Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Epidemiological characteristics and factors associated with critical time intervals of COVID-19 in eighteen provinces, China: A retrospective study

Feng Zhou\textsuperscript{a}, Chong You\textsuperscript{b}, Xiaoyu Zhang\textsuperscript{d}, Kaihuan Qian\textsuperscript{a}, Yan Hou\textsuperscript{a}, Yanhui Gao\textsuperscript{c}, Xiao-Hua Zhou\textsuperscript{a,b,c,*}

\textsuperscript{a} Department of Biostatistics, School of Public Health, Peking University, Beijing, China
\textsuperscript{b} Beijing International Center for Mathematical Research, Peking University, Beijing, China
\textsuperscript{c} School of Mathematical Sciences, Peking University, Beijing, China
\textsuperscript{d} School of Medicine, Macau University of Science and Technology, Macau, China
\textsuperscript{*} Department of Epidemiology and Medical Statistics, School of Public Health, Guangdong Pharmaceutical University, Guangzhou, China

\textbf{A R T I C L E  I N F O}

Article history:
Received 9 September 2020
Received in revised form 23 September 2020
Accepted 30 September 2020

Keywords:
COVID-19
Epidemiology
Time interval
Imported cases
Household transmission

\textbf{A B S T R A C T}

\textbf{Background:} As COVID-19 ravages continuously worldwide, more information on the epidemiological characteristics and factors associated with the time intervals between critical events is needed to contain the pandemic and assess the effectiveness of interventions.

\textbf{Methods:} Individual information on confirmed cases from January 21 to March 2 was collected from provincial or municipal health commissions. We identified the difference between imported and local cases in the epidemiological characteristics. Two models were established to estimate the factors associated with the time interval from symptom onset to hospitalization (TOH) and length of hospital stay (LOS), respectively.

\textbf{Results:} Among 7042 cases, 3392 (48.17\%) were local cases, and 3304 (46.92\%) were imported cases. Since the first intervention was adopted in Hubei on January 23, the daily reported imported cases reached a peak on January 28 and gradually decreased since then. Imported cases were on average younger (41 vs. 48) and had more males (58.66\% vs. 47.53\%) than local cases. Furthermore, imported cases had more contacts with other confirmed cases (2.80 \(\pm\) 2.33 vs. 2.17 \(\pm\) 2.10), which were mainly within family members (2.26 \(\pm\) 2.18 vs. 1.57 \(\pm\) 2.06). The TOH and LOS were 2.67 \(\pm\) 3.69 and 18.96 \(\pm\) 7.63 days, respectively, and a longer TOH was observed in elderly living in the provincial capital cities that had a higher migration intensity with Hubei.

\textbf{Conclusions:} Measures to restrict traffic can effectively reduce imported spread. However, household transmission is still not controlled, particularly for the infection of imported cases to elderly women. It is still essential to surveil and educate patients about early admission or isolation.

\textsuperscript{©} 2020 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\textbf{Introduction}

As of September 20, 2020, a total of more than 30 million confirmed cases of coronavirus disease 2019 (COVID-19), as well as more than 900,000 deaths, had been reported by World Health Organization (WHO) worldwide (Organization, 2020a). At the same time, China had reported 85,291 lab-confirmed cases with 4634 deaths (China National Health Commission of the People's Republic 2020a). Despite the WHO and the international community making many efforts to control this pandemic, our knowledge about the COVID-19 is still very limited, and the number of daily reported cases is still increasing sharply worldwide (Organization, 2020b).

In the context of the rapid spread of COVID-19, a full understanding of this infectious disease’s epidemiological characteristics is crucial for epidemic control and public policy practices. Several studies conducted in China, Italy, and the United States have reported some epidemiological characteristics of COVID-19 in the initial phase (Grasselli et al., 2020; Liang et al., 2020; Price-Haywood et al., 2020; Richardson et al., 2020; Wu and McGoogan, 2020). However, there is still a lack of research on the...
space-time characteristics in the populations of imported and local cases, which is of great significance. Imported cases play a critical role in disease spread, mainly because it is an indicator for predicting new clusters of infections. Understanding its epidemiological characteristics would help us to assess the possible effect of non-pharmaceutical interventions (NPIs), such as travel restrictions (Desjardins et al., 2020; Gilbert et al., 2020).

Furthermore, considering the changes in susceptible populations, exposure opportunity, and disease intervention over the epidemic's progresses and its locations, the epidemiological characteristics of disease should be estimated spatiotemporally to better describe the epidemic (Zhang et al., 2020a). For example, the space-time characteristics of COVID-19 revealed by previous studies can prioritize locations and the best time for different NPIs (Desjardins et al., 2020; Lai et al., 2020; Masur et al., 2020). Therefore, exploring the epidemiological characteristics of imported cases from a space-time perspective is critical and provides guidance for countries on interventions taken at different periods and regions, specifically in resource-scarce regions and countries.

As a highly contagious disease, early detection, isolation, hospitalization, and diagnosis of COVID-19 are also crucial for control, and they can effectively reduce the risk of disease transmission (Bi et al., 2020; Rong et al., 2020; Thompson, 2020). Delay in hospitalization or isolation may lead to prolonged periods of infectiousness and increase the difficulty and burden of infectious disease control. Previous studies have described some characteristics of patients with COVID-19, including the time interval between key events (Liang et al., 2020; Tian et al., 2020; Yu et al., 2020). The existing literature also highlights the reduction in the time interval from symptom onset to hospitalization/isolation after various interventions (Li et al., 2020; Zhang et al., 2020b).

However, little is known about individual-level influence factors associated with delaying hospital admission and the length of hospital stay. Identifying these factors would not only help us predict the medical burden and reasonably allocate medical resources but also would inform response efforts across the world.

In this study, we described the spatiotemporal distribution of COVID-19 in eighteen provinces of China (outside Hubei province), and we investigated the epidemiological characteristics of imported cases and local cases from the beginning of this epidemic until it was under reasonable control. We further assessed the critical influence factors associated with the time interval from symptom onset to hospitalization (TOH) and length of hospital stay (LOS), including demographic and temporal and spatial characteristics.

Materials and methods

Data source and study population

We constructed a retrospective cohort study of COVID-19 confirmed cases, based on the detailed information published by the provincial or municipal health commissions in eighteen provinces of China (outside Hubei province) from January 21 to March 2. The details of sampling and data collection are shown in Figure 1. Data collectors were trained and divided into five groups of two, according to provinces, to collect timely epidemiological data of confirmed cases. LinkMed EDC was used for data entry, the two collectors in each group entered the same data, and we conducted data verification and consistency tests in real-time.

Definition of available variables

Specifically, demographic characteristics, epidemiological history, and date of the critical event were extracted from the confirmed case details' official report. (1) Demographic information, including age, gender, residence at the time of diagnosis, and type of symptoms, were included in our analysis. (2) Epidemiological history includes a history of travel or residence in other regions and a contact history of confirmed cases. According to whether the patient had a travel or residence history in other regions within 14 days before diagnosis and likely exposure to pathogens in those regions, patients were divided into imported and local cases. Similarly, we can identify whether patients had contact with confirmed cases of family and non-family members. (3) The dates of events include the date of symptoms onset, hospitalization/isolation, CDC diagnosis, and recovery/death. Hospitalization/isolation is defined as a patient receiving regular hospital treatment (this does not include small medical institutions such as clinics and community health service centers) or a mandatory isolation measure implemented by the community. In this study, we used the time interval between two events to analyze this data, including time interval from symptom onset to hospitalization (TOH) and length of hospital stay (LOS).

![Flowchart of population selection.](image-url)
Additionally, we also collected information on the intensity of migration from Hubei to these 18 provinces in the week before January 23, which was obtained from the Baidu migration map (Baidu, 2020). Migration intensity between provinces and Hubei was categorized into four levels: strong connection (≥0.15%), medium connection (0.05%–0.15%), weak connection (0.03–0.05%) and a very weak connection (<0.03%). Finally, according to the daily trend of new cases and date of intervention, we divided the entire epidemic into five periods from the beginning of the epidemic (January 21) to March 2. Period 1: January 21–23; Period 2: January 24–30; Period 3: January 31–February 6; Period 4: February 7–13; Period 5: February 14–March 2. Period 1: The period, before interventions were taken, is the natural state of the epidemic; Period 2–4: Each week after the interventions were taken is a period. Period 5: Taking into account the gradual decrease in the number of new cases per day after February 14, especially the number of new cases per day after February 22 is very small (below 10 cases). Therefore, the period after February 14 is divided into one period, which indicates a recession of this epidemic.

Statistical analyses

We described the epidemic scale in 18 provinces and the proportion of imported cases spatiotemporally. The demographic characteristics of imported and local cases were reported. Two models were established to identify and quantify the relevant sociodemographic factors, the TOH and LOS, respectively. In the first model, we estimated the TOH factors using a general linear model with a Poisson distribution and a log link. As well, the odds ratio (OR) and their 95% confidence intervals (CI) were calculated after incorporating multiple variables (Coxe et al., 2009; SAS, 2016). In the second model, an accelerated failure time (AFT) model was used to handle the survival data with both left and right censored (Kalbfleisch, 2002; Paul, 2010). In our study of analyzing factors associated with LOS, left censoring would occur if we know that a patient recovered before March 2, but the exact time cannot be obtained. Similarly, right censoring would occur for patients who are diagnosed with COVID-19 in the later phase of the epidemic. Moreover, we included the TOH in the model and used the hazard ratio (HR) and their 95% CIs to identify the difference in LOS among recovered patients with different characteristics. Based on the distribution of LOS, which is denoted by T, we established the Weibull model, written as,

$$\log(T_i) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \sigma e$$

where $e_i$ is a random disturbance term, and $\beta_0, \ldots, \beta_k$, and $\sigma$ are parameters to be estimated. Then we applied a likelihood function with censored data to estimate the parameter values.

$$L = \prod_{i=1}^{n} f(t_i)^{\delta_i} S_i(t_i)^{1-\delta_i}$$

Here, it is the time of the event or the time of censoring, $\delta_i$ is an indicator variable with a value of 1 if $t_i$ is uncensored or 0 if censored. The term $f(t_i)$ denotes the individual's probability density function, $S_i(a) - S_i(b)$ is a survivor function representing the probability of an outcome event time in the interval $(a, b)$.

For left censoring data, $S_i(a) = 1$; while for right censoring data, $S_i(b) = 0$. All statistical analyses were conducted using SAS 9.4 (SAS Institute Inc., North Carolina, USA). $P < 0.05$ was considered statistically significant.

Figure 2. Distribution of the number of COVID-19 cases by region and calendar date from January 21 to March 2 in 18 provinces in China.
Panel A. The map shows the influence of all imported cases in 18 provinces as of March 2, and these provinces have different migration identity from Hubei.
Panel B. The daily number of new cases in 18 provinces divided into imported cases and local cases.
### Table 1
Demographic characteristics in imported and local patients.

| Characteristics | All Patients (n = 7042) | Imported Patients (n = 3392) | Local Patients (n = 3304) | Unknown Patients (n = 346) | Odds Ratio (95% CI)* |
|-----------------|------------------------|-----------------------------|--------------------------|--------------------------|-------------------|
| no./no. of patients with data (%) | | | | | |
| Male | 3593/6763 (53.13) | 1566/2674 (58.56) | 1940/3307 (58.66) | | |
| Age group | | | | | |
| < 19 | 363/6684 (5.43) | 146/2633 (5.55) | 167/3259 (5.12) | | |
| 20 – 29 | 832/6684 (12.45) | 425/2633 (16.14) | 513/3259 (15.74) | | |
| 30 – 39 | 1421/6684 (21.26) | 694/2633 (26.36) | 861/3259 (26.42) | | |
| 40 – 49 | 144/6684 (21.57) | 592/2633 (22.48) | 739/3259 (22.68) | | |
| 50 – 59 | 1399/6684 (20.93) | 471/2633 (17.89) | 598/3259 (18.35) | | |
| >60 | 1227/6684 (18.36) | 305/2633 (11.58) | 381/3259 (11.69) | | |
| Means ± SD | 44.67 ± 10.57 | 41.12 ± 14.87 | 42.61 ± 13.83† | | |
| Median (P25-P75) | 45 (33–56) | 40 (31–52) | 40 (31–52) | | |
| City of residence at the time of diagnosis | | | | | |
| Other city | 5810/7062 (82.27) | 2265/2753 (82.27) | 2752/3392 (81.13) | | |
| Provincial capital | 1252/7062 (17.73) | 488/2753 (17.73) | 2746/3392 (83.11) | | |
| Symptoms | | | | | |
| Fever† | 2938/6673 (79.99) | 1168/2633 (43.19) | 1462/3259 (45.62) | | |
| Cough† | 1293/6684 (19.36) | 505/2633 (19.38) | 645/3259 (19.84) | | |
| Fatigue† | 353/2280 (15.41) | 144/2280 (6.32) | 179/3235 (5.52) | | |
| Cold† | 320/2280 (14.37) | 120/2280 (5.31) | 170/3235 (5.27) | | |
| Sore throat† | 277/2280 (12.10) | 123/2280 (5.42) | 152/3235 (4.68) | | |
| Headache† | 238/2280 (10.39) | 106/2280 (4.67) | 127/3235 (3.94) | | |
| Muscle or joint pain† | 161/2280 (7.03) | 54/2280 (2.38) | 55/3235 (1.70) | | |
| Digestive symptoms† | 156/2280 (6.81) | 50/2280 (2.23) | 67/3235 (2.08) | | |
| Dyspnea† | 67/2280 (2.93) | 19/2280 (0.83) | 10/3235 (0.31) | | |
| Epidemiological contact history† | | | | | |
| unclear | 4242/7062 (60.07) | 2081/2753 (75.59) | 2438/3392 (71.88) | | |
| clear | 2820/7062 (39.93) | 672/2753 (24.41) | 954/3392 (28.13) | | |
| Family-confirmed patient contact history† | | | | | |
| unclear | 4976/7062 (70.46) | 2174/2753 (78.97) | 2596/3392 (76.53) | | |
| clear | 2086/7062 (29.54) | 579/2753 (21.03) | 796/3392 (23.47) | | |
| Non-family-confirmed patient contact history† | | | | | |
| unclear | 6055/7062 (85.74) | 2602/2753 (94.52) | 2539/3392 (76.85)† | | |
| clear | 1007/7062 (14.26) | 151/2753 (5.48) | 126/3392 (3.75) | | |

* no./no. of patients with symptoms and reported.  
† no./no. of patients with symptoms and reported (Except for patients with only fever symptoms).  
We can trace the history of contact with the confirmed patient.  
P = 0.05, Chi-square test;  
P = 0.05, Student t-test;  
P = 0.05, Wilcoxon rank-sum test. These three methods were used to compare the difference between imported and local cases and the difference between imported from Hubei province and other provinces outside Hubei.  
Odds Ratio represents the comparison between imported and local cases after adjusting for region, diagnosis date, and other confounders in this table. CI denotes confidence interval.

### Results

**Spatiotemporal characteristics of the epidemic**

Among 7042 cases, 3392 (48.17%) of patients were local cases, and 3304 (46.92%) of patients were imported cases; fewer than 5% (346) of other patients were unable to confirm their travel history within 14 days before diagnosis. The temporal and spatial distribution of imported and local cases is shown in Figure 2. From panel A, we can see that the greater the intensity of migration to and from Hubei, the more cases in the province. For provinces with a migration intensity greater than 0.03%, the proportion of imported cases to total cases was about 50%. However, for provinces including Tianjin, Ningxia, and Hebei with very weak connections (<0.03%) with Hubei, they had more local cases than imported cases. Since the first intervention was adopted in Hubei on January 23, the daily reported imported cases reached the highest count on January 28, and the proportion of imported cases to the total cases gradually decreased over time, reaching 50% on February 2 (Figure 2).

**Demographic characteristics and the time interval of key events**

Table 1 presents the characteristics of analytically confirmed patients. Among all patients included in the analysis, the mean age was 45 years, and 53.13% were male. Imported cases were on average younger (41 vs. 48), had more males (58.66% vs. 47.53%), had more provincial capital residents (18.87% vs. 16.89%), and had a
lower proportion of clearly confirmed case contact history (28.13% vs. 56.48%), whether a family member confirmed case (23.47% vs. 39.04%) or non-family (7.13% vs. 23.15%), compared to local cases. After adjustment for measured confounders by logistic regression, similar results were still found. The most common symptoms included fever (79.99%) and cough (56.46%). These patterns are similar in both imported cases and local cases. Compared to imported cases from other provinces (outside Hubei), cases imported from Hubei had fewer provincial capital residents (17.73% vs. 23.79%), and a lower proportion of clearly contacted with a confirmed case (24.41% vs. 44.13%). For the time interval, the frequency and best-fitting probability density function for TOH and LOS are presented in Figure 3, respectively. As shown in Table 2 bottom, the mean TOH was 2.67 ± 3.69 days; imported cases had a shorter TOH than local cases (2.48 ± 3.55 vs. 2.89 ± 3.87). In addition, the mean LOS was 18.96 ± 7.63 days, and imported cases had a similar LOS to local cases (19.22 ± 7.84 vs. 18.65 ± 7.32).

Contact history of confirmed cases

The top half of Table 2 shows the contact history of confirmed cases, divided into contact with confirmed cases of family members and non-family members. The overall median number of confirmed cases, the patients who have been contacted, was 2 (interquartile range, 1–3) in this epidemic. Among them, the patient had contact with 1.80 confirmed cases of family members and 0.58 confirmed cases of non-family members on average. Imported cases had contacted more confirmed cases than local cases (2.80 ± 2.33 vs. 2.17 ± 2.10). Furthermore, imported cases had contact with confirmed cases of family members more than local cases (2.26 ± 2.18 vs. 1.57 ± 2.06), while those had contact with confirmed cases of non-family members was equal to local cases (0.53 ± 1.32 vs. 0.60 ± 1.00). In addition, imported cases from Hubei had contacted more confirmed cases than imported cases from other provinces (2.86 ± 2.26 vs. 2.66 ± 2.51), especially in contact with confirmed cases of family members.

Factors associated with the time interval from symptom onset to hospitalization

The left panel of Table 3 shows the results of the first model for the influence factors of TOH. A longer TOH was observed in older and provincial capital cases. The older the case is, the longer the TOH. As compared with the cases younger than 20, especially for cases older than 60 years old, the TOH increased by approximately twice (OR = 1.87; 95% CI: 1.63, 2.13). Patients in the provincial capital had a 1.08-fold longer TOH than patients in other cities. Moreover, significant risk factors for longer TOH also were identified in the middle and later periods of the epidemic (OR = 1.40; 95% CI: 1.21,1.61, and OR = 1.88; 95% CI: 1.50,2.34, respectively). Patients with fever had a shorter TOH than those without fever (OR = 0.84; 95% CI: 0.79,0.89). Patients who clearly contacted family-confirmed cases shortened this time significantly (OR = 0.89; 95% CI: 0.85,0.93). Furthermore, patients who lived in regions with a lower migration intensity with Hubei province had shorter TOH. Notably, for patients living in regions where the migration intensity was more than 0.15, migration intensity (1) between 0.05% and 0.15%, had an 0.87 times decreased risk of a longer time, (2) between 0.03% and 0.05%, had down to 0.74 times, (3) less than 0.03%, had down to 0.69 times. Also, there were no significant differences in TOH between imported and local cases.

Factors associated with length of hospital stay

The right panel of Table 3 gives the HR estimates of related factors associated with LOS. There were no significant differences in LOS among different gender or age groups. It was also shown that differences in LOS relative to city type and fever symptoms were not statistically significant. Patients who clearly contacted with a family-confirmed case had a longer LOS (HR = 1.05; 95% CI: 1.01,1.09) than patients who did not clearly have such contact. Moreover, we found that local patients had a shorter hospital stay than imported cases (HR = 0.95; 95% CI: 0.91,0.99).

Furthermore, patients reported in the later periods of this epidemic had a shorter hospital stay than patients in the initial epidemic (HR = 0.66; 95% CI: 0.57,0.77). Compared with patients whose TOH was less than or equal to one day, the LOS of patients whose TOH was more than four days was reduced by 0.05 percentage. And a similar result appeared in patients whose TOH was 2–3 days (HR = 0.94; 95% CI: 0.89,0.99).

Discussion

Comprehensive epidemiological characteristics of COVID-19 covering the entire period of the epidemic and summaries of China’s experience are useful for public health control. This study described the epidemiological characteristics of imported and local cases, including temporal and spatial characteristics. Indeed, regions with greater migration intensity with Hubei had more imported cases. After the lockdown measures taken by cities in Hubei after January 23, with the goal to interrupt sustained COVID-19 transmission outside Hubei Province (Nie et al., 2020), we found the daily reported imported cases reached a peak on January 28 and gradually decreased after that. This data suggests that traffic restrictions or lockdown in the epicenter can effectively reduce the export of cases (Islam et al., 2020; Zhang et al., 2020b).

Moreover, outside of the epicenter, it is also apparent that timely restriction and quarantine of suspicious imported cases.
Table 2

Description of confirmed case contact history of patients and several time intervals.

| Characteristics | All Patients (n = 7042) | Imported Patients (n = 3392) | From other provinces (n = 1639) | All (n = 3392) | Local Patients (n = 3304) | Unknown Patients (n = 346) |
|-----------------|-------------------------|-----------------------------|--------------------------------|----------------|---------------------------|---------------------------|
| Number of confirmed cases the patient has contacteda | 1                      | 1275/2820(45.21)            | 202/672(30.06)                 | 106/282(37.59)  | 308/954(32.29)        | 967/1866                  | 42/105/40.00             |
|                | 2                      | 759/2820(26.91)             | 223/672(33.18)                 | 79/282(28.01)   | 302/954(31.66)        | 457/1866(24.49)           | 29/105/27.62             |
|                | ≥3                     | 786/2820(27.87)             | 247/672(36.76)                 | 97/282(34.40)   | 344/954(36.06)        | 442/1866                  | 34/105/32.38             |
| Number of confirmed cases of family membersb | 0.25                   | 2.38 ± 0.20                 | 2.86 ± 0.26                   | 2.66 ± 2.51b    | 2.80 ± 0.23           | 2.17 ± 1.20               | 0.49 ± 1.26              |
| Number of confirmed cases of non-family membersc | 0                      | 734/2820(26.03)             | 93/672(13.84)                  | 65/282(23.05)   | 158/954(16.56)        | 576/1866                  | 21/105/20.00             |
| Intervals from symptom onset (days) | To hospitalization, n | 334                        | 1337                          | 370             | 1707                     | 1546                      | 90                      |
| Means ± SD     | 2.67 ± 3.69            | 2.25 ± 3.50                 | 3.33 ± 3.58d                  | 2.48 ± 3.55     | 2.89 ± 3.87d           | 2.57 ± 3.12               |                        |
| Median(P25-P75) | 1(0–4)                 | 1(0–4)                      | 3(0–6)d                       | 1(0–4)         | 2(0–5)d                 | 2(0–4)                   |                        |
| To CDC diagnosis, n | 4697                   | 1869                        | 439                           | 2299           | 2277                     | 121                      |                        |
| Means ± SD     | 5.92 ± 4.16            | 5.39 ± 3.89                 | 6.95 ± 4.31b                  | 5.69 ± 4.02     | 6.15 ± 4.28b           | 5.90 ± 4.08               |                        |
| Median(P25-P75) | 5(3–8)                 | 4(3–7)                      | 7(4–9)c                       | 5(3–9)c        | 5(3–8)c                  | 5(3–8)c                  |                        |
| To recovery, n | 1066                   | 529                         | 114                           | 643            | 390                      | 33                       |                        |
| Means ± SD     | 21.74 ± 8.13           | 21.43 ± 8.21                | 23.73 ± 8.95d                 | 21.84 ± 8.38   | 21.62 ± 7.80            | 21.33 ± 7.01              |                        |
| Median(P25-P75) | 20(16–26)              | 20(16–25)                   | 22(17–25)c                    | 20(16–25)      | 20(16–25)               | 21(18–24)                |                        |
| Intervals from hospitalization (days) | To CDC diagnosis, n | 4667                       | 1910                          | 486            | 2396                     | 2157                      | 114                     |
| Means ± SD     | 3.35 ± 2.98            | 2.24 ± 2.85                 | 3.41 ± 3.09                   | 3.27 ± 2.90     | 3.44 ± 3.06             | 3.32 ± 2.82               |                        |
| Median(P25-P75) | 2(1–4)                 | 2(1–4)                      | 2(1–4)                        | 2(1–4)         | 2(1–5)                   | 2(2–4)                   |                        |
| To recovery, n | 1229                   | 607                         | 127                           | 734            | 463                      | 32                       |                        |
| Means ± SD     | 18.96 ± 7.63           | 19.14 ± 8.77                | 19.61 ± 8.16                  | 19.22 ± 7.84   | 18.65 ± 7.32            | 17.59 ± 7.32              |                        |
| Median(P25-P75) | 17(14–22)              | 18(14–23)                   | 19(14–23)                     | 18(14–23)      | 17(14–22)               | 17(13–21)                |                        |
| Intervals from CDC diagnosis (days) | To recovery, n | 1422                       | 700                           | 139            | 839                      | 545                      | 48                      |
| Means ± SD     | 15.70 ± 7.30           | 15.82 ± 7.53                | 16.53 ± 7.76                  | 15.94 ± 7.57   | 15.39 ± 6.82            | 15.10 ± 7.70              |                        |
| Median(P25-P75) | 14(11–19)              | 14(11–19)                   | 15(11–20)                     | 14(11–19)      | 14(11–18)               | 13(10–19)                |                        |

a no./no. of patients with clear contact history (%).  
b P < 0.05, Chi-square test;  
c P < 0.05, Student t-test;  
d P < 0.05, Wilcoxon rank-sum test. These three methods were used to compare the difference between imported and local cases and the difference between imported from Hubei province and from other provinces outside Hubei.

individuals with a travel history from the epicenter can effectively reduce the local transmission of imported cases (Cui et al., 2020; Kwok et al., 2020; Lai et al., 2020b). Even in the provinces that were not in close contact with Hubei, the surveillance of imported cases could not be overlooked. Taking Tianjin, Ningxia, and Hubei province as examples, local cases were twice as numerous as imported cases, which was related to several local gathering events of imported cases (Chen et al., 2020; Dong et al., 2020; Zhang et al., 2020b).

This study confirms previously described characteristics (Liang et al., 2020; Wu and McGoogan, 2020) and highlights the difference between imported and local cases. Throughout this epidemic, imported patients were younger, had a higher proportion of males, and had more provincial capital residents than local cases. This may match the situation that labor exports are mainly young and middle-aged males in China. This result also implies that older women living in non-provincial capital cities were at greater risk of exposure when the epidemic spreads to the local region. A study on household transmission also found similar results (Xu et al., 2020).

The proportion of patients whose contact history can be can be traced back was higher in local cases, compared with imported cases. This may be due to the complicated epidemic chain in Hubei Province in the initial phase of the epidemic, making it difficult to track the contact history of imported cases. Nonetheless, approximately 40% of local cases may be attributed to household transmission. Among the patients who were clearly exposed to confirmed cases, imported cases had more contacts with other
confirmed cases than local cases on average, and contacts were mainly family members. Although we are unable to determine the infectious relationship between them, it might partly explain why household transmission caused by imported cases was more prominent. This suggests that after NPIs were taken, such as restricting population movement, more effective interventions were still needed to be taken to control household transmission simultaneously, in particular for the infection of imported cases to elderly women in non-provincial capital cities. Indeed, the Chinese government encouraged people to stay at home as much as possible (Lai et al., 2020a). The cases that migrated out from Hubei before January 23 still had the risk of household transmission in the local context, therefore, emergency measures were taken by local governments across China to strengthen the tracking and isolation of recent travelers from Hubei (China National Health Commission of the People’s Republic, 2020b, China The State Council of the People’s Republic, 2020), which reduced this risk to a certain extent. Moreover, our study showed that daily local cases reached a peak on the 14th day (February 6) after the lockdown and then gradually declined. This also illustrates that the government's early response was significant for containing the local spread of imported cases.

Our findings show that there was a lag of 2.67 days from symptom onset to hospital admission, and the average length of hospital stay was about 19 days, which were similar to previous studies conducted in China (Khaliqi et al., 2020; Liang et al., 2020; Linton et al., 2020). Surprisingly, we found that the older the patients are, the longer the hospitalization is delayed. Considering the situation that medical resources outside Hubei Province had not reached saturation, this might be related to the hospitalization days of the patients.
admission pattern of viral respiratory diseases or the lack of recognition of the disease in elderly patients (Petrelli et al., 2020). Besides, the TOH at the later phase of the epidemic showed a rebound trend. Cases reported in the later phase of the epidemic had a careless attitude in seeking medical resources; the decline in control efforts was a possible reason. However, research in China (outside Hubei province) from January 21 to February 17 demonstrated a shorter hospital admission delay from January 28 to February 17 (4.4 vs. 2.6 days) (Zhang et al., 2020a). Before adjusting for other factors, our research also showed slightly shorter hospital admission delays in the week after January 23. Except for the different study population and period, we suggest that this result may be affected by the confound, such as age, gender, the city at the time of diagnosis, etc. Our research included the later phase of the epidemic and adjusted other demographic factors.

This study also confirms that patients living in a provincial capital closely connected to the epicenter had a longer TOH. This provides new demands on epidemic prevention and control, that is, in provincial capital cities close to the epicenter, case tracking, surveillance, and education of immediate admission/isolation should be emphasized. A mathematical model study showed that if the mean time from symptom onset to hospitalization can be halved by surveillance, then the probability that a case leads to transmission is low (Thompson, 2020). Interestingly, we found associations of clear family-confirmed patient contact history with early hospital admission. This finding makes up for the above result, namely, household transmission caused by imported cases still needs attention. Although we cannot cut off the transmission of the virus among family members by restricting population movement or even a city lockdown, it may effectively compensate by reducing the TOH for patients with family-confirmed patient contact history. This reduction may be related to the early detection and isolation of imported cases and their possible close contacts, especially family members (China National Health Commission of the People's Republic, 2020b). Our results also found that the average LOS of 19 days will not decrease by early admission. Perhaps it is related to the characteristics of viral infectious disease. By contrast, the decrease in LOS in the later phase of the epidemic may be due to the continuous improvement of medical technology for this disease.

This study included an extensive study of cases from the beginning of this epidemic until it was under reasonable control and used a novel methodology. However, there are some limitations. First, as a retrospective study, since the date of symptom onset is self-report-based, there may be recall bias. Second, although we made an effort to collect patient discharge information, we still could not obtain some patients’ discharge data. Fortunately, nearly 90% of patients were discharged from the hospital at the end-point of observation on March 2, which provides an opportunity for the statistical methodology using survival data with left censoring. Third, given that the proportion of deaths in the study population was unusually small, less than 1%, the impact of death truncation was not considered when analyzing the length of hospitalization. Finally, our study did not include the southeast provinces. Still, Henan and Zhejiang province were similar to those provinces in the intensity of migration and scale of the epidemic, and our results are also consistent with several studies conducted in Shenzhen and Hong Kong in epidemiological characteristics during the same period (Bi et al., 2020; Lai et al., 2020b).

In conclusion, through retrospective analysis of each phase of this epidemic’s epidemiological characteristics, this study confirms the effectiveness of policies and provides a reference for the parameters of mathematical modeling. At the beginning of the epidemic, measures to restrict traffic and even lock down the city could effectively reduce imported spread. However, household transmission is not yet controlled, particularly for the infection of imported cases being transmitted to elderly women living in non-provincial capital cities. Fortunately, our results identified patients who clearly had contact with a confirmed case of family members and who were admitted to the hospital earlier during the entire epidemic. Even so, surveillance and patients’ education about early admission or isolation should still be given great importance in future prevention and control, especially for the elderly living in provincial capital cities that were more closely connected with the epicenter.

Conflicts of interest
No potential conflict of interest exists in submitting this manuscript, and the manuscript is approved by all authors for publication.

Authors’ contribution
Feng Zhou: Conceptualization, Visualization, Writing - original draft. Chong You: Formal analysis, Writing - review & editing. Xiaoyu Zhang: Validation. Kaihuan Qian: Investigation, Data curation. Yan Hou: Data curation, Supervision. Yanhui Gao: Software, Validation. Xiao-Hua Zhou: Methodology, Funding acquisition, Project administration.

Ethical approval
Not required. The study was anonymous, and individual information was collected from provincial or municipal health commissions, which is public data to help control this epidemic.

Acknowledgments
We thank Xueqing Liu, Yuying Li, Kaihuan Qian, Quishin Lin, Taojun Hu, Meihao Wan, Weijian Ye, Rui Wang, Tianzong Zhang, Qindan Zheng, Mingjia Cai, Niannian Peng, Mengqi Miao, and Qian Yu from Peking University for the assistance of data collection. And thanks to Prof. Yan Hou and her team for their help on the data entry.

This work was supported by the National Natural Science Foundation of China [Grant number 82041023] Zhejiang University special scientific research fund for COVID-19 prevention and control [Grant number 2020GZX016].

References
Baidu Map. Baidu Migration Map. 2020. . . (accessed January 23, 2020) https://qianxi.baidu.com/2020/.
Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect Dis 2020;20(8):911–9. doi:https://dx.doi.org/10.1016/S1473-3099(20)30287-5.
Chen Y, Zhang K, Zhu G, Liu L, Yan X, Cai Z, et al. Clinical characteristics and treatment of critically ill patients with COVID-19 in Hebei. Ann Palliat Med 2020;9(4):2118–30. doi:https://dx.doi.org/10.21037/apm-20-1273.
National Health Commission of the People's Republic of China. Update on Pneumonia of New Coronavirus Infection as of 24:00 on September 20, 2020. . (accessed September 20, 2020) http://www.nhc.gov.cn/jyqfkdt/202009/03382201308b1b4b64ba911314d51621.shtml.
National Health Commission of the People's Republic of China. Prevention and control plan for new coronavirus pneumonia. second edition 2020. . (accessed September 20, 2020) http://www.nhc.gov.cn/jyqfkdt/202009/03382201308b1b4b64ba911314d51621.shtml.
The State Council of the People's Republic of China. The announcement on strengthening community prevention and control of pandemic pneumonia epidemic situation of a new coronavirus infection. 2020. . (accessed September 20, 2020) http://www.gov.cn/zhengce/2020-01/27/content_5472516.htm.
Coxe S, West SG, Aiken LS. The analysis of count data: a gentle introduction to Poisson regression and its alternatives. J Pers Assess 2009;91(2):121–36, doi: http://dx.doi.org/10.1080/00223890802634175.
Cui Q, Hu Z, Li Y, Han J, Teng Z, Qian J. Dynamic variations of the COVID-19 disease at different quarantine strategies in Wuhan and mainland China. J Infect Public Health. 2020;13(6):849–55. doi:http://dx.doi.org/10.1016/j.jiph.2020.05.014.

Dejardins MR, Holt A, Delmelle EM. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. Appl Geogr 2020;118:102202. doi:http://dx.doi.org/10.1016/j.apgeog.2020.102202.

Dong XC, Li JM, Bai YJ, Liu ZQ, Zhou PH, Gao L, et al. [Epidemiological characteristics of confirmed COVID-19 cases in Tianjin]. Zhonghua Liu Xing Bing Xue Za Zhi 2020;41(5):303–11. doi:http://dx.doi.org/10.3760/cma.j.cn112338-20200221-00146.

Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boelle P-Y, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. Lancet 2020;395(10227):871–7. doi:http://dx.doi.org/10.1016/S0140-6736(20)30411-6.

Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabirol I, Castelli A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. JAMA 2020;323(16):1574–81. doi:http://dx.doi.org/10.1001/jama.2020.3593.

Islam N, Sharp SJ, Chowell G, Shabnam S, Kawachi I, Lacey B, et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. BMJ 2020;370:m2743. doi:http://dx.doi.org/10.1136/bmj.m2743.

Kalbfleisch JDA P, RL. The Statistical Analysis of Failure Time Data. second edition. John Wiley & Sons, Inc; 2002.

Khalili M, Karamouzian M, Nasiri N, Javadi S, Mirzadeh A, Sharifi H. Epidemiological characteristics of COVID-19: a systematic review and meta-analysis. Epidemic Infect 2020:148:e130. doi:http://dx.doi.org/10.1016/S0956-2228(20)301430.

Kwok KO, Wong VWY, Wei WI, Wong SY, Tang JW. Epidemiological characteristics of the first 51 laboratory-confirmed cases of COVID-19 epidemic in Hong Kong, February 13, 2020. Euro Surveill 2020;25(20):2000155. doi:http://dx.doi.org/10.2807/1560-7917.ES.2020.25.16.2000155.

Lai KC, Ng RWY, Wong MCS, Chung KC, Yeoh YK, Chan Z, et al. Epidemiological characteristics of the first 100 cases of coronavirus disease 2019 (COVID-19) in Hong Kong Special Administrative Region, China, a city with a stringent containment policy. Int J Epidemiol 2020a; doi:http://dx.doi.org/10.1093/ije/dyaa106 [published online ahead of print].

Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luso W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 2020b;585 (7825):410–3. doi:http://dx.doi.org/10.1038/s41586-020-2293-x.

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 2020;382(13):1199–207. doi:http://dx.doi.org/10.1056/NEJMoa2001316.

Li PH, Guan WX, Li CC, Li YM, Liang HR, Zhao Y, et al. Clinical characteristics and outcomes of hospitalised patients with COVID-19 treated in Hubei (epicentre) and outside Hubei (non-epicentre): a nationwide analysis of China. Eur Respir J 2020;55(6):2000562. doi:http://dx.doi.org/10.1183/13993003.00562-2020.

Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung SM, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. J Clin Med 2020;9(3):538. doi:http://dx.doi.org/10.3390/jcm9030538.

Masur A, Yu M, Luo W, Dewan A. Space-Time patterns, change, and propagation of COVID-19 risk relative to the intervention scenarios in Bangladesh. Int J Environ Res Public Health 2020;17(16); doi:http://dx.doi.org/10.3390/ijerph17165911.

Nie X, Fan L, Mu G, Tan Q, Wang M, Xie Y, et al. Epidemiological characteristics and incubation period of 7015 confirmed cases with coronavirus disease 2019 outside Hubei Province in China. J Infect Dis 2020;222(1):26–33. doi:http://dx.doi.org/10.1093/infdis/ijza211.

World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. 2020 data last updated: 2020/9/20; 12:00pm. https://covid19.who.int/, 2020 (accessed September 20, 2020).

World Health Organization. World Health Organization, WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020. (accessed on 12 March 2020) https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11—March-2020.

Paul D Allison. Survival Analysis Using SAS: A Practical Guide. second edition SAS Publishing; 2010.

Petrelli CM, Jones SA, Yang J, Rajagopalan H, D’Ommern L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ 2020;369:m1966. doi:http://dx.doi.org/10.1136/bmj.m1966.

Price-Haywood EG, Burton J, Fort D, Sceace L. Hospitalization and mortality among black patients and white patients with Coronavirus-19. N Eng J Med 2020;382(26). doi:http://dx.doi.org/10.1056/NEJMc2016866 2534-43.

Richardson S, Hirsch JS, Narasimhan M, Crawford MJ, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5708 patients hospitalized with COVID-19 in the New York City area. JAMA 2020;323(20). doi:http://dx.doi.org/10.1001/jama.2020.6775 2052-9.

Rong YM, Yang L, Chu HD, Fan M. Effect of delay in diagnosis on transmission of COVID-19. Math Biosci Eng 2020;17(3). doi:http://dx.doi.org/10.3934/mbbe.2020149 2725-40.

SAS Publishing. S.A.S./STAT 14.2 User’s Guide: The GLIMMIX Procedure (Chapter). SAS Publishing; 2016.

Thompson RN. Novel Coronavirus Outbreak in Wuhan, China, 2020: Intense surveillance is vital for preventing sustained transmission in new locations. J Clin Med 2020;9(2):498: doi:http://dx.doi.org/10.3390/jcm9020498.

Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, et al. Characteristics of COVID-19 infection in Beijing. J Infect 2020;80(4). doi:http://dx.doi.org/10.1016/j.jinf.2020.02.018 401-6.

Wu Z, McGowan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 7234 cases from the Chinese center for disease control and prevention. JAMA 2020;323 (13):1239–42. doi:http://dx.doi.org/10.1001/jama.2020.2648.

Xu X, Liu X, Wang L, Ali ST, Du Z, Bosetti P, et al. Household transmissions of SARS-CoV-2 in the time of unprecedented travel lockdown in China. medRxiv 2020; doi:http://dx.doi.org/10.1101/2020.03.02.20029868 03.02.20029868 [Preprint].

Yu X, Sun X, Cui P, Pan H, Lin S, Han R, et al. Epidemiological and clinical characteristics of 333 confirmed cases with coronavirus disease 2019 in Shanghai, China. Transbound Emerg Dis 2020;67(4). doi:http://dx.doi.org/10.1111/tbed.13604 1697-707.

Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology of novel coronavirus diseases 2019 and possible interruption of local transmission outside Hubei Province in China: a descriptive and modeling study. medRxiv 2020a;20026328. doi:http://dx.doi.org/10.1101/2020.02.21.20026328 [Preprint].

Zhang SX, Li J, Zhou P, Na JR, Liu BF, Zheng XW, et al. [The analysis of clinical characteristics of 34 novel coronavirus pneumonia cases in Ningxia Hui autonomous region]. Zhonghua Jie He He Hu Xi Za Zhi 2020b;43(5):431–6. doi:http://dx.doi.org/10.3760/cma.j.cn112147-20200219-00121.