

# Basic epidemiological parameter values from data of real-world in mega-cities: the characteristics of COVID-19 in Beijing, China

**Xiaoli Wang**

Beijing Center for Disease Prevention and Control

**Yang Pan**

Beijing Center for Disease Prevention and Control

**Daitao Zhang**

Beijing center for disease prevention and control

**Lijuan Chen**

Beijing Center for Disease Prevention and Control

**Lei Jia**

Beijing Center for Disease prevention and control

**Xinyu Li**

Beijing Center for Disease Prevention and Control

**Peng Yang** (✉ [yangpengcdc@163.com](mailto:yangpengcdc@163.com))

**Quanyi Wang**

Beijing Center for Disease Prevention and Control

**C. Raina Macintyre**

University of New South Wales

---

## Research article

**Keywords:** COVID-19, epidemiology, transmissibility, secondary attack rate

**Posted Date:** July 13th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-23937/v2>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on July 20th, 2020. See the published version at <https://doi.org/10.1186/s12879-020-05251-9>.

# Abstract

**Background.** With the spread of SARS-CoV-2 worldwide, understanding the basic epidemiological parameter values of COVID-19 from real-world data in mega-cities is essential for disease prevention and control.

**Methods.** To investigate the epidemiological parameters in SARS-CoV-2 infected cases in Beijing, we studied all confirmed cases and close contacts in Beijing from Jan 1st to Apr 3rd 2020. The epidemiological and virological characteristics of SARS-CoV-2 were analyzed.

**Results.** A total of 602 cases were positive for SARS-CoV-2, including 585 confirmed patients and 17 asymptomatic infections. The imported cases were mainly from Wuhan initially and then from abroad. Among 585 confirmed case-patients, the median age was 39 years old. The mean incubation period was 6.3 days. The secondary attack rate among households was higher than social contacts (15.6 vs 4.6%). The secondary attack rate of healthcare workers (HCWs) was higher than non-HCWs' (7.3 vs 4.2%). The basic reproduction number was 2.0, and the average serial interval was 7.6 days. No significant genetic variant was identified.

**Conclusions.** The transmissibility of SARS-CoV-2 was relatively high, especially among households and from HCWs, which draws specific public health attention. So far, no evidence of widespread circulation of SARS-CoV-2 in communities in Beijing was found.

## Background

In December 2019, a new beta coronavirus (CoV) was first detected in Wuhan, Hubei Province of China, causing clusters of unknown pneumonia patients. It spread rapidly throughout Hubei and then to other provinces and abroad. The World Health Organization (WHO) declared the outbreak a public health emergency of international concern on January 30th, 2020[1]. The novel coronavirus has been named SARS-CoV-2 by the International Committee on Taxonomy of Viruses and the disease it causes has been named "coronavirus disease 2019" (COVID-19) [2]. As of Apr 12th, 2020, the global number of confirmed cases of COVID-19 has surpassed 1,600,000, including more than 100,000 deaths due to acute respiratory failure or other related complications [3].

Currently, the minority of the population of China has been infected, which leaves the majority of China susceptible to infection. Beijing, the capital of China, which has a population of 21 million residents including 8 millions of migrants at particular risk. Since the first case detected in Beijing on Jan 19th, the number of cases imported from Wuhan increased soon. To prevent further spread, the Chinese government announced the lockdown (forced quarantine or in-and-outflow restriction) of Wuhan City on Jan 23rd before the upcoming spring festival during which massive population movements were expected to take place. Since it is of vital importance to get knowledge of basic epidemiological parameter values from data of real-world in mega-cities, we report the epidemiological and virological

findings of 602 cases with SARS-CoV-2 infection and the outbreak response conducted in Beijing from Jan 1st to Apr 3rd 2020.

## Methods

### Outbreak Response

Protocols for COVID-19 diagnosis and treatment, surveillance, epidemiological investigation, management of close contacts, and laboratory testing were formulated, and relevant surveillance activities and epidemiological investigations conducted. All individuals with suspected SARS-CoV-2 infection were isolated in designated hospitals for treatment. The epidemiological investigation, contact tracing, and quarantine were conducted by the Centers for Disease Prevention and Control (CDCs) at district level or municipal level. Close contacts were defined based on the protocol released by National Health Commission of the People's Republic of China (NHCPRC). Without any efficient personal protective equipment, a person who had contact within one meter with a confirmed case after onset of symptoms was defined as close contact. The definition changed on Feb 10th in accordance with the WHO guidelines [4], from the onset to four days prior to symptoms, given evidence of presymptomatic transmission. Close contacts were required to be quarantined at home or in designated sites for cases who have no separate room or imported cases from overseas for medical observation till 14 days after the last exposure. Besides traditional contact tracing, big data and artificial intelligence (AI) were used to strengthen contact tracing. For enhanced surveillance to detect COVID-19 in influenza-like illness (ILI) cases, all samples from ILI cases reported in routine influenza virological surveillance system [5] were tested for SARS-CoV-2 using real-time reverse transcription-polymerase chain reaction (Real-time RT-PCR) since January 28th.

### Case definitions

The case definitions of suspected and confirmed cases of COVID-19 were based on the protocol released by the NHCPRC [6]. Some modifications in case definitions were made according to recent studies in China. The definition modifications of suspected cases and confirmed cases are shown in Supplementary Table 1. Among all confirmed cases, mild cases were defined cases without evidence of pneumonia, and moderate cases as those with pneumonia but no oxygen therapy required. Severe cases were defined as those with dyspnea, respiratory frequency  $\geq 30$ /minute, blood oxygen saturation  $\leq 93\%$ ,  $PaO_2/FiO_2$  ratio  $< 300$ , and/or lung infiltrates  $> 50\%$  within 24-48 hours. Critical cases were those that exhibited respiratory failure, septic shock, and/or multiple organ dysfunction/failure.

Asymptomatic laboratory-confirmed cases were defined a person who were positive for SARS-CoV-2 while showing no clinical symptoms and signs.

### Data collection

The demographic and epidemiological information, clinical symptoms and outcomes were obtained using a standardized questionnaire by interviewing patients and/or their family members/relatives, attending doctors and other health care providers, supplemented by patient medical records [7]. All data were checked by two public health professionals.

### Laboratory testing

Upper and lower respiratory tract specimens, urine, blood, stool, and sputum specimens were obtained from patients. For all RNA extractions, RNA was extracted from 200 µL of sample and eluted in 90 µL elution buffer by KingFisher Flex Purification System (Thermo Fisher, USA). The real-time RT-PCR assay was performed by TaqMan Fast Virus 1-Step Master Mix (ThermoFisher) in ABI 7500 fast system (ABI, USA). 0.5 µM of forward primer, 0.5 µM of reverse primer, 0.25 µM of probe, and 5 µl of RNA sample were mixed in a 25µL monoplex quantitative RT-PCR reaction. The primer and probe were generated following the national guideline [7], and the reaction condition was set according to the manufacturer's protocol. Cases testing positive for both target genes (open reading frame 1ab and nucleocapsid protein) were determined as laboratory-confirmed cases. Meanwhile, commercial kit (Bio-germ Inc., Shanghai, China and ABT Inc., Beijing, China), which were approved by China Food and Drug Administration (CFDA), were also used in testing. For those cases tested by a commercial kit, the results were determined following the manufacturer's instructions. Upper and lower respiratory tract specimens from confirmed cases were used for viral genome sequencing. Next-generation sequencing (NGS) was achieved based on a metagenomics strategy followed by Illumina sequencing (Illumina Inc., San Diego, CA, USA). Output data were assembled by a viral genome-targeted assembly pipeline with a homology search E-value of  $1e^{-10}$ . Full genome sequences used in the current study are available on request. Sequence alignment was achieved by MAFFT [8]. The Neighbor-Joining (N-J) method with a bootstrap of 1000 was used for phylogenetic analysis.

### Statistical analysis

We used descriptive statistical methods to analyze the epidemiological characteristics of confirmed cases with SARS-CoV-2 infection. Monte Carlo approach was applied to simulate 1000 times and estimate the average incubation period and its confidential interval. Data analysis was performed using SPSS statistical software package version 20.0 (IBM SPSS Inc., Chicago, IL, USA) and GraphPad Prism version 7.0 (GraphPad Software., San Diego, CA, USA). All statistical tests were 2-sided, and statistical significance was set at *P* value less than 0.05.

### Ethics Approval

Data collection and analysis of cases and close contacts were determined by the NHCPRC to be part of a continual public health outbreak investigation, and as such was granted exemption from the institutional review board.

## **Results**

## Epidemiological characteristics of confirmed cases

From Jan 1st 2020 to Apr 3rd 2020, respiratory specimens from 7,432 suspected cases were tested for SARS-CoV-2, of which 602 (8.1%) were positive. Among the laboratory-confirmed SARS-CoV-2 infections, 585 were confirmed case-patients with symptoms and signs and 17 were asymptomatic COVID-19 cases. The proportion of asymptomatic infection was 2.8% (95% CI: 1.5-4.2).

The epidemic curve in Beijing by date of confirmation and onset is shown in Figure 1A. The first case was confirmed on Jan 19th. The number of infection rapidly grew to Jan 30th and peaked between Jan 31st and Feb 1st. The infections steadily declined between Jan 30th and Feb 18th. However, the number of reported cases fluctuated due to a local cluster (14 cases) that occurred in a company on Feb 25th in Beijing and imported cases from abroad.

Since the first imported case from Iran on Feb 29th, the number of cases from overseas increases rapidly. To effectively treat and manage these cases, Xiaotangshan Hospital was launched on Mar 16th. On Mar 23rd, nine cities were designated as the first entry sites for international passenger flights bound for Beijing. To detect more potential cases, all international passengers arriving in Beijing were required to be tested and quarantined at designated facilities on Mar 25th. The number of imported case from overseas in Beijing ranked first in China on Mar 29th. As of Apr 3rd, a total of 169 cases had been identified in Beijing. However, only 4 local cases were identified at the same time. And besides one local case was related to one imported case from abroad, no local cases had been reported for 27 days.

As shown in Figure 1B, among 585 confirmed case-patients, 124 (21.2%) were imported from Wuhan, 76 (13.0%) from other regions of China, 169 (28.9%) from overseas, 1 (0.2%) was a secondary case of imported cases from overseas, 201 (34.4%) were local cases and 14 (2.4%) were under investigation. Among the 169 cases from overseas, most were from the United Kingdom (55), Spain (47), Italy (18), the United States of America (17) and France (7). The epidemic in Beijing has been undertaken in two main phases. The first phase started on Jan 19th when the first two cases were confirmed and ended on Feb 28th. The second phase started on Feb 29th when the first imported case from overseas was identified. At the beginning of the first phase (before Feb 1st), the majority of confirmed cases were imported from Wuhan (61.0%, 97/159), while the proportion of imported cases from Wuhan had decreased since Feb 1st (one week after the lockdown of Wuhan) (16.1%, 27/168) ( $\chi^2=70.065$ ,  $P<0.001$ ). At the beginning of the second phase (before Mar 6th when the last local cases identified), imported cases from overseas and local cases coexisted, and then the imported cases from overseas were dominated.

Of 585 confirmed case-patients, 268(45.8%) cases were male. The male-to-female ratio was 0.9:1. The incidence of males was slightly lower than females, 2.5/100,000 and 2.9/100,000, respectively (Figure 2). The median age of confirmed cases was 39 years old (range, 0.5 to 94; interquartile range, 27 to 56). About 19% (114/585) were 60 or above, and 7.9% (46/585) were children <18 years old, among which 34.8% (16/46) were children under five years old (Table 1). A total of 17 (2.9%) cases aged 80 years or above. The incidence rate of the population 60 years old or above (6.9/100,000) was the highest

compared to the other three groups, followed by the 18-59 years old group (5.6/100,000), the 5-17 years old group (0.5/100,000), and children under five years old (0.3/100,000).

Table 1

Epidemiologic Characteristics of 585 confirmed case-patients with SARS-CoV-2 Infection in Beijing, China.

Characteristic	Cases (n = 585) No./total no.(%)
Sex	
Male	268/585(45.81)
Female	317/585(54.19)
Age group	
0-	16/585(2.74)
5-	30/585 (5.13)
18-	427/585 (72.99)
60-	112/585 (19.15)
Imported cases	368/585 (62.91)
Exposure history	
Huabei Seafood Wholesale Market	0/585(0.00)
History of residence or travel	368/585(62.91)
History of residence or travel in Wuhan	124/585(21.20)
History of residence or travel in other regions outside Wuhan in China	76/585(12.99)
History of residence or travel	169/585(28.89)
Contact with confirmed or suspected cases	201/585(34.36)
Contact with confirmed or suspected cases from abroad	1/585(0.17)
Under investigation	14/585(2.39)
Health care worker	17/585(2.91)
Case of severity	
Mild	212/585(36.24)
Moderate	291/585(49.74)

Characteristic	Csese (n = 585) No./total no.(%)
Severe	66/585(11.28)
Critical	16/585(2.74)

The median age of 46 children <18 years was 7 years old (interquartile range, 3 to 13). The sex ratio was 0.8:1. The proportion of asymptomatic infection among children <18 years and adults was 4.2% and 2.7% ( $\chi^2=0.343$ ,  $P=0.558$ ). Among 585 confirmed case-patients, 17 cases (2.9%) were healthcare workers (HCW). Epidemiological investigations suggested that 7 cases were infected due to health care activities and the remaining 10 were infected due to close contact with household cases rather than in a health care setting according to data from epidemiological investigation.

Of all confirmed case-patients, 36.3% were mild, 49.7% were moderate, 11.3% (64/585) were severe cases and 2.7% (16/585) were critically ill. The proportion of severe and critical cases decreased from 21.4% before Feb 1st to 7.2% after Feb 1st (Figure 3A). The association between illness severity and age was shown in Figure 3B. It was shown that illness severity aggravated with age (Supplementary Table 2,  $\chi^2=50.576$ ,  $P<0.001$ ). A total of 8 cases deceased (with a crude case-fatality rate of 1.4%), among which 7 deaths were elderly adults over 60 years and 1 death was a 50-year-old man. The case fatality rate for males was 1.9% (5/268) and 1.0% (3/317) for females. The case fatality rate for  $\geq 80$  age group was 29.4% (5/17). All of the deaths had comorbid conditions, of which 75.0% (6/8) had hypertension or cardiovascular disease. Among 46 children <18 years, 28 (60.9%) were mild, 17 (37.0%) were moderate and 1 (2.2%) was severe. The severe case was a three-year-old child with leukemia.

### **Medical care timelines**

The median duration from onset of symptoms to their first medical visit was estimated to be 3 days ( $P_{2.5}$ - $P_{97.5}$ : 0-17). The median duration from onset to case confirmation was estimated to be 5 days ( $P_{2.5}$ - $P_{97.5}$ : 1.0-19.5).

### **Incubation estimates**

We reviewed the records of the confirmed cases and found 37.4% (219/585) had clear history of contacts with cases prior to symptom onset; based on which we estimated that the mean incubation period was 6.3 days (95% CI: 6.0-6.6) and the median was 5.7 days ( $P_{2.5}$ - $P_{97.5}$ : 5.2-6.1).

### **Close contacts**

By Apr 3rd, a total of 4,007 close contacts were quarantined, 186 were confirmed with SARS-CoV-2 infection, with an overall secondary attack rate of 4.6% (95% CI: 4.0-5.3). The secondary rate was higher

among family members or relatives (15.6%, 111/714) than that among social contacts (2.2%, 75/3363) ( $\chi^2=239.852$ ,  $P<0.001$ ).

Among 441 close contact of HCWs, 32 were confirmed with SARS-CoV-2 infection, with an overall secondary attack rate of 7.3%, which was higher than that of non-HCWs', with a secondary attack rate of 4.2% (154/3636) ( $\chi^2=8.243$ ,  $P=0.004$ ).

## **Clusters**

Till Apr 3rd, a total of 117 clusters occurred, involving 391 confirmed cases. Among 391 cases, 246 (66.3%) occurred in family, 56 from abroad (15.1%), 28 (7.6%) in health care facilities, 28 (7.6%) in public areas and 13 (3.5%) in mixed areas. Before Feb 18th, clusters mainly occurred in family and then were predominantly from abroad after Feb 28th (Figure 4A).

There were 246 confirmed cases were involved in 91 family clusters. The median number of involved cases of family clusters was 3 (Range, 2-7; IQR, 2-3). To estimate the basic reproduction number and serial interval in families, we identified 38 family clusters where one single index case was introduced. Since the 38 index cases caused 76 secondary cases among 193 family close contacts, we estimated that  $R_0$  in family clusters was 2.0 (95% CI: 1.6-2.4). And the average serial interval was estimated at 7.6 days (95% CI: 6.4-8.9). The median serial interval was 6.0 days (range, 0-34; IQR, 3.3-11.0) (see Figure 4B).

Two clusters occurred in hospitals, involving 38 confirmed cases. One cluster occurred in the cardiac intensive care unit (CICU) and intensive care unit (ICU) of a general hospital in Beijing, involving 35 cases, among which 7 were HCWs. Another clusters in hospital involved 3 cases including one hospitalized patient and her two relatives providing paramedical assistants.

## **SARS-CoV-2 testing among ILI cases**

From Jan 28th to Apr 3rd 2020, a total of 3,267 specimens were collected from 3,267 individuals, and no SARS-CoV-2 viral RNA was identified.

## **Viral RNA in clinical samples**

Among all 585 confirmed case-patients, 936 specimens from 243 cases were available for analysis, including 612 pharyngeal swabs, 7 nasal swabs, 5 saliva specimens, 219 sputa specimens, 2 serum/plasma specimens, 22 urine specimens, and 69 fecal specimens. 354 pharyngeal swabs, 7 nasal swabs, 156 sputa specimens, and 28 fecal specimens showed positive results in real-time RT-PCR tests. The RNA positive rate of throat swabs in mild, ordinary, severe and critical cases was 58.73%, 55.17%, 62.07%, and 76.47%, respectively ( $\chi^2=4.050$ ,  $P=0.256$ ). In contrast to these, the RNA positive rate of sputum in mild, ordinary, and severe cases was 63.08%, 75.86%, and 70.27%, respectively ( $\chi^2=3.329$ ,

$P=0.189$ ). No viral RNA was found in all 22 urine specimens, 2 serum/plasma specimens and 5 saliva specimens.

### **Phylogenetic analysis of SARS-CoV-2 in Beijing**

A total of five viral full genomes were obtained during the study period, including four sequences from imported cases from Whuan and one from a secondary case. Phylogenetic analysis suggested all tested viruses belonged to lineage B of the genus beta-coronavirus, and is genetically closely related to SARS-CoV-2 isolates in Wuhan (Figure 5). It showed high genetic similarity among all tested viruses of 99.97% - 99.99%. It is worth noting that all five viruses carried 442L, 472F, 479Q, 487N, and 491Y in viral S gene receptor-binding subdomain.

## **Discussion**

The transmission of COVID-19 in Beijing was initially imported from Wuhan, Hubei of China, then generated subsequent community clusters in Beijing, dominated by intrafamilial transmission. Later, after the epidemic was controlled in China, imported cases from abroad were predominantly. Our results could provide important epidemiological and genetic parameters from the real-world settings outside Wuhan for further analysis, including evaluations of the impact of control strategy and predictions of COVID-19 dissemination in large cities. We also showed the importance of enhanced surveillance and testing capacity in a mega-city where a very small proportion of the population are immune to the virus. Until a vaccine is available, this is critical to prevent new epidemics from arising.

We estimated basic epidemiological parameters from family clusters where exposure and infection details were well characterized. We estimated an average serial interval of 7.6 days based on the data of 38 family clusters, which was similar to Li's [9]. We further bring up the 95% CI from 5.3-19.0 to a higher precision of 6.4-8.9 due to the larger sample size. We also obtained a similar estimate of  $R_0$  approximately 2 (95% CI: 1.6-2.4) to Li's (indicating that each patient was able to further infect 2 persons on average) [9]. Some estimates of  $R_0$  used the report date of cases, and are not as reliable as  $R_0$  estimates from the date of symptom onset. The  $R_0$  for SARS-CoV-2 was similar to SARS. In contrast, the estimated  $R_0$  of SARS coronavirus ranged from 1.1 to 4.2 with most estimates between 2 and 3 [10-12]. However, the major difference with COVID 19 is pre- and asymptomatic transmission, which makes the disease control practice far more challenging. The previous models highly relied on the assumptions underpinning the models, the timing of diagnosing and reporting of confirmed cases.

We estimated the secondary transmission of COVID-19 in 2,902 close contacts. The secondary attack rate of household was statistically higher than that of social contact (16.6% vs. 3.4%,  $P<0.001$ ). This suggests that disease control around contact tracing should focus on spread from person to person among household close contacts as the highest risk. Thus, the implementation of prevention and control measures in household is particularly required. The household secondary attack rate of COVID-19 was slightly higher than SARS in Beijing, 2003 [13,14]. This might imply that the transmissibility of SARS-CoV-

2 was higher than SARS. Due to the atypical or unspecific presence of mild infection or asymptomatic infection, the family members had been fully exposed to cases before they were confirmed, and this increased the risk of infection among household close contacts. In addition, we found that the secondary attack rate of HCWs was relative higher than that of non-HCWs' which implied that we should focus on the prevention among HCWs who might cause potential clusters or outbreaks both in healthcare facilities and families.

On Jan 23rd, the Chinese government officially announced the implementation of lockdown of Wuhan City. We showed that from Feb 1st, about 1 incubation period from the lockdown, the proportion of cases with a travel history to Wuhan progressively declined, indicating the impact of lockdown of Wuhan on the epidemic in Beijing.

The laboratory findings for SARS-CoV-2 infection pointed out the preference of lower respiratory tract samples in real-time RT-PCR. The positive rate of sputum was slightly higher than that of pharyngeal swab in moderate and severe cases. Our previous quantitative study also showed the relatively high viral loads in sputum with a median of  $7.52 \times 10^5$  [15]. Recent reports also suggested some other types of sample for SARS-CoV-2 RNA testing [16,17], including nasopharyngeal swabs, saliva samples and tear samples. The performances of these specimens need further assessment since the viral RNA was not found in urine samples and saliva samples in our study.

None of the 3,267 samples from ILI cases based on routine ILI surveillance was tested positive for SARS-CoV-2. This provides confidence that COVID-19 has not spread widely in Beijing, China. Little genetic variance was found in SARS-CoV-2 viruses in Beijing. When the virus jumped directly from bats to human or via some unknown intermediate hosts is still unclear [18,19], the genetic analysis suggested the SARS-CoV-2 had a relatively low estimated mean evolutionary rate of  $1.79 \times 10^{-3}$  to  $1.82 \times 10^{-3}$  substitutions per site per year [20], which was in consistent with our results. In this study, a high similarity was observed in tested strains collected at the early stage of COVID-19 outbreak when compared with reference strain (NC\_045512). However, numerous novel mutations, such as S-D614G, were identified in current strains worldwide according to recent studies [21]. The increased genetic diversity of SARS-CoV-2 in human hosts over time cannot be neglected.

Our analysis had several limitations. First, if the close contacts of confirmed cases traveled to other provinces outside Beijing, some of them might be lost to follow-up. This might bring bias to the secondary attack rate of close contact and thus lead to lower  $R_0$  estimates. Second, detailed information on such exposure history as frequency, intensity, and duration for all cases was not available. This may influence the estimates of the incubation period. Third, the relationship between index cases and their quarantined close contact was not well recorded, which limited the analysis of secondary attack rates by different relationships. A strength of the study is the well-established surveillance systems such as ILI surveillance, which were able to provide enhanced capability to study the epidemiology and exclude substantial community transmission.

## Conclusion

The transmissibility of SARS-CoV-2 was relatively high, especially among households and from HCWs, which should be focused on. The lockdown of Wuhan City was effective in halting the spread of the COVID-19 to Beijing.

## List Of Abbreviations

CoV: coronavirus; COVID-19: coronavirus disease 2019; AI: artificial intelligence; ILI: influenza-like illness; RT-PCR: reverse transcriptase polymerase chain reaction; NGS: Next-generation sequencing; N-J: Neighbor-Joining; CI: confidential interval; HCWs: healthcare workers;

## Declarations

### Ethics approval and consent to participate

Data collection and analysis were determined by the National Health Commission of the People's Republic of China to be part of the continual public health investigation of an emerging outbreak and thus the informed consent and ethics approval was waived by institutional review board and human research ethics committee of Beijing Center for Disease Prevention and Control.

### Consent for publication

Not applicable.

### Availability of data and materials

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

### Competing Interests.

The authors declare no competing interests.

### Authors Contributions

Concept and design: PY and QW.

Acquisition, analysis, or interpretation of data: All Authors.

Drafting of the manuscript: XW, YP and CRM.

Critical revision of the manuscript: PY and QW.

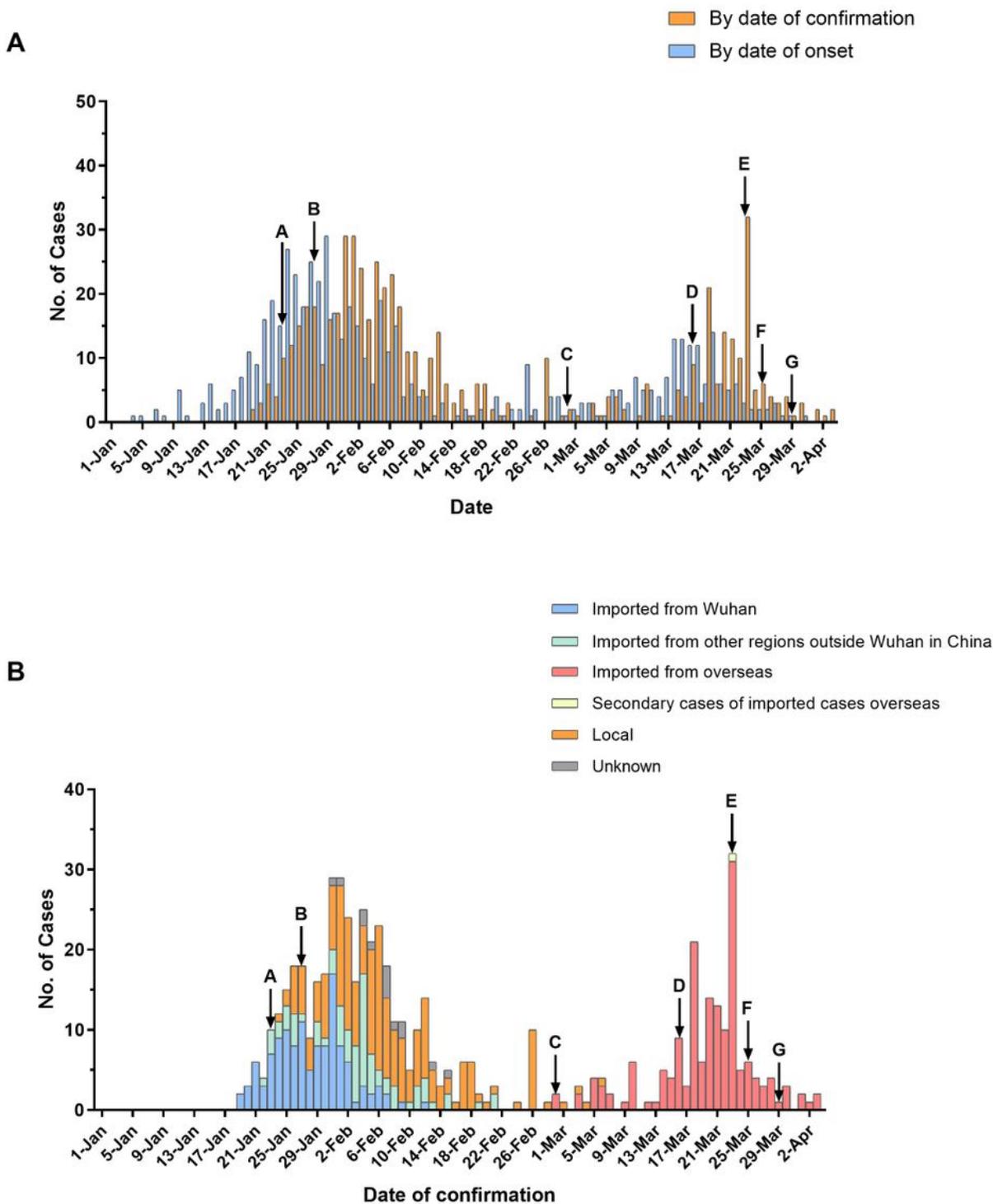
Laboratory testing: YP, DZ, LC.

## References

1. World Health Organization. Public Health Emergency of International Concern declared. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>. Published 31 Jan 2020. Accessed 7 Apr 2020.
2. Gorbalenya AE, Baker SC, Baric RS, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 2020; 5:536–44.
3. World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report –46. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf_4). Published 06 Mar 2020. Accessed 13 Apr 2020.
4. World Health Organization. The First Few X (FFX) Cases and contact investigation protocol for 2019-novel coronavirus (2019-nCoV) infection. [https://www.who.int/publications-detail/the-first-few-x-\(ffx\)-cases-and-contact-investigation-protocol-for-2019-novel-coronavirus-\(2019-ncov\)-infection](https://www.who.int/publications-detail/the-first-few-x-(ffx)-cases-and-contact-investigation-protocol-for-2019-novel-coronavirus-(2019-ncov)-infection). Published 2020. Accessed 7 Apr 2020.
5. Yang P, Duan W, Lv M, et al. Review of an influenza surveillance system, Beijing, People's Republic of China. *Emerg Infect Dis* 2009; 15(10):1603-8.
6. National Health Commission (NHC) of the PRC. Protocol for COVID-19 Prevention and Control, (2nd Edition). 2020. <http://www.nhc.gov.cn/xcs/zhengcwj/202001/c67cfe29ecf1470e8c7fc47d3b751e88.shtml>. Published 23 Jan 2020. Accessed 7 Apr 2020.
7. National Health Commission (NHC) of the PRC. Protocol for COVID-19 diagnosis and treatment (7th Edition). 2020. <http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml>. Accessed 7 July 2020.
8. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 2019; 20(4):1160-1166.
9. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020; 382:1199-207.
10. Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003; 300(5627):1966-1970.
11. Organization WH. Consensus document on the epidemiology of severe acute respiratory syndrome (SARS). WHO. <https://apps.who.int/iris/handle/10665/70863>. Published 2003. Accessed 7 Apr 2020.
12. Bauch CT, Lloyd-Smith JO, Coffee MP, Galvani AP. Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology* 2005;16(6):791-801.
13. Pang X, Zhu Z, Xu F, et al. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. *JAMA* 2003;290(24):3215-3221.

14. Goh DL, Lee BW, Chia KS, et al. Secondary household transmission of SARS, Singapore. *Emerg Infect Dis*. 2004;10(2):232-234.
15. Pan Y, Zhang D, Yang P, Poon LLM, Wang Q. Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect Dis* 2020; (20)30113-4.
16. To KK, Tsang OT, Chik-Yan Yip C, et al. Consistent detection of 2019 novel coronavirus in saliva. *Clin Infect Dis* 2020; [Epub ahead of print].
17. Xia J, Tong J, Liu M, Shen Y, Guo D. Evaluation of coronavirus in tears and conjunctival secretions of patients with SARS-CoV-2 infection. *J Med Virol* 2020;[Epub ahead of print].
18. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020; 579:270-3.
19. Ceraolo C, Giorgi FM. Genomic variance of the 2019-nCoV coronavirus. *J Med Virol* 2020; 92(5):522-8.
20. Li X, Wang W, Zhao X, et al. Transmission dynamics and evolutionary history of 2019-nCoV. *J Med Virol* 2020; 92:501-11.
21. Korber B, Fischer W.M, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* 2020; [Epub ahead of print].

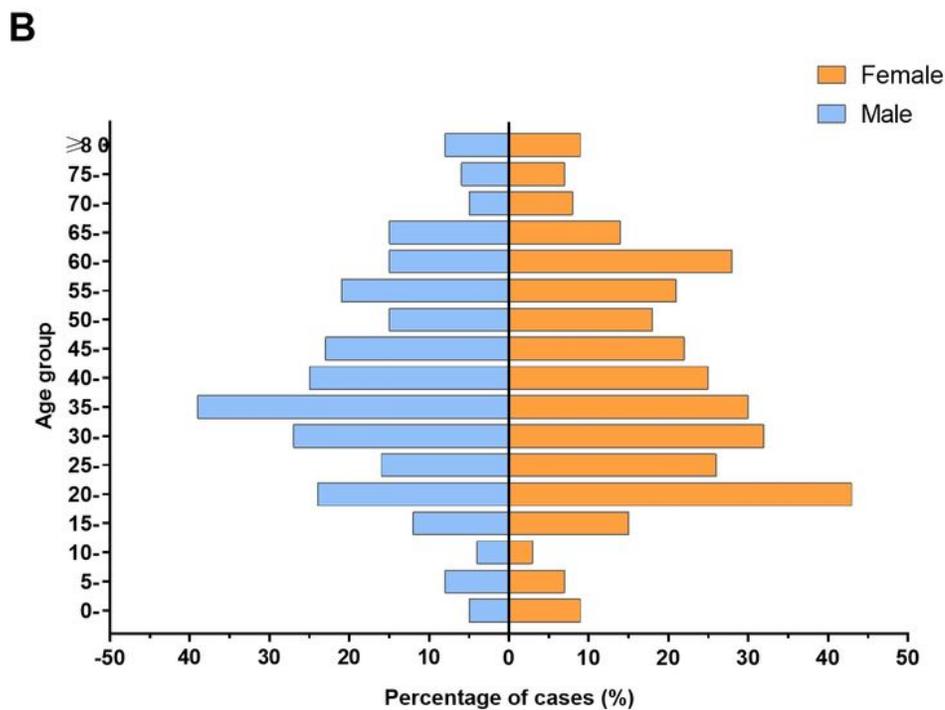
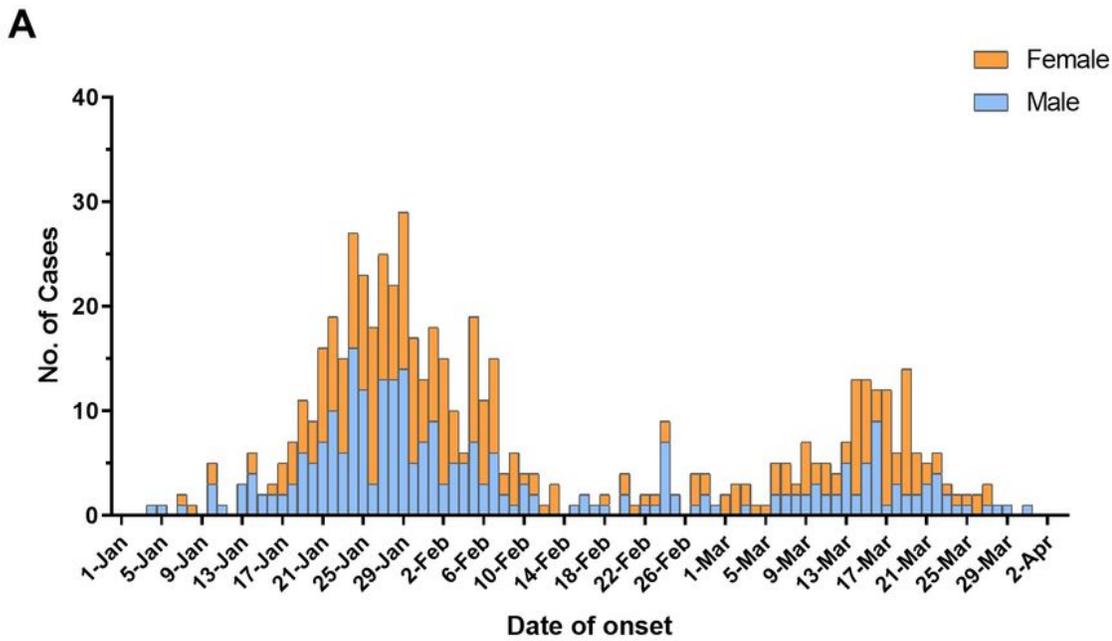
## Figures



**Figure 1**

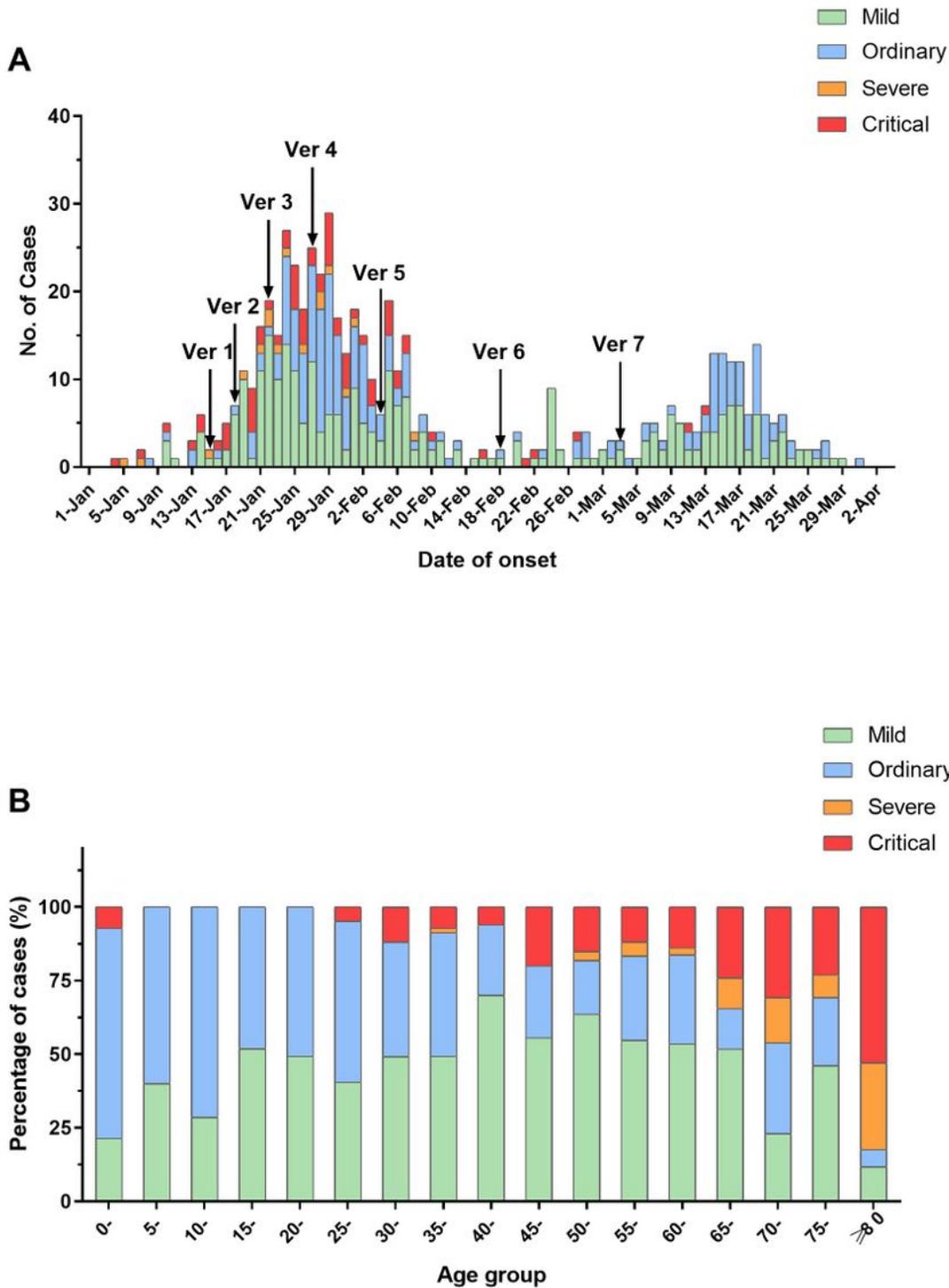
Epidemiologic Characteristics of 585 confirmed cases with SARS-CoV-2 infection Note: Local cases include the secondary cases of imported case outside in China, excluding secondary cases of imported case from abroad. A: Wuhan was locked down on 23 Jan. B: First fatal cases occurred on 27 Jan. C: First imported case from overseas identified on 29 Feb. D: Xiaotangshan Hospital launched for imported cases from overseas on 16 Mar. E: Nine cities designated as points of first entry for international passenger

flights bound for Beijing on 23 Mar. F: All international passengers arriving in Beijing must be tested and quarantined at designated facilities on 25 Mar. G: The number of imported case from overseas in Beijing ranked first in China on 29 Mar,



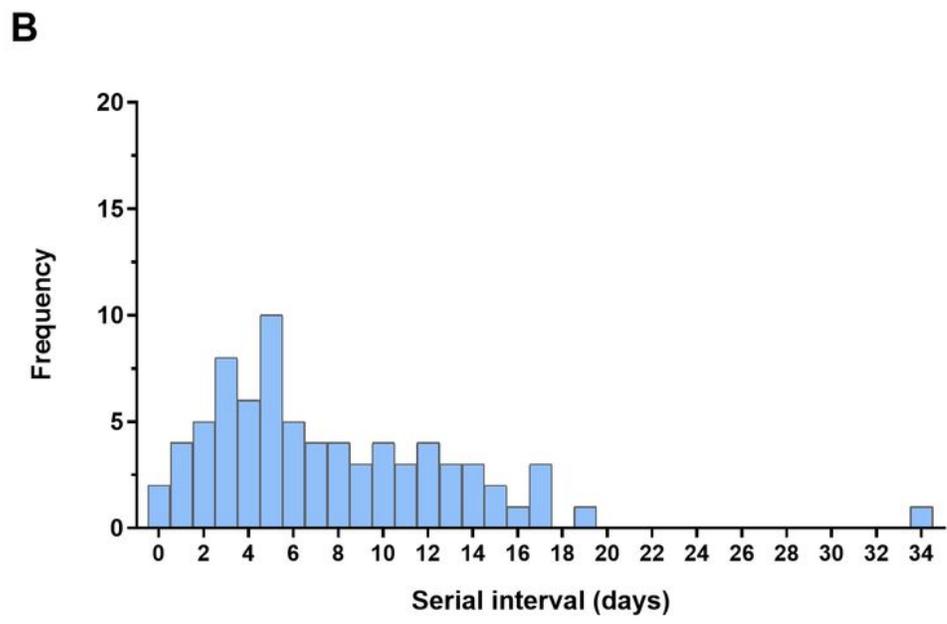
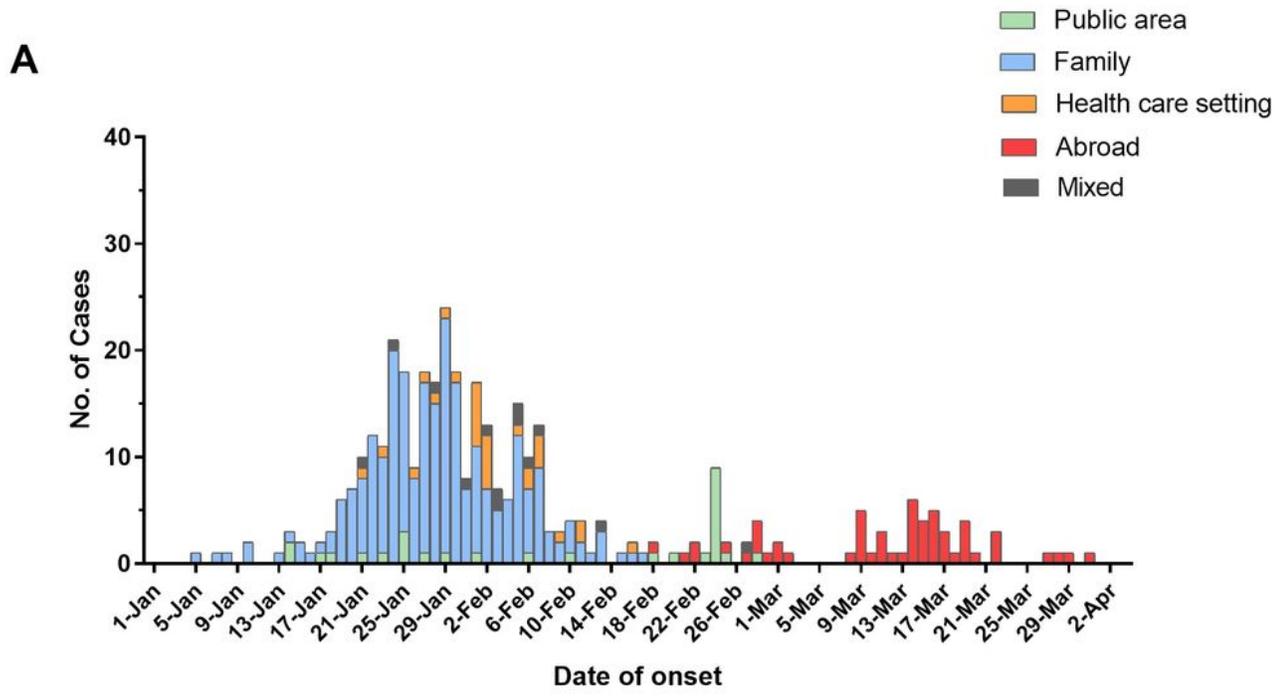
**Figure 2**

The age and sex distribution of 585 confirmed case-patients with SARS-CoV-2 infection



**Figure 3**

The clinical severity of 585 confirmed case-patients with SARS-CoV-2 infection Note: Ver: Version. The different version of definition for confirmed cases were marked.



**Figure 4**

The source and distribution of clusters and the serial interval for family cluster introduced one single index case

