Epidemiological characteristics of COVID-19 clusters in Hainan, China

Sha Xiao, MMa, Yunru Liu, PhDa, Fang Liu, BMb, Hanxi Zhang, PhDc, Fan Zhang, PhDa, Lu Wang, PhD, MDa,c,*

Abstract
To analyze the epidemiological characteristics of coronavirus disease 2019 (COVID-19) clusters in Hainan, and to provide a basis for the prevention and control of disease clusters.

Descriptive epidemiology was used to retrospectively analyze the characteristics of disease clusters in 168 cases of COVID-19.

Of the 168 COVID-19 cases, 99 (58.93%) comprised 29 clusters, 22 (75.86%) of which were imported and included 63 cases (63.64%), while 7 clusters (24.14%) were local and included 36 cases (36.36%). Of the cluster cases, 49 were men (49.49%) and 50 were women (50.50%), the median age was 52 years, and the maximum number of cases from 41 to 60 was at 37 years (37.37%).

There were 67 first generation cases (67.68%), 28 (28.28%) second generation, and 4 (4.04%) third generation. Of the clusters, 68.97% occurred from January 31 to February 7, with the highest peak on February 6. The local disease clusters occurred with a time lag. The 2 cities with the most reported incidents were Sanya (10 cases, 34.48%) and Haikou (5 cases, 17.24%). Family clusters were most frequent, with 18 clusters (62.07%) involving 62 cases (62.63%), followed by social clusters, with 3 clusters (10.34%). The most complex clusters involved 3 cluster types (family, travel, and community). There was a statistically significant difference in the infectivity of the imported clusters versus the local clusters, with imported clusters being lower (Z = -2.851, P = .004). The infectivity of all cases or family members was highest in Haikou and lowest in Sanya. The infectivity of all cases with an incubation period of ≤7 days was 1.53 ± 1.01, in which the infectivity of family members was 1.29 ± 1.10. The infectivity of all cases with an incubation period of ≤14 days was 1.89 ± 1.23, in which the infectivity of family members was 1.43 ± 1.37.

COVID-19 clusters in Hainan mainly occurred in families, and local clusters had high infectivity. Therefore, key populations and regions should be monitored, and targeted preventive measures should be carried out to provide a reference for the prevention and control of disease clusters.

Abbreviations: COVID-19 = Corona Virus Disease 2019, WHO = World Health Organization.

Keywords: cluster, coronavirus disease 2019, epidemiological characteristics, Hainan

1. Introduction
An outbreak of infectious viral pneumonia of unknown origin occurred worldwide in 2019.[1] This disease, which was officially given the name coronavirus disease 2019 (COVID-19) on February 11, 2020, has now spread to more than 200 other countries.[2] The uncertainty and unpredictability of this new and highly infectious disease led to the inability to make timely decisions regarding specific preventive and control measures, which seriously affected social stability and economic development worldwide.[3,4] Therefore, the World Health Organization (WHO) declared the COVID-19 pandemic to be a global public health emergency.[5]

To control the spread of COVID-19, the Hainan province of China launched a first-level response to this major public health emergency on January 25, 2020. Previous research has reported many clusters of COVID-19,[6,7] but the epidemiological characteristics of COVID-19 clusters in Hainan are still unknown. The present study aimed to investigate the epidemiological characteristics of COVID-19 disease clusters in Hainan. We also aimed to contribute information urgently needed to key populations and areas, and to provide practices for the prevention of disease clusters in Hainan.

2. Methods
2.1. Study design
This retrospective study included all patients diagnosed with COVID-19 in Hainan from January 22 to February 19, 2020. All patients enrolled in the present study were diagnosed with COVID-19 according to the WHO interim guidance, and were
divided into 2 groups, import cases and local cases, based on epidemiological data.\[8\]

2.2. Data sources and collection
All data regarding 168 cases of COVID-19 in Hainan through 00:00 on February 19, 2020, were obtained from the official website of the Health Committee of Hainan Province (http://wst.hainan.gov.cn/yqfk/index/index/qianyi.html). We obtained the following data for all cases: sex, age, report area, cluster type (family, social, travel, work, community, and vehicle), exposure period, date of disease onset, date of first admission, and date of confirmation. All data were checked by 2 researchers (Sha Xiao and Hanxi Zhang). The population characteristics, cluster characteristics, timeliness of diagnosis and treatment, and infectivity of all cases or family members were described. We used publicly available data for our study, so we did not include ethical review.

2.3. Definitions
We divided the cases into 3 types: imported, local, and unknown. The imported cases were those that had an epidemiological history of living, traveling, doing business, or working in other provinces prior to their diagnosis, while the local cases had no epidemiological history of living, traveling, doing business, or working in other provinces prior to their diagnosis.

The confirmed time of the clusters was the date of diagnosis of the second confirmed case in each cluster. Imported clusters were defined as those in which all cases involved were imported, whereas local clusters were defined as those that included more than 1 local case.

For imported cases, the exposure period was the time between the date of onset and the date of return to Hainan; first admission period after onset = first admission date – onset date; confirmed period after first admission = confirmed date – first admission date; confirmed period after onset = confirmed date – onset date.

2.4. Statistical analysis
We analyzed the basic social demographic and epidemiological characteristics of different cluster types and geographically different distributions. Infectivity was calculated based on the number of individuals infected by the first confirmed case. We set the first case of each disease cluster as the indicator case, and set the number of family members (or other close contacts) confirmed after contact with the first case as the continuous cases. A 2-tailed P value of <.05, was considered statistically significant. Statistical analyses were performed using SPSS 21.0.

3. Results
3.1. Characteristics of disease clusters in Hainan
A total of 29 clusters, involving 99 cases, were identified, of which 22 (75.86%) were imported, including 63 (63.64%) cases. Of the imported cases, 20 were imported from Hubei and 2 from Thailand. There were 7 (24.14%) local clusters involving 36 (36.36%) cases, 5 of which were caused by imported cases, while 2 were due to unknown reasons. The largest cluster involved 9 cases (Table 1).

### Table 1
Characteristics of cluster epidemic in Hainan.

| Characteristics | Clusters number (%) | Cases number (%) |
|-----------------|---------------------|-----------------|
| Clusters        | 29 (100.00)         | 99 (100.00)     |
| Imported clusters |                    |                 |
| From Hubei      | 22 (75.86)          | 63 (63.64)      |
| From Thailand   | 2 (9.09)            | 7 (11.11)       |
| Local clusters  |                    |                 |
| From Wuhan      | 7 (24.14)           | 36 (36.37)      |
| From Zhengzhou  | 3 (42.85)           | 16 (44.44)      |
| From Wulumuqi   | 4 (14.29)           | 14 (38.89)      |
| Unknown reasons | 1 (3.45)            | 9 (0.09)        |
| Cases           |                     |                 |
| Number of cases in each cluster |               |                 |
| 6               | 2 (6.90)            | 12 (12.12)      |
| 5               | 3 (10.34)           | 15 (15.15)      |
| 4               | 6 (20.69)           | 24 (24.24)      |
| 3               | 5 (17.24)           | 15 (15.15)      |
| 2               | 12 (41.38)          | 24 (24.24)      |

3.2. Characteristics of cases involved in clusters
There were 99 cases involved in 29 clusters, including 49 men (49.49%) and 50 women (50.50%), with ages ranging from 3 months to 79 years, with a median age of 52 years, and the range of 41 to 60 years old was the largest percentage (37.37%). There were 67 cases (67.68%) occurred in only 1 generation, 28 (28.28%) of 99 were in the second generation, and 4 (4.04%) in the third generation (Table 2).

### Table 2
Characteristics of cases involved in clusters in Hainan.

| Characteristics | Cases N (%) |
|-----------------|-------------|
| Gender          |             |
| Men             | 49 (49.49)  |
| Women           | 50 (50.50)  |
| Age             |             |
| 3 mo–20         | 8 (8.08)    |
| 21–40           | 25 (25.25)  |
| 41–60           | 37 (37.37)  |
| 61–80           | 29 (29.29)  |
| Generation      |             |
| One             | 67 (67.68)  |
| Two             | 28 (28.28)  |
| Three           | 4 (4.04)    |

3.3. Time distribution of disease clusters
The first cluster occurred on January 23, 2020. A total of 68.97% of clusters occurred between January 31 and February 7, 2020, with February 6 being the peak of the disease clusters. The last confirmed time of a disease cluster was on February 17. The first occurrence time of the imported clusters was earlier than that of the local clusters, whereas the last end time of the local clusters was later than that of the imported clusters (Fig. 1).

3.4. Regional distribution of disease clusters in Hainan
The regional distributions of 29 clusters showed that Sanya had the largest number of clusters (10/29, 34.48%), followed by Haikou (5/29, 17.24%). The regional distribution of the 99
cluster cases showed that Sanya also had the largest number of cases involved in clusters (30/99, 30.30%), followed by Haikou (22/99, 22.22%) (Table 3).

3.5. Types of disease clusters in Hainan

The main types of clusters were simple family clusters, of which there were 18 (62.07%), involving 62 (62.63%) cases, followed by social clusters, of which there were 3 (10.34%). The most complex disease clusters involved 3 cluster types (family, travel, and community) (Table 4).

3.6. Timeliness of diagnosis and treatment of disease cluster cases in Hainan

Overall, the average exposure period for 83 cases was 9.22 days (0–21 days), the average first admission period after onset for 86 cases was 1.14 days (0–12 days), the average confirmed period after first admission for 92 cases was 3.43 days (0–15 days), and the average confirmed period after onset for 86 cases was 4.81 days (0–16 days).

Of the imported clusters, the mean interval exposure period for 55 cases was 8.07 days (0–21 days), the first admission period after onset for 55 cases was 0.6 days (0–8 days), the mean confirmed period after first admission for 59 cases was 3.37 days (0–15 days), and the mean confirmed period after onset for 55 cases was 4.22 days (0–15 days).

Of the local clusters, the average exposure period for 28 cases was 11.46 days (1–21 days), the average first admission period after onset for 31 cases was 2.10 days (0–12 days), the average confirmed period after first admission for 33 cases was 3.55 days (0–16 days), and the average confirmed period after onset for 31 cases was 5.87 days (0–16 days).

The average exposure period, first admission period after onset, confirmed period after first admission, and confirmed period after onset were all lower for the imported clusters than the local clusters ($Z = –2.600, –2.600, –0.066, and –1.191$, respectively). There were statistically significant differences in both the exposure period and confirmed period after onset between the input clusters and local clusters ($P < .05$).

3.7. Analysis of infectivity

The infectivity of all cases in imported clusters ($1.86 \pm 1.08$) was lower than that of local clusters ($4.14 \pm 2.12$), and the difference was statistically significant ($Z = –2.851, P = .004$). The infectivity of all cases and family members in Haikou was the highest, while Sanya was the lowest. There was no significant difference among the 4 groups ($P > .05$). The infectivity of all cases with an exposure period $\leq 7$ days was $1.53 \pm 1.01$, and family members was $1.29 \pm 1.10$. The infectivity of all cases with an exposure period $\leq 14$ days was $1.89 \pm 1.23$, and family members was $1.43 \pm 1.37$ (Table 5).

4. Discussion

As a new infectious disease, COVID-19 has had strong infectious, pathogenic, and clustering characteristics. It has seriously affected people’s physical and mental health.[9] While the Chinese government provided all relevant information to the WHO at the beginning of the pandemic, no country can fully prepare the resources to detect and triage the infected patients. For the sake of transparency, we must first have all the relevant information. Until now, there have been many unknowns

| Region  | Clusters number (%) | Cases number (%) |
|---------|----------------------|------------------|
| Sanya   | 10 (34.48)           | 20 (30.30)       |
| Haikou  | 5 (17.24)            | 22 (22.22)       |
| Wanning | 3 (10.34)            | 10 (10.10)       |
| Dazhou  | 2 (6.90)             | 9 (9.10)         |
| Qionghai| 2 (6.90)             | 5 (5.05)         |
| Lingshi | 1 (3.45)             | 5 (5.05)         |
| Changjiang | 1 (3.45)       | 5 (5.05)         |
| Other regions | 5 (17.24) | 13 (13.13)       |
| Total   | 29 (100.00)          | 99 (100.00)      |
regarding COVID-19. The present study showed that 58.93% of the cases in Hainan were involved in disease clusters. A total of 29 clusters occurred, comprised of 22 (75.86%) imported and 7 (24.14%) local clusters, and the largest disease cluster involved 9 cases. Of the 99 cluster cases, 49.49% were men and 50.50% were women, aged 3 to 79 years (median 52 years), and 91.9% were 21 to 80 years old, which provided evidence for general susceptibility. The age range was relatively wide, which may be related to the characteristics of the “migrant population,” in which family members spent the winter in Hainan, particularly the older population.

The results of the present study showed that Sanya and Haikou had the most disease clusters, followed by Wanning. The regional distribution of the cases involved was the same as that of the clusters. Studies have found a correlation between population migration and COVID-19 morbidity.[10] We felt that the regional distribution characteristics of Hainan’s clusters may be related to these 3 eastern coastal cities, which are hot areas for the “migratory bird” population, with convenient transportation. The first confirmed time for a disease cluster was on January 23, 2020, and 68.97% of the clusters occurred between January 31 and February 7, 2020, peaking on February 6, with the last confirmed cluster on February 17, 2020. The first occurrence time for the imported clusters was earlier than that of the local clusters, while the last end time of the local clusters was later than that of the import clusters. The occurrence of local clusters had a lag time, since most of the local clusters were caused by the imported cases, who had gone to Hainan before the Spring Festival of China, and 85.86% of the cases went to Hainan between January 16 and January 25, 2020. The first level response to major public health emergencies was launched on January 25 by the Hainan Provincial Government, which played a key role in the detection of the disease clusters.

The present study also described 2 local clusters with unknown origins, one of which involved 9 cases, all without a history of contact with confirmed cases. The first case in this cluster was imported from Wuhan to Haikou on January 2, 2020, with an onset time of February 5. The remaining 8 cases were all family members, but had no epidemiological history prior to February 5. The other local cluster of unknown origin involved 5 cases, none of which had a history of contact with another confirmed case. The first case in this cluster was imported from Jiangxi on November 24, 2019, and the other cases, family members or colleagues of the first case, had no epidemiological history before the onset time of January 29, 2020.

The disease clusters in Hainan were mainly family clusters, followed by social clusters. The most complex cluster involved 3 cluster types (family, travel, and community). The results of the present study were the same as those reported by Zhao Shan-lu et al[11] in Hunan. Ozdemir et al[12] reported that 37.2% of the confirmed COVID-19 cases had family clusters from Turkey, but in our study, the ratio is 62.07%. In all clusters, the average exposure period for 83 cases was 9.22 days (0–21 days), which was higher than that of the disease estimated by the state (about

| Clusters | Family members number | All cases number | Infected (mean ± standard deviation) |
|----------|-----------------------|-----------------|-------------------------------------|
| Sources  |                       |                 | Family members number | All cases number |
| Input    | 22                    | 32              | 41                    | 1.45±1.14         | 1.86±1.08* |
| Local    | 7                     | 21              | 29                    | 3.00±2.94         | 4.14±2.12* |
| Regions  |                       |                 |                       |                   |
| Sanya    | 10                    | 14              | 20                    | 1.40±0.70         | 2.00±0.94 |
| Haikou   | 5                     | 13              | 17                    | 2.60±3.21         | 3.40±2.89 |
| Wanning  | 3                     | 6               | 7                     | 2.00±2.65         | 2.33±2.31 |
| Others   | 11                    | 20              | 26                    | 1.82±1.66         | 2.36±1.43 |
| Exposure period |               |                 |                       |                   |
| ≤7 d     | 17                    | 22              | 26                    | 1.29±1.10         | 1.53±1.01 |
| ≤14 d    | 28                    | 40              | 53                    | 1.43±1.37         | 1.89±1.23 |

*Z=-2.851, P=.004.
5 days).\[13\] There were statistically significant ($P < .05$) differences in both exposure period and confirmed period after onset between import and local clusters.

Infectivity is an important indicator for the evaluation of new infectious diseases. Most studies use basic reproduction number ($R_0$) to calculate infectivity, but the disease cluster factors were not considered.\[14\] The present study calculated infectivity based on the disease clusters. The infectivity of all cases in imported clusters (1.86 ± 1.08) was lower than that of local clusters (4.14 ± 2.12), and the difference was statistically significant ($Z = -2.851$, $P = .004$). The infectivity of all cases and family members in Haikou was the highest, while Sanya was the lowest. The infectivity of all cases and family members with an incubation period ≤7 days was lower than that of those with an incubation period ≤14 days, indicating that the shorter the incubation period, the stronger the infectivity.

There are some limitations to the present study. First, asymptomatic patients and those with mild symptoms might have recovered by themselves. These cases might be missed by the government due to the lack of medical contact, which can be a worldwide pattern of oversight. If a person has symptoms, they are limited to only over-the-counter drugs that can be self-administered, otherwise they have to seek medical attention to rule out COVID-19 or be isolated for treatment for COVID-19. Second, the number of cases in Hainan province is less than that in most regions due to the strict measures applied to curb the spread of COVID-19 in Hainan. Being an isolated island, such as Hainan, affords the advantage of mitigating further transmission.

The present study demonstrated that the COVID-19 disease clusters in Hainan are dominated by family aggregation, and that all cases of local clusters were highly infectious. These data could help inform and scale the COVID-19 epidemic and spatiotemporal distribution of new coronavirus pneumonia in Guangdong Province [in Chinese]. J Trop Med 2020;20:427–30 +571http://kns.cnki.net/kcms/detail/44.1503.R.20200213.0806.002.html.

Zhao SL, Gao LD, Zhan KW, et al. Analysis of the cluster epidemic characteristics of coronavirus disease 2019 in Hunan Province [in Chinese]. Pract Prev Med 2020;27http://kns.cnki.net/kcms/detail/43.1223.R.20200306.0825.002.html.

Ozdemir YE, Balkan II, Bayramlar OF, et al. Clinical characteristics of mild-moderate COVID-19 patients and risk factors for the development of pneumonia. Find Microbiome 2021;5:342–56.

Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 7). Chin Med J 2020;133:1087–95.

Huang LL, Shen SP, Yu P, Wei YY. Dynamic basic reproduction number based evaluation for current prevention and control of COVID-19 outbreak in China [in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2020;41:466–9.

**Author contributions**

Data curation: Sha Xiao, Yunru Liu, Fang Liu.

Formal analysis: Sha Xiao.

Writing – original draft: Sha Xiao.

Writing – review & editing: Sha Xiao, Yunru Liu, Fang Liu, Hanxi Zhang, Fan Zhang, Lu Wang.

**References**

[1] Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 2020;395:507–13.

[2] Holsmere ML, DeBolt C, Lindquist S, et al. First case of 2019 novel coronavirus in the United States. N Engl J Med 2020;382:929–36.

[3] Hou YD. Strategies and effects of prevention of major emerging infectious diseases [in Chinese]. Electron J Emerg Infect Dis 2019;4:129–32.

[4] Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhuan, China. Lancet 2020;395:497–506.

[5] World Health Organization. Novel coronavirus (2019-nCoV)-situation report-19-8 29 January 2020 [EB/OL]. Available at: https://apps.who.int/iris/bitstream/handle/10665/330774/nCoVsitrep29Jan2020-eng.pdf?sequence=1&isAllowed=y.

[6] Sun QL, Li ZC, Tan XL, et al. A clustering epidemic of pneumonia caused by 2019-nCoV [in Chinese]. Pract Prev Med 2020;27:389–92.

[7] Zhou H, Zhu HW, Chen BT, et al. Epidemiological survey on a family aggregation COVID-19 in Y County, Chenzhou City, Hunan [in Chinese]. Shanghai J Prev Med 2020;20:929–36.

[8] World Health Organization. Clinical Management of Severe Acute Respiratory Infection When Novel Coronavirus (nCoV) Infections Suspected: Interim Guidance. 2020.

[9] Singh AK, Gillies CL, Singh R, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. Diabetes Obes Metab 2020;22:1915–24.

[10] Wang XZ, Liao CH, Li ZH, et al. Preliminary analysis on the early epidemiological and spatiotemporal distribution of new coronavirus pneumonia in Guangdong Province [in Chinese]. J Trop Med 2020;20:427–30 +571http://kns.cnki.net/kcms/detail/44.1503.R.20200213.0806.002.html.

[11] Zhao SL, Gao LD, Zhan KW, et al. Analysis of the cluster epidemic characteristics of coronavirus disease 2019 in Hunan Province [in Chinese]. Pract Prev Med 2020;27http://kns.cnki.net/kcms/detail/43.1223.R.20200306.0825.002.html.

[12] Ozdemir YE, Balkan II, Bayramlar OF, et al. Clinical characteristics of mild-moderate COVID-19 patients and risk factors for the development of pneumonia. Find Microbiome 2021;5:342–56.

[13] Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 7). Chin Med J 2020;133:1087–95.

[14] Huang LL, Shen SP, Yu P, Wei YY. Dynamic basic reproduction number based evaluation for current prevention and control of COVID-19 outbreak in China [in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2020;41:466–9.