# Earlier diagnosis improves COVID-19 prognosis: a nationwide retrospective cohort analysis

Yi-Jun Chen[1#], Wen-Hua Jian[1#], Zhen-Yu Liang[1#], Wei-Jie Guan[2#], Wen-Hua Liang[3#], Ru-Chong Chen[2], Chun-Li Tang[2], Tao Wang[2], Heng-Rui Liang[3], Yi-Min Li[4], Xiao-Qing Liu[4], Ling Sang[4], Lin-Ling Cheng[1], Feng Ye[1], Shi-Yue Li[1], Nuo-Fu Zhang[1], Zhe Zhang[1], Ying Fang[1], Jian-Xing He[3], Nan-Shan Zhong[1], Jin-Ping Zheng[1]

[1]National Clinical Research Center for Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou Institute of Respiratory Health, Guangzhou Medical University, Guangzhou, China; [2]State Key Laboratory of Respiratory Disease & National Clinical Research Center for Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou Medical University, Guangzhou, China; [3]Department of Thoracic Oncology and Surgery, China State Key Laboratory of Respiratory Disease & National Clinical Research Center for Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China; [4]Department of Pulmonary and Critical Care Medicine, China State Key Laboratory of Respiratory Disease & National Clinical Research Center for Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

*Contributions:* (I) Conception and design: YJ Chen, WH Jian, ZY Liang, WJ Guan, WH Liang, JP Zheng; (II) Administrative support: JP Zheng, JX He, NS Zhong; (III) Provision of study materials or patients: RC Chen, CL Tang, T Wang, HR Liang, YM Li, XQ Liu, L Sang, LL Cheng, F Ye, SY Li, NF Zhang, Z Zhang, Y Fang, JP Zheng, JX He, NS Zhong; (IV) Collection and assembly of data: RC Chen, CL Tang, T Wang, HR Liang, YM Li, XQ Liu, L Sang, LL Cheng, F Ye, SY Li, NF Zhang, Z Zhang, Y Fang; (V) Data analysis and interpretation: YJ Chen, WH Jian, ZY Liang, HR Liang, JP Zheng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Jin-Ping Zheng. State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, First Affiliated Hospital of Guangzhou Medical University, 151 Yanjiang Road, Guangzhou 510120, China. Email: jpzhenggy@163.com.

**Background:** Risk of adverse outcomes in COVID-19 patients by stratifying by the time from symptom onset to confirmed diagnosis status is still uncertain.

**Methods:** We included 1,590 hospitalized COVID-19 patients confirmed by real-time RT-PCR assay or high-throughput sequencing of pharyngeal and nasal swab specimens from 575 hospitals across China between 11 December 2019 and 31 January 2020. Times from symptom onset to confirmed diagnosis, from symptom onset to first medical visit and from first medical visit to confirmed diagnosis were described and turned into binary variables by the maximally selected rank statistics method. Then, survival analysis, including a log-rank test, Cox regression, and conditional inference tree (CTREE) was conducted, regarding whether patients progressed to a severe disease level during the observational period (assessed as severe pneumonia according to the Chinese Expert Consensus on Clinical Practice for Emergency Severe Pneumonia, admission to an intensive care unit, administration of invasive ventilation, or death) as the prognosis outcome, the dependent variable. Independent factors included whether the time from symptom onset to confirmed diagnosis was longer than 5 days (the exposure) and other demographic and clinical factors as multivariate adjustments. The clinical characteristics of the patients with different times from symptom onset to confirmed diagnosis were also compared.

**Results:** The medians of the times from symptom onset to confirmed diagnosis, from symptom onset to first medical visit, and from first medical visit to confirmed diagnosis were 6, 3, and 2 days. After adjusting for age, sex, smoking status, and comorbidity status, age [hazard ratio (HR): 1.03; 95% CI: 1.01–1.04], comorbidity (HR: 1.84; 95% CI: 1.23–2.73), and a duration from symptom onset to confirmed diagnosis of >5 days (HR: 1.69; 95% CI: 1.10–2.60) were independent predictors of COVID-19 prognosis, which echoed the CTREE models, with significant nodes such as time from symptom onset to confirmed diagnosis, age,

and comorbidities. Males, older patients with symptoms such as dry cough/productive cough/shortness of breath, and prior COPD were observed more often in the patients who procrastinated before initiating the first medical consultation.

**Conclusions:** A longer time from symptom onset to confirmed diagnosis yielded a worse COVID-19 prognosis.

**Keywords:** COVID-19; early consultation; early diagnosis; time from symptom onset to confirmed diagnosis; survival analysis

## Introduction

COVID-19 has been spreading in many countries, and the World Health Organization announced that this disease will persist for a long time (1). Recent studies have indicated that procrastination before receiving a COVID-19 diagnosis may cause delayed admission and treatment, leading to higher mortality. Several studies have documented that from the perspective of the imaging examination, late-phase patients (imaged 6–14 days after symptom onset) are more likely to have aggravated abnormal lung CT findings and lesions (2,3). Whether earlier diagnosis can improve COVID-19 prognosis is a critical issue to be addressed. To date, the association between COVID-19 prognosis and time from symptom onset to confirmed diagnosis has received scant in-depth analysis and is still ambiguous.

Shortening the time from symptom onset to confirmed diagnosis includes two aspects: the duration from symptom onset to initial medical contact and the duration from initial visit to a confirmed diagnosis. First, we need patients to identify the earliest symptoms and visit the clinic as promptly as possible (4), but this time interval varies in persons at present (5,6). In addition, hospitals must improve diagnostic efficiency, that is, the time to confirm a COVID-19 diagnosis. Notably, the number of reverse-transcription polymerase chain reaction (RT-PCR) tests among patients can be different; for instance, some patients must undergo ≥2 tests to confirm a diagnosis, and although most are diagnosed in 2 days, up to 7 days may elapse before a confirmed diagnosis (7).

There is no clear consensus among nations on attention to early diagnosis and treatment. The urgency level of shortening the time from symptom onset to confirmed diagnosis varies in countries; that is, some countries may greatly value improving the efficiency of diagnosis, while others may not, and some may change their guidelines and attach importance to it though underrating it at first. Japan, Thailand, Italy, England, Russia, New Zealand, etc., have been gradually liberalizing the conditions for applying PCR tests, such as lessening the symptom requirements for testing or offering free testing (8-13). For instance, given Japan's previous high threshold for COVID-19 testing and some patients' delayed conditions, on 6 May 2020, Japan's Ministry of Health, Labour, and Welfare lowered the threshold for testing, and the new standard did not mention specific fever degrees. However, some countries hold a different opinion; for example, Sweden introduced a new strategy in May 2020 to stop counting confirmed cases and to stop testing people not in hospitals or those not representing high-prevalence populations (14).

We hypothesized that earlier diagnosis might produce better prognosis and that a delayed diagnosis would be associated with poor prognosis as an independent factor. To test this hypothesis, we collected data for COVID-19 patients in China nationwide and analyzed the distribution and cutoffs of time from symptom onset to confirmed diagnosis, including from symptom onset to first medical visit and from first medical visit to confirmed diagnosis. Survival analysis and evaluation models as well as comparisons between different diagnosis efficiency groups were conducted to remedy the neglected aspects of other studies, better notarize the role of time from symptom onset to confirmed diagnosis confirmed and determine who is truly in need of earlier consultation and diagnosis to avoid overwhelming the health services system.

We present the following article in accordance with the Materials Design Analysis Reporting (MDAR) checklist and the STROBE reporting checklist (available at http://dx.doi.org/10.21037/atm-20-7210).

## Methods

### Design, data sources, and data extraction

This was a multicenter retrospective observational cohort study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of the First Affiliated Hospital of Guangzhou Medical University (No.: 2020-92). As this is a retrospective observational study, informed consent was waived. Ratified by the National Health Commission of the People's Republic of China, a national cohort of COVID-19 patients at 575 hospitals across 31 provinces/autonomous regions/provincial municipalities in mainland China was retrospectively established, covering 31.7% of accredited hospitals admitting COVID-19 patients (15,16). According to the WHO interim guidance, cases were confirmed by real-time RT-PCR assay or high-throughput sequencing of pharyngeal and nasal swab specimens. The cohort comprised 1,590 patients hospitalized with COVID-19 from December 11th to Jan 31st and accounted for 13.5% of the total cases at that time. The clinical data, examined and extracted as a computerized database by experienced respiratory clinicians, were verified by double entry prior to analysis.

### Data collection

Demographic, clinical, and prognostic characteristics (i.e., sex, age, smoking status, primary preexisting chronic diseases, symptoms) and the first inspection and thoracic image results on admission were all collected. For chronic comorbidities, we recorded chronic obstructive pulmonary disease (COPD), diabetes, hypertension, coronary heart disease, cerebrovascular disease, hepatitis B infection, malignancy, chronic renal diseases, and immunodeficiency.

The dates of symptom onset, first medical visit, and confirmed diagnosis were also recorded. We calculated the time from symptom onset to confirmed diagnosis, time from symptom onset to first medical visit, and time from first medical visit to confirmed diagnosis; any values less than 1.0 day were deleted from our analysis. A longer time from symptom onset to confirmed diagnosis was deemed exposure in the cohort.

The prognosis outcome was reaching severe disease level, consisting of patients assessed as having severe pneumonia according to the Chinese Expert Consensus on Clinical Practice for Emergency Severe Pneumonia (see Online Supplement for details), those admitted to an intensive care unit (ICU), those who received invasive ventilation, and those who died; the first date of reaching one of these occurrences was used to determine the duration between admission and reaching a severe disease level.

The prognosis outcome and duration between admission and reaching a severe disease level were recorded. Patients with durations between admission and reaching a severe disease level of more than 31 days (more than 95% of patients with a complete end point had a duration of less than 31 days) were treated as right-censored data in prognosis analysis.

### Statistical analysis

We conducted a summary of demographic information on the total sample, and the data were compared between severe and nonsevere cases. Descriptive analysis for the time from symptom onset to confirmed diagnosis (including from symptom onset to first medical visit and from first medical visit to confirmed diagnosis) and survival time (duration between admission and reaching a severe disease level) was performed, and a frequency density map was drawn. A demographic and clinical data comparison between the high and low diagnosis efficiency groups, depending on the cutoff value of the time from symptom onset to confirmed diagnosis, was also processed. Continuous data are presented as the means and ranges. The Wilcoxon rank-sum test was used to compare nonparametric values. Categorical data are presented as counts and percentages and were compared using chi-square and Fisher's exact tests. In the case of missing values (considered missing completely at random (MCAR)), no method of replacement was used. A complete case analysis was performed (listwise deletion) (17).

To turn the continuous variables into categorical ones, reducing disproportionate impact from extreme values to stabilize the calculated model ("Days", as a unit of measurement, is too precise for evaluating prognosis results), maximally selected rank statistics (Maxstat) analysis (18) was performed to determine the appropriate cutoff time from symptom onset to confirmed diagnosis. Survival curves from different diagnosis efficiency groups were prepared by using the Kaplan-Meier (KM) method.

Total survival curves of severe-level patients were drawn. The prognostic significance of the prehospitalization factors (time from symptom onset to confirmed diagnosis, sex,

**Table 1** Time from symptom onset to diagnosis confirmed and from admission to reaching sever level

| Time | N | Mean | Standard deviation | Min | 1st Qu. | Median | 3rd Qu. |
|---|---|---|---|---|---|---|---|
| From symptom onset to diagnosis confirmed | 905 | 6.97 | 5.33 | 0 | 3 | 6 | 9 |
| From symptom onset to first medical visit | 1,385 | 3.67 | 4.12 | 0 | 0 | 3 | 6 |
| From first medical visit to diagnosis confirmed | 882 | 3.73 | 5.03 | 0 | 1 | 2 | 5 |
| From admission to reaching sever level | 1,240 | 9.34 | 5.94 | 0 | 7 | 8 | 11 |

In this cohort, 1,590 cases in total, there are 1,501 cases with symptom onset date, 1,512 cases with first medical visit date, 935 cases with diagnosis confirmation date, 1,246 cases with admission date, 234 cases developed to severe level and with the exact date. The missing data in 4 kinds of time length above met the assumption for Missing Completely at Random (MCAR), and no method of replacement was used, according to logit statistics method considering age, pre-existing diseases, and endpoint factors.

age, smoking status, and presence of preexisting chronic diseases) was analyzed via a log-rank test and multivariate Cox proportional hazard regression when the proportional hazard assumption was not violated. The hazard ratio (HR), along with the 95% confidence interval (95% CI), were described.

Conditional inference tree (CTREE) analysis, a machine learning method, was performed as a complementary analysis with significant factors (probability values < 0.1 in the multivariate Cox proportional hazard regression), which often utilizes multiple significance tests on the permutations of the features on the tree nodes and information measures such as the Gini coefficient to partition the predictors most associated with the outcome and split the tree recursively, dividing the patients into subsamples with different severe disease level risks (19).

Additionally, we compared the prognostic characteristics and first inspection results on admission between patients visiting clinics 4 days after symptom onset and the others. We also analyzed how diagnostic efficiency improved with time in China (see Online Supplement for details). Probability values <0.05 were considered statistically significant, and statistical analyses were performed using R software (R version 4.0.0 https://www.r-project.org/).

## Results

### Demographic and clinical characteristics

By 31 January 2020, we had included 1,590 patients; 237 of these patients reached a severe disease level, accounting for 14.9%, of which 50 (3.1% of the total patients, 21.0% of the severe disease-level patients) died, 102 (6.4%, 43.0%) were admitted to an ICU, 51 (3.2%, 21.5%) received invasive ventilation, and 185 (11.64%, 78.1%) were evaluated as

having severe disease by the Chinese Expert Consensus on Clinical Practice for Emergency Severe Pneumonia (20). A summary of the demographic information for the total cohort is shown in Table S1, and the data were compared between severe and nonsevere levels.

Of these 1,590 patients, the median age was 48.0 years old, and only 647 patients (40.6%) were females. Except for fever on or after hospitalization (88.0%), the most common symptom was dry cough (70.2%). We identified 399 (25.1%) patients with at least one chronic preexisting disease. More than 85% of the patients had at least one abnormal chest CT or X-ray manifestation. Significantly considerable differences in diagnosis efficiency and outcome-related durations were also observed.

### Diagnosis time-efficiency or outcome-related duration characteristics and cut points

We further analyzed and described the durations of diagnosis-related time efficiency and outcome (*Table 1*), and a frequency density figure was drawn (Figure S1). For these 1,590 patients, the medians of duration between symptom onset and confirmed diagnosis, between symptom onset and first medical visit, between first medical visit and confirmed diagnosis, and between admission and reaching a severe disease level were 6, 3, 2, and 8 days, respectively.

To determine the duration cutoffs, 350 observations were deleted due to missing outcome durations (lack of either admission date or the date when the condition turned to a severe disease level or calculated as negative numbers). In addition, 11 patients' (5%, 11/237) numbers of days between admission and reaching a severe disease level were significant outliers (far more than the third quartile plus 1.5 times of differentials between the first and third

**Table 2** Cut-offs and Log-rank test results of diagnosis time-efficiency

| Durations | N | Cut point | Statistic | Log-rank | |
| --- | --- | --- | --- | --- | --- |
| | | | | Chisq | P value |
| From symptom onset to diagnosis confirmation | 866 | 5 | 4.14 | 21.3 | <0.0001 |
| From symptom onset to first medical visit | 1,113 | 4 | 3.72 | 11.9 | 0.00016 |
| From first medical visit to diagnosis confirmation | 844 | 3 | 2.99 | 12.4 | 0.00043 |

quartile, and the authenticity failed to be traced back) and were excluded. Hence, 1229 patients remained for Maxstat analysis to determine the optimal thresholds.

The cutoffs of durations between symptom onset and confirmed diagnosis, between symptom onset and first medical visit, and between first medical visit and confirmed diagnosis were 5, 4, and 3 days, respectively (*Table 2*), according to the optimal log-rank test P value (*Figures 1,2*).

### *Prognostic analysis*

We conducted further prognostic analysis on the 1,229 patients with complete prognosis result indexes (admission date, endpoint date, endpoint result). An overall survival curve was drawn (Figure S2). Univariate log-rank analysis showed that age, sex, comorbidities, and time from symptom onset to confirmed diagnosis were doubtful influencing factors on COVID-19 prognosis (P<0.05, *Table 3*). After adjusting for age, sex, smoking status, and comorbidity status, multivariate Cox proportional hazard regression analysis showed that age [hazard ratio (HR): 1.03; 95% CI: 1.01–1.04], comorbidity status (HR: 1.84; 95% CI: 1.23–2.73), and duration from symptom onset to confirmed diagnosis (HR: 1.69; 95% CI: 1.10–2.60) were strong independent predictors of COVID-19 severity. When considering the duration from symptom onset to the first medical visit and from the first medical visit to confirmed diagnosis, age (HR: 1.03; 95% CI: 1.01–1.04), comorbidity status (HR: 1.77; 95% CI: 1.18–2.68), and duration from symptom onset to the first medical visit (HR: 1.56; 95% CI: 1.07–2.26) were significantly associated with COVID-19 severity (*Table 3*).

CTREE was used to further analyze the association between prognostic outcome and significant factors in the multivariate Cox regression (P<0.1), determining risk thresholds and relations in differentiating overall survival. The CTREE model with age, chronic comorbidity status, and duration from symptom onset to diagnosis demonstrated that patients with preexisting chronic diseases and aged over 61 years were more likely to reach a severe disease level. Patients without chronic comorbidities could also have a similar risk of reaching severity when diagnosed >5 days after symptom onset and aged over 51 years (*Figure 3A*).

Another CTREE model with age, chronic comorbidity status, duration from symptom onset to a first medical visit, and duration from first medical visit to confirmed diagnosis revealed that patients without chronic comorbidities who received a diagnosis >3 days after the first medical consultation and aged over 55 years had the worst prognosis. Patients with chronic comorbidities visiting the clinic >4 days after symptom onset took second place (*Figure 3B*). Both CTREE prognosis models showed that diagnosis-related time-efficiency durations play an important role in disease progression and that individuals with older age and more comorbidities are more likely to be at risk.

Furthermore, by comparing the prognostic characteristics and first inspection results on admission between patients visiting the clinic 4 days after symptom onset and others, we found that the incidence of acute respiratory distress syndrome (ARDS) was higher in patients visiting clinics more than 4 days after symptom onset. Male sex, older age, dry cough, productive cough, shortness of breath, and COPD were more common in the same group. Lower oxygen saturation and albumin levels as well as higher WBC counts, C-reactive protein levels, glutamic oxaloacetic transaminase levels, lactate dehydrogenase levels, direct bilirubin levels, and D-dimer levels for the first admission test were found in the procrastination group (P<0.05) (*Table 4*).

Additionally, the overall diagnosis efficiency increased over time in China, especially the duration from a first medical visit to confirmed diagnosis, which improved from 16 to 23 January. The duration from symptom onset to the first medical visit was noticeably shortened and became
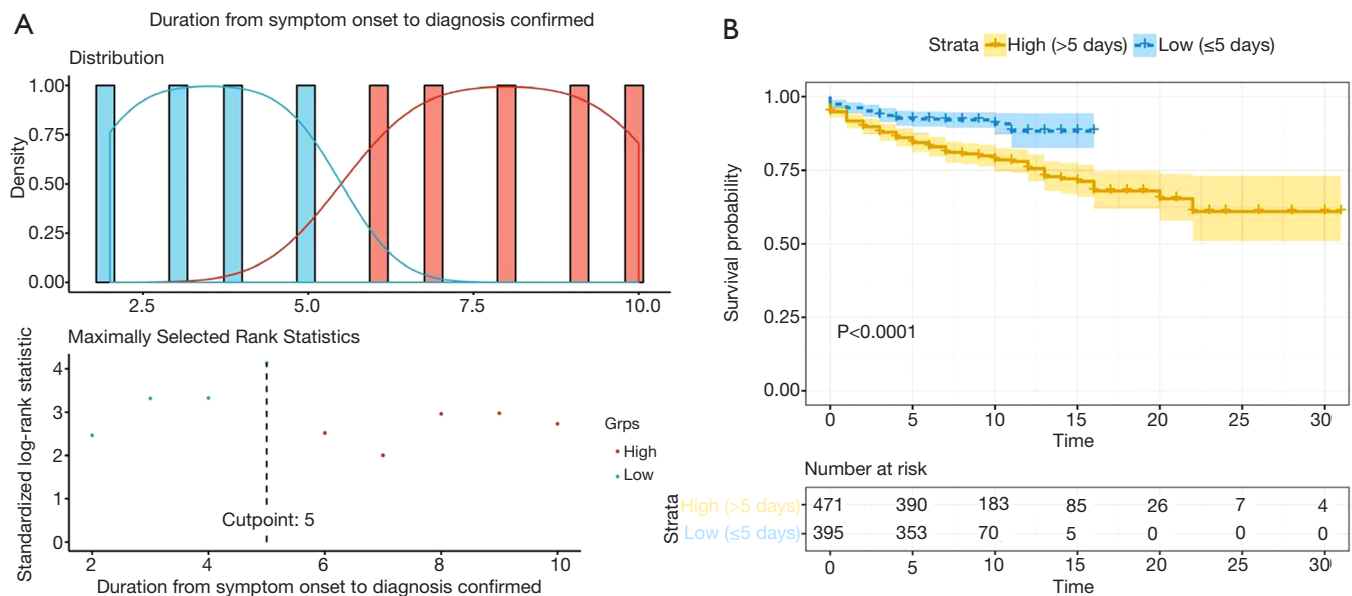
Page 6 of 15

Chen et al. Association between COVID-19 prognosis and earlier diagnosis



**Figure 1** Maximally selected log-rank statistics for the cutoff point of duration from symptom onset to confirmed diagnosis. (A) Patients divided into two groups, high (right red part) and low (left blue part), based on their duration from symptom onset to confirmed diagnosis; the cutoff point (the dotted line, showing the highest point) defined by maximally selected rank statistics. (B) The time-dependent risk of reaching a severe disease level between patients with "high" duration from symptom onset to confirmed diagnosis (yellow curve) and "low" duration from symptom onset to confirmed diagnosis (blue curve); transparent parts indicate the 95% confidence interval (95% CI). Maximally selected rank statistics allow the evaluation of cutoff points, which provide the classification of observations into two groups by a continuous or ordinal predictor variable. The computation of the exact distribution of a maximally selected rank statistic is discussed, and a new lower bound of the distribution is derived based on an extension of an algorithm for the exact distribution of a linear rank statistic.

more consistent from 5 to 20 January (Figure S3).

## Discussion

Our research is the first study to thoroughly investigate the impact of time from symptom onset to confirmed diagnosis on prognosis and clinical characteristics in COVID-19 patients in China nationwide. Together with the time from symptom onset to confirmed diagnosis, prehospital factors among patients with COVID-19, such as comorbidities, age, sex, and smoking status, were also considered in the comprehensive risk assessment and survival models of prognosis. Our findings suggested that the time from symptom onset to confirmed diagnosis could play a crucial role in COVID-19 prognosis with emphasis on older patients and those with chronic comorbidities; these findings are notable and should receive attention from both hospitals and especially patients. COPD patients are more likely to procrastinate before initiating the first medical consultation.

### The commonality with other studies

According to our results, there are advantages to decreasing the total time between symptom onset and confirmed diagnosis to less than 5 days. Patients with chronic comorbidities are recommended to visit the clinic no more than 4 days after symptom onset. Patients over 55 years old should be diagnosed 3 days after medical consultation. Therefore, patients with older age and chronic comorbidities are more sensitive to diagnostic efficiency (21,22).

Our research is consistent with the latest references in the field of COVID-19 progression. Notwithstanding the variations in different studies owing to the sample size and where the patients were treated, it has been reported that COVID-19 symptom aggravation could occur 1–20 days, mainly 7–14 days, after symptom onset (21,23,24). The viral loads in sputum samples and throat swabs peak at approximately 5–10 days after symptom onset, especially 5–7 days (25-27). A high viral load and subsequent viremia may increase the severity of illness (28). Older patients
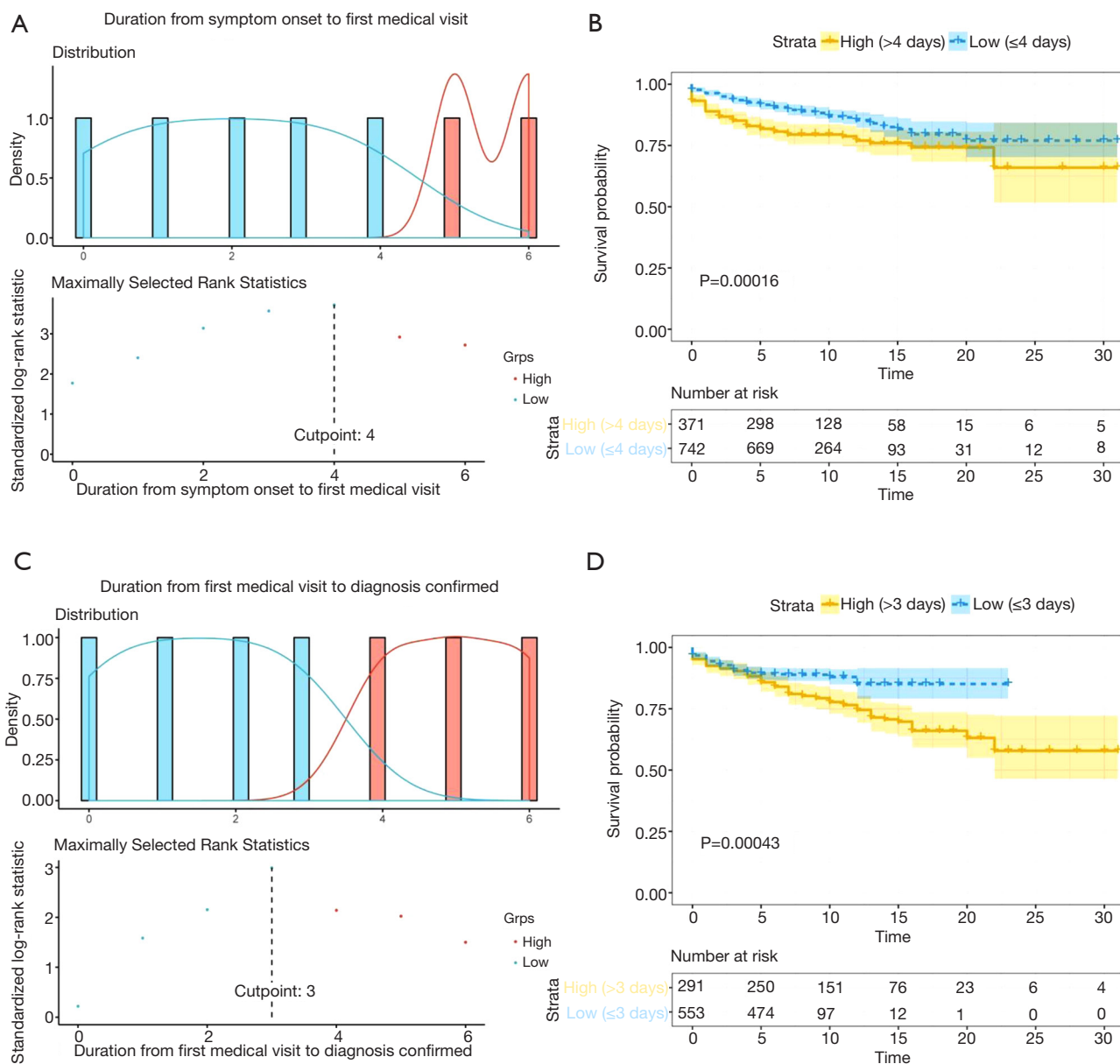
**Figure 2** Maximally selected log-rank statistics for cutoff points of durations from symptom onset to first medical visit and from first medical visit to confirmed diagnosis. (A) Patients divided into two groups, high (right red part) and low (left blue part), based on their duration from symptom onset to first medical consultation; the cutoff point (the dotted line, showing the highest point) defined by maximally selected rank statistics. (B) The time-dependent risk of reaching a severe disease level between patients with "high" duration from symptom onset to first medical consultation (yellow curve) and "low" duration from symptom onset to first medical consultation (blue curve); transparent parts indicate the 95% confidence interval (95% CI). (C) Patients divided into two groups, high (right red part) and low (left blue part), based on their duration from first medical consultation to confirmed diagnosis; the cutoff point (the dotted line, showing the highest point) was defined by maximally selected rank statistics. (D) The time-dependent risk of reaching a severe disease level between patients with "high" duration from first medical consultation to confirmed diagnosis (yellow curve) and "low" duration from symptom onset to confirmed diagnosis (blue curve); transparent parts indicate the 95% confidence interval (95% CI). Maximally selected rank statistics allow the evaluation of cutoff points, which provide the classification of observations into two groups by a continuous or ordinal predictor variable. The computation of the exact distribution of a maximally selected rank statistic is discussed, and a new lower bound of the distribution is derived based on an extension of an algorithm for the exact distribution of a linear rank statistic.

**Table 3** Univariate log-rank analysis and multivariate Cox proportional hazard regression analysis

| | Log-rank test | multivariate COX proportional hazard regression | | | |
|---|---|---|---|---|---|
| | P value | coef | exp(coef) (95% CI) | z | P value |
| **Model 1** | | | | | |
| age | <0.01 | 0.03 | 1.03 (1.01, 1.04) | 3.86 | <0.01*** |
| gender | 0.03 | −0.30 | 0.74 (0.51, 1.08) | −1.56 | 0.12 |
| Smoking status | 0.9 | −0.34 | 0.71 (0.37, 1.35) | −1.05 | 0.30 |
| Comorbidities | <0.01 | 0.61 | 1.84 (1.23, 2.73) | 3.00 | <0.01** |
| Duration from symptom onset to confirmation >5 days | <0.01 | 0.53 | 1.69 (1.10, 2.60) | 2.40 | 0.01* |
| **Model 2** | | | | | |
| age | <0.01 | 0.03 | 1.03 (1.01, 1.04) | 3.91 | <0.01*** |
| gender | 0.03 | −0.29 | 0.75 (0.51, 1.11) | −1.45 | 0.14 |
| Smoking status | 0.90 | −0.47 | 0.63 (0.31, 1.26) | −1.31 | 0.19 |
| Comorbidities | <0.01 | 0.57 | 1.77 (1.18, 2.68) | 2.73 | <0.01** |
| Duration from symptom onset to first visit >4 days | <0.01 | 0.44 | 1.56 (1.07, 2.26) | 2.34 | 0.02* |
| Duration from first visit to confirmation >3 days | <0.01 | 0.37 | 1.45 (0.99, 2.13) | 1.89 | 0.06[#] |

***$P<0.001$; **$P<0.01$; *$P<0.05$; [#]$P<0.1$.

usually have a shorter period from symptom onset to adverse outcomes, ranging from 6 to 41 days and have a higher risk of symptomatic infection (22,29,30).

### Characteristics of patients procrastinating before first medical consultation

Notably, there were significantly more COPD patients in the group with a longer duration from symptom onset to a first medical visit as well as patients with symptoms such as dry cough, productive cough, and shortness of breath, which are often seen in COPD patients. Common symptoms of COVID-19 include fever, dry cough, fatigue, productive cough, and shortness of breath, which are similar to the symptoms of COPD. Consequently, we can infer that COPD patients may have difficulties identifying COVID-19 symptoms, and guidance for them should be made. In addition, several researchers have demonstrated that COPD could be a risk factor for worse prognosis in COVID-19 patients (31,32).

We also found more severe disease, worse prognosis, or ARDS cases in the procrastination group, which echoes

many other types of research. Severe patients always suffer from dyspnea and/or hypoxemia 7 days after symptom onset and then quickly progress to coagulopathy, septic shock, ARDS, irreformable metabolic acidosis and multiple organ failure, especially in older patients (22,29,30).

The procrastination group had a higher incidence of total abnormalities, ground-glass opacities, and pulmonary infiltrates. Many studies on the time course of lung aggregation on chest images echo our results, revealing that abnormalities reach the greatest severity approximately 6–14 days after initial symptom onset (33-37). Patients scanned within 2–4 days after symptom onset show no or fewer abnormalities. Our study also verifies that visiting the clinic later may be associated with worse lung images both on radiographs and computed tomography.

In the first admission test, oxygen saturation and albumin levels were found to be lower in patients who procrastinated before receiving a medical consultation, indicating their worse status. The higher glutamic oxaloacetic transaminase, lactate dehydrogenase, and direct bilirubin levels observed in these patients may be related to the expression patterns of angiotensin-converting enzyme 2 (ACE2). Some studies
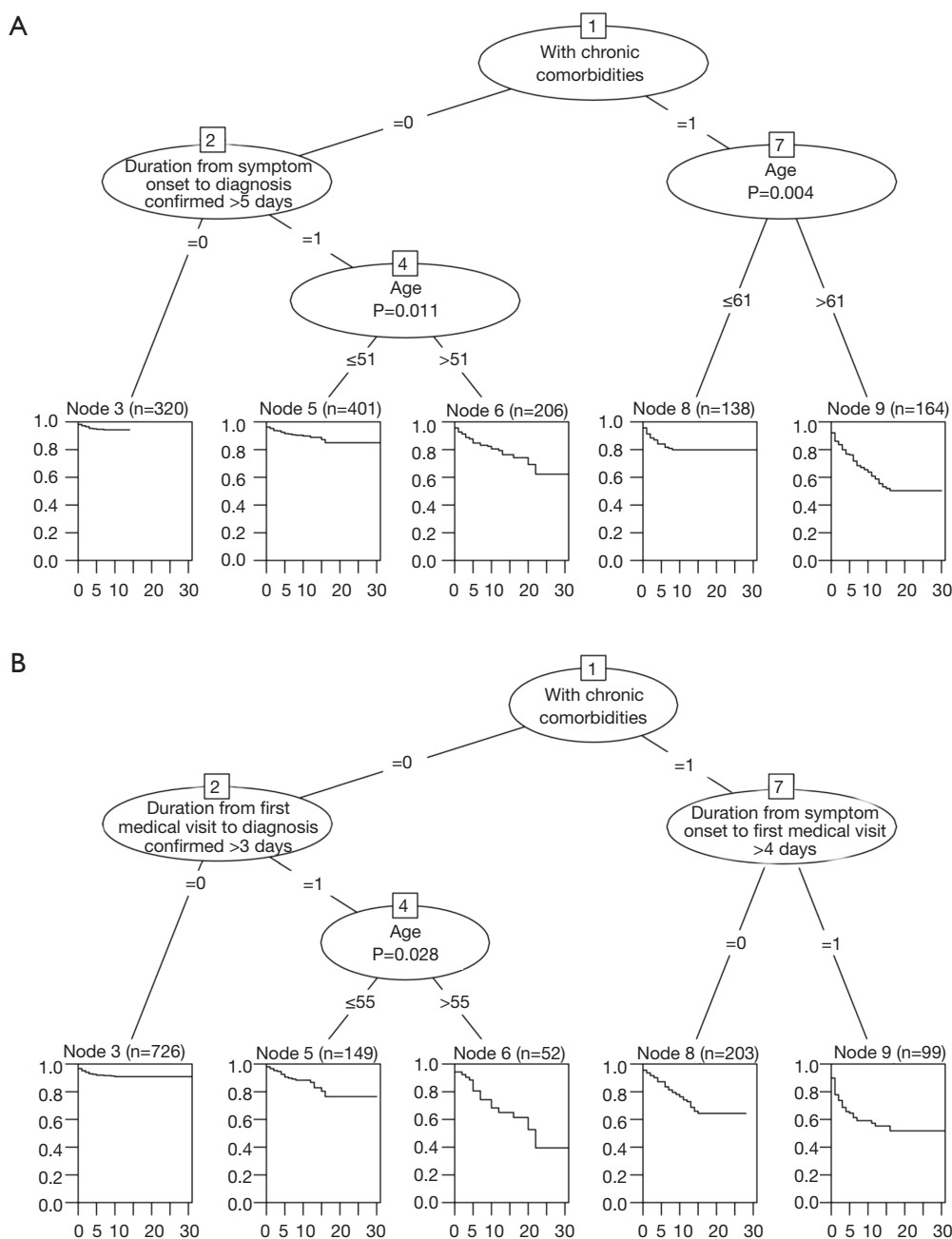
**Figure 3** Conditional inference tree models for COVID-19 prognosis with time from symptom onset to confirmed diagnosis and other prehospital factors. (A) The time-dependent risk of reaching a severe disease level divided into 5 sections according to the significantly separated nodes in the model tree; p values were calculated by the corresponding time series test (log-rank test); the model included duration from symptom onset to confirmed diagnosis, age, and chronic comorbidity status. (B) The time-dependent risk of reaching a severe disease level divided into 5 sections according to the significant separated nodes in the model tree; p values were calculated by corresponding time series test (log-rank test); the model included duration from symptom onset to first medical consultation, duration from first medical consultation to confirmed diagnosis, age, and chronic comorbidity status. The conditional inference tree (CTREE) recursively performs univariate splits of the dependent variable based on values on a set of covariates. CTREE tends to select variables that have many possible splits or many missing values using a significance test procedure to select variables instead of selecting the variable that maximizes an information measure (e.g., Gini coefficient).

*Ann Transl Med* 2021;9(11):941 | http://dx.doi.org/10.21037/atm-20-7210

Page 10 of 15

Chen et al. Association between COVID-19 prognosis and earlier diagnosis

**Table 4** Comparison of prognosis characteristics and first inspection results on admission between patients visiting the clinic in 4 days after symptom onset and the others

| Variables | Total | Time from symptom onset to first medical visit ≤4 days | Time from symptom onset to first medical visit >4 days | P value |
|---|---|---|---|---|
| Outcomes | | | | |
| Sever level | 209 (15.09) | 118 (12.81) | 91 (19.61) | <0.01 |
| septic shock | 24 (3.31) | 13 (2.73) | 11 (4.42) | 0.23 |
| Secondary bacterial or fungal infection | 62 (9.73) | 42 (10.02) | 20 (9.17) | 0.73 |
| ARDS | 72 (9.99) | 38 (8.02) | 34 (13.77) | 0.02 |
| Acute renal failure | 15 (2.10) | 10 (2.13) | 5 (2.05) | 0.95 |
| DIC | 6 (0.84) | 2 (0.43) | 4 (1.63) | 0.09 |
| Rhabdomyolysis | 1 (0.14) | 1 (0.21) | 0 (0.00) | 0.47 |
| Demographic characteristic | | | | |
| Age | 48.0 (36.00–61.00) | 47.0 (35.00–60.00) | 51.0 (40.00–63.00) | <0.01 |
| Gender, male | 795 (57.65) | 514 (55.99) | 281 (60.95) | 0.08 |
| Former/current smoker | 102 (7.36) | 63 (6.84) | 39 (8.41) | 0.29 |
| Symptoms | | | | |
| Dry cough | 972 (72.65) | 617 (69.56) | 355 (78.71) | <0.01 |
| Pharyngodynia | 171 (14.58) | 116 (14.93) | 55 (13.89) | 0.63 |
| Conjunctival congestion | 8 (0.67) | 6 (0.76) | 2 (0.49) | 0.58 |
| Nasal congestion | 65 (5.62) | 38 (4.99) | 27 (6.82) | 0.20 |
| Headache | 190 (16.06) | 134 (17.07) | 56 (14.07) | 0.18 |
| Productive cough | 479 (37.57) | 281 (33.22) | 198 (46.15) | <0.00 |
| Fatigue | 542 (44.54) | 347 (43.27) | 195 (46.99) | 0.22 |
| Hemoptysis | 13 (1.11) | 9 (1.15) | 4 (1.02) | 0.83 |
| Shortness of breath | 303 (21.88) | 158 (17.16) | 145 (31.25) | <0.01 |
| Nausea/vomiting | 75 (6.12) | 50 (6.17) | 25 (6.02) | 0.92 |
| Diarrhea | 55 (4.53) | 34 (4.23) | 21 (5.13) | 0.47 |
| Myalgia/arthralgia | 223 (18.61) | 138 (17.38) | 85 (21.04) | 0.12 |
| Chill | 154 (12.94) | 96 (12.15) | 58 (14.50) | 0.25 |
| Comorbidities | | | | |
| Chronic comorbidities | 357 (25.78) | 236 (25.62) | 121 (26.08) | 0.86 |
| COPD | 21 (1.52) | 8 (0.87) | 13 (2.80) | <0.01 |
| Diabetes | 116 (8.38) | 77 (8.36) | 39 (8.41) | 0.98 |
| Hypertension | 249 (17.98) | 162 (17.59) | 87 (18.75) | 0.60 |
| Coronary heart disease | 49 (3.54) | 35 (3.80) | 14 (3.02) | 0.46 |
| Cerebrovascular disease | 27 (1.95) | 15 (1.63) | 12 (2.59) | 0.22 |

**Table 4** (*continued*)

Table 4 (*continued*)

| Variables | Total | Time from symptom onset to first medical visit ≤4 days | Time from symptom onset to first medical visit >4 days | P value |
|---|---|---|---|---|
| Hepatitis B | 23 (1.66) | 16 (1.74) | 7 (1.51) | 0.75 |
| Malignancy | 13 (0.94) | 9 (0.98) | 4 (0.86) | 0.83 |
| Chronic renal diseases | 17 (1.23) | 15 (1.63) | 2 (0.43) | 0.06 |
| Immunodeficiency | 3 (0.22) | 1 (0.11) | 2 (0.43) | 0.22 |
| Radiological parameters | | | | |
| Abnormality | 1,094 (85.07) | 684 (81.33) | 410 (92.13) | <0.01 |
| Abnormality in X-ray | 219 (63.66) | 128 (57.92) | 91 (73.98) | <0.01 |
| Abnormality in CT | 1,019 (85.41) | 636 (82.28) | 383 (91.19) | <0.01 |
| Having chest X-ray | 1193 (91.56) | 773 (90.73) | 420 (93.13) | 0.14 |
| Having chest CT | 344 (29.50) | 221 (28.63) | 123 (31.22) | 0.36 |
| Ground-glass opacities | 711 (55.29) | 427 (50.77) | 284 (63.82) | <0.01 |
| Local pulmonary infiltrates | 551 (42.85) | 349 (41.50) | 202 (45.39) | 0.18 |
| Pulmonary infiltrates | 686 (53.34) | 409 (48.63) | 277 (62.25) | <0.01 |
| Interstitial disorders | 195 (15.16) | 115 (13.67) | 80 (17.98) | 0.04 |
| Ground-glass opacities in X-ray | 80 (23.26) | 47 (21.27) | 33 (26.83) | 0.24 |
| Local pulmonary infiltrates in X-ray | 108 (31.40) | 71 (32.13) | 37 (30.08) | 0.70 |
| Pulmonary infiltrates in X-ray | 153 (44.48) | 84 (38.01) | 69 (56.10) | <0.01 |
| Interstitial disorders in X-ray | 26 (7.56) | 11 (4.98) | 15 (12.20) | 0.02 |
| Ground-glass opacities in CT | 686 (57.50) | 412 (53.30) | 274 (65.24) | <0.01 |
| Local pulmonary infiltrates in CT | 497 (41.66) | 311 (40.23) | 186 (44.29) | 0.18 |
| Pulmonary infiltrates in CT | 613 (51.38) | 366 (47.35) | 247 (58.81) | <0.01 |
| Interstitial disorders in CT | 184 (15.42) | 111 (14.36) | 73 (17.38) | 0.168 |
| First inspection results on admission | | | | |
| Temperature on admission (℃) | 37.2 (36.70–38.00) | 37.3 (36.70–38.00) | 37.2 (36.70–38.00) | 0.41 |
| $PaO_2$ (mmHg) | 81.34 (62.95–97.75) | 83.0 (65.10–98.00) | 77.0 (58.98–96.00) | 0.08 |
| $FiO_2$ (%) | 21.0 (21.00–29.00) | 21.0 (21.00–29.00) | 21.0 (21.00–29.00) | 0.33 |
| Oxygen saturation under air (%) | 96.0 (94.00–98.00) | 97.0 (95.00–98.00) | 95.0 (93.00–97.85) | <0.01 |
| WBC ($\times 10^9$/L) | 4.92 (3.60–6.43) | 4.84 (3.53–6.27) | 5.21 (3.89–6.62) | 0.01 |
| Lymphocyte ($\times 10^9$/L) | 0.97 (0.70–1.34) | 0.99 (0.70–1.31) | 0.95 (0.70–1.40) | 0.91 |
| Blood platelet ($\times 10^9$/L) | 169.0 (132.00–213.00) | 169.0 (132.00–213.00) | 170.0 (133.00–213.00) | 0.70 |
| Hemoglobin (g/dL) | 132.0 (117.00–145.00) | 132.0 (117.10–146.00) | 132.0 (116.00–143.00) | 0.74 |
| C-reactive protein (mg/L) | 15.63 (6.20–44.10) | 12.88 (4.85–35.03) | 24.75 (10.00–55.41) | <0.01 |
| Procalcitonin (ng/mL) | 0.05 (0.04–0.11) | 0.05 (0.04–0.11) | 0.05 (0.04–0.11) | 0.96 |
| Lactic dehydrogenase (U/L) | 247.0 (191.00–338.20) | 230.0 (180.00–312.00) | 274.0 (224.00–387.25) | <0.01 |

Table 4 (*continued*)

**Table 4** (*continued*)

| Variables | Total | Time from symptom onset to first medical visit ≤4 days | Time from symptom onset to first medical visit >4 days | P value |
|---|---|---|---|---|
| Glutamic oxalacetic transaminase (U/L) | 29.75 (22.00–42.00) | 28.0 (21.91–40.00) | 32.0 (24.00–43.85) | <0.01 |
| Glutamic-pyruvic transaminase (U/L) | 26.0 (17.00–41.00) | 25.0 (16.23–40.00) | 26.0 (18.00–43.00) | 0.09 |
| Direct bilirubin (µmol/L) | 3.3 (2.40–4.80) | 3.2 (2.40–4.80) | 3.5 (2.50–4.90) | 0.05 |
| Indirect bilirubin (µmol/L) | 6.5 (4.51–8.80) | 6.5 (4.50–8.70) | 6.7 (4.60–9.17) | 0.26 |
| Total bilirubin (µmol/l) | 9.98 (7.40–13.40) | 9.8 (7.20–13.10) | 10.2 (7.69–13.62) | 0.10 |
| Creatine kinase (U/L) | 85.2 (54.80–142.70) | 83.0 (56.00–138.75) | 93.0 (52.00–155.00) | 0.43 |
| Creatinine (µmol/L) | 68.0 (54.42–82.00) | 68.0 (54.45–82.25) | 68.0 (54.70–80.40) | 0.89 |
| Hypersensitive troponin I (pg/mL) | 2.2 (0.01–9.45) | 2.0 (0.01–9.96) | 2.75 (0.01–8.72) | 0.78 |
| Albumin (g/L) | 38.8 (33.40–43.30) | 39.5 (33.80–44.10) | 37.3 (32.00–41.40) | <0.01 |
| D-Dimer (mg/L) | 0.53 (0.23–1.42) | 0.49 (0.21–1.06) | 0.59 (0.28–1.72) | <0.01 |
| Prothrombin time (s) | 12.0 (11.00–13.00) | 12.0 (11.00–13.00) | 12.0 (11.00–13.00) | 0.92 |
| Activated partial thromboplastin time (s) | 31.0 (26.00–35.00) | 31.0 (26.00–35.00) | 30.5 (26.00–35.00) | 0.77 |

Data were expressed as means (standard deviation), for parametric continuous data, as median (first quartile; third quartile), for parametric continuous data, or as n (%), where n is the sample number of patients, and % is the proportion with available data. ARDS, acute respiratory distress syndrome; DIC, disseminated intravascular coagulation; COPD, chronic obstructive pulmonary disease.

have revealed that 2019 novel coronavirus (SARS-CoV-2) may directly infect and impair bile duct cells and cause bile duct dysfunction by using ACE2. In contrast, bile duct epithelial cells play a key role in liver regeneration and the immune response (38). Higher D-dimer levels in the procrastination group could also worsen their prognosis, as a novel study confirmed an independent association between thrombosis (D-dimer) and mortality (4). Higher white blood cell (WBC) counts and C-reactive protein (CRP) levels were also discovered, and higher CRP concentrations were associated with COVID-19 severity (39,40). Delayed initiation of supportive care for COVID-19 patients might prolong the illness duration, affecting host immune-inflammatory and thrombotic responses and even clinical outcomes (4).

### *Limitations*

First, recall bias could be a limitation. The date of symptom onset may not be precise, depending on patients' recall after admission for COVID-19. Second, not all patients were followed entirely to the outcome, truncating the correlation with disease course for some patients to some degree.

Third, some patients were excluded from the analysis because of incomplete clinical histories or characteristics. Fourth, our cohort was formed in the very early outbreak phase of the epidemic, and the treatment factors and patient admission diagnosis were not completely collected in their totality. We only took prehospital factors, which are more related to clinical triage, into consideration in our survival analysis, although treatment strategy factors could be crucial for prognosis. In addition, only the presence or absence of chronic comorbidities described above was recorded. Fifth, there are reports of asymptomatic COVID-19 among patients with COVID-19 infection, especially younger patients (41). They may have negative results in the PCR test, chest radiographs, or CT scans. This subset of patients does not fit the models we produced in this research.

### Conclusions

Among laboratory-confirmed cases of COVID-19 nationwide, patients with a longer duration from symptom onset to confirmed diagnosis yielded worse prognosis, especially those who procrastinated before receiving a medical consultation. Time from symptom onset to

confirmed diagnosis shows more sensitivity in older patients, and a delayed diagnosis confirmation (of more than 5 days after symptom onset) could worsen their prognosis. Moreover, a diagnosis within 3 days should be made for older patients over 55 years. Patients with chronic comorbidities should consult doctors 4 days after symptom onset.

Proper triage should be applied by more carefully inquiring about the illness history to identify patients' risk stratification, recognizing patients with a delayed condition who would be more likely to develop an adverse prognosis, namely, older patients with comorbidities. Patients with COPD or any other respiratory disease should be more aware of changes in their symptoms, seek prompt medical attention, and be better guided.

## Footnote

*Reporting Checklist:* The authors have completed the MDAR reporting checklist and the STROBE reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-7210

*Data Sharing Statement:* Available at http://dx.doi.org/10.21037/atm-20-7210

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-7210). Dr. JH serves as an Editors-in-Chief of *Annals of Translational Medicine*. The other authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work and ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of the First Affiliated Hospital of Guangzhou Medical University (No.: 2020-92). As this is a retrospective observational study, informed consent was waived.

## References

1. WHO Timeline - COVID-19 [Internet]. Available online: https://www.who.int/emergencies/diseases/novel-

**Page 14 of 15**

Chen et al. Association between COVID-19 prognosis and earlier diagnosis

coronavirus-2019/situation-reports

2. Bernheim A, Mei X, Huang M, et al. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. Radiology 2020;295:200463.

3. Zhang L, Zhu F, Xie L, et al. Clinical characteristics of COVID-19-infected cancer patients: a retrospective case study in three hospitals within Wuhan, China. Ann Oncol 2020;31:894-901.

4. Cummings MJ, Baldwin MR, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. medRxiv 2020;6736:2020.04.15.20067157.

5. Sarah W, Natasha S, Beth S, et al. Comparisons in early and late presentation to hospital in COVID-19 patients. In: British Thoracic Society Winter Meeting, Programme and Abstracts. 2021.

6. Moser DK, Kimble LP, Alberts MJ, et al. Reducing Delay in Seeking Treatment by Patients With Acute Coronary Syndrome and Stroke A Scientific Statement From the American Heart Association Council on Cardiovascular Nursing and Stroke Council. Circulation 2006;114:168-82.

7. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR Yicheng. Radiology 2020;296:E115-7.

8. Current Situation of the New Coronavirus Infection and the Response of the Ministry of Health, Labour and Welfare (May 7, 2020 edition) [Internet]. Available online: https://www.mhlw.go.jp/stf/newpage_11189.html

9. The COVID-19 situation and measures in Thailand [Internet]. Available online: https://ddc.moph.go.th/viralpneumonia/eng/ind_situation.php

10. Sobyanin told the Muscovites about the free test for antibodies to the coronavirus [Internet]. Available online: https://handofmoscow.com/2020/05/17/sobyanin-told-the-muscovites-about-the-free-test-for-antibodies-to-the-coronavirus/

11. Cuomo Announces Expanded Virus Testing in New York [Internet]. Available online: https://www.lamayor.org/mayor-garcetti-announces-expansion-free-covid-19-testing

12. Mayor Garcetti Announces Expansion of Free COVID-19 Testing [Internet]. Available online: https://www.lamayor.org/mayor-garcetti-announces-expansion-free-covid-19-testing

13. COVID-19 modelling and other commissioned reports [Internet]. Available online: https://www.health.govt.nz/publication/covid-19-modelling-and-other-commissioned-reports

14. Fact check: Has Sweden stopped testing people for the coronavirus? [Internet]. Available online: https://www.thelocal.se/20200320/fact-check-has-sweden-stopped-testing-people-for-the-coronavirus

15. Liang W, Guan W, Chen R, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. Lancet Oncol.2020;21:335-7.

16. Guan WJ, Liang WH, Zhao Y, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. Eur Respir J 2020;55:2000547.

17. Jaber S, Quintard H, Cinotti R, et al. Risk factors and outcomes for airway failure versus non-airway failure in the intensive care unit: A multicenter observational study of 1514 extubation procedures. Crit Care 2018;22:236.

18. Hothorn T, Lausen B. On the exact distribution of maximally selected rank statistics. Comput Stat Data Anal.2003;43:121-37.

19. Sessa M, Rasmussen DB, Jensen MT, et al. Metoprolol Versus Carvedilol in Patients With Heart Failure, Chronic Obstructive Pulmonary Disease, Diabetes Mellitus, and Renal Failure. Am J Cardiol 2020;125:1069-76.

20. Physicians AEPS of C. Chinese Expert Consensus on Clinical Practice for Emergency Severe Pneumonia. 2nd ed. Association Emergency Physician Session of Chinese Physicians; 2016. 97–107 p.

21. Kang S, Peng W, Zhu Y, et al. Recent progress in understanding 2019 novel coronavirus (SARS-CoV-2) associated with human respiratory disease: detection, mechanisms and treatment. Int J Antimicrob Agents 2020;55:105950.

22. Bulut C, Kato Y. Epidemiology of covid-19. Turk J Med Sci 2020;50:563-70.

23. Li T, Lu H, Zhang W. Clinical observation and management of COVID-19 patients. Emerg Microbes Infect 2020;9:687-90.

24. 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.

25. Pan Y, Zhang D, Yang P, et al. Viral load of SARS-CoV-2 in clinical samples. Lancet Infect Dis 2020;20:411-2.

26. Bhat TA, Kalathil SG, Bogner PN, et al. An animal model of inhaled Vitamin E acetate and Evali-like lung injury. N Engl J Med 2020;382:1175-7.

27. Liu Y, Yan LM, Wan L, et al. Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis 2020;20:656-7.

28. Henry BM, de Oliveira MHS, Benoit J, et al.

Gastrointestinal symptoms associated with severity of coronavirus disease 2019 (COVID-19): a pooled analysis. Intern Emerg Med 2020;15:857-9.

29. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. 2020; (January).

30. Wu JT, Leung K, Bushman M, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nat Med 2020;26:506-10.

31. Zhao Q, Meng M, Kumar R, et al. The impact of COPD and smoking history on the severity of Covid-19: A systemic review and meta-analysis. J Med Virol 2020;92:1915-21.

32. Leung JM, Yang CX, Tam A, , et al. ACE-2 expression in the small airway epithelia of smokers and COPD patients: implications for COVID-19. Eur Respir J 2020;55:2000688.

33. Shi H, Han X, Jiang N, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan , China : a descriptive study. Lancet Infect Dis 2020;20:425-34.

34. Pan F, Ye T, Sun P, et al. Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19) Pneumonia. Radiology 2020;2019:200370.

35. Kanne JP, Little BP, Chung JH, et al. Essentials for Radiologists on COVID-19: An Update—Radiology Scientific Expert Pane. Radiology 2020;296:E113-E114.

36. Wong HYF, Lam HYS, Fong AH-T, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients. Radiology 2020;296:E72-E78.

37. Chen J, Qi T, Liu L, et al. Clinical progression of patients with COVID-19 in Shanghai, China. J Infect 2020;80:e1-6.

38. Fan C, Li K, Ding Y, et al. ACE2 Expression in Kidney and Testis May Cause Kidney and Testis Damage After 2019-nCoV Infection. medRxiv. 2020;2020.02.12.20022418.

39. Wang L. C-reactive protein levels in the early stage of COVID-19. Med Mal Infect 2020;50:332-4.

40. Kermali M, Khalsa RK, Pillai K, et al. The role of biomarkers in diagnosis of COVID-19 – A systematic review. Life Sci 2020;254:117788.

41. Kim H. Outbreak of novel coronavirus (COVID-19): What is the role of radiologists? Eur Radiol 2020;30:3266-7.