever and dry cough were prevalent in all patient cohorts as stratified by

COVID-19 severity. Fatigue and expectoration were enriched symptoms in the severe and normal groups (Figure 2A).

The symptom of fever occurred in 603/655 (92.1%) of the patients (Table 1) but was not associated with disease severity. A higher "highest temperature of fever" was typically associated with more severe SARS-CoV-2 infection; however, it is difficult to precisely define high and low highest temperatures of fever given the high dependence of this parameter on the time slot when it was measured and high individual heterogeneity among patients. The mean highest temperature in the light group was 38.0°C (range 37.6-38.5°C), that in the normal group was 38.0°C (range 37.8-38.5°C), and that in the severe group was 38.4°C (range 38.0-38.9°C) ( $\chi^2_2=15.5$ ,  $P<.001$ ).

Other symptoms, including dry cough, fatigue, expectoration, chest distress, muscle or joint ache, shortness of breath, and multiple symptoms, all convey significant prognostic value on disease severity. Fatigue was the most prevalent symptom (184/655, 28.1%) among patients (Table 1) and could significantly ( $\chi^2_2=20.8$ ,  $P<.001$ ) stratify and linearize patients into light, normal, and severe groups regarding clinical severity (Table 2, Figure 2A). Expectoration was observed in 169/655 (25.8%) of the patients, with significant ( $\chi^2_2=14.5$ ,  $P=.001$ ) prognostic value on disease severity (Table 1, Table 2). Expectoration significantly ( $\chi^2_2=6.6$ ,  $P=.04$ ) differed among patients with CVD and patients with complications other than CVD as well as patients without coexisting disorders (Multimedia Appendix 1), and it was the sole explanatory symptom included in the linear regression among CVD patients ( $\beta=0.310$ , 95% CI 0.042-0.579,  $P=.02$ , Table 3). Shortness of breath was significantly ( $\chi^2_2=18.3$ ,  $P<.001$ ) enriched in the severe group (Table 2) and was observed in 33/655 (5.0%) of the patients infected with COVID-19 (Table 1). Chest distress occurred in 12.7% of this cohort (Table 1), and the sample group was significantly enriched with severely infected patients ( $\chi^2_2=11.367$ ,  $P=.003$ , Table 2). Muscle or joint ache was documented in 78/655 (11.9%) of the patients and was significantly enriched in severe patients ( $\chi^2_2=7.7$ ,  $P=.02$ , Table 2) and female patients ( $\chi^2_1=4.5$ ,  $P=.03$ , Multimedia Appendix 1). Dry cough, although it could stratify patients according to COVID-19 severity with statistical significance ( $\chi^2_2=7.4$ ,  $P=.03$ , Table 2), did not show good discrimination power (Figure 2B); this may be due to the high prevalence of dry cough among patients (276/655, 42.1%, Table 1). Headache was enriched in the 15 to 39 years age group, followed by the 40 to 64 years age group, with statistical significance ( $\chi^2_3=11.6$ ,  $P=.009$ , Multimedia Appendix 1); however, headache could not differentiate disease severity. The majority of the patients (501/655, 76.5%) reported multiple symptoms (Table 1).

**Table 1.** Documented symptoms and general clinical characteristics of 655 patients with COVID-19 (N=655).

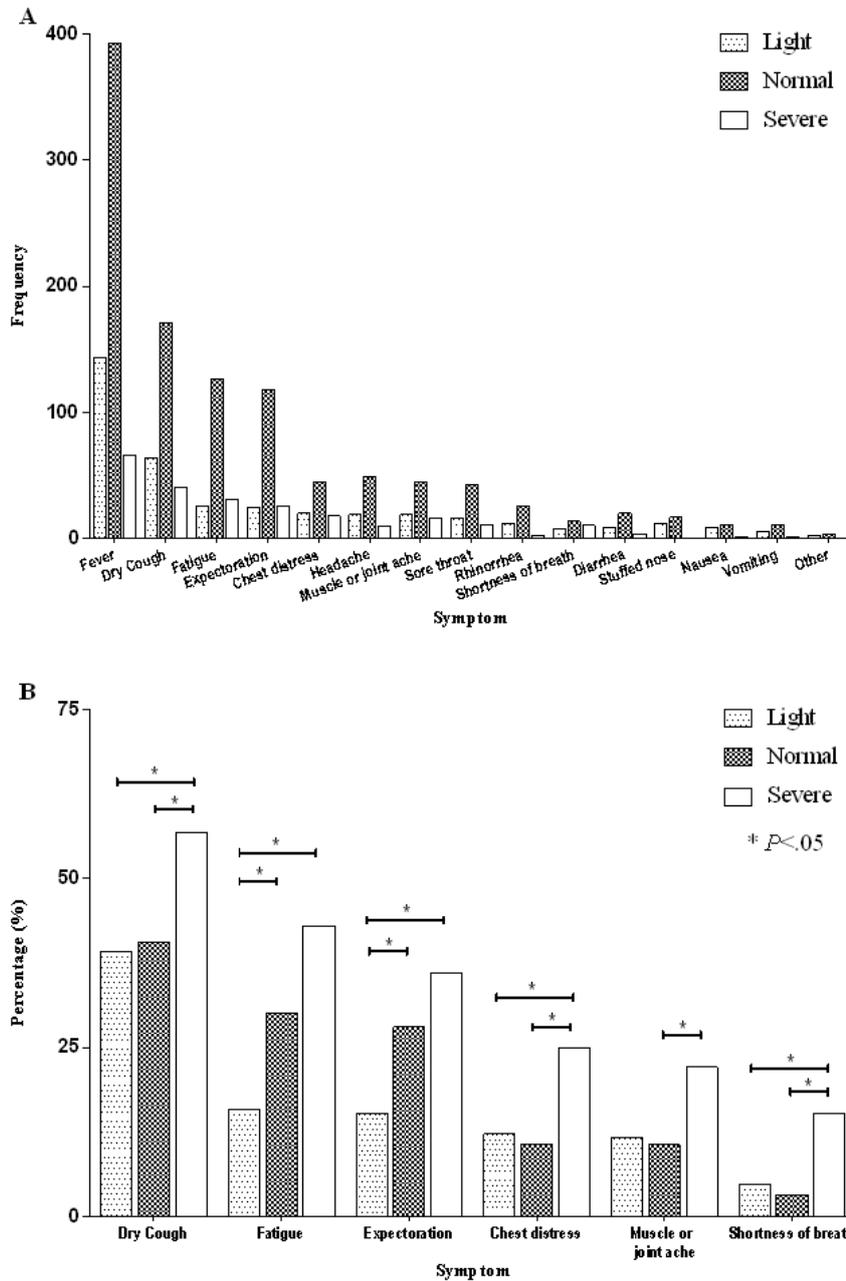
Characteristic	n (%)
<b>Disease severity</b>	
Light	163 (24.9)
Normal	420 (64.1)
Severe	72 (11)
<b>Symptoms</b>	
<b>Fever (°C)</b>	
<37.3	23 (3.5)
37.3-38.0	287 (43.8)
38.1-39.0	241 (36.8)
>39.0	37 (5.6)
Temperature not documented	15 (2.1)
Dry cough	276 (42.1)
Fatigue	184 (28.1)
Expectoration	169 (25.8)
Chest distress	83 (12.7)
Headache	80 (12.2)
Muscle or joint ache	78 (11.9)
Sore throat	70 (10.7)
Rhinorrhea	41 (6.3)
Shortness of breath	33 (5.0)
Diarrhea	33 (5.0)
Stuffed nose	30 (4.6)
Nausea	22 (3.4)
Vomiting	19 (2.9)
Other symptoms	7 (1.1)
Multiple symptoms	501 (76.5)
<b>Age (years)</b>	
<15	12 (1.8)
15-39	265 (40.5)
40-64	322 (49.2)
≥65	56 (8.5)
<b>Gender</b>	
Male	367 (56.0)
Female	288 (44.0)
<b>Coexisting disorders</b>	
CVD <sup>a</sup>	89 (59.3)
Disorders other than CVD	61 (40.7)
None	505 (77.1)

<sup>a</sup>CVD: cardiovascular disease (including hypertension).

**Figure 1.** Chest computed tomographic images of a female patient aged 56 years infected with coronavirus disease. The images show ground glass opacity in both lungs on day 5 after symptom onset.



**Figure 2.** Summarized report of documented symptoms. Plots of (A) patient number of documented symptoms and (B) patient percentage of prognostic symptoms in cohorts stratified by coronavirus disease severity.



**Table 2.** Symptoms stratified by disease severity among patients with COVID-19 (N=655).

Symptom	Light (n=163), n (%)	Normal (n=420), n (%)	Severe (n=72), n (%)	Chi-square (df)	P value
<b>Fever</b>					
All fever	144 (88.3)	393 (93.6)	66 (91.7)	4.4 (2)	.11
<37.3°C	4 (2.9)	18 (4.6)	1 (1.6)	21.5 (6) <sup>b</sup>	.002
37.3-38.0°C	80 (58.4) <sup>a</sup>	187 (48.1)	20 (32.3) <sup>a</sup>		
38.1-39.0°C	43 (31.4) <sup>a</sup>	166 (42.7)	32 (51.6) <sup>a</sup>		
>39.0°C	10 (7.3)	18 (4.6) <sup>a</sup>	9 (14.5) <sup>a</sup>		
Temperature not documented	7 (4.3)	4 (0.1)	4 (5.6)		
Dry cough	64 (39.3)	171 (40.7)	41 (56.9) <sup>c</sup>	7.4 (2)	.03
Fatigue	26 (16.0) <sup>c</sup>	127 (30.2)	31 (43.1)	20.8 (2)	<.001
Expectoration	25 (15.3) <sup>c</sup>	118 (28.1)	26 (36.1)	14.5 (2)	.001
Chest distress	20 (12.3)	45 (10.7)	18 (25.0) <sup>c</sup>	11.4 (2)	.003
Headache	20 (12.3)	49 (11.7)	10 (13.9)	0.3 (2)	.86
Muscle or joint ache	19 (11.7)	45 (10.7) <sup>a</sup>	16 (22.2) <sup>a</sup>	7.7 (2)	.02
Sore throat	16 (9.8)	43 (10.2)	11 (15.3)	1.8 (2)	.41
Rhinorrhea	12 (7.4)	26 (6.2)	3 (4.2)	0.9 (2)	.65
Shortness of breath	8 (4.9)	14 (3.3)	11 (15.3) <sup>c</sup>	18.3 (2)	<.001
Diarrhea	9 (5.5)	20 (4.8)	4 (5.6)	0.2 (2)	.91
Stuffed nose	12 (7.4)	17 (4.0)	1 (1.4)	4.8 (2)	.09
Nausea	9 (5.5)	11 (2.6)	2 (2.8)	3.1 (2)	.21
Vomiting	6 (3.7)	11 (2.6)	2 (2.8)	0.7 (1)	.78
Other	3 (1.8)	4 (1.0)	0 (0)	1.3 (1)	.49
Multiple symptoms	111 (68.1)	327 (77.9)	63 (87.50)	11.7 (2)	.003

<sup>a</sup>Pairwise significance.

<sup>b</sup>Calculated with the R by C chi-square test.

<sup>c</sup>Significant compared with the other two groups.

Linear models were established to explain disease severity by including all patients or cohort groups stratified by age, gender, and coexisting disorder. All models constructed were significant (Table 3). Fatigue was the most frequently present symptom in these equations, followed by expectoration, and the coefficients of both symptoms were positive (Table 3). While expectoration was the sole symptom associated with patients with CVD, fatigue was linked to patients with coexisting disorders other than CVD (Table 3). Excluding patients with comorbidities did not change the variables included in the model except for slight variations in the coefficients. Stuffed nose was a negative

explanatory variable in the models including all patients, individuals without comorbidities, and male patients only (Table 3). Nausea was a negative explanatory variable and played a dominant role in the equation established for male patients (Table 3).

The symptoms of highest temperature of fever (correlation 0.166, 95% CI 0.082-0.251,  $P<.001$ ), expectoration (correlation 0.104, 95% CI 0.022-0.184,  $P=.012$ ) and shortness of breath (correlation 0.125, 95% CI 0.001-0.041,  $P=.002$ ) were significantly correlated with days from illness onset to diagnosis (Table 4).

**Table 3.** Linear models of clinical severity as explained by symptoms.

Stratification factor	Parameter from stepwise modeling	Regression coefficient	95% CI	P value	Model <sup>a</sup>	
All	Constant	1.755	1.697 to 1.812	<.001	Y=0.237*F+0.206*E-0.291*S+1.755	
	Fatigue	0.237	0.141 to 0.334	<.001		
	Expectoration	0.206	0.107 to 0.305	<.001		
	Stuffed nose	0.291	-0.499 to -0.082	.006		
Age	Constant	1.499	1.272 to 1.725	<.001	Y=0.460*F+0.656*E+1.499	
	≥65 years	Expectoration	0.656	0.253 to 1.059		.002
	Fatigue	0.460	0.066 to 0.854	.02		
40-64 years	Constant	1.820	1.739 to 1.900	<.001	Y=0.202*F+0.137*E+1.82	
	Fatigue	0.202	0.072 to 0.331	.002		
	Expectoration	0.137	0.001 to 0.274	.048		
15-39 years	Constant	1.789	1.710 to 1.868	<.001	Y=0.166*F+1.789	
	Fatigue	0.166	0.008 to 0.324	.04		
<15 years	Constant	1.20	0.918 to 1.482	<.001	Y=0.8*F+1.2	
Gender	Constant	1.783	1.706 to 1.860	<.001	Y=0.219*F+0.212*E-0.513*N-0.343*S+0.276*B+1.783	
	Male gender	Fatigue	0.219	0.091 to 0.346		.001
	Expectoration	0.212	0.083 to 0.340	.001		
	Nausea	-0.513	-0.880 to -0.147	.006		
	Stuffed nose	-0.343	-0.602 to -0.084	.01		
	Shortness of breath	0.276	0.027 to 0.525	.03		
	Constant	1.718	1.632 to 1.803	<.001		
Female gender	Fatigue	0.252	0.100 to 0.404	.001	Y=0.252*F+0.164*E+1.718	
	Expectoration	0.164	0.006 to 0.322	.04		
Coexisting disorder	Constant	1.864	1.727 to 2.000	<.001		
Coexisting CVD <sup>b</sup>	Expectoration	0.310	0.042 to 0.579	.02	Y=0.31*E+1.864	
	Constant	1.725	1.548 to 1.902	<.001		
Coexisting disorder other than CVD	Fatigue	0.418	0.116 to 0.720	.008	Y=0.418*F+1.725	
	Constant	1.736	1.672 to 1.800	<.001		
No coexisting disorder	Fatigue	0.264	0.152 to 0.376	<.001	Y=0.264*F+0.227*E-0.314*S+1.736	
	Expectoration	0.227	0.111 to 0.343	<.001		
	Stuffed nose	-0.314	-0.548 to -0.080	.009		

<sup>a</sup>B: shortness of breath. E: expectoration. F: fatigue. N: nausea. S: stuffed nose.

<sup>b</sup>CVD: cardiovascular disease.

**Table 4.** Spearman correlation analysis of symptoms and days from illness onset to diagnosis.

Symptoms	Days from illness onset to diagnosis		
	Correlation	95% CI	<i>P</i> value
Fever	−0.039	−0.119 to 0.065	.35
Highest temperature of fever	0.166 <sup>a</sup>	0.082 to 0.251	<.001
Temperature <37.3°C	−0.130	−0.492 to 0.306	.55
Temperature 37.3-38.0°C	0.006	−0.116 to 0.124	.91
Temperature 38.1-39.0°C	0.009	−0.121 to 0.139	.89
Temperature >39.0°C	0.109	−0.190 to 0.403	.52
Dry cough	0.058	−0.021 to 0.132	.15
Fatigue	0.023	−0.067 to 0.101	.58
Expectoration	0.104 <sup>b</sup>	0.022 to 0.184	.012
Chest distress	0.070	−0.011 to 0.157	.09
Headache	−0.037	−0.117 to 0.044	.37
Muscle or joint ache	0.013	−0.069 to 0.093	.76
Sore throat	0.008	−0.071 to 0.080	.85
Rhinorrhea	−0.041	−0.114 to 0.033	.33
Shortness of breath	0.125 <sup>a</sup>	0.001 to 0.041	.002
Diarrhea	−0.004	−0.086 to 0.087	.92
Stuffed nose	−0.065	−0.152 to 0.033	.12
Nausea	0.018	−0.072 to 0.103	.66
Vomiting	−0.032	−0.131 to 0.063	.43
Multiple symptoms	0.079	−0.002 to 0.158	.056

<sup>a</sup>*P*<.01 (two-tailed).

<sup>b</sup>*P*<.05 (two-tailed).

## Discussion

### Principal Findings

The main findings of this study were that fatigue and expectoration are signs of severe COVID-19 infection and that stuffed nose and nausea are favorable prognostic factors of disease severity.

Two symptoms, fatigue and expectoration, showed linear associations with COVID-19 severity (Figure 2B). Four symptoms, namely shortness of breath, chest distress, muscle or joint ache, and dry cough, were more commonly present in severe patients. Thus, if more patients are documented with these symptoms, more resources for intensive medical care should be administrated.

These prognostic symptoms were interconnected with gender, CVD, and age. The enrichment of shortness of breath in severe patients was the most evident, especially among male patients; this suggests its feasibility as a sign (with relatively low type I error) for intensive care among confirmed cases. Stuffed nose and nausea were associated with less severe COVID-19, especially among male patients. Expectoration was significantly associated with CVD complications, suggesting a correlation between CVD and lower respiratory tract infection. An age of

40 years was shown to be a breakpoint for symptoms prognostic of disease severity. While fatigue conveyed prognostic value for all age groups, expectoration showed significance when the patients' age exceeded 40 years and dominated the model when their age exceeded 65 years. This may be due to the increased likelihood of developing CVD with increasing age and explainable by the strong association between expectoration and CVD comorbidity. As a clinical suggestion, individuals concomitantly having these characteristics and symptoms should be suspected for infection and given immediate quarantine and potential intensive medical care during the COVID-19 epidemic.

It is known that immune response to virus infection plays a vital role in the inflammation involved in heart diseases such as myocarditis, atherosclerosis, and cardiac insufficiency, and it constitutes the pathogenesis of cardiac disorder in humans [13]. Shortness of breath is a typical sign of heart and lung conditions, and fever is a symptom of stimulated immune response to infection. It was observed from our study that expectoration was associated with CVD and male gender; expectoration, shortness of breath, and fever were significantly correlated with days from illness onset to diagnosis. Therefore, we expect that patients with comorbidities, especially CVD, would experience exacerbated COVID-19 severity.

It is worth mentioning that symptoms prognostic of COVID-19 severity differ from symptoms for early COVID-19 diagnosis. While fever and dry cough were the most prevalent symptoms among infected individuals, fatigue, expectoration, and chest distress conveyed prognostic value on disease severity. This can be explained by the small percentage of patients represented in the whole case series (72, 11%) who could not be predicted from prevalent symptoms.

Although many case series including higher numbers of cases than that in this study have been reported [4,14-16], given the rapid changing global situation of COVID-19 and the sharp rise in the number of infected cases during the past few months, relatively little research has been devoted to studying the prognostic value of symptoms on severity. The sole relevant study was reported by Dong et al [17], who analyzed data from 663 patients. Compared with Dong's study, which analyzed two symptoms (expectoration and muscle ache), one laboratory

test index (albumin), and one patient characteristic (gender), our study concentrates on a more focused and complete list of characteristics that encompasses 14 symptoms. However, this study is limited by its high dependence on the accuracy and completeness of the symptoms recorded for each patient.

## Conclusions

Our study provides a statistical analysis of documented symptoms of 655 confirmed COVID-19 patients from Henan Province to aid SARS-CoV-2 diagnosis and prognosis. We conclude that fatigue and expectoration are the most important symptoms prognostic of severe COVID-19, and gender, age, and CVD comorbidity are factors associated with these symptoms; muscle or joint ache commonly occurs in female patients, and younger patients are likely to develop headache; and high temperature in fever, expectoration, and shortness of breath are typically associated with delayed diagnosis.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant No. 81972789), the National Science and Technology Major Project (Grant No. 2018ZX10302205-004-002), and the Major Project of TCM Research in Henan Province (Grant No. 2018ZYD01).

## Authors' Contributions

DXF is the principal investigator and guarantor of the paper, had full access to all the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis. DXF and LJT obtained the funding. DXF designed the research and drafted the manuscript. LJT and MY performed the statistical analysis. DXF, LJT, CZ, NYF, MY, and GQY contributed to the acquisition, analysis, or interpretation of data and critically reviewed and revised the article for important intellectual content. All authors approved the final manuscript and decided to submit the article for publication. LJT, ZC, and XFD contributed equally as co-first authors.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Symptoms stratified by CVD disorder, gender, and age in a case series of COVID-19 patients from Henan Province (N=655). [[DOCX File, 25 KB-Multimedia Appendix 1](#)]

## References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020 Feb 20;382(8):727-733. [doi: [10.1056/nejmoa2001017](https://doi.org/10.1056/nejmoa2001017)] [Medline: [31978945](https://pubmed.ncbi.nlm.nih.gov/31978945/)]
2. Guo Y, Cao Q, Hong Z, Tan Y, Chen S, Jin H, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil Med Res* 2020 Mar 13;7(1):11 [FREE Full text] [doi: [10.1186/s40779-020-00240-0](https://doi.org/10.1186/s40779-020-00240-0)] [Medline: [32169119](https://pubmed.ncbi.nlm.nih.gov/32169119/)]
3. 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 Feb 15;395(10223):497-506. [doi: [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)] [Medline: [31986264](https://pubmed.ncbi.nlm.nih.gov/31986264/)]
4. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: [10.1056/NEJMoa2002032](https://doi.org/10.1056/NEJMoa2002032)] [Medline: [32109013](https://pubmed.ncbi.nlm.nih.gov/32109013/)]
5. Novel Coronavirus (2019-nCoV) Situation Report-15. World Health Organization. 2020 Feb 04. URL: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200204-sitrep-15-ncov.pdf?sfvrsn=88fe8ad6\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200204-sitrep-15-ncov.pdf?sfvrsn=88fe8ad6_4) [accessed 2020-02-15]
6. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020 Mar 13;367(6483):1260-1263 [FREE Full text] [doi: [10.1126/science.abb2507](https://doi.org/10.1126/science.abb2507)] [Medline: [32075877](https://pubmed.ncbi.nlm.nih.gov/32075877/)]

7. Coronavirus Disease (COVID-19) Situation Report – 150. World Health Organization. 2020 Jun 18. URL: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200618-covid-19-sitrep-150.pdf?sfvrsn=aa9fe9cf\\_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200618-covid-19-sitrep-150.pdf?sfvrsn=aa9fe9cf_2) [accessed 2020-06-19]
8. New Coronavirus Infection Pneumonia Prevention and Control Plan (Second Edition). Webpage in Chinese. National Health Commission of the People's Republic of China. 2020 Jan 22. URL: <http://www.nhc.gov.cn/jkj/s3577/202001/c67cfe29ecf1470e8c7fc47d3b751e88.shtml> [accessed 2020-02-15]
9. Peiris J. Severe Acute Respiratory Syndrome (SARS). *J Clin Virol* 2003 Dec;28(3):245-247 [FREE Full text] [doi: [10.1016/j.jcv.2003.08.005](https://doi.org/10.1016/j.jcv.2003.08.005)] [Medline: [14522062](https://pubmed.ncbi.nlm.nih.gov/14522062/)]
10. Patel A, Jernigan DB, 2019-nCoV CDC Response Team. Initial Public Health Response and Interim Clinical Guidance for the 2019 Novel Coronavirus Outbreak - United States, December 31, 2019-February 4, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Feb 07;69(5):140-146 [FREE Full text] [doi: [10.15585/mmwr.mm6905e1](https://doi.org/10.15585/mmwr.mm6905e1)] [Medline: [32027631](https://pubmed.ncbi.nlm.nih.gov/32027631/)]
11. Novel coronavirus diagnostic and therapeutic plan (Seventh Edition). Webpage in Chinese. National Health Commission of the People's Republic of China. URL: <http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989/files/ce3e6945832a438eaae415350a8ce964.pdf> [accessed 2020-06-11]
12. Specific primers and probes for detection 2019 novel coronavirus. Webpage in Chinese. Chinese National Institute for Viral Disease Control and Prevention. 2020 Jan 21. URL: [http://ivdc.chinacdc.cn/kyjz/202001/t20200121\\_211337.html](http://ivdc.chinacdc.cn/kyjz/202001/t20200121_211337.html) [accessed 2020-06-18]
13. Sengur A. An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases. *Comput Biol Med* 2008 Mar;38(3):329-338. [doi: [10.1016/j.combiomed.2007.11.004](https://doi.org/10.1016/j.combiomed.2007.11.004)] [Medline: [18177849](https://pubmed.ncbi.nlm.nih.gov/18177849/)]
14. Korean Society of Infectious Diseases, Korean Society of Pediatric Infectious Diseases, Korean Society of Epidemiology, Korean Society for Antimicrobial Therapy, Korean Society for Healthcare-associated Infection Control and Prevention, Korea Centers for Disease Control and Prevention. Report on the Epidemiological Features of Coronavirus Disease 2019 (COVID-19) Outbreak in the Republic of Korea from January 19 to March 2, 2020. *J Korean Med Sci* 2020 Mar 16;35(10):e112 [FREE Full text] [doi: [10.3346/jkms.2020.35.e112](https://doi.org/10.3346/jkms.2020.35.e112)] [Medline: [32174069](https://pubmed.ncbi.nlm.nih.gov/32174069/)]
15. Hao S, Zhang S, Lian J, Jin X, Ye C, Cai H, et al. Liver Enzyme Elevation in Coronavirus Disease 2019: A Multicenter, Retrospective, Cross-Sectional Study. *Am J Gastroenterol* 2020 Jun 01:preprint. [doi: [10.14309/ajg.0000000000000717](https://doi.org/10.14309/ajg.0000000000000717)] [Medline: [32483005](https://pubmed.ncbi.nlm.nih.gov/32483005/)]
16. Cariou B, Hadjadj S, Wargny M, Pichelin M, Al-Salameh A, Allix I, CORONADO investigators. Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: the CORONADO study. *Diabetologia* 2020 May 29:preprint [FREE Full text] [doi: [10.1007/s00125-020-05180-x](https://doi.org/10.1007/s00125-020-05180-x)] [Medline: [32472191](https://pubmed.ncbi.nlm.nih.gov/32472191/)]
17. 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 Jun;26(6):767-772 [FREE Full text] [doi: [10.1016/j.cmi.2020.04.012](https://doi.org/10.1016/j.cmi.2020.04.012)] [Medline: [32304745](https://pubmed.ncbi.nlm.nih.gov/32304745/)]

## Abbreviations

**COVID-19:** coronavirus disease

**C<sub>t</sub>:** cycle threshold

**CT:** computed tomography

**CVD:** cardiovascular disease

**Henan CD:** Center for Disease Control and Prevention of Henan Province

**IQR:** interquartile range

**RT-PCR:** real time reverse transcription–polymerase chain reaction

**SARS:** severe acute respiratory syndrome

**SARS-CoV-2:** severe acute respiratory syndrome coronavirus 2

**WHO:** World Health Organization

*Edited by G Eysenbach; submitted 26.04.20; peer-reviewed by P Banik, E van Gorp, A Bhagavathula; comments to author 06.06.20; revised version received 14.06.20; accepted 15.06.20; published 30.06.20*

*Please cite as:*

*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

URL: <https://www.jmir.org/2020/6/e19636>

doi: [10.2196/19636](https://doi.org/10.2196/19636)

PMID:

©Jitian Li, Zhe Chen, Yifei Nie, Yan Ma, Qiaoyun Guo, Xiaofeng Dai. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 30.06.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.