



# Clinical characteristics of Coronavirus Disease 2019 and development of a prediction model for prolonged hospital length of stay

Yucai Hong<sup>1#</sup>, Xinhui Wu<sup>2#</sup>, Jijing Qu<sup>3</sup>, Yuandi Gao<sup>2</sup>, Hao Chen<sup>2</sup>, Zhongheng Zhang<sup>1</sup>

<sup>1</sup>Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China; <sup>2</sup>Department of Emergency Medicine, <sup>3</sup>Department of Critical Care Medicine, Yueqing People's Hospital, Yueqing 325600, China

*Contributions:* (I) Conception and design: Y Hong, X Wu; (II) Administrative support: J Qu, X Wu; (III) Provision of study materials or patients: Y Gao, H Chen; (IV) Collection and assembly of data: Y Gao, H Chen; (V) Data analysis and interpretation: Z Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

*Correspondence to:* Jijing Qu, MD. Department of Critical Care Medicine, Yueqing People's Hospital, Yueqing 325600, China. Email: qjj201010@163.com; Zhongheng Zhang, MD. No 3, East Qingchun Road, Hangzhou 310016, China. Email: zh\_zhang1984@zju.edu.cn.

**Background:** The epidemic of Coronavirus Disease 2019 (COVID-19) has become a global health emergency, but the clinical characteristics of COVID-19 are not fully described. We aimed to describe the clinical characteristics of COVID-19 outside of Wuhan city; and to develop a multivariate model to predict the risk of prolonged length of stay in hospital (ProLOS).

**Methods:** The study was conducted in a tertiary care hospital in Zhejiang province from January to February 20, 2020. Medical records of all confirmed cases of COVID-19 were retrospectively reviewed. Patients were categorized into the ProLOS and non-ProLOS groups by hospital length of stay greater and less than 14 days, respectively. Conventional descriptive statistics were applied. Multivariate regression model was built to predict the risk of ProLOS, with variables selected using stepwise approach.

**Results:** A total of 75 patients with confirmed COVID-19 were included for quantitative analysis, including 25 (33%) patients in the ProLOS group. ProLOS patients were more likely to have history of traveling to Wuhan (68% *vs.* 28%;  $P=0.002$ ). Patients in the ProLOS group showed lower neutrophil counts [median (IQR):  $2.50 (1.77-3.23) \times 10^9/L$  *vs.*  $2.90 (2.21-4.19) \times 10^9/L$ ;  $P=0.048$ ], higher partial thrombin time (PT) ( $13.42 \pm 0.63$  *vs.*  $13.10 \pm 0.48$  s;  $P=0.029$ ), lower D-Dimer [ $0.26 (0.22-0.46)$  *vs.*  $0.44 (0.32-0.84)$  mg/L;  $P=0.012$ ]. There was no patient died and no severe case in our cohort. The overall LOS was 11 days (IQR, 5–15 days). The median cost for a hospital stay was 7,388.19 RMB (IQR, 5,085.39–11,145.44). The prediction model included five variables of procalcitonin, heart rate, epidemiological history, lymphocyte count and cough. The discrimination of the model was 84.8% (95% CI: 75.3% to 94.4%).

**Conclusions:** Our study described clinical characteristics of COVID-19 outside of Wuhan city and found that the illness was less severe than that in the core epidemic region. A multivariate model was developed to predict ProLOS, which showed good discrimination.

**Keywords:** Novel coronavirus; clinical characteristics; cost; prediction; length of stay (LOS)

Submitted Mar 04, 2020. Accepted for publication Mar 27, 2020.

doi: 10.21037/atm.2020.03.147

View this article at: <http://dx.doi.org/10.21037/atm.2020.03.147>

## Introduction

Coronavirus Disease 2019 (COVID-19) was first found in Wuhan city, Hubei Province, China, which expanded rapidly and caused an epidemic in China and other countries (1-4). The COVID-19 outbreak has been declared as a global health emergency due to the report of thousands of cases and the evidence of human-to-human transmission. The complete clinical course and prognosis with regard to COVID-19 is not fully understood due to limited cases being reported (2,5). Reported illnesses have ranged from mild to severe, including death (5-7). Currently, most studies describing the clinical picture of COVID-19 were from Wuhan city because there were large number of cases in that city. However, the crowdsourcing data from epidemic monitoring suggested some evidence of differences in mortality rate between Wuhan and other cities (8). We proposed that the clinical picture of COVID-19 outside Wuhan had some distinct features that were different from current reports. The present study aimed to describe clinical characteristics of COVID-19 out of Wuhan city.

Furthermore, early prediction of disease course is important for the management of COVID-19. Risk stratification can help to accurately allocate medical resources and inform medical decision making. However, there has been no study focused on the development of a prediction tool for the management of COVID-19. Thus, the study aimed to develop a prediction model for the risk of prolonged length of stay in the hospital. The third novelty in our study was that we estimated financial cost for the hospital stay due to COVID-19.

## Methods

### *Study design and setting*

The study was conducted in a tertiary care hospital in Yueqing city, Zhejiang Province from January to February 20, 2020. Patients were followed until February 20, 2020. Medical records of confirmed COVID-19 cases were retrospectively reviewed and relevant clinical data were extracted. The study was approved by the ethics committee of Yueqing People's Hospital (YQYY20200001). Informed consent was waived due to retrospective nature of the study. The study was conducted in accordance of the Helsinki Declaration.

### *Participants*

A subject was suspected to have COVID-19 if he/she had both clinical features and epidemiological risk meeting the following criteria. Clinical features included fever and signs/symptoms of a lower respiratory illness (e.g., cough or shortness of breath) requiring hospitalization. Epidemiologic risk included: (I) any person, including health care workers, who has had close contact with a laboratory-confirmed 2019 novel coronavirus (2019-nCoV) patient within 14 days of symptom onset; (II) a history of travel from Hubei Province, China within 14 days of symptom onset and (III) a history of travel from mainland China within 14 days of symptom onset. A suspected subject could be confirmed to had COVID-19 if he/she met one of the following criteria: (I) novel coronavirus nucleic acid was positive as confirmed by real time (RT)-PCR in respiratory or blood specimen; and (II) genetic sequencing showed highly homogenous sequence with the known novel coronavirus (9).

All patients with confirmed COVID-19 were included for analysis. Patients would be excluded if they were: (I) younger than 18 years old, (II) pregnant and (III) had severe trauma.

### *Variables*

Demographic data including age, sex, height, weight and career were extracted from medical records. Comorbidities included diabetes mellitus, hypertension, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), cardiovascular disease, liver disease and immunodeficiency. Symptoms were recorded on admission to the hospital, including fever, fatigue, cough, anorexia, dyspnea, diarrhea, nausea, vomiting, headache and abdominal pain. Vital signs were also extracted on admission including blood pressure, heart rate, temperature. The time from symptom onset to hospital admission was recorded in days. Laboratory variables included white blood cell count (WBC), neutrophil count, lymphocyte count, platelet count, coagulation profile, procalcitonin (PCT), C-reactive protein. Blood gas included arterial partial pressure of oxygen ( $\text{PaO}_2$ ), arterial partial pressure of carbon dioxide ( $\text{PaCO}_2$ ) and pH value. Missing data were handled by single imputation with regression models to predict missing values.

Patients were categorized into prolonged length of stay in hospital (ProLOS) and non-ProLOS groups at the LOS of 14 days.

### Statistical methods

Continuous variables were tested for normality. Normal data were expressed as mean and standard deviation (SD) and were compared between ProLOS and non-ProLOS groups with *t*-test. Skewed data were expressed as median and interquartile range (IQR); and were compared between groups using rank sum test. Categorical variables were expressed as the number and percentage, were compared using Chi-square or Fisher's exact test (10).

Variables with P value less than 0.1 were included to build a full model, with the LOS group as the binary

response variable (ProLOS *vs.* non-ProLOS). Then, the stepwise selection and elimination method was applied to the full model. The inclusion or exclusion of a variable was based on the Akaike Information Criterion (AIC) (11). The model performance was explored with the receiver operating characteristic curve and relevant the area under the curve (AUC). Finally, the utility of the prediction model was demonstrated using a nomogram.

All statistical analyses were performed using R (version 3.6.1). A two-tailed P value of less than 0.05 was considered statistically significant.

### Results

A total of 75 patients with confirmed COVID-19 were included for quantitative analysis (*Table 1*). Twenty-five

**Table 1** Baseline clinical characteristics of COVID-2019 patients

Variables	Total (n=75)	Non-ProLOS (n=50)	ProLOS (n=25)	P value
Male, n (%)	41 [55]	30 [60]	11 [44]	0.286
Age (years), mean ± SD	46.37±13.34	47.50±14.21	44.12±11.33	0.269
Height (cm), median [IQR]	162 [154, 171]	165 [160, 174]	156 [154, 165]	0.028
Weight (kg), mean ± SD	64.20±19.15	63.77±18.77	65.06±20.24	0.791
Epidemiological history, n (%)				0.002
Wuhan traveling	31 [41]	14 [28]	17 [68]	
Patient contact	44 [59]	36 [72]	8 [32]	
Career, n (%)				0.059
Unknown	5 [7]	3 [6]	2 [8]	
Self-employed	3 [4]	3 [6]	0 [0]	
Public functionary	1 [1]	1 [2]	0 [0]	
Farmer	2 [3]	2 [4]	0 [0]	
Householder	21 [28]	12 [24]	9 [36]	
Industry worker	11 [15]	3 [6]	8 [32]	
Business	8 [11]	6 [12]	2 [8]	
Freelancer	23 [31]	19 [38]	4 [16]	
Retired	1 [1]	1 [2]	0 [0]	
Hypertension, n [%]	10 [13]	7 [14]	3 [12]	1.000
Diabetes mellitus, n [%]	5 [7]	4 [8]	1 [4]	0.659
Malignancy, n [%]	0 [0]	0 [0]	0 [0]	1.000

**Table 1** (continued)

Table 1 (continued)

Variables	Total (n=75)	Non-ProLOS (n=50)	ProLOS (n=25)	P value
Cardiovascular disease, n [%]	1 [1]	1 [2]	0 [0]	1.000
COPD, n [%]	0 [0]	0 [0]	0 [0]	1.000
CKD, n [%]	0 [0]	0 [0]	0 [0]	1.000
Liver disease, n [%]	1 [1]	1 [2]	0 [0]	1.000
Immunodeficiency, n [%]	0 [0]	0 [0]	0 [0]	1.000
Symptom from onset to admission (days), median [IQR]	4 [2, 7.5]	5 [3, 8.75]	4 [1, 5]	0.044
Fever, n [%]	53 [71]	39 [78]	14 [56]	0.088
Fatigue, n [%]	13 [17]	7 [14]	6 [24]	0.338
Cough, n [%]	49 [65]	38 [76]	11 [44]	0.013
Anorexia, n [%]	38 [51]	32 [64]	6 [24]	0.003
Dyspnea, n [%]	1 [1]	1 [2]	0 [0]	1.000
Diarrhea, n [%]	6 [8]	4 [8]	2 [8]	1.000
Nausea, n [%]	1 [1]	1 [2]	0 [0]	1.000
Vomiting, n [%]	0 [0]	0 [0]	0 [0]	1.000
Dizziness, n [%]	3 [4]	3 [6]	0 [0]	0.546
Headache, n [%]	7 [9]	3 [6]	4 [16]	0.213
Abdominal pain, n [%]	0 [0]	0 [0]	0 [0]	1.000
RR, median [IQR]	20 [20, 20]	20 [20, 20]	20 [20, 20]	0.482
SBP, mean ± SD	126.13±17.31	128.70±17.54	121.00±15.95	0.062
DBP, mean ± SD	81.52±12.01	81.76±11.93	81.04±12.40	0.811
HR, mean ± SD	87.17±13.83	89.22±13.07	83.08±14.66	0.083
Temperature, median (IQR)	37.20 (36.85, 38.05)	37.35 (36.92, 38.10)	37.00 (36.8, 37.90)	0.239

COVID-2019, Coronavirus Disease 2019; ProLOS, prolonged length of stay in hospital; SD, standard deviation; IQR, interquartile range; COPD, chronic obstructive pulmonary disease; CKD, chronic kidney disease; RR, respiratory rate; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, hear rate.

(33%) patients were in the ProLOS group. There was no significant difference between ProLOS and non-ProLOS groups in age, sex, weight and career. ProLOS patients were more likely to have a history of traveling to Wuhan (68% *vs.* 28%;  $P=0.002$ ). There was no significant difference between the two groups in most comorbidities. Patients in the non-ProLOS group were more likely to report cough as initial symptom (76% *vs.* 44%;  $P=0.013$ ), had higher SBP (128.70±17.54 *vs.* 121.00±15.95 mmHg;  $P=0.062$ ).

Patients in the ProLOS group showed lower neutrophil counts [median (IQR): 2.50 (1.77–3.23)  $\times 10^9/L$  *vs.* 2.90 (2.21–4.19)  $\times 10^9/L$ ;  $P=0.048$ ], higher partial thrombin time

(PT) (13.42±0.63 *vs.* 13.10±0.48 s;  $P=0.029$ ), lower D-Dimer [0.26 (0.22–0.46) *vs.* 0.44 (0.32–0.84) mg/L;  $P=0.012$ ]. However, many laboratory tests were within normal range for the majority of patients and the difference within normal range may not be meaningful; thus, we categorized these laboratory values by its normal range cutoff points. The results showed that elevated PCT was associated with higher risk of ProLOS (64% *vs.* 22%;  $P<0.001$ , Table 2).

There was no patient died and no severe case in our cohort (Table 3). The overall LOS was 11 days (IQR, 5–15 days). The ProLOS group showed significantly longer stay in hospital [16 (IQR, 15–17) *vs.* 7 (IQR, 4–11) days;

**Table 2** Laboratory findings of COVID-2019 patients on admission

Variables	Total (n=75)	Non-ProLOS (n=50)	ProLOS (n=25)	P value
WBC ( $\times 10^9/L$ ), median (IQR)	4.64 (3.62, 5.76)	4.75 (3.83, 6.10)	4.32 (3.27, 5.00)	0.096
WBC less than $4 \times 10^9/L$ , n [%]	26 [35]	16 [32]	10 [40]	0.668
Neutrophils ( $\times 10^9/L$ ), median (IQR)	2.65 (2.01, 3.63)	2.90 (2.21, 4.19)	2.50 (1.77, 3.23)	0.048
Neutrophil category, n [%]				0.176
[0, 2]	19 [25]	10 [20]	9 [36]	
[2, 5]	52 [69]	36 [72]	16 [64]	
[5, 9]	4 [5]	4 [8]	0 [0]	
Lymphocyte ( $\times 10^9/L$ ), mean $\pm$ SD	1.33 $\pm$ 0.42	1.29 $\pm$ 0.44	1.40 $\pm$ 0.39	0.311
Lymphocyte less than $1 \times 10^9/L$ , n [%]	17 [23]	13 [26]	4 [16]	0.495
Monocyte ( $\times 10^9/L$ ), median (IQR)	0.46 (0.36, 0.57)	0.47 (0.38, 0.58)	0.42 (0.34, 0.56)	0.451
Platelet ( $\times 10^9/L$ ), median (IQR)	188.00 (158.50, 222.00)	179.50 (154.25, 221.00)	192.00 (168.00, 222.00)	0.500
PT (s), mean $\pm$ SD	13.21 $\pm$ 0.55	13.10 $\pm$ 0.48	13.42 $\pm$ 0.63	0.029
APTT (s), mean $\pm$ SD	39.30 $\pm$ 5.39	40.02 $\pm$ 5.69	37.86 $\pm$ 4.51	0.078
D-Dimer (mg/L), median (IQR)	0.36 (0.26, 0.78)	0.44 (0.32, 0.84)	0.26 (0.22, 0.46)	0.012
CK (U/L), median (IQR)	63.00 (50.00, 112.50)	76.50 (55.25, 118.25)	51.00 (45.00, 107.00)	0.026
Elevated CK, n [%]	9 [12]	6 [12]	3 [12]	1.000
CKMB (U/L), median (IQR)	17.00 (13.50, 21.00)	18.00 (15.25, 24.75)	14.00 (10.00, 17.00)	<0.001
Elevated CKMB, n [%]	12 [16]	12 [24]	0 [0]	0.006
LDH (U/L), median (IQR)	230.00 (190.00, 298.00)	262.00 (207.25, 315.75)	203.00 (181.00, 244.00)	<0.001
AST (U/L), median (IQR)	22.00 (19.00, 32.00)	26.00 (20.25, 36.75)	20.00 (16.00, 21.00)	<0.001
Elevated AST, n [%]	11 [15]	11 [22]	0 [0]	0.012
ALT (U/L), mean $\pm$ SD	16.65 $\pm$ 9.26	18.82 $\pm$ 9.20	12.32 $\pm$ 7.87	0.002
TB (mmol/L), median (IQR)	10.00 (6.70, 14.35)	10.35 (7.25, 14.55)	7.80 (5.70, 12.80)	0.069
Creatinine (mmol/L), mean $\pm$ SD	63.77 $\pm$ 15.80	64.02 $\pm$ 14.84	63.28 $\pm$ 17.88	0.860
PCT (ng/mL), median (IQR)	0.25 (0.25, 0.32)	0.25 (0.25, 0.25)	0.29 (0.25, 0.36)	0.003
Elevated PCT, n [%]	27 [36]	11 [22]	16 [64]	<0.001
PaO <sub>2</sub> , mean $\pm$ SD	99.95 $\pm$ 26.53	99.53 $\pm$ 25.39	100.80 $\pm$ 29.20	0.854
FiO <sub>2</sub> , median [IQR]	33 [33, 33]	33 [33, 33]	33 [29, 33]	0.055
PaCO <sub>2</sub> , mean $\pm$ SD	35.80 $\pm$ 4.83	34.91 $\pm$ 4.01	37.57 $\pm$ 5.85	0.049
HCO <sub>3</sub> , mean $\pm$ SD	23.47 $\pm$ 2.56	23.43 $\pm$ 2.39	23.54 $\pm$ 2.93	0.862
BE, median (IQR)	0 (-2, 2)	0 (-1.75, 2)	0 (-2, 1)	0.409
pH, median (IQR)	7.43 (7.41, 7.46)	7.44 (7.41, 7.47)	7.41 (7.4, 7.43)	0.016

COVID-2019, Coronavirus Disease 2019; ProLOS, prolonged length of stay in hospital; WBC, white blood count; IQR, interquartile range; SD, standard deviation; PT, partial thrombin time; APTT, activated partial thrombin time; CK, creatine kinase; CKMB, creatine kinase myocardial band; LDH, lactate dehydrogenase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; TB, total bilirubin; PCT, procalcitonin; BE, base excess.

**Table 3** Treatment and clinical outcomes

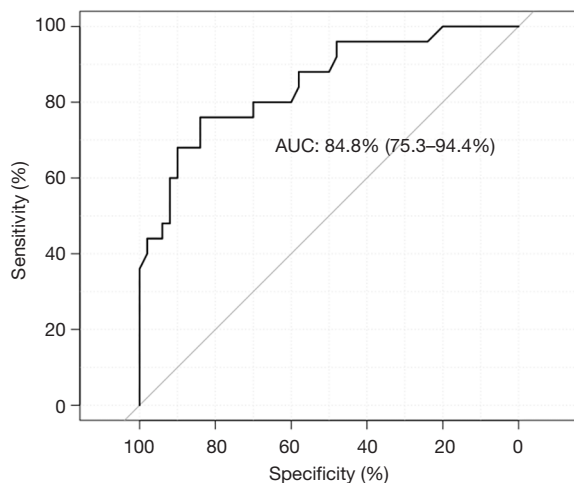
Variables	Total (n=75)	Non-ProLOS (n=50)	ProLOS (n=25)	P value
LOS in hospital (days), median [IQR]	11 [5, 15]	7 [4, 11]	16 [15, 17]	<0.001
Cost (RMB), median (IQR)	7,388.19 (5,085.39, 11,145.44)	5,855.64 (4,758.35, 8,885.84)	10,250.85 (9,197.60, 25,316.12)	<0.001
Antiviral, n [%]	75 [100]	50 [100]	25 [100]	1.000
Corticosteroids, n [%]	17 [23]	8 [16]	9 [36]	0.097
Antibiotic, n [%]	75 [100]	50 [100]	25 [100]	1.000

ProLOS, prolonged length of stay in hospital; LOS, length of stay; IQR, interquartile range.

**Table 4** Prediction model developed by stepwise selection and elimination procedure

Variable	OR (95% CI)	P value
Epidemiological history (patient contact versus Wuhan traveling)	0.21 (0.05, 0.71)	0.015
Cough	0.36 (0.09, 1.32)	0.125
HR (with each 1 beat increase)	0.95 (0.90, 1.00)	0.053
Lymphocyte less than $1 \times 10^9/L$	0.14 (0.02, 0.74)	0.033
Elevated PCT	8.23 (2.14, 38.16)	0.004

PCT, procalcitonin; OR, odds ratio; HR, hear rate; CI, confidence interval.



**Figure 1** Discrimination of the prediction model. The area under curve (AUC) was 84.8% (95% CI: 75.3% to 94.4%).

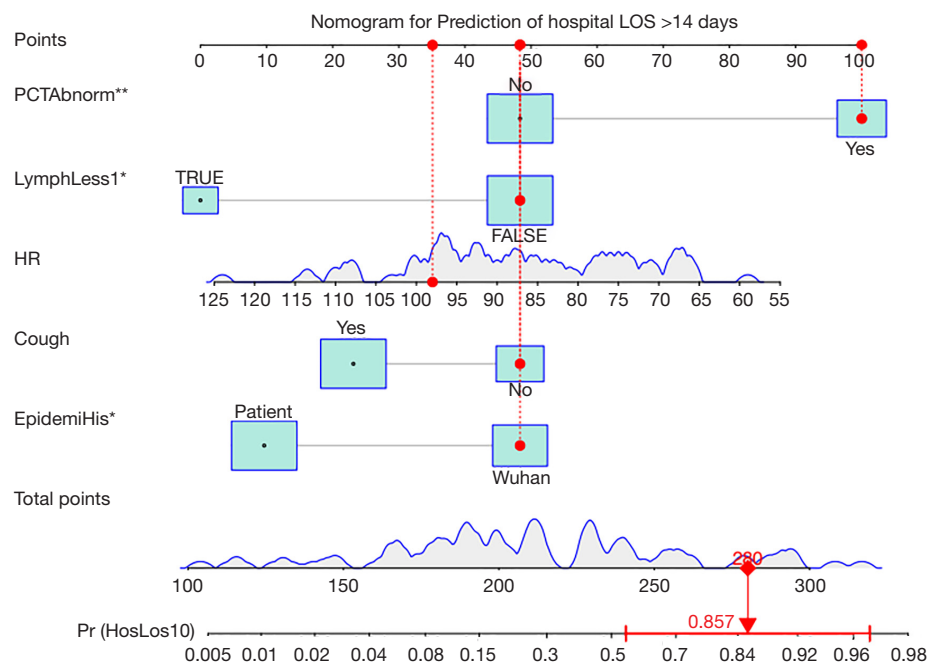
$P < 0.001$ ]. The median cost for a hospital stay was 7,388.19 RMB (IQR, 5,085.39–11,145.44 RMB). ProLOS group had significantly greater cost than the non-ProLOS group [10,250.85 (IQR, 9,197.60–25,316.12) *vs.* 5,855.64 (IQR, 4,758.35–8,885.84) RMB;  $P < 0.001$ ]. All patients received

antiviral and antibiotic treatment. Patients who received corticosteroids tended to have longer LOS in hospital (36% *vs.* 16%;  $P = 0.097$ ).

A final model comprising 5 predictors was obtained by iteratively applying the stepwise variable selection procedure. Note that not all variables need to have a P value less than 0.05, but this combination achieved a smallest AIC. Elevated PCT was significantly associated with ProLOS (OR: 8.23; 95% CI: 2.14 to 38.16;  $P = 0.004$ ). Interestingly, lymphocyte count less than  $1 \times 10^9/L$  was associated with non-ProLOS (OR: 0.14; 95% CI: 0.02 to 0.74;  $P = 0.033$ , *Table 4*). The prediction performance of the model was good as represented by an AUC of 84.8% (95% CI: 75.3% to 94.4%, *Figure 1*). The use of the model can be demonstrated with a nomogram (*Figure 2*).

## Discussion

The study described clinical characteristics of COVID-19 patients outside of Wuhan city. There was no severe case or fatality in our cohort. Elevated PCT was a strong and independent predictor of ProLOS. The median cost for a hospital stay was 7,388.19 RMB (IQR, 5,085.39–



**Figure 2** Nomogram demonstrating the use of the prediction model. Each patient can have a point in each of the items and then the points are summed together to get a total point. In the above example, the patient has a total point of 280, which corresponds to the probability of 87.5% to prolonged hospital stay. \*, significance with  $P < 0.05$ ; \*\*, significance with  $P < 0.01$ . PCTAbnorm, abnormal procalcitonin; LymphLess1, lymphocyte count less than  $\times 10^9/L$ ; HR, heart rate; EpidemiHis, epidemiological history; LOS, length of stay in hospital.

11,145.44 RMB). We further developed a prediction model for mild cases to predict ProLOS. The model had good discrimination power as demonstrated by an AUC  $> 80\%$ . The model can be employed for risk stratification of COVID-19 patients, which can assist medical decision making and accurate allocation of medical resources.

It is interesting to note that only 70% patients reported fever as their initial symptoms, which challenges the current diagnostic criteria that clinical features should include fever and signs/symptoms of a lower respiratory illness, otherwise COVID-19 will not be suspected (9). Our result was consistent with a large study by Guan and colleagues that only around 50% had fever on their initial presentation (12). Thus, cautions should be raised in case of atypical presentations, since false negative cases can be dangerous from the perspective of epidemic control (13). Previous studies have consistently reported that comorbidities were strong risk factors for adverse clinical outcomes (6,14,15). However, our study did not find the association of comorbidity and ProLOS. This could be explained by the fact that our patients had little burden of comorbidity (e.g., two patients had chronic organ dysfunctions, and 20 patients had hypertension or diabetes mellitus). Probably, uncomplicated

hypertension and diabetes mellitus did not impact short term clinical outcome. Thus, the study was underpowered to detect a difference in comorbidity distribution between ProLOS and non-ProLOS groups. Our hospitalized COVID-19 patients were significantly less severely ill as compared to those reported in Wuhan city. For instance, Wang and colleagues reported 36 of the 138 included patients required intensive care unit (ICU) admission due to acute respiratory distress syndrome [22 (61.1%)], arrhythmia [16 (44.4%)], and shock [11 (30.6%)] (3). However, this figure was much lower in a multicenter study involving all provinces of the mainland China, in which only 5% patients required ICU admission (12). A study from Zhejiang province showed that only 1 of the 62 included patients required ICU admission (16), which was consistent with our findings.

Medical cost associated with COVID-19 has long been overlooked in previous reports, and our study was the first to estimate the cost for hospitalization due to COVID-19. Such information can be importance for policy making. Another novelty of the study was the development of a prediction model by using variables collected on hospital admission. The model had good discrimination power.

Longer hospital stay was associated with significantly higher medical cost and thus early prediction of ProLOS patient who were at risk of ProLOS was important for early medical decision making and resource allocation.

The treatment of COVID-19 was generally non-specific and supportive. Many clinical trials are still ongoing to investigate efficacy and safety of drugs such as antiviral therapy and corticosteroids (17). Although our study was not designed to investigate the effectiveness of treatment, the descriptive data provided some practice pattern in the treatment of COVID-19. Our data showed that all patients received antibiotics and antiviral therapy; and 17 patients (23%) received systematic corticosteroids. Given that our subjects were all mild, corticosteroids may not be very effective in shortening LOS in the hospital (18,19). Our study showed that greater proportion of subjects used corticosteroids in the ProLOS group, which could be explained by the selection bias (i.e., more severely ill patients were more likely to receive corticosteroids) (20).

Several limitations must be acknowledged. First, the study was performed in single center and the results may not be extrapolated to other regions. Comparing to the data from Wuhan city, we have already identified many differences between different cohorts. Second, although our prediction model showed good discrimination, external validation was not performed (21). However, we tried to prevent overfitting by reducing the number of predictors. Third, the subjects were followed until February 20, 2020, and thus the long-term outcome was unknown. For example, the pulmonary function might be compromised due to COVID-19. Thus, we plan to follow up this cohort for pulmonary function. Finally, the study could only describe the epidemiological data on the treatment of COVID-19, no comparative analysis could be made due to inherent bias from observational studies. For instance, the observed association of more corticosteroid use with longer hospital stay can be confounded by the selection bias.

In conclusion, our study described clinical characteristics of COVID-19 outside of Wuhan city and found that the illness was less severe than that in the core epidemic region. A multivariate model was developed to predict prolonged hospital stay, which showed good discrimination.

## Acknowledgments

*Funding:* None.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm.2020.03.147>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the ethics committee of Yueqing People's Hospital (No. YQYY202000001). Informed consent was waived due to retrospective nature of the study. The study was conducted in accordance of the Helsinki Declaration.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Kickbusch I, Leung G. Response to the emerging novel coronavirus outbreak. *BMJ* 2020;368:m406.
2. Chen L, Liu HG, Liu W, et al. Analysis of clinical features of 29 patients with 2019 novel coronavirus pneumonia. *Zhonghua Jie He He Hu Xi Za Zhi* 2020;43:E005-5.
3. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020. [Epub ahead of print].
4. Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395:514-23.
5. Li J, Li S, Cai Y, et al. Epidemiological and Clinical Characteristics of 17 Hospitalized Patients with 2019 Novel Coronavirus Infections Outside Wuhan, China. *medRxiv* 2020.02.11.20022053. doi: <https://doi.org/10.1101/2020.02.11.20022053>.
6. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.



- Lancet 2020;395:497-506.
7. Ren LL, Wang YM, Wu ZQ, et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. *Chin Med J (Engl)* 2020. [Epub ahead of print].
  8. Sun K, Chen J, Viboud C. Early epidemiological analysis of the 2019-nCoV outbreak based on a crowdsourced data. medRxiv 2020.01.31.20019935. doi: <https://doi.org/10.1101/2020.01.31.20019935>.
  9. Jin Y, Cai L, Cheng Z, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Mil Med Res* 2020;7:4-23.
  10. Zhang Z, Gayle AA, Wang J, et al. Comparing baseline characteristics between groups: an introduction to the CBCgrps package. *Ann Transl Med* 2017;5:484.
  11. Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med* 2016;4:136.
  12. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020. [Epub ahead of print].
  13. Zhou L, Liu HG. Early detection and disease assessment of patients with novel coronavirus pneumonia. *Zhonghua Jie He He Hu Xi Za Zhi* 2020;43:167-70.
  14. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507-13.
  15. Liu J, Liu Y, Xiang P, Pu L, Xiong H, Li C, et al. Neutrophil-to-Lymphocyte Ratio Predicts Severe Illness Patients with 2019 Novel Coronavirus in the Early Stage. medRxiv 2020.02.10.20021584. doi: <https://doi.org/10.1101/2020.02.10.20021584>.
  16. Xu XW, Wu XX, Jiang XG, et al. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series. *BMJ* 2020;368:m606. Erratum in: *BMJ*. 2020 Feb 27;368:m792.
  17. Du B, Qiu HB, Zhan X, et al. Pharmacotherapeutics for the New Coronavirus Pneumonia. *Zhonghua Jie He He Hu Xi Za Zhi* 2020;43:E012. [Epub ahead of print].
  18. Cao B, Gao H, Zhou B, et al. Adjuvant Corticosteroid Treatment in Adults With Influenza A (H7N9) Viral Pneumonia. *Crit Care Med* 2016;44:e318-28.
  19. Chen RC, Tang XP, Tan SY, et al. Treatment of severe acute respiratory syndrome with glucocorticoids: the Guangzhou experience. *Chest* 2006;129:1441-52.
  20. Corraini P, Olsen M, Pedersen L, et al. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clin Epidemiol* 2017;9:331-8. Erratum in: *Clin Epidemiol*. 2018 Mar 01;10 :223. *Clin Epidemiol*. 2019 Apr 09;11:245.
  21. Riley RD, Ensor J, Snell KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016;353:i3140. Erratum in: *BMJ*. 2019 Jun 25;365:l4379.

**Cite this article as:** Hong Y, Wu X, Qu J, Gao Y, Chen H, Zhang Z. Clinical characteristics of Coronavirus Disease 2019 and development of a prediction model for prolonged hospital length of stay. *Ann Transl Med* 2020;8(7):443. doi: [10.21037/atm.2020.03.147](https://doi.org/10.21037/atm.2020.03.147)