

## RESEARCH ARTICLE

# Spatiotemporal analysis and the characteristics of the case transmission network of 2019 novel coronavirus disease (COVID-19) in Zhejiang Province, China

Haocheng Wu<sup>1,2</sup>, Chen Wu<sup>1</sup>, Qinbao Lu<sup>1</sup>, Zheyuan Ding<sup>1</sup>, Ming Xue<sup>3</sup>, Junfen Lin<sup>1,2\*</sup>

**1** Zhejiang Province Center for Disease Control and Prevention, Hangzhou, Zhejiang Province, China, **2** Key Laboratory for Vaccine, Prevention and Control of Infectious Disease of Zhejiang Province, Hangzhou, Zhejiang Province, China, **3** Hangzhou Centre for Disease Control and Prevention, Hangzhou, Zhejiang Province, China

\* [hchwu@cdc.zj.cn](mailto:hchwu@cdc.zj.cn)



## OPEN ACCESS

**Citation:** Wu H, Wu C, Lu Q, Ding Z, Xue M, Lin J (2021) Spatiotemporal analysis and the characteristics of the case transmission network of 2019 novel coronavirus disease (COVID-19) in Zhejiang Province, China. PLoS ONE 16(9): e0257587. <https://doi.org/10.1371/journal.pone.0257587>

**Editor:** Ka Chun Chong, The Chinese University of Hong Kong, HONG KONG

**Received:** March 31, 2021

**Accepted:** September 3, 2021

**Published:** September 17, 2021

**Copyright:** © 2021 Wu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** This work was supported by the Major Science and Technology Project of the Science and Technology Department of Zhejiang Province (2021C03038). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

### Background

Zhejiang Province is one of the five provinces in China that had the highest incidence of novel coronavirus disease (COVID-19). Zhejiang, ranked fourth highest in COVID-19 incidence, is located in the Yangtze River Delta region of southeast China. This study was undertaken to identify the space-time characteristics of COVID-19 in Zhejiang.

### Methods

Data on COVID-19 cases in Zhejiang Province from January to July 2020 were obtained from this network system. Individual information on cases and deaths was imported, and surveillance information, including demographic characteristics and geographic and temporal distributions, was computed by the system. The Knox test was used to identify possible space-time interactions to test whether cases that are close in distance were also close in time. Network analysis was performed to determine the relationship among the cases in a transmission community and to try to identify the key nodes.

### Results

In total, 1475 COVID-19 cases and 1 fatal case were reported from January to July 2020 in Zhejiang Province, China. Most of the cases occurred before February 15<sup>th</sup>, which accounted for 90.10%. The imported cases increased and became the main risk in Zhejiang Province after February 2020. The risk areas showed strong heterogeneity according to the Knox test. The areas at short distances within 1 kilometer and at brief periods within 5 days presented relatively high risk. The numbers of subcommunities for the four clusters were 12, 9, 6 and 4. There was obvious heterogeneity in the modularity of subcommunities. The maximum values of the node centrality for the four clusters were 2.9474, 4.3706, 4.1080 and 2.7500.

**Competing interests:** The authors have declared that no competing interests exist.

## Conclusions

COVID-19 was brought under control over a short period in Zhejiang Province. Imported infections from outside of mainland China then became a new challenge. The effects of spatiotemporal interaction exhibited interval heterogeneity. The characteristics of transmission showed short range and short term risks. The importance to the cluster of each case was detected, and the key patients were identified. It is suggested that we should focus on key patients in complex conditions and in situations with limited control resources.

## Introduction

In December 2019, a cluster of cases of pneumonia caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was reported in Wuhan, Hubei Province, China [1, 2]. Subsequently, novel coronavirus disease (COVID-19) spread into many counties [3]. On January 30, 2020, the World Health Organization (WHO) declared the outbreak a public health emergency of international concern [4]. As of August 5, COVID-19 has infected more than 18.35 million individuals worldwide and caused more than 696 000 deaths [5].

Currently, the basic epidemiological characteristics of COVID-19 are clear. The original host of SARS-CoV-2 may be bats, which spread the virus to humans through some intermediate animal hosts and caused an epidemic of COVID-19 [6]. The main source of transmission of COVID-19 is infected symptomatic patients, and carriers can also be infectious before they develop symptoms or without ever developing symptoms [6–8]. Current evidence suggests that the virus is transmitted mostly via droplets or contact [6]. The virus can also spread in poorly ventilated and/or crowded indoor settings, where people tend to spend longer periods of time. This is because aerosols remain suspended in the air or travel farther than 1 metre [9]. Fecal-oral transmission can also be a possible route [10]. The population is generally susceptible to the virus. The average incubation period is 5.2 days (95% CL: 4.1~7.0) [6].

COVID-19 remains a severe public health issue in mainland China. The top five provinces with the highest incidence are Hubei, Guangdong, Henan, Zhejiang and Hunan. Zhejiang, ranked fourth highest in COVID-19 incidence, is located in the Yangtze River Delta region of southeast China. The objective of this study is to identify the space-time characteristics of COVID-19 in Zhejiang. The Knox test was applied to identify space-time clusters and suggested a relatively high-risk boundary among a series of space-time intervals, which was compared with those in the condition of random permutation [11]. It also helps to identify the serial intervals in the epidemic with cases continually occurring [12]. We used network analysis to determine the relationships among the cases in a transmission community and tried to identify the nodes that belong to the most communities (key infection). Furthermore, the mechanism of transmission can be revealed by comparing the characteristics between different transmission communities.

## Materials and methods

### Ethical review

This study was reviewed and approved by the Ethics Committee of the Zhejiang Provincial Centers for Disease Control and Prevention. All of the data of the individuals were kept confidential. Written informed consent was obtained from all of the patients before diagnosis and reporting the data to the China Information Network System of Disease Prevention and

Control. All of the methods employed in the study were in accordance with the applicable guidelines and regulations.

## Profile of Zhejiang Province

Zhejiang Province is located in southeast China between longitudes 118°E-123°E and latitudes 27°N-32°N. There are two subprovincial cities (Hangzhou and Ningbo) and nine prefecture-level cities, including Wenzhou, Huzhou, Jiaxing, Shaoxing, Jinhua, Zhoushan, Quzhou, Taizhou and Lishui, which cover 90 counties.

## The data collection

Any human COVID-19 case diagnosed in a hospital must be reported through the China Information Network System of Disease Prevention and Control by the medical staff. The data of the COVID-19 cases in Zhejiang Province from January to July 2020 were obtained from this network system. Individual information on cases and deaths was imported, and surveillance information, including demographic characteristics and geographic and temporal distributions, was computed by the system. The definition of reporting cases refers to the 'Diagnosis and treatment of COVID-19 (trial version 7)' [13].

## Knox test

The Knox test was carried out to detect the space-time interactions [14]. In epidemiological terms, this is usually interpreted as a test of the level of spatial clustering of infectious disease at sequential overlapping time intervals. The base Knox test statistic,  $X$ , is the number of pairs of cases that are close in both space and time. The statistic is calculated as:

$$X(s, t) = \sum_{i=2}^N \sum_{j=1}^{i-1} a_{ij}^s a_{ij}^t \quad (1)$$

where  $N$  is the number of cases,  $a_{ij}^s$  is equal to 1 if cases  $i$  and  $j$  are close in space, and 0 otherwise,  $a_{ij}^t$  is equal to 1 if cases  $i$  and  $j$  are close in time and 0 otherwise, and  $s$  and  $t$  represent pre-specified spatial and temporal distances [15, 16]. To identify the serial interval of an infectious disease, the incremental Knox test (IKT) was implemented. The interval Knox statistic is formulated as:

$$IK(s, t) = \sum_{i=2}^N \sum_{j=1}^{i-1} a_{ij}^s b_{ij}^t \quad (2)$$

where  $b_{ij}^t$  is equal to 1 if cases  $i$  and  $j$  occur  $t$  units apart and 0 otherwise. This result indicates which time intervals are within  $t$  at which spatial clustering occurs [15, 16]. This statistic is useful for testing the null hypothesis of no space-time interactions versus the alternative hypothesis of a contagious process [16]. Monte Carlo hypothesis testing, which was suggested by Mantel, is the most commonly used method for this test statistic [17]. In the Monte Carlo test, the temporal distances are randomly permuted with the spatial distances unchanged among the cases to estimate the distribution of the statistic under the null hypothesis of no space-time interaction [16]. This permutation was computed  $M$  times to get  $M$  statistics. At the same time, the  $P$ -value is obtained through Monte Carlo hypothesis testing, by comparing the rank of the maximum likelihood from the real data set with the maximum likelihoods from the

random data sets. The  $P$  value is computed as:

$$\frac{R}{M + 1} \tag{3}$$

where  $R$  is the rank of the real data and  $M$  is the number of simulations. In order for  $P$  to be a ‘good-looking’ number, the number of simulations is set to 999 or some other number ending in 999 such as 1999, 9999 or 99999. That way it is always clear whether to reject or not reject the null hypothesis for typical cut-off values such as 0.05, 0.01 and 0.001. Here the number of simulation was restricted to 999 and the significance was set at  $p < 0.05$  correspondingly.

Based on our knowledge of the COVID-19 transmission pattern and previous cluster analyses of COVID-19 transmission, the IKT is used to examine space-time interactions from 0 meters to 1000 kilometers at time intervals from 1 to 20 days [18]. The spatial distance was divided into 12 intervals. The shortest interval is 10 meters, and the longest interval is 1000 kilometers. The other intervals are 50 meters, 100 meters, 500 meters, 1 kilometer, 5 kilometers, 10 kilometers, 50 kilometers, 100 kilometers, 300 kilometers and 500 kilometers.

Furthermore, an epidemiologically meaningful notion of excess risk (ER) is calculated as the ratio of the observed statistic divided by the permutation mean  $\mu_x(s, t)$ .

$$ER(s, t) = \frac{IK(s, t)}{\mu_x(s, t)} \tag{4}$$

### Transmission network analysis

Networks have become a key approach to understanding systems of interacting objects, unifying the study of diverse phenomena, including biological organisms and human society [19]. The essential element when analyzing the structure, function, and dynamics of biological networks is the identification of communities of related nodes in our study [20, 21]. This approach is greatly enhanced by clustering the links between nodes rather than the nodes themselves. With this method, it is possible for nodes to belong to multiple communities, which in turn reveals the overlapping and nested structure of the network while simultaneously identifying key nodes with membership across several communities [19, 22]. Similarities between links,  $e_{ik}$  and  $e_{jk}$ , that share a node ( $k$ ) are assigned using the Jaccard coefficient, which is computed as:

$$S(e_{ik}, e_{jk}) = \frac{|n+(i) \cap n+(j)|}{|n+(i) \cup n+(j)|} \tag{5}$$

where  $n+(i)$  refers to the first-order node neighborhood of node  $i$ . After assigning pairwise similarities to all of the links in the network, the links are hierarchically clustered [20]. The resulting dendrogram is cut at a point that maximizes the density of the links within the clusters after normalizing against the maximum and minimum numbers of links possible in each cluster [20].

The index of node centrality is used to measure the importance of a node in a network [21]. The node centrality can be written as:

$$C_c(i) = \sum_{i \in j}^N \left( 1 - \frac{1}{m} \sum_{i \in j \cap k}^m S(j, k) \right) \tag{6}$$

where the main sum is over the  $N$  communities to which node  $i$  belongs, and  $S(j, k)$  refers to

the similarity between communities  $j$  and  $k$ , calculated as the Jaccard coefficient for the number of shared nodes between each community pair, and this is averaged over the  $m$  communities paired with community  $j$  and in which node  $i$  jointly belongs.

Furthermore, the modularity of communities—the relative number of links within the community versus links outside of the community—is computed. The modularity of community  $i$  can be written as:

$$M_i = \left( \frac{e_w(i)}{n_i(n_i - 1)/2} \right) * \left( \frac{e_b(i)}{n_i \bar{d}} \right)^{-1} \quad (7)$$

where  $e_w(i)$  and  $e_b(i)$  are the number of links within and without community  $i$ , respectively,  $n_i$  is the number of nodes in community  $i$ , and  $\bar{d}$  is the average degree of nodes in the network.

## Results

### The temporal trend of the COVID-19 epidemic

From January to July 2020, there were a total of 1475 COVID-19 cases and only 1 fatal case reported in Zhejiang Province, China. The fatality rate was 0.068%. The first case was diagnosed on January 5<sup>th</sup>. The number of cases increased markedly starting on January 15<sup>th</sup>. The highest daily number of cases was 94 on January 26<sup>th</sup>. The incidence subsequently decreased from the beginning of February. There were 951, 398, 77, 40, 4, 3 and 2 cases identified in each month during this period. Most of the cases occurred before February 15<sup>th</sup>, which accounted for 90.10%. There were 1368 cases of infection inside mainland China. There were 107 cases infected outside of mainland China. The peak of the epidemic before March was attributed to mainland infections. After that, the imported cases increased and became the main risk in Zhejiang Province (Fig 1).

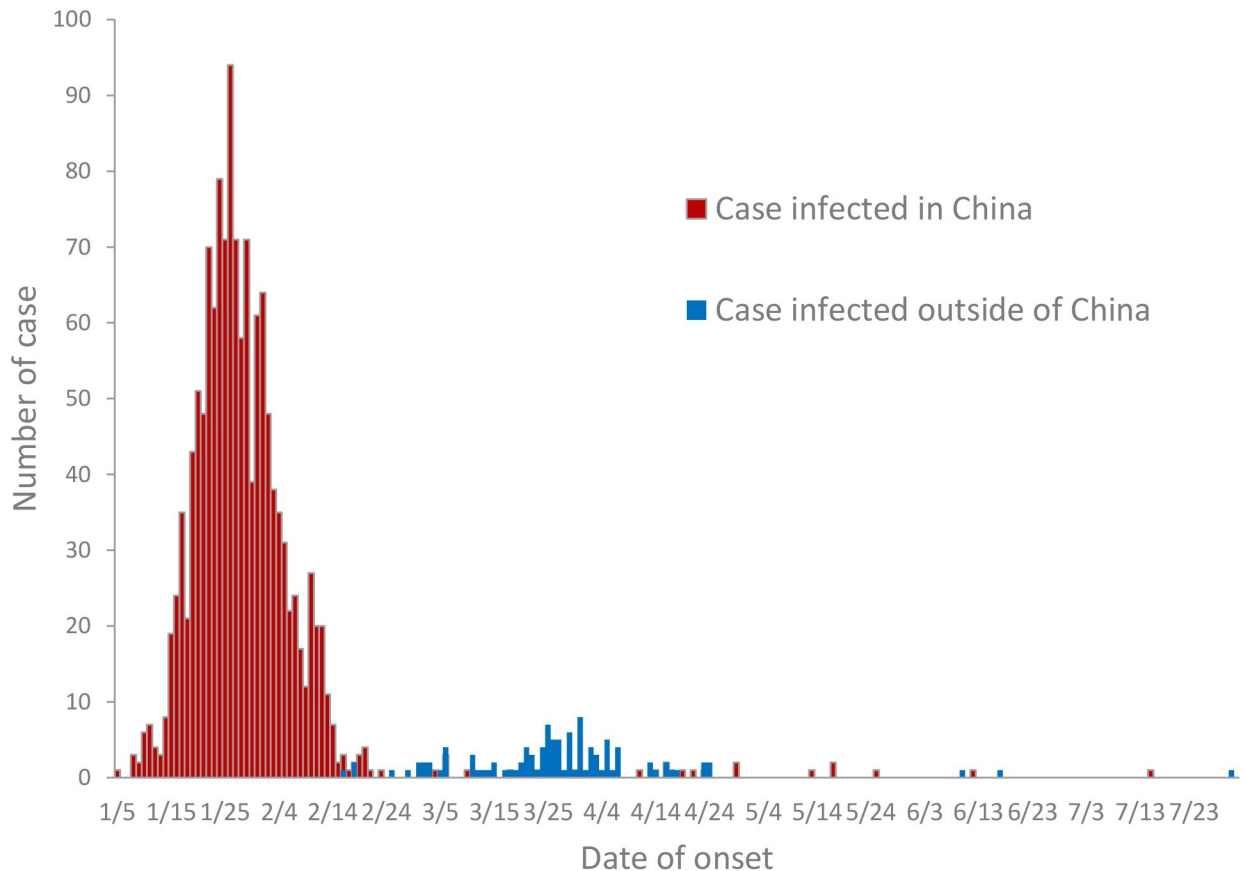
### Spatiotemporal patterns analysis

According to the characteristics of the temporal trend of this epidemic, only the mainland cases that occurred before March were analyzed for the spatiotemporal pattern.

Imported cases infected overseas and sporadic endemic cases after February were not included for analysis. The total number of cases used in the analysis was 1344.

The median distance between the case pairs was 161.2 kilometers. The distance of most of the case pairs was within 200 kilometers. The time distances were divided into 20 intervals. The interval was 1 day. The median time interval between the case pairs was 7.23 days. The time interval of most of the case pairs was within 20 days (Fig 2).

The excess risk (ER) values are reported in Table 1. Most of the values are significantly different from zero at the  $\alpha = 0.05$  level. It should be noted that the values for the distance interval beyond 100 kilometers are not significant. The values from 1 to 5 kilometers at time intervals from 19 to 20 days are also not significant. The risk areas show a strong heterogeneity. The highest levels of elevated risk occurred at intervals of 0 and 1 day and within 10 m. The areas at short distances within 1 kilometer and at brief periods within 5 days presented relatively high risk ( $ER > 1.5$ , most of the values were greater than 2). Beyond these space-time limits, the excess risk values rapidly decreased. When the distance between two cases increased, the ER values were smaller. At the same time, the ER values also gradually decreased as the onset interval between the two cases was prolonged. However, particular patterns, such as temporal periodicity of the sharp increase in risk, were not observed, the serial interval of this epidemic can't be confirmed (Fig 3).



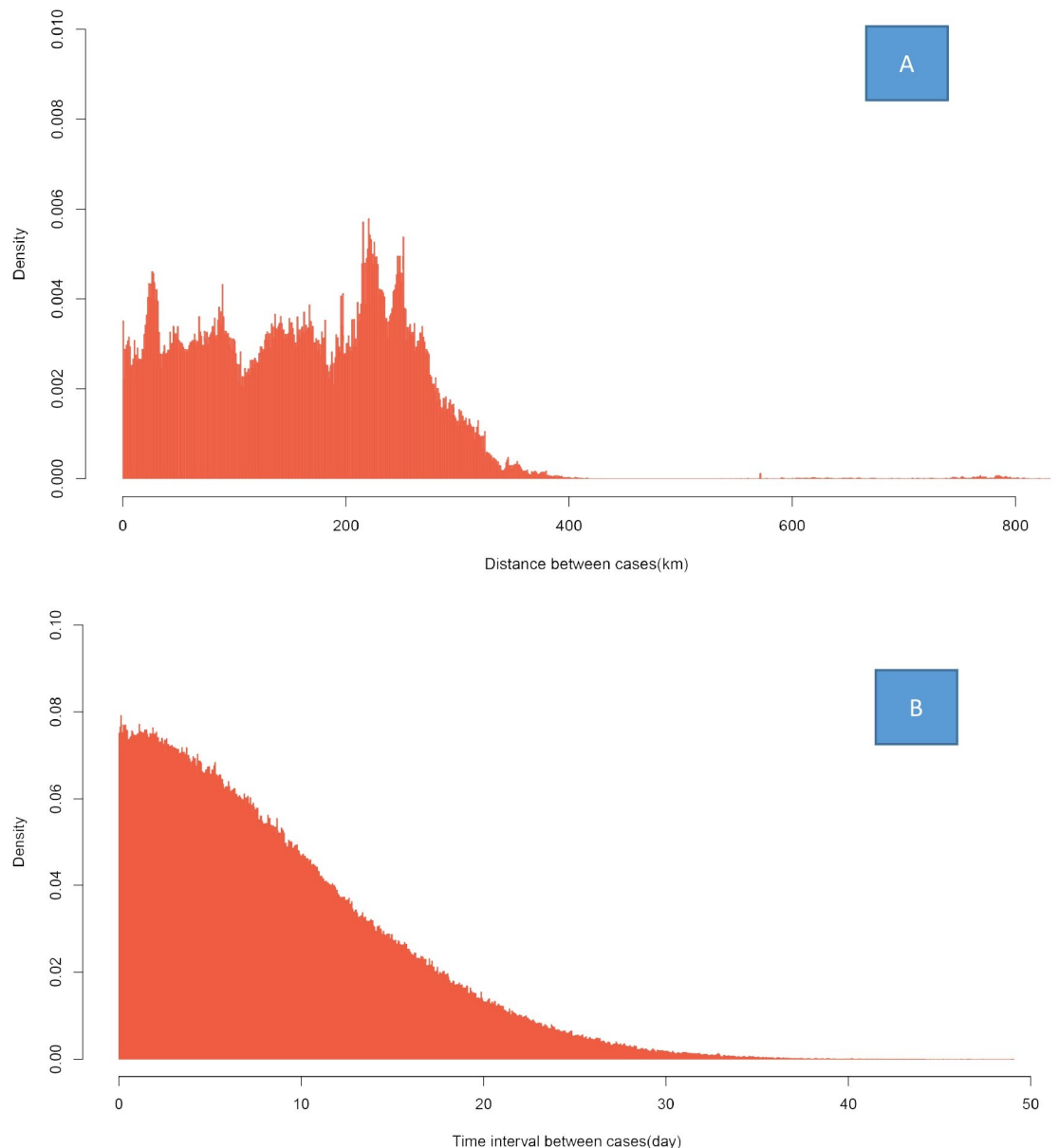
**Fig 1. Onset of COVID-19 from January to July 2020 in Zhejiang Province, China.**

<https://doi.org/10.1371/journal.pone.0257587.g001>

### Transmission network analysis

There were 198 clusters from January to July 2020 in Zhejiang Province. The number of infections in most of the clusters (51.51%) was only 2. The number of cases in seven clusters was more than 15. Furthermore, the transmission relationship among the cases was explicit only in four clusters; thus, transmission networks were built for these four clusters. The numbers of affected cases in these four clusters were 66, 23, 23 and 19, respectively.

In the first step, dendrograms were built by the link clustering algorithm. This plot shows the link communities that result from cutting the dendrogram at a point where the partition density is maximized (Fig 4). In this step, each cluster was divided into several subcommunities. It will produce the network of the clusters. In cluster one, which included 66 infections, 104 edges and 66 nodes were identified in this network. This cluster was finally divided into 12 subcommunities, and the largest subcommunities contained 34 nodes (Fig 4A). In cluster two, which included 23 infections, there were 38 edges and 23 nodes identified in this network. This cluster was finally divided into 9 subcommunities, and the largest subcommunities contained 5 nodes (Fig 4B). In cluster three, which included 23 infections, 68 edges and 23 nodes were identified in this network. This cluster was finally divided into 6 subcommunities, and the largest subcommunities contained 10 nodes (Fig 4C). In cluster four, which included 19 infections, there were 33 edges and 19 nodes identified in this network. This cluster was finally divided into 4 subcommunities, and the largest subcommunities contained 5 nodes (Fig 4D).



**Fig 2. The density distribution of the interval between two COVID-19 cases during the epidemic.** (A) The density of the distribution of distance between two cases. (B) The density of the distribution of the onset interval between two cases.

<https://doi.org/10.1371/journal.pone.0257587.g002>

In the second step, networks were built, and the relationship between the cases was plotted (Fig 5). In cluster one, the maximum value of node centrality was identified in case A8, with a centrality of 2.9474 (Fig 5A). The node centralities of cases A4, A9, A10 and A3 were 2.9444, 2.9444, 2.9444 and 2.9459, respectively. The node centrality of these cases mentioned above was almost equal, which suggested that these cases were similarly important for transmission in cluster 1. The maximum value of the subcommunities to which these cases belong is 2. In cluster two, the maximum value of node centrality was identified in case B2, with a centrality of 4.3706, followed by B4 (3.5644), B1 (2.6667) and B8 (2.6667). The maximum number of subcommunities to which these cases belong is 4 (Fig 5B). Case B2 is the most important transmission node in cluster 2. In cluster three, the maximum value of node centrality was identified

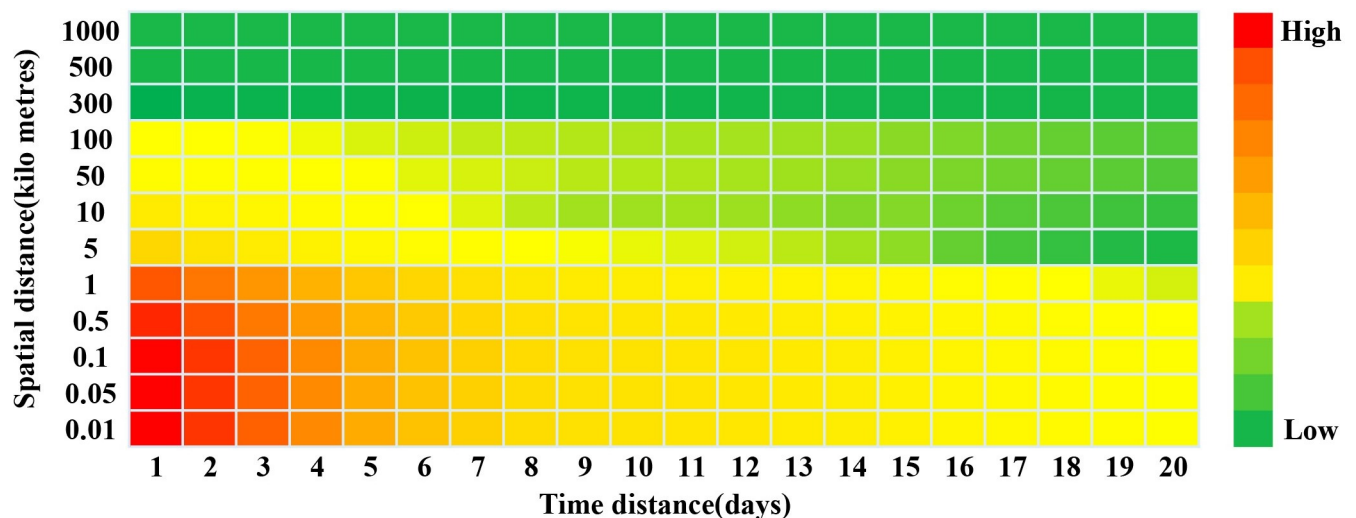
**Table 1.** The ER values for the epidemic of COVID-19 in Zhejiang Province using a range of distances and time intervals.

Time interval(days)	Distance(kilometers)											
	0.01	0.05	0.1	0.5	1	5	10	50	100	300	500	1000
1	3.11	3.10	3.09	2.81	2.42	1.38	1.21	1.08	1.05	0.99	1.00	1.00
2	2.69	2.68	2.68	2.46	2.16	1.28	1.15	1.07	1.05	1.00	1.00	1.00
3	2.34	2.33	2.32	2.15	1.90	1.21	1.12	1.06	1.05	1.00	1.00	1.00
4	2.01	2.00	2.00	1.87	1.68	1.16	1.09	1.06	1.05	1.00	1.00	1.00
5	1.73	1.73	1.73	1.64	1.50	1.12	1.07	1.05	1.04	1.00	1.00	1.00
6	1.55	1.55	1.55	1.49	1.39	1.09	1.05	1.04	1.04	1.00	1.00	1.00
7	1.43	1.43	1.43	1.39	1.30	1.07	1.04	1.04	1.04	1.00	1.00	1.00
8	1.34	1.34	1.34	1.31	1.24	1.06	1.04	1.04	1.04	1.00	1.00	1.00
9	1.29	1.30	1.30	1.27	1.21	1.05	1.03	1.03	1.03	1.00	1.00	1.00
10	1.28	1.28	1.28	1.26	1.20	1.05	1.03	1.03	1.03	1.00	1.00	1.00
11	1.27	1.27	1.27	1.24	1.19	1.04	1.03	1.03	1.03	1.00	1.00	1.00
12	1.25	1.25	1.25	1.22	1.17	1.04	1.03	1.03	1.03	1.00	1.00	1.00
13	1.22	1.22	1.22	1.20	1.15	1.04	1.03	1.03	1.03	1.00	1.00	1.00
14	1.20	1.20	1.20	1.18	1.13	1.03	1.02	1.03	1.03	1.00	1.00	1.00
15	1.17	1.17	1.17	1.15	1.11	1.03	1.02	1.02	1.02	1.00	1.00	1.00
16	1.14	1.14	1.14	1.12	1.09	1.02	1.02	1.02	1.02	1.00	1.00	1.00
17	1.11	1.12	1.11	1.10	1.07	1.01	1.01	1.02	1.02	1.00	1.00	1.00
18	1.09	1.09	1.09	1.08	1.05	1.01	1.01	1.02	1.02	1.00	1.00	1.00
19	1.08	1.08	1.08	1.07	1.05	1.00	1.01	1.01	1.01	1.00	1.00	1.00
20	1.06	1.07	1.07	1.06	1.04	1.00	1.01	1.01	1.01	1.00	1.00	1.00

The number in italic font indicates that the ER value is significantly different from zero at the alpha = 0.05 level.

<https://doi.org/10.1371/journal.pone.0257587.t001>

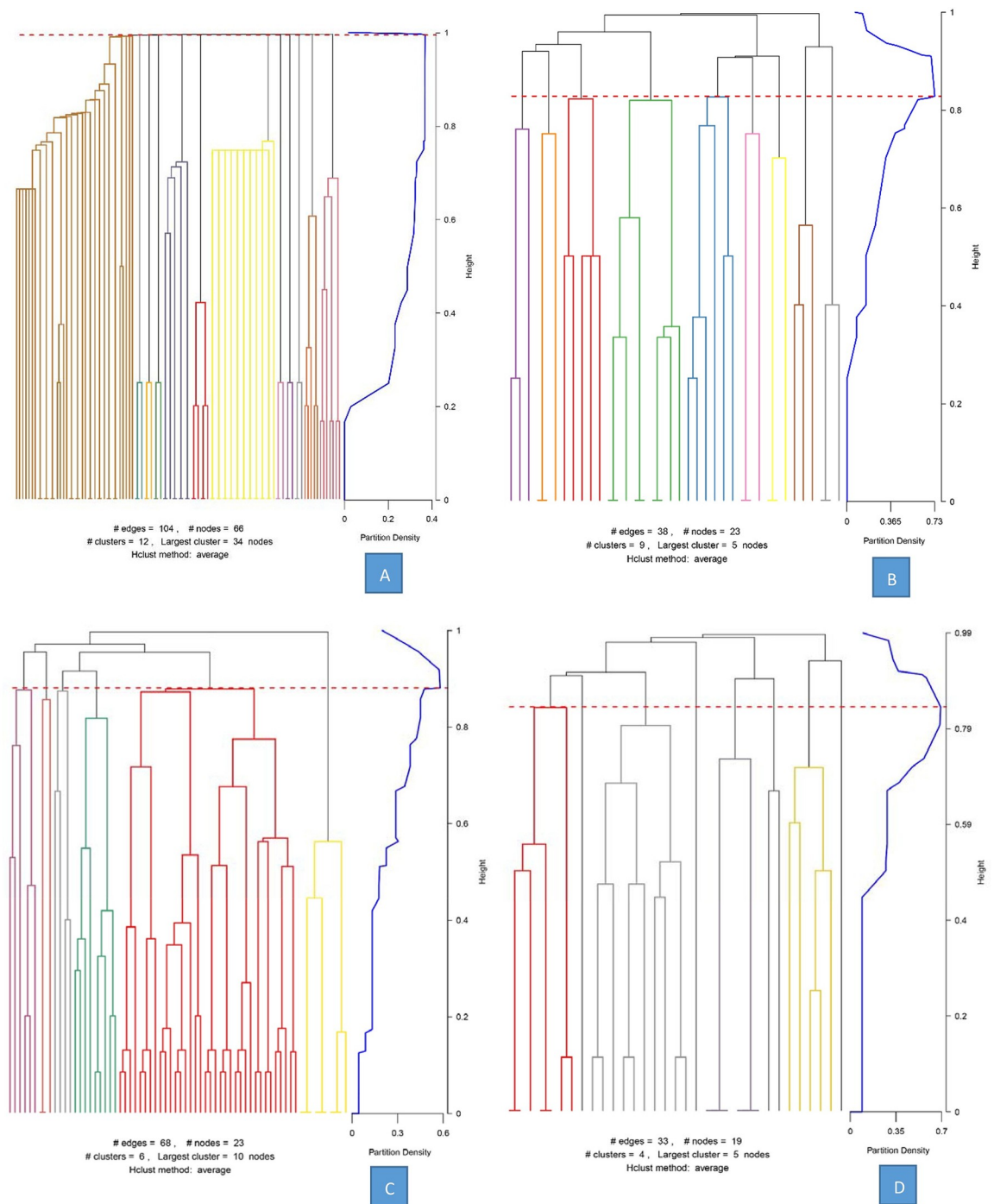
for case C5, with a centrality of 4.1080, followed by C15 (2.9500) and C16 (2.9500). Case C5 also had the maximum value of subcommunities (4), which means that C5 is the most important node in cluster 3 (Fig 5C). In cluster four, cases D1 and D5 had the maximum value of node centrality (2.7500), followed by cases D10 (2.5000) and D12 (2.5000). The maximum number of subcommunities to which these cases belong is 2 (Fig 5D).



**Fig 3.** Heatmap of the excess risk due to space-time interactions for the COVID-19 epidemic in Zhejiang Province from January to July 2020.

<https://doi.org/10.1371/journal.pone.0257587.g003>

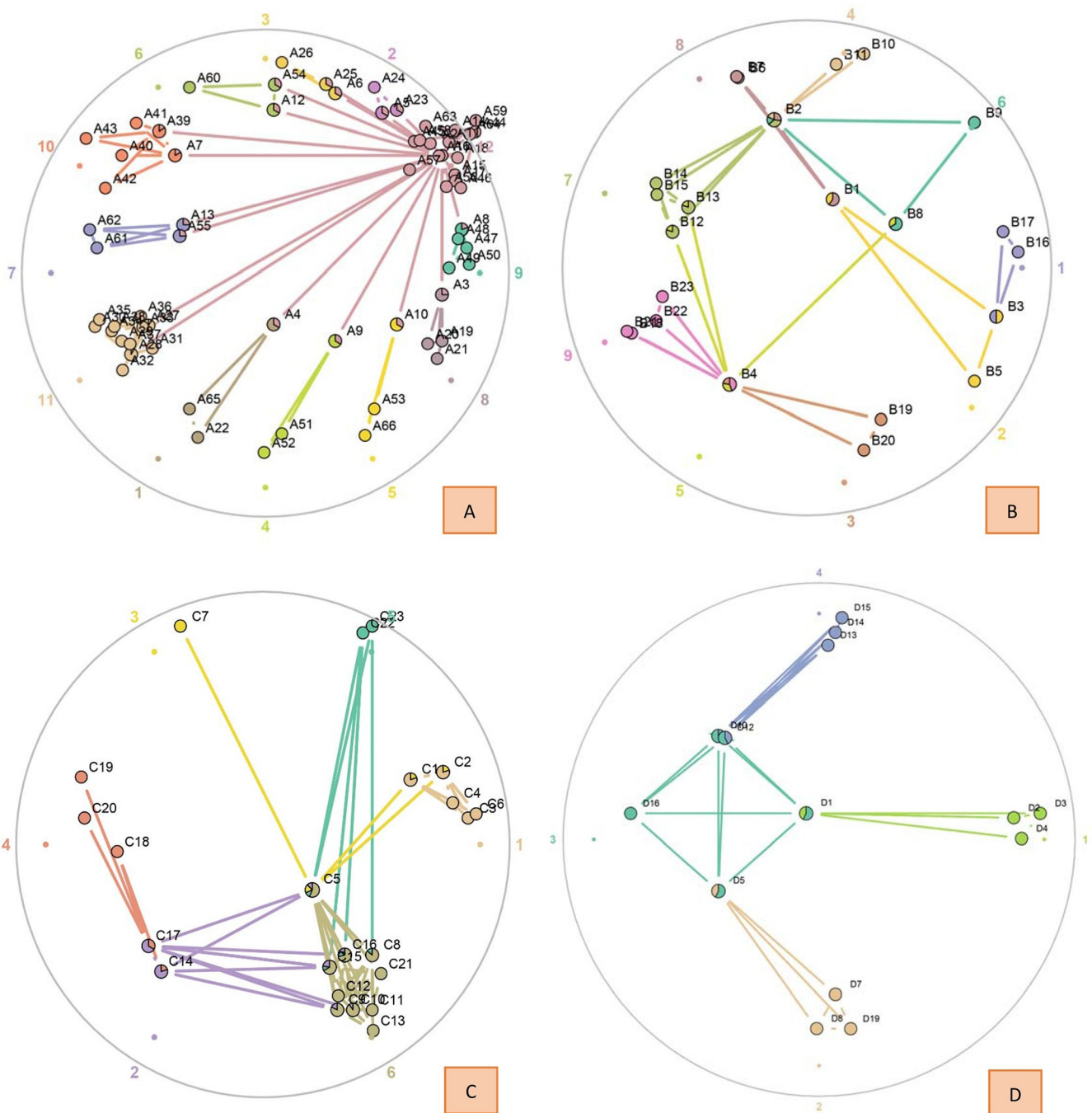




**Fig 4. The dendrogram of link clustering for the clusters.** (A) The dendrogram of cluster 1. (B) The dendrogram of cluster 2. (C) The dendrogram of cluster 3. (D) The dendrogram of cluster 4.

<https://doi.org/10.1371/journal.pone.0257587.g004>

In the third step, the modularity of the subcommunity was calculated. There was obvious heterogeneity in the modularity of the subcommunities. In cluster one, the minimum value of the modularity of the subcommunities was 0.13 for subcommunity 12. This result suggested that the number of case links related to other subcommunities is greater than the links inside



**Fig 5. The network of transmission between cases.** (A) The network of cluster 1. (B) The network of cluster 2. (C) The network of cluster 3. (D) The network of cluster 4. The fraction of the total number of edges that a node has in each community is depicted using a pie chart. The colorful numbers outside of the circle indicate the serial number of each subcommunity.

<https://doi.org/10.1371/journal.pone.0257587.g005>

this subcommunity. In cluster two, the minimum value of the modularity of subcommunities was 0.44 for subcommunity 5. In cluster three, the minimum value of the modularity of subcommunities was 0.45 for subcommunity 2. Furthermore, the minimum value of the modularity of subcommunities was 1.09 for subcommunity 3 in cluster four (Table 2).

Table 2. The modularity of the subcommunity of the COVID-19 clusters in Zhejiang Province.

Cluster	Subcommunity											
	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12
No.1	9.45	4.73	4.73	9.45	9.45	4.73	6.30	10.51	11.03	5.67	6.02	0.13
No.2	4.96	1.98	1.42	1.10	0.44	0.99	1.65	1.10	1.98	-	-	-
No.3	14.78	0.45	0.59	1.31	0.47	2.96	-	-	-	-	-	-
No.4	2.32	2.32	1.09	1.52	-	-	-	-	-	-	-	-

<https://doi.org/10.1371/journal.pone.0257587.t002>

## Discussion

According to the temporal characteristics of the seven-month epidemic, we found that COVID-19 was brought under control over a short period in Zhejiang Province. Actually, this disease was incorporated as a notifiable disease in the Infectious Disease Law and Health and Quarantine Law in China on January 20<sup>th</sup> [2]. The first case in Zhejiang Province was reported on January 21<sup>st</sup>, and then the control measures were strengthened after that date [23]. Therefore, took nearly one month from January 21<sup>st</sup> to February 15<sup>th</sup> to eliminate the local epidemic. The result of Knox test suggested that the transmission pattern showed short-range and short-term risk. This trend was similar to the epidemic curve in other provinces of China [24]. There were no effective pharmacological interventions or vaccines available at that time, so reducing the rate of infection was a priority and the prevention of infection was the best approach to flatten the epidemic curve [25]. According to the experience of disease control measures implemented in Zhejiang Province and other provinces of China, the most effective approach to preventing disease transmission is expanding the physical or social distance and the use of face masks, which allowed the epidemic to be slowed down over a short period [24]. The result of Zhejiang Province give an strong evidence that timely non-pharmaceutical interventions are essential to quickly control the epidemic. The fatality rate in Zhejiang Province was lower than the average rate in China [24]. The main reason may be the high urbanization rate and GDP (gross domestic product) per capita in Zhejiang Province (ranked fourth in China with 15,755 dollars in 2019). These also reflect high-level medical conditions that led to a low fatality rate. After containing the endemic infections, imported infections from outside of mainland China became another challenge. The subsequent preparedness and response to this pandemic should continue in all countries/territories/areas worldwide [26].

The spatiotemporal analysis of the Knox test has been widely applied to dengue fever, cerebrospinal meningitis, and hand, foot, and mouth disease [11, 12, 27, 28]. This method helped us to identify the boundaries in space and time of the maximum contagious disease transmission. It also allowed us to identify high-risk areas with precise distances and time intervals. Furthermore, temporal periodicity can be detected during continuous transmission conditions [11, 12]. In our study, the effect of a spatiotemporal interaction was obvious and exhibited interval heterogeneity. The principal finding in our study is that the areas with the highest level of elevated risk were all within a distance of 1 kilometer. The elevated risk of clustering in this period (January 1<sup>st</sup> to March 31<sup>st</sup>) of the epidemic in Zhejiang Province was relatively high within 5 days. The highest risk value occurred at intervals of 0 and 1 day and gradually decreased as the onset interval was prolonged. The characteristics of transmission in these periods showed short-range risk, and the risk of remote transmission was relatively low. This result suggested that COVID-19 infection from one case to other individuals is inclined to occur in nearby persons. According to previous studies, the main reason for this short-range high risk is that family clusters accounted for a significant number of infections during that period of the COVID-19 epidemic in China [29–31]. Most of the cases came from the same

family that lived in one house or near each other. Furthermore, the timing of the outbreak, prior to the annual Chinese Lunar New Year holiday, resulted in a higher frequency of having dinner parties among family members [4]. The cases who returned from Wuhan city took part in family dinners and then transmitted the disease to other family members who usually became the secondary case.

As the intervention measures were only implemented in late January, at that time, the social distance was lengthened and social transmission was interrupted since the risk of remote range transmission was lower than that of short range transmission. At the same time, the characteristics of the heatmap of the excess risk also indicated that the epidemic curve declined rapidly over one month. Another reason may be that individuals often wear face masks to protect themselves against infection by this virus when they are outside. However, they generally do not use them at home or in their own neighborhood, which leads to more short range transmission. Unlike previous studies on dengue fever, temporal periodicity was not identified in our study. The risk values decreased as the onset interval was prolonged. The high-risk areas were within one serial interval. The reason for this should be attributed to timely interventions to block the spread of the disease at the social level and to confine most cases within family clusters. This result also suggested that the timely non-pharmaceutical interventions are essential to quickly control the epidemic.

According to the results of the transmission network analysis, we found that the importance to the cluster of each case was different. A few cases became important transmission nodes due to their high value of node centrality. This result suggested that these cases played a key role in the cluster, and some of them, such as case C5, may be a superspreader who transmits the virus to more than 10 persons. Why did these cases become key nodes in the transmission network? Careful investigation should be launched in the future to address this problem. Nevertheless, based on previous studies, the factors for a case becoming a source of infection of COVID-19 include the occupation of the case, their behavior and living habits, the range of their activity, their state of illness and so on [32]. Furthermore, an increasing number of studies have indicated that 30~60% of infections are asymptomatic and mild cases [33, 34]. These asymptomatic and mild cases can also transmit the virus to other people [33–35]. Because it is difficult to distinguish these cases from healthy persons, the chance of spreading the virus by asymptomatic and mild cases would be greater than that of severe patients. Therefore, these cases could also become key nodes in the transmission network.

In addition to the key nodes in the network, an ability to spread the virus to other subcommunities or subclusters was detected. The heterogeneity in the modularity of subcommunities suggested that a group of cases played a more important role in the network to spread the virus to other persons and lead to secondary infections. Based on the importance of transmission within the cluster for each case, we should focus on the key patients (key nodes in the network) in complex conditions or situations of limited control resources. When key patients are brought under control, the main transmission chain can be interrupted, which will help to mitigate the epidemic.

Several limitations should be noted within our study. First, in the spatiotemporal analysis, the distance between the cases was computed by their usual living address. However, the actual site where the case contracted the virus may be a working location or a public place, such as mass transit and restaurants. Therefore, the risk areas of the distance interval may be biased. Second, our sample included cases of COVID-19 reported from a passive surveillance system. Future studies should consider mild and asymptomatic cases that do not seek medical care. Third, data regarding the population characteristics, including socioeconomic status, human activities and clinical progress, were not collected. Consequently, the difference between the case as the key node and the other cases in the transmission network could not be compared.

To achieve an accurate forecast and timely intervention for the possible key patient in transmission, future studies should incorporate these factors into the analysis.

In conclusion, COVID-19 was brought under control over a short period in Zhejiang Province. The fatality rate in Zhejiang Province was lower than the average rate in China. Imported infections from outside of mainland China then became a new challenge. Based on the Knox test, the effect of spatiotemporal interactions exhibited interval heterogeneity. The areas with the highest level of elevated risk were within a distance of 1 kilometer and a period of 5 days. The characteristics of transmission showed short-range and short-term risk. The pattern of non-temporal periodicity suggested that timely non-pharmaceutical interventions are essential to quickly control the epidemic. Using transmission network analysis, the importance to the cluster of each case was detected, and the key patients were identified. It is suggested that we should focus on key patients in complex conditions or in situations with limited control resources.

## Supporting information

**S1 File. The database of the COVID-19 cases in Zhejiang Province.**  
(CSV)

## Acknowledgments

We thank all of the clinical personnel for investigating and reporting the information about the COVID-19 cases. We very much appreciate the 90 county-level CDCs for their investigation and correction of the epidemiological information.

## Author Contributions

**Conceptualization:** Junfen Lin.

**Data curation:** Qinbao Lu, Zheyuan Ding.

**Formal analysis:** Haocheng Wu.

**Investigation:** Zheyuan Ding.

**Validation:** Chen Wu, Ming Xue.

**Writing – original draft:** Haocheng Wu.

**Writing – review & editing:** Chen Wu.

## References

1. Zhu N, Zhang DY, Wang WL, Li XW, Yang B, Song JD, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med.* 2020 Feb 20; 382(8):727–733. <https://doi.org/10.1056/NEJMoa2001017> PMID: 31978945
2. Li Q, Guan XH, Wu P, Wang XY, Zhou L, Tong YQ, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med.* 2020 Mar 26; 382(13):1199–1207. <https://doi.org/10.1056/NEJMoa2001316> PMID: 31995857
3. Vincent J, Munster Marion, Koopmans, Neeltje van Doremalen, Debby van Riel, Emmie de Wit. A Novel Coronavirus Emerging in China -Key Questions for Impact Assessment. *N Engl J Med.* 2020 Feb 20; 382(8):692–694. <https://doi.org/10.1056/NEJMp2000929> PMID: 31978293
4. Huang Y, Wu QQ, Wang P, Xu Y, Wang L, Zhao YS, et al. Measures Undertaken in China to Avoid COVID-19 Infection: Internet-Based, Cross-Sectional Survey Study. *J Med Internet Res.* 2020 May 12; 22(5):e18718. <https://doi.org/10.2196/18718> PMID: 32396516

5. World Health Organization. Coronavirus disease(COVID-19) Situation Report-198[EB/OL]. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200805-covid-19-sitrep-198.pdf?sfvrsn=f99d1754\\_2\\_5](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200805-covid-19-sitrep-198.pdf?sfvrsn=f99d1754_2_5) August,2020.
6. Special Expert Group for Control of the Epidemic of Novel Coronavirus Pneumonia of the Chinese Preventive Medicine Association. An update on the epidemiological characteristics of novel coronavirus pneumonia(COVID-19). *Chin J Epidemiol.*2020 Feb 10; 41(2):139–144. <https://doi.org/10.3760/cma.j.issn.0254-6450.2020.02.002> PMID: 32057211
7. Jasper Fuk-Woo Chan Yuan SF, Kok Kin-Hang, Kelvin Kai-Wang To Hin Chu, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person to person transmission: a study of a family cluster. *Lancet.*2020 Feb 15; 395(10223):514–523. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9) PMID: 31986261
8. Mahase Elisabeth. China coronavirus: mild but infectious cases may make it hard to control outbreak, report warns. *BMJ.*2020 Jan 28; 368:m325. <https://doi.org/10.1136/bmj.m325> PMID: 31992570
9. Coronavirus disease (COVID-19): How is it transmitted?[EB/OL]. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted>
10. Holshue Michelle L, Chas DeBolt, Lindquist Scott, Lofy Kathy H, John Wiesman, Bruce Hollianne, et al. First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med.*2020 Mar 5; 382(10):929–936. <https://doi.org/10.1056/NEJMoa2001191> PMID: 32004427
11. Tran Annelise, Deparis Xavier, Dussart Philippe, Morvan Jacques, Rabarison Patrick, Remy Franck, et al. Dengue Spatial and Temporal Patterns, French Guiana, 2001. *Emerg Infect Dis.* 2004 Apr; 10(4):615–21. <https://doi.org/10.3201/eid1004.030186> PMID: 15200850
12. Fine Paul E M. The interval between successive cases of an infectious disease. *Am J Epidemiol.*2003 Dec 1; 158(11):1039–47. <https://doi.org/10.1093/aje/kwg251> PMID: 14630599
13. National Health Commission. Diagnosis and treatment of COVID-19(trial version 8 revised edition)[EB/OL].(2021-04-14)[2021-04-15]. <http://www.nhc.gov.cn/yzygj/s7653p/202104/7de0b3837c8b4606a0594aeb0105232b.shtml>
14. Knox EG. Detection of space-time interactions. *Appl Stat.*1964.13:25–29. <https://doi.org/10.2307/2985220>
15. Aldstadt Jared. An incremental Knox test for the determination of the serial interval between successive cases of an infectious disease. *Stoch Environ Res Risk Assess.* 2007 Apr 17.21:487–500. <https://doi.org/10.1007/s00477-007-0132-3>
16. Aldstadt Jared, Yoon In-Kyu, Tannitisupawong Darunee, Jarman Richard G, Thomas Stephen J, Gibbons Robert V, et al. Space-time analysis of hospitalized dengue patients in rural Thailand reveals important temporal intervals in the pattern of dengue virus transmission. *Trop Med Int Health.*2012 Sep; 17(9):1076–85. <https://doi.org/10.1111/j.1365-3156.2012.03040.x> PMID: 22808917
17. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 1967 Feb; 27(2):209–20.PMID: 6018555
18. Sun WW, Ling F, Pan J R, Cai J, Miao Z P, Liu S Let al. Epidemiological characteristics of COVID-19 family clustering in Zhejiang Province. *Chin J Prev Med.*2020 Jun 6; 54(6):625–629. <https://doi.org/10.3760/cma.j.cn112150-20200227-00199> PMID: 32171192
19. Ahn Yong-Yeol, James P Bagrow Sune Lehmann. Link communities reveal multiscale complexity in networks. *Nature.*2010 Aug 5; 466(7307):761–4. <https://doi.org/10.1038/nature09182> PMID: 20562860
20. Kalinka Alex T, Pavel Tomancak. linkcomm: an R package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. *Bioinformatics.* 2011 Jul 15; 27(14):2011–2. <https://doi.org/10.1093/bioinformatics/btr311> PMID: 21596792
21. Radicchi Filippo, Castellano Claudio, Cecconi Federico, et al. Dening and identifying communities in networks. *Proc. Natl Acad. Sci USA.*2004 Apr; 101(9), 2658–2663. <https://doi.org/10.1073/pnas.0400054101> PMID: 14981240
22. Evans T S, Lambiotte R. Line graphs, link partitions and overlapping communities. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2009 Jul; 80(1 Pt 2):016105. <https://doi.org/10.1103/PhysRevE.80.016105> PMID: 19658772
23. Lin JF, Wu MN, Wu HC, Zhang T, Wu C, Li FD. Epidemiological characteristics of coronavirus disease 2019 in Zhejiang Province. *Preventive Medicine.* 2020 Mar 19; 32(3): 217–222. <https://doi.org/10.1136/bmj.m1443> PMID: 32317267
24. Epidemiology Working Group for NCIP Epidemic Response. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases(COVID-19) in China. *Chin J Epidemiol.*2020 Feb 10; 41(2):145–151. <https://doi.org/10.3760/cma.j.issn.0254-6450.2020.02.003> PMID: 32064853

25. Chu Derek K, Akl Elie A, Stephanie Duda, Karla Solo, Sally Yaacoub, Schünemann Holger Jet al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet*.2020 Jun 27; 395(10242):1973–1987. [https://doi.org/10.1016/S0140-6736\(20\)31142-9](https://doi.org/10.1016/S0140-6736(20)31142-9) PMID: 32497510
26. Guo Y, Huang Y M, Huang J, Jin Y Z, Jiang W, Liu P L, et al. COVID-19 pandemic: global epidemiological trends and China's subsequent preparedness and responses. *Chin J Epidemiol*.2020 May 10; 41(5):642–647. <https://doi.org/10.3760/cma.j.cn112338-20200301-00222> PMID: 32164401
27. Liu QL, Li XS, Feng ZJ, Ma JQ. Study on the application of Knox method to temporal-spatial cluster for infectious disease. *Chin J Epidemiol*.2007 Aug; 28(8): 802–5. PMID: 18080571
28. Chen SY, Sun LM, Tan XH, Xie RS, Lin HL, Liu T, et al. Space-time interaction of hand, foot, and mouth disease in fine scale in Yuexiu District, Guangzhou. *South China J Prev Med*.2017 May 8; 43(2):106–111. <https://doi.org/doi:10.13217/j.scjpm.2017.0106>
29. Gan H, Zhang Y, Yuan M, Wu XY, Liu ZR, Liu M, et al. Epidemiological analysis on 1052 cases of COVID-19 in epidemic clusters. *Chin J Epidemiol*.2020 Jul 10; 41(7):1004–1008. <https://doi.org/10.3760/cma.j.cn112338-20200301-00223> PMID: 32213270
30. Yang HY, Xu J, Li Y, Liang X, Jin YF, Chen SY, et al. The preliminary analysis on the characteristics of the cluster for the COVID-19. *Chin J Epidemiol*.2020 May 10; 41(5):623–628. <https://doi.org/10.3760/cma.j.cn112338-20200223-00153> PMID: 32145716
31. Dong XC, Li JM, Bai JY, Liu ZQ, Zhou PH, Gao L, et al. Epidemiological characteristics of confirmed COVID-19 cases in Tianjin. *Chin J Epidemiol*.2020 May 10; 41(5):638–641. <https://doi.org/10.3760/cma.j.cn112338-20200221-00146> PMID: 32164400
32. Gao WJ, Li LM. Advances on presymptomatic or asymptomatic carrier transmission of COVID-19. *Chin J Epidemiol*.2020 Apr 10; 41(4):485–488. <https://doi.org/10.3760/cma.j.cn112338-20200228-00207> PMID: 32141279
33. Gao WJ, Zheng K, Ke J, Li LM. Advances on the asymptomatic infection of COVID-19. *Chin J Epidemiol*.2020 Jul 10; 41(7):990–993. <https://doi.org/10.3760/cma.j.cn112338-20200404-00514> PMID: 32741160
34. Qiu Jane. Covert coronavirus infections could be seeding new outbreaks. *Nature*.2020 Mar 20. <https://doi.org/10.1038/d41586-020-00822-x> PMID: 32203376
35. Wu ZY. Contribution of asymptomatic and pre-symptomatic cases of COVID-19 in spreading virus and targeted control strategies. *Chin J Epidemiol*.2020 Jun 10; 41(6):801–805. <https://doi.org/10.3760/cma.j.cn112338-20200406-00517> PMID: 32274917