Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The construction and visualization of the transmission networks for COVID-19: A potential solution for contact tracing and assessments of epidemics

Caiying Luo
West China School of Public Health and West China Fourth Hospital, Sichuan University
Yue Ma
West China School of Public Health and West China Fourth Hospital, Sichuan University
Pei Jiang
West China School of Public Health and West China Fourth Hospital, Sichuan University
Tao Zhang
West China School of Public Health and West China Fourth Hospital, Sichuan University
Fei Yin (scupbyff@163.com)
West China School of Public Health and West China Fourth Hospital, Sichuan University

Research Article

Keywords: COVID-19, Epidemiological Data, Transmission Network, Visualization Technique, Framework, Evolving Epidemiology, Global Response

Posted Date: June 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32972/v1

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 at Scientific Reports on April 21st, 2021. See the published version at https://doi.org/10.1038/s41598-021-87802-x.

Title Page

The construction and visualization of the transmission networks for COVID-19: A potential solution for contact tracing and assessments of epidemics

Caiying Luo, M.D. [1, #], Yue Ma, Ph.D. [1,#], Pei Jiang, M.D. [1], Tao Zhang, Ph.D. [1], Fei Yin, Ph.D. [1,*]

1: West China School of Public Health and West China Fourth Hospital, Sichuan University.

#: These authors contributed equally to this work.

*: Author for correspondence: Email: scupbyff@163.com

Abstract

The WHO described coronavirus disease 2019 (COVID-19) as a pandemic due to the speed and scale of its transmission. Without effective interventions, the rapidly increasing number of COVID-19 cases would greatly increase the burden of clinical treatments. Identifying the transmission sources and pathways is of vital importance to block transmission and allocate limited public health resources. According to the relationships among cases, we constructed disease transmission network graphs for the COVID-19 epidemic through a visualization technique based on individual reports of epidemiological data. We proposed an analysis strategy of the transmission network with the epidemiological data in Tianjin and Chengdu. The transmission networks showed different transmission characteristics. In Tianjin, an imported case can produce an average of 2.9 secondary infections and ultimately produce up to 4 generations of infections, with a maximum of 6 cases generated before being identified. In Chengdu, 45 noninformative cases and 24 cases with vague exposure information made it difficult to provide accurate information by the transmission network. The proposed analysis framework of visualized transmission networks can trace the transmission source and contacts, assess the current situation of transmission and prevention, and provide evidence for the global response and control of the COVID-19 pandemic.

1. Introduction

Since December 2019, an epidemic of viral pneumonia caused by a novel zoonotic coronavirus developed in Wuhan, the capital city of Hubei Province in China.[1] The

World Health Organization (WHO) named the disease caused by the novel coronavirus "coronavirus disease 2019" (COVID-19) on February 11, 2020. COVID-19 often spreads by person-to-person transmission via respiratory droplets and close contact. The clinical features of COVID-19 are mainly fever, dry cough, and fatigue. However, some cases may progress to severe viral pneumonia with acute respiratory distress syndrome or even death.[2] The epidemic of COVID-19 spread rapidly to all provinces of China as well as numerous other countries.[3] By the end of March 2020, China had reported more than 82 thousand cases. Over 823 thousand infections and 40 thousand deaths have been confirmed worldwide.[3] The WHO described COVID-19 as a pandemic due to the speed and scale of transmission.[4] Therefore, a coordinated global response to control this pandemic is urgently needed to meet the unprecedented challenge for global public health.

Without effective population-based public health interventions, the rapidly increasing number of COVID-19 cases will greatly increase the burden of clinical treatments. At that point, massive severe cases would exceed the capacity of the health system, resulting in a sharp rise in mortality.[5-7] However, due to the shortage of health resources caused by the global pandemic, it is difficult to implement large-scale screening or even censuses. Hence, identifying the transmission sources and pathways is of vital importance to allocate limited public health resources. The integrated transmission chains and networks of COVID-19 transmission are revealed using epidemiological investigations. For example, based on the epidemiological data of 5830 confirmed cases in the early stage of the COVID-19 outbreak in Italy, researchers constructed transmission chains through contact tracing. The results show that the outbreak started in the northern region of Italy as early as January 2020 rather than February 20, 2020, when the first COVID-19 case was confirmed in the Lombardy region.[8] In addition, close contacts indicated by transmission chains can also be used to identify potential susceptible groups, which contributes to clarifying the emphasis of prevention, narrowing the scope and reducing the pressure of prevention and control, thus improving the efficiency of allocating limited health resources. Therefore, such comprehensive insight into the transmission network would be of great importance for local prevention and control policymakers.

Nevertheless, most current epidemiological investigation studies based on individual data primarily focus on describing the characteristics of COVID-19 infections, such as mortality, age distribution and sex ratio.[9-12] Information about exposure behaviors and contacts is normally provided by unstructured reports; it is heavily time-consuming to extract useful information from these reports to construct disease transmission networks. Additionally, such reports without a unified and structured form would be difficult to include for further analysis. To the best of our knowledge, no literature has visualized the relationship among COVID-19 infection cases using epidemiological data. In this study, the epidemiological data of each case, officially published from January 21 to February 22, 2020, in Chengdu and Tianjin, China, were used to construct visualized transmission networks according to the relationship among cases. Based on that, the transmission characteristics of COVID-19 were visually presented and analyzed using measures of transmission networks. In addition, we explore the value of epidemiological investigation reports to provide evidence for the global response and control of the COVID-19 pandemic.

2. Results

2.1 Characteristics of COVID-19 infection

A total of 135 and 143 confirmed COVID-19 patients in Tianjin and Chengdu, respectively, were included. Four and 45 cases were noninformative cases in Tianjin and Chengdu, respectively. The 131 and 98 cases with valid information about exposure history in Tianjin and Chengdu, respectively, were used to construct transmission networks, of which 70 (53.43%) and 44 (44.89%) cases, respectively, were males. The median ages were both 49 years (ranging from 9 to 90 years in Tianjin and 3 months to 88 years in Chengdu). The median time from symptom onset to hospital admission was 2 and 3 days in Tianjin and Chengdu, respectively (Table 1). The median time from symptom onset to be defined as a confirmed case was 4.5 and 6 days, respectively (Table 1).

2.2 Visualized transmission networks

Figure 1 shows the COVID-19 transmission network graph in Tianjin. In the initial cluster of 18 (13.74%), cases had a history of exposure to Hubei Province, and each case infected an average of 2.9 contacts. In the component that started with the central node of infections directly related to Hubei Province, one case directly infected 0-4 contacts, with 0.79 as the average. At most, 6 patients were infected by one or more generations, with 5.68 as the average chain size for the initial cluster of imported cases.

Twenty cases infected their relatives in the household or colleagues in the workplace, forming several multinode transmission chains with a maximum length of 4. Another 9 (6.87%) cases were employees of the Tianjin high-speed train administration. The transmission network of Tianjin consists of a total of 73 chains, of which 23 (36.99%) have a length greater than 1. There are 4 transmission chains with a maximum length of 4 in the component, starting with the central node of imported cases as the source of infection (Table 2).

Figure 2 shows the transmission network graph of Chengdu. In the 98 (68.53%) informative cases in the transmission network, 30 (30.61%) cases had a history of exposure to Hubei Province. Each case generated at most 3 direct secondary infections and a maximum of 4 patients by one or more generations. In addition, for 24 cases with vague information, it was uncertain whether they were infected by contact with imported cases or by contact with other categories of cases. A total of 88 transmission chains were constructed in Chengdu. Twenty-one chains started from the central node of the unclear exposures except Hubei Province. Nineteen chains had a length of 1. The length of one chain was up to 3 (Table 2).

2.3 Applications of transmission network graphs

In this study, over 85% of patients in Tianjin were nonimported cases from the Tianjin high-speed train administration and multiple families. The longest length of the transmission chain of imported cases reached 4, which suggested that the spread of COVID-19 in Tianjin became dominated by community transmission. Not only strictly managing imported populations but also preventing community transmission were necessary to prevent the spread of COVID-19. In addition, the Tianjin high-speed train administration and related people should be a key target in further prevention and control measures. Similarly, the proportion of cases in community transmission in Chengdu obviously exceeded that of the imported cases. Thirteen transmission chains in Tianjin with lengths greater than 3 suggested that these cases could spread for 3 generations before being confirmed. Therefore, the timeliness of case detection needs improvement. By contrast, the length of considerable transmission chains in Chengdu was lower than 3. However, 24 cases with vague information may be infected by two or more generations. One transmission chain with a length of 3 could be considered a chain with a length of up to 6 if the starting node is an imported case. Therefore, considerable cases with vague information lead to the fact that the length and number of transmission chains cannot accurately assess the spread of COVID-19.

There were only 4 noninformative cases in Tianjin, with a coverage rate of 97%. In Chengdu, up to one-third of cases lacked exposure history. Compared with Chengdu, Tianjin had a higher quality of epidemiological investigations and a more complete transmission network. In Chengdu, the large number of noninformative cases and cases with vague information made it difficult to provide accurate information by the transmission network. Although the transmission chains in Chengdu were relatively short, many undiscovered and uncontrolled risks may exist that had not been revealed in the transmission network. Once community transmission accelerates, new transmission chains might rapidly appear and extend before cases are detected. Hence, in Chengdu, more healthcare workers need to be allocated to conduct epidemiological investigations to improve the coverage and quality of epidemiological investigations, the timeliness of detecting cases, and the rapidity of diagnosis.

3. Discussion

Considering the relationships among cases, we constructed the disease transmission networks and presented the transmission network graphs for the COVID-19 epidemic through a visualization technique based on the individual reports of epidemiological data. Then, in a framework of intuitive and quantitative analysis, we compared the transmission characteristics of COVID-19 of Tianjin and Chengdu in China. This valuable application of the visualization technique was further explored, including tracing the source of infections, discovering potential super-spreaders, and evaluating prevention and control measures. Meanwhile, we discussed the potential insufficiency in the current form of individual epidemiological data. Our research may provide an important basis for jointly constructing multiregional and large-scale disease transmission networks.

Finding "patient zero" plays an important role in preventing the COVID-19 epidemic, such as identifying the origin and further spreading, as well as in studying the transmission characteristics,[13-15] as does the identification of the super-spreading event (i.e., one COVID-19 case produces at least 8[16], 10[17], or many more than the average number of secondary patients).[18,19] However, epidemiological case reports are written by health workers after investigations with confirmed cases. In each report, only the patient-related source of infections can be obtained from the collected exposure information, which cannot provide enough information to trace back to "patient zero" and present the full transmission chains. Thus, such reports need to be integrated using contact tracing analysis to construct transmission graphs, which provide crucial clues for the identification of "patient zero" and super-spreaders.[20,21] For example, by contact tracing and constructing disease transmission chains, researchers found that the epidemic of COVID-19 in Italy had spread much earlier than February 20, 2020, when the first case was confirmed.[8] In our study, we found that one COVID-19 patient generally directly produces 0 to 4 infections in Tianjin, while in Chengdu, the number is up to 3. There was no evidence of super-spreaders in the two cities. Terminal nodes of the transmission chains can be applied in several aspects for prevention and control, including identifying potential high-risk populations, determining priorities, and narrowing the scope of quarantine and thus allocating limited health resources effectively.

Currently, the available epidemiological data of each individual case vary from city to city. The exact exposure history can be extracted from released epidemiological data in Tianjin, while in Chengdu, only whether some of the cases have been in close contact with confirmed cases can be extracted, and the relationships among cases are not clear. From January 21 to February 22, 131 (97.04%) patients in Tianjin can be integrated into the transmission network, in which the source, relevant infections and terminal nodes can be revealed. In Chengdu, only 98 (68.53%) cases had infection pathways. The other 45 patients were noninformative cases, and thus, the sources and potential infectious ranges remain unclear, as these noninformative cases cannot be integrated into any transmission chains. By comparing the results of the disease transmission networks of the two cities, we found that the transmission network in Tianjin was more complete with clear transmission chains. Preventive measures can be carried out mainly by focusing on the close contacts of each node, which could reduce the consumption of limited health resources. Noninformative cases and cases with vague information also suggest the risk from unclear transmission chains. In Chengdu, the tracing transmission chains of approximately one-third of the cases were not available. Therefore, the transmission network graph of the COVID-19 epidemic in Chengdu had less coverage and provided less information. The number of nodes in each pathway and the close contacts of each node in the transmission network cannot be determined, indicating higher unpredictable risks in Chengdu. This result indicates that, in epidemiological investigations, the exposure history of each infected should be collected as completely as possible. Traceable transmission chains of each case would greatly reduce the unpredictable risks for prevention and control and avoid the waste of health resources. In addition, the composition of the transmission chains presents the main type of local transmission, which suggests that further prevention and control should focus more on imported or community-spread cases. The length of transmission chains partly suggests that the timeliness of case detection, as well as the quality of the epidemiological investigation reflected by the rate of case coverage, provides a valuable index for evaluation of the efficiency of the local control measures. The relatively poor quality of epidemiological data in Chengdu may suggest the shortage of public health manpower, which may provide evidence for adjusting control and prevention strategies and allocating resources.

Currently, most epidemiological investigation reports with information on exposure behaviors and contacts are provided by unstructured reports with different forms in different cities. Therefore, it is difficult to construct an integrated cross-regional transmission network and gain full use of the epidemiological data for COVID-19 prevention and control. Thus, we suggest that health administrations develop a standard guideline for epidemiological data collection, and all such data should be managed and released in a timely manner.[22] On the one hand, researchers can jointly construct a multiregional transmission network to trace the spreading of COVID-19. On the other hand, integrated transmission networks can improve public awareness of COVID-19 epidemics, enhance public compliance with control measures, and reduce the difficulty of implementation and resource consumption. Moreover, with transmission networks, network-based analysis can be carried out to evaluate the transmission rates and the complexity of network structures, which may provide clues for large-scale interventions. In addition, for emerging infectious diseases, constructing transmission chains through contact tracing can estimate infectivity at an early stage to quantify the risk and trends of infectious diseases.[23-25]

It is worth noting that COVID-19 cases in one category might be included in another category in transmission network graphs. For instance, some cases in the category of family clusters are contained in the category of exposure to infections directly related to Hubei Province. To fully demonstrate the information contained in the epidemiological data, however, this study classified these patients into another category, with family aggregation as a vital feature of infectious disease transmission. Better classification strategies are needed to discuss the local transmission characteristics in different cities. Meanwhile, regional unification should be considered for the classification standard to guarantee the exchangeability of data when cross-regional transmission networks are constructed.

For infectious diseases with stronger infectivity, longer incubation periods and higher fatality risks, such as COVID-19, if public health interventions were not carried out in a timely and effective manner, cases would increase rapidly, consume limited clinical resources quickly and lead to high mortality.[26] Therefore, the emphasis of control measures should not only focus on clinical treatments but also ensure sufficient resources in epidemiological investigations.[27] These measures could contribute to controlling the source of infections, reducing the risk of exposure, decreasing the incidence by shortening transmission chains, and easing the pressure of clinical treatments. The epidemiological data of the COVID-19 epidemic in Tianjin and Chengdu were used to propose an analysis framework for the individual epidemiological data. Our results illustrated the importance of visualized epidemiological transmission networks in preventing and controlling the epidemic of COVID-19. Currently, the content and format of epidemiological data are not unified, causing the transmission network graphs of Tianjin and Chengdu to show different performances in risk assessment. Therefore, the collection, management, and release of epidemiological data should be improved for the joint construction of large-scale and multiregional disease transmission networks to provide a better understanding of the COVID-19 epidemic and to provide evidence for local prevention and control policymakers.

4. Methods

4.1 Data sources and collection

Since January 21, 2020, the official websites of several municipal health commissions in mainland China have successively released individual records of confirmed COVID-19 cases. As of February 22, 23 (74.19%) of the 31 capitals/municipalities in mainland China had begun to publish individual reports, most of which are unstructured reports with different forms and content. Detailed relationships among individual cases, such as relatives, colleagues, or other contacts, can be obtained in some cities, such as Tianjin, Chongqing and Xinyang. In other cities, such as Chengdu, Beijing, and Shanghai, only limited exposure information can be extracted for a few cases. For instance, some cases in Chengdu were reported to be related to other confirmed cases, but no detailed information was available to indicate which specific confirmed cases were related. Daily individual records of the confirmed COVID-19 cases in Tianjin and Chengdu from January 21 to February 22 were used in our analysis; these data were collected from the websites of municipal health commissions. Information was extracted from the individual records to build a structured database including 3 sections: demographic characteristics (sex, age, and district); key timelines (date of symptom onset, date of hospital admission and date of confirmation as a case) and exposure history (exposure to Hubei Province and relationship among cases).[28] We provided an example of the unstructured individual reports and the structured individual line list in Figure 3.[29]

4.2 Construction of transmission networks

We constructed the transmission networks in three steps:

First, cases were categorized according to the types of exposure history.

There were two main common categories of exposure history in cities besides the foci: exposed to Hubei Province of imported cases and exposed to the imported cases of other cases. Other cases without exposure history were defined as noninformative cases. In addition, there were a group of clustering cases in Tianjin. These cases worked in the Tianjin high-speed train administration and had a common exposure history at the same workplace. Likewise, some of the cases in Chengdu with a history of exposure to confirmed cases but without detailed information to identify the specific related cases, were defined as cases with vague information. For example, one new case in Chengdu confirmed on February 9 was reported to be relevant to another case confirmed on February 2. However, on February 2, 4 cases were confirmed, of which the new case was unknown.

Second, the central nodes were set based on the categories of exposure history.

According to the categories of exposure history in the previous step, three central nodes were set to represent the sources of exposure. Hubei Province, infections directly related to Hubei Province, and the Tianjin high-speed train administration were set as the starting central nodes of the transmission network in Tianjin. Similarly, for the transmission network of Chengdu, three central nodes were set to present Hubei Province, infections directly related to Hubei Province, and unclear exposures except for those in Hubei Province.

Finally, cases were integrated as nodes into the transmission networks by the source of exposure.

In the transmission networks, the nodes other than starting nodes represented confirmed cases. Those cases that had clear contact histories with specific confirmed cases were linked with directional edges. The directions of edges denote the direction of COVID-19 transmissions between cases. The cases without related cases in the exposure history were directly linked to the corresponding central node of the source of exposure. Nodes of noninformative cases without the exposure history scatter outside the transmission network and are not part of the components in the transmission network. Starting with nodes of the earliest traceable sources of infection, the transmission chain was composed of corresponding directional edges, and all nodes of secondary cases were linked by one or more generations of transmission.

4.3 Statistical analysis

Different characteristics of the transmission can be described by measures of the transmission networks, i.e., number of chains, chain sizes, maximum lengths of chains, average chain size and average number of nodes linked to each generation of cases. The definitions and implications in COVID-19 are shown in Table 3 and illustrated using a simplified sample in Figure 4.

These measures were then summarized to compare the transmission characteristics in Tianjin and Chengdu. The average chain size and number of nodes linked to each generation of cases were quantified in Tianjin's transmission network, as in Chengdu, approximately one-third of the confirmed cases of COVID-19 cannot be integrated into the transmission network due to uncertain exposure history.

All statistical analyses were performed with R3.5.1 using the package *statnet* to visualize the transmission networks.

4.4 Ethics statement

All patient information of COVID-19 epidemiological data were collected from the official websites of municipal health commissions and this study was approved by the institutional review board of the School of Public Health, Sichuan University. All methods were carried out in accordance with relevant guidelines and regulations. All data were collected from publicly available sources. Data were deidentified, and informed consent was waived.

5. Data Availability

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

References

1 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* **382**, 1199-1207, doi:10.1056/NEJMoa2001316 (2020).
2 National Health Commission of the People's Republic of China. Diagnosis and treatment guideline on pneumonia infection with 2019 novel coronavirus (7th

trial edn). (2020) Available at, http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989/files/ce3e6945832a438eaae415350a8ce964.pdf (Accessed: 8th Apr 2020).

3 World Health Organization. Coronavirus disease 2019 (COVID-19) situation report–72. (2020) Available at, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200401-sitrep-72-covid-19.pdf?sfvrsn=3dd8971b_2 (Accessed: 8th Apr 2020).

4 World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. (2020) Available at, https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (Accessed: 11th Mar 2020).

5 Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*, doi:10.1016/S1473-3099(20)30243-7 (2020).

6 Lancet, T. COVID-19: protecting health-care workers. *Lancet* **395**, 922, doi:10.1016/S0140-6736(20)30644-9 (2020).

7 Lancet, T. COVID-19: learning from experience. *Lancet* **395**, 1011, doi:10.1016/S0140-6736(20)30686-3 (2020).

8 Cereda, D. *et al.* The early phase of the COVID-19 outbreak in Lombardy, Italy. Preprint at https://arxiv.org/ftp/arxiv/papers/2003/2003.09320.pdf (2020).

9 Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. Vital surveillances: The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. (2020) Available at, http://weekly.chinacdc.cn/en/article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51 (Accessed: 20th Mar 2020).

10 Guan, W.-J. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*, doi:10.1056/NEJMoa2002032 (2020).

11 Dong, X. C. *et al.* Epidemiological characteristics of confirmed COVID-19 cases in Tianjin. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 638-642 (in Chinese), doi:10.3760/cma.j.cn112338-20200221-00146 (2020).

12 Zheng, Y. *et al.* Epidemiological characteristics and clinical features of 32 critical and 67 noncritical cases of COVID-19 in Chengdu. *J Clin Virol* **127**, 104366, doi:10.1016/j.jcv.2020.104366 (2020).

13 Porcheddu, R., Serra, C., Kelvin, D., Kelvin, N. & Rubino, S. Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. *J Infect Dev Ctries* **14**, 125-128, doi:10.3855/jidc.12600 (2020).

14 Carinci, F. Covid-19: preparedness, decentralisation, and the hunt for patient zero. *BMJ* **368**, bmj.m799, doi:10.1136/bmj.m799 (2020).

15 McKay, R. A. "Patient Zero": the absence of a patient's view of the early North American AIDS epidemic. *Bull Hist Med* **88**, 161-194, doi:10.1353/bhm.2014.0005 (2014).

16 Shen, Z. *et al.* Superspreading SARS events, Beijing, 2003. *Emerging Infect Dis* **10**, 256-260 (2004).

17 Centers for Disease Control and Prevention (CDC). Severe acute respiratory syndrome--Singapore, 2003. *MMWR Morb Mortal Wkly Rep* **52**, 405-411 (2003).

18 Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355-359 (2005).

19 Stein, R. A. Super-spreaders in infectious diseases. *Int J Infect Dis* **15**, e510-e513, doi:10.1016/j.ijid.2010.06.020 (2011).

20 European Centre for Disease Prevention and Control. Contact tracing: public health management of persons, including healthcare workers, having had contact with COVID-19 cases in the European Union – second update. (2020) Available at, https://www.ecdc.europa.eu/sites/default/files/documents/Contact-tracing-Public-health-management-persons-including-healthcare-workers-having-had-contact-with-COVID-19-cases-in-the-European-Union%E2%80%93second-update_0.pdf (Accessed: 14th Apr 2020).

21 Hu, K. *et al.* Identification of a super-spreading chain of transmission associated with COVID-19. *medRxiv* (2020).

22 Xu, B. & Kraemer, M. U. G. Open access epidemiological data from the COVID-19 outbreak. *Lancet Infect Dis*, doi:10.1016/S1473-3099(20)30119-5 (2020).

23 Soetens, L., Klinkenberg, D., Swaan, C., Hahné, S. & Wallinga, J. Real-time Estimation of Epidemiologic Parameters from Contact Tracing Data During an Emerging Infectious Disease Outbreak. *Epidemiology* **29**, 230-236, doi:10.1097/EDE.0000000000000776 (2018).

24 De, P., Singh, A. E., Wong, T., Yacoub, W. & Jolly, A. M. Sexual network analysis of a gonorrhoea outbreak. *Sex Transm Infect* **80**, 280-285 (2004).

25 Bjørnstad, O. N. *Epidemics: models and data using R*. 49-50 (Springer, 2018).

26 Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689-697, doi:10.1016/S0140-6736(20)30260-9 (2020).

27 Wang, F.-S. & Zhang, C. What to do next to control the 2019-nCoV epidemic? *Lancet* **395**, 391-393, doi:10.1016/S0140-6736(20)30300-7 (2020).

28 Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *Lancet Infect Dis*, doi:10.1016/S1473-3099(20)30230-9 (2020).

29 Tianjin Municipal Health Commission. Announcements. (2020) Available at, http://wsjk.tj.gov.cn/art/2020/1/27/art_14_70224.html (Accessed: 22th Apr 2020).

Acknowledgments

We thank Yue Ma, Fei Yin and Tao Zhang for providing suggestions and comments for this study. This work was supported by the National Natural Science Foundation of China (Grant No. 81872713 and No. 81803332) and the Sichuan Science & Technology Program (Grant No. 2019YFS0471 and 2018SZ0284).

Author information

These authors contributed equally: Caiying Luo and Yue Ma.

Affiliations

West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, China

Caiying Luo, Yue Ma, Pei Jiang, Tao Zhang & Fei Yin

Contributions

FY and YM designed the study and contributed to data analysis. CL and PJ contributed to the data collection, literature search, data analysis, data interpretation, figures, and writing. CL, FY, YM, and TZ contributed to data interpretation. All authors contributed to writing the manuscript and revising the final version.

Corresponding author

Correspondence to Fei Yin.

Additional information

Competing interests

The authors declare that they have no competing interests.

Table 1. Characteristics of COVID-19 cases in Tianjin and Chengdu.

	Tianjin	Chengdu
Age, median(IQR), years	49(36-61)	49(33-59)
Male, *n*/*n*(%)	70/131(53.43%)	44/98(44.89%)
Time from symptom onset to be defined as a confirmed case, median (IQR), days*	4.5(2-8)	6(3-11)
Time from symptom onset to hospital admission, median (IQR), days*	2(1-5)	3(0-7)
Cases with exposure history, *n*/*n*(%)	131/135(97.04%)	98/143(68.53%)
Imported cases, *n*/*n*(%)	18/131(13.74%)	30/98(30.61%)
Nonimported cases, *n*/*n*(%)	113/131(86.26%)	68/98(69.39%)

* Four noninformative cases in Tianjin had information about sex and age, while 42 of the 45 noninformative cases in Chengdu had no individual information. To keep the statistical analysis consistent, the descriptions of all variables (except exposure history) were based only on the cases with valid information about exposure history.

*Of the 131 cases with an exposure history in Tianjin, 13 cases lacked information about key timelines (date of symptom onset, date of hospital admission, and date of confirmation as a case). Forty-four of 98 cases with an exposure history in Chengdu lacked information about key timelines. These cases were excluded only when the time from symptom onset to hospital admission was analyzed, as well as the time from symptom onset to confirmation as a case.

Table 2. Distribution of transmission chains for COVID-19 cases in Tianjin and Chengdu.

Central node	Chain size	Tianjin		Chengdu	
		Maximum length of chains	Number of chains	Maximum length of chains	Number of chains
Hubei Province	1	1	13	1	29
	2	2	4	0	0
	3	3	1	2	1
Tianjin high-speed train administration	1	1	7	-	-
	3	3	1	-	-
	4	3	1	-	-
Infections directly related to Hubei Province	1	1	26	1	34
	2	2	7	2	3
	3	3	4	0	0
	4	4	4	0	0
	5	3	2	0	0
	6	4	3	0	0
Unclear exposures except Hubei Province	1	-	-	1	19
	3	-	-	3	1
	4	-	-	2	1
Total	-	-	73	-	88

Table 3. The detailed description for 5 indexes applied for assessing the evolving epidemiology of COVID-19.

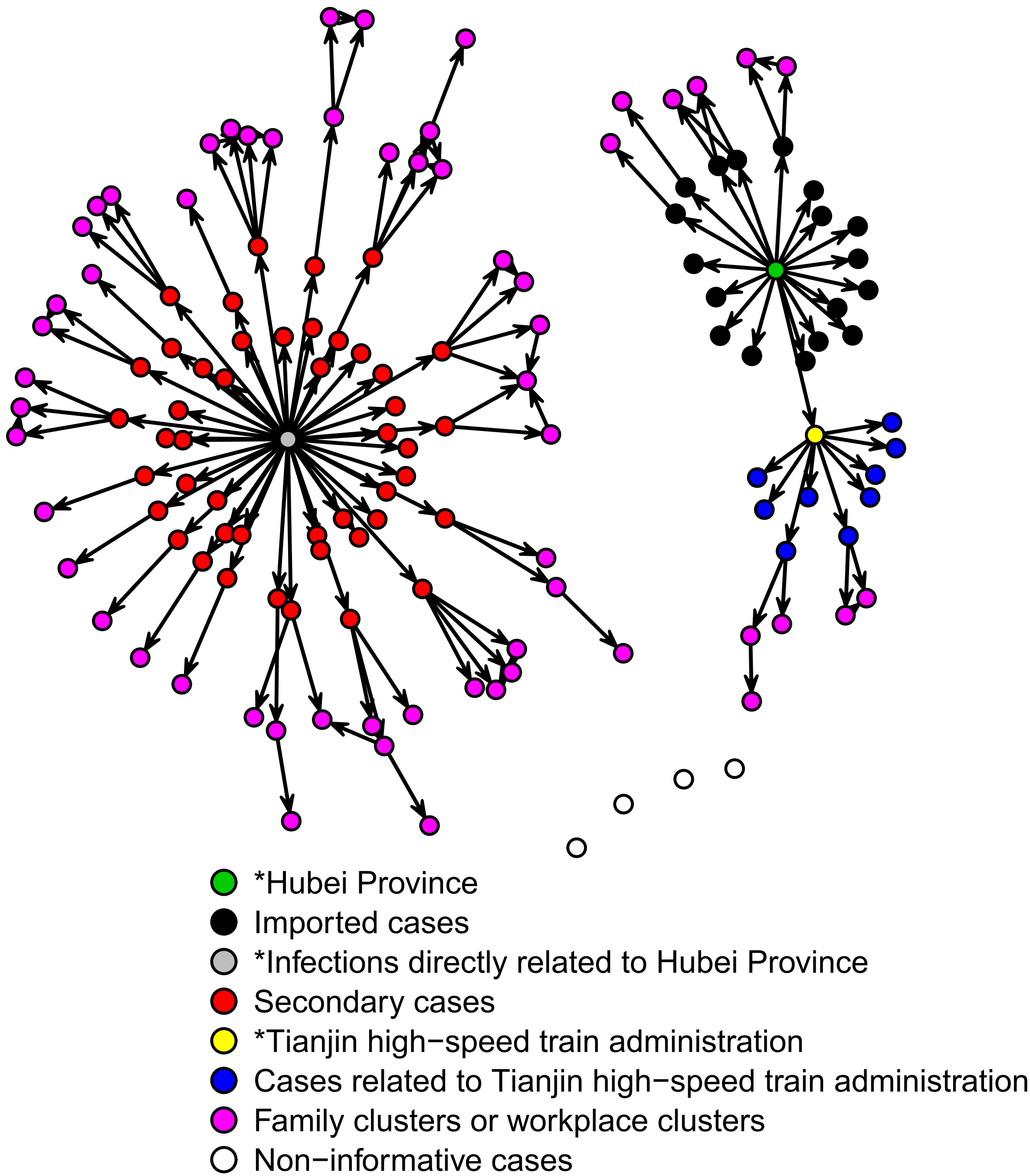
Index	Definition	Implication in COVID-19	Example in Figure 2
Number of chains	The number of chains starting with a central node	The spread of transmission through the source of infection	2
Chain size	The number of nodes in each transmission chain except the central node	The number of cases in a chain and the scope of a transmission chain of COVID-19	Chain I: 1 Chain II: 5
Maximum length of chains	The maximum number of directional edges in each chain	The maximum generations of transmission before the secondary case is detected and controlled	Chain I: 1 Chain II: 3
Average chain size	Dividing the summation of chain sizes starting with same central node by the number of cases in the central node	The average reproductive number of cases from specific exposures	6/2=2.5
Average number of nodes linked to each generation of cases	Dividing the total number of nodes in the same distance with central node by the total number of front-end nodes	The infectivity of different generations of cases	First generation:2/2=1 Second generation:3/2=1.5 Third generation:1/3=0.3

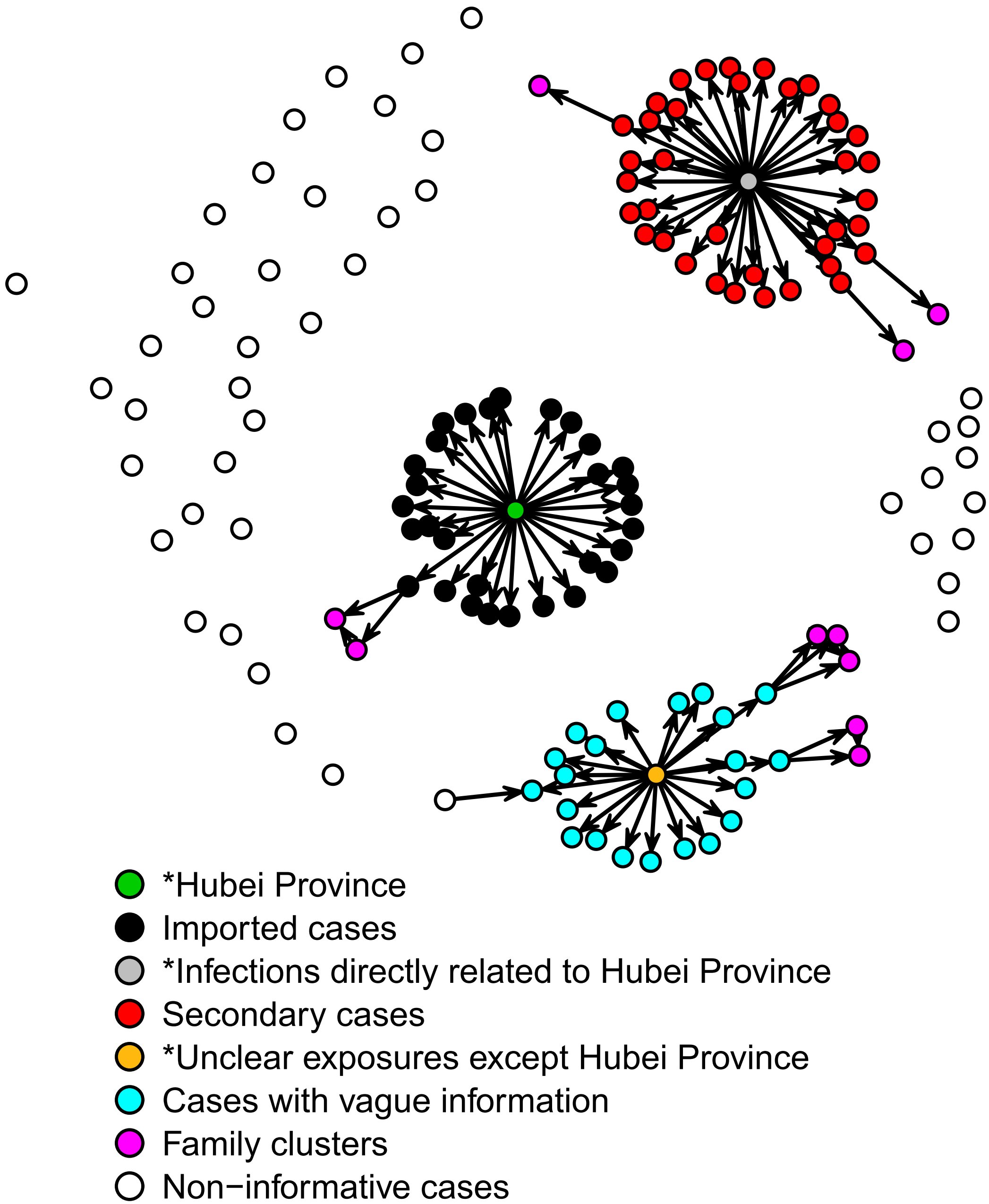
Figure 1. Transmission network graph for confirmed COVID-19 cases in Tianjin.

Figure 2. Transmission network graph for confirmed COVID-19 cases in Chengdu.

Figure 3. Example of unstructured individual reports and structured databases of Tianjin and Chengdu.

Figure 4. Example of a simplified transmission network with a group of 2 cases as the central node and a total of 2 chains in the network.


*Hubei Province
Imported cases
*Infections directly related to Hubei Province
Secondary cases
*Tianjin high−speed train administration
Cases related to Tianjin high−speed train administration
Family clusters or workplace clusters
Non−informative cases


*Hubei Province
Imported cases
*Infections directly related to Hubei Province
Secondary cases
*Unclear exposures except Hubei Province
Cases with vague information
Family clusters
Non−informative cases

Three new COVID-19 cases confirmed in Tianjin increasing to 17

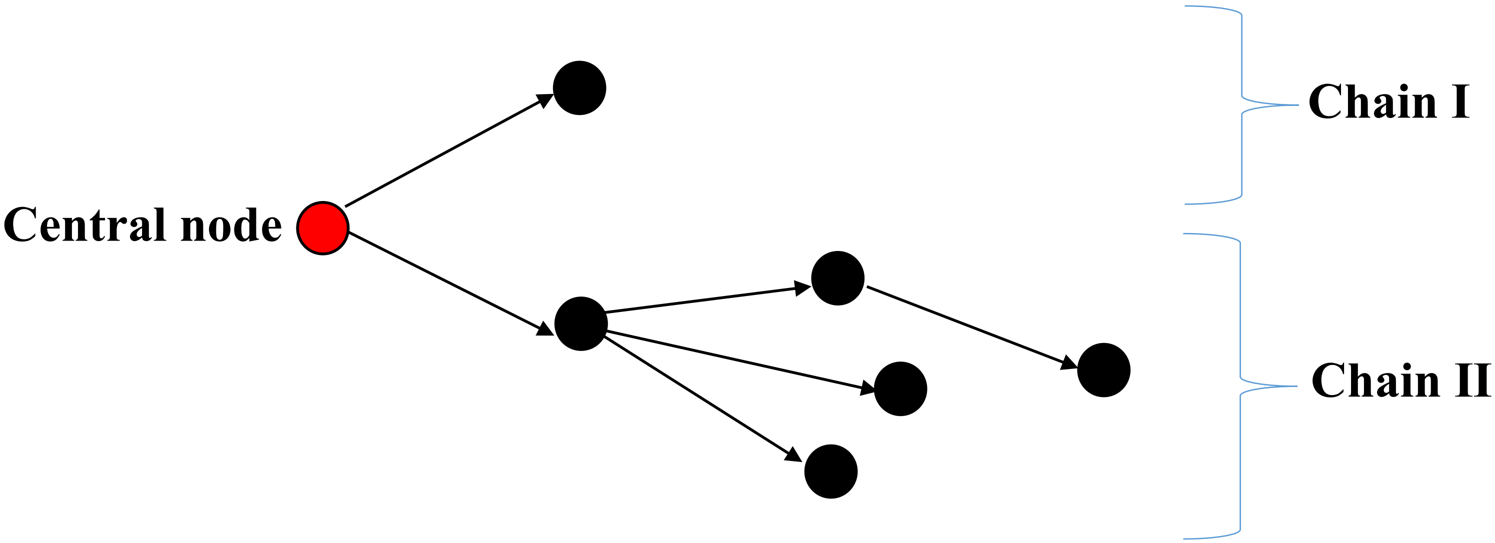
Release Date: 2020-01-27 12:23 Source: Committee Office Views: 2110 share to: 0

As of 6 a.m. on January 27, Tianjin City added three new COVID-19 cases, and the number of confirmed cases increased to 17 cases.

…

Chen, male, 78 years old, lives in Hebei District, Tianjin. He had a Fever on January 21, and then he went to the Fourth Central Hospital of Tianjin on January 23 and 24. The RT-PCR test result was negative, so he was isolated at home. On January 26, he fevered again and went to the Fourth Central Hospital of Tianjin for medical treatment. The RT-PCR test was positive, and the expert team confirmed him, a severe case, as the 16th case in Tianjin. He was transferred to Haihe Hospital for treatment and his physical state is stable. This case is the relative of the 5th confirmed case, Chen.

Location	Case ID	Gender	Age	District	Date of symptoms onset	Date of hospital admission	Date of defined as a confirmed case	Exposure to Hubei province	Relationship among cases
Tianjin	4	1	40	Heping district	2020/1/14	2020/1/14	2020/1/22	1	
Tianjin	5	1	46	Hebei district	2020/1/15	2020/1/17	2020/1/23	0	employee of Tianjin high-speed train administration
Tianjin	11	2	55	Hebei district	2020/1/23	2020/1/23	2020/1/26		
Tianjin	16	1	78	Hebei district	2020/1/21	2020/1/23	2020/1/27	0	relative of the 5th case
Tianjin	32	1	28	Ninghe district	2020/1/23	2020/1/26	2020/1/31	0	
…									
Chengdu	73	2	55	Jinniu district	2020/1/26	2020/1/31	2020/2/1	0	
Chengdu	74	2	33	Wuhan city			2020/2/2	1	
Chengdu	75	2	76	Tianfu district			2020/2/2		
Chengdu	76	2	22	Qingbaijiang district	2020/1/31	2020/1/31	2020/2/2	0	relative of one confirmed case
Chengdu	77	1	50	Gaoxin district			2020/2/2		
Chengdu	97	2	47	Wuhou district	2020/2/4	2020/2/4	2020/2/5	0	relative of one case confirmed on Feb 2
Chengdu	98	1	49	Wuhou district	2020/1/27	2020/2/5	2020/2/6	0	
Chengdu	99	2	46	Wuhou district	2020/1/26	2020/2/5	2020/2/6	0	relative of the 98th case


Central node
Chain I
Chain II

Figures

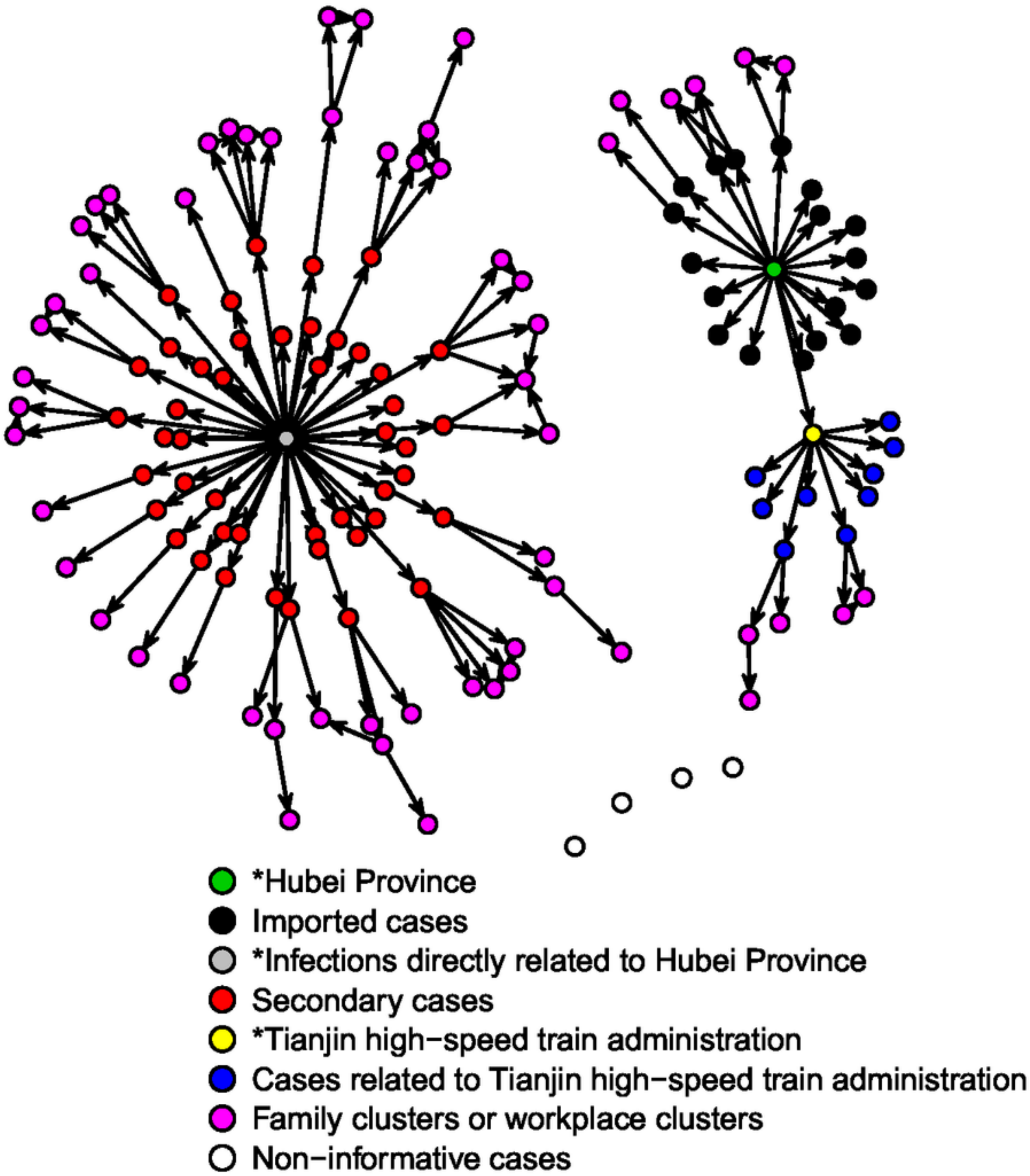


Figure 1

Transmission network graph for confirmed COVID-19 cases in Tianjin.

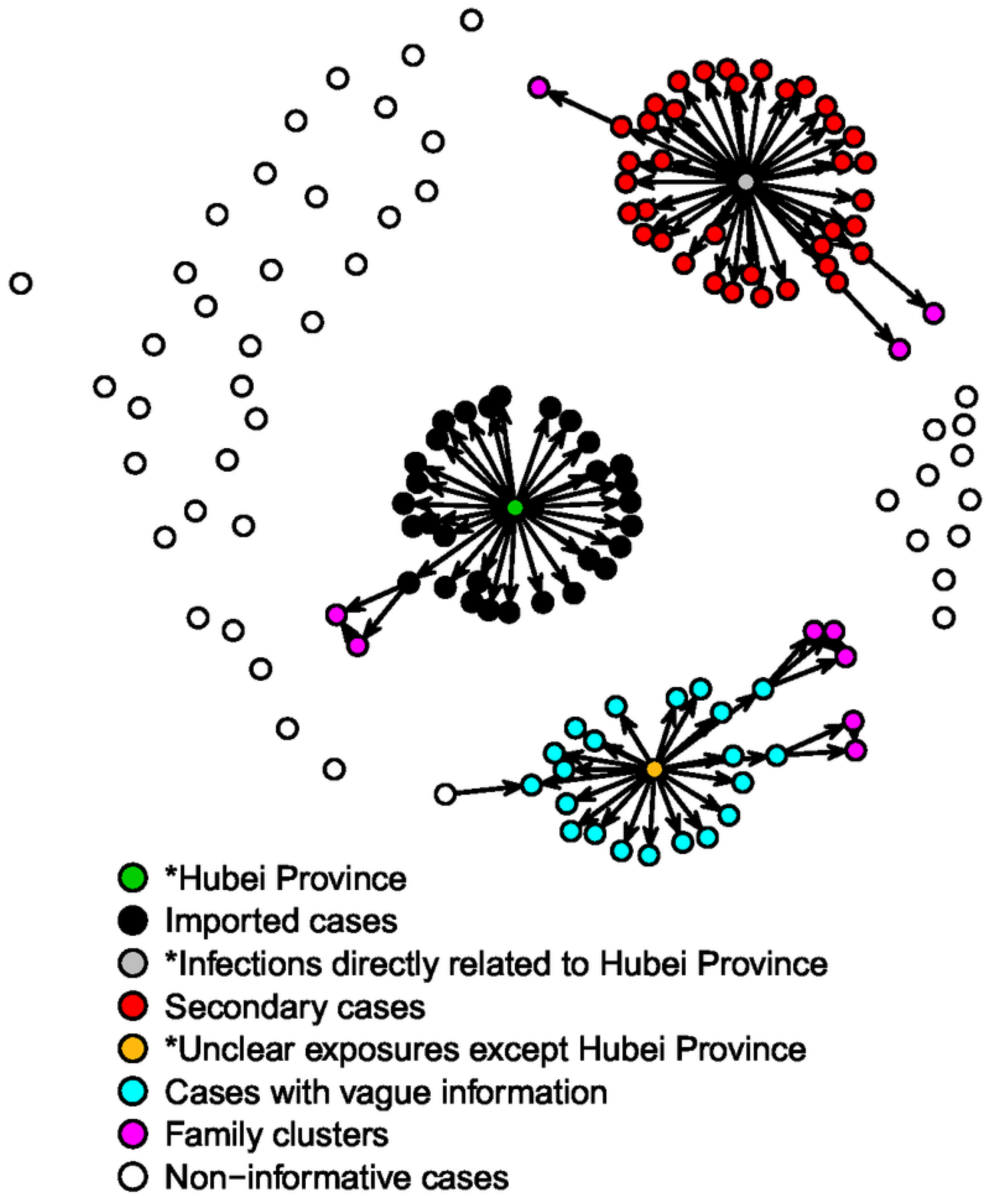


Figure 2

Transmission network graph for confirmed COVID-19 cases in Chengdu.

Three new COVID-19 cases confirmed in Tianjin increasing to 17

Release Date: 2020-01-27 12:23 Source: Committee Office Views: 2110 share to: 0

As of 6 a.m. on January 27, Tianjin City added three new COVID-19 cases, and the number of confirmed cases increased to 17 cases.

...

Chen, male, 78 years old, lives in Hebei District, Tianjin. He had a Fever on January 21, and then he went to the Fourth Central Hospital of Tianjin on January 23 and 24. The RT-PCR test result was negative, so he was isolated at home. On January 26, he fevered again and went to the Fourth Central Hospital of Tianjin for medical treatment. The RT-PCR test was positive, and the expert team confirmed him, a severe case, as the 16th case in Tianjin. He was transferred to Haihe Hospital for treatment and his physical state is stable. This case is the relative of the 5th confirmed case, Chen.

Location	Case ID	Gender	Age	District	Date of symptoms onset	Date of hospital admission	Date of defined as a confirmed case	Exposure to Hubei province	Relationship among cases
Tianjin	4	1	40	Heping district	2020/1/14	2020/1/14	2020/1/22	1	
Tianjin	5	1	46	Hebei district	2020/1/15	2020/1/17	2020/1/23	0	employee of Tianjin high-speed train administration
Tianjin	11	2	55	Hebei district	2020/1/23	2020/1/23	2020/1/26		
Tianjin	16	1	78	Hebei district	2020/1/21	2020/1/23	2020/1/27	0	relative of the 5th case
Tianjin	32	1	28	Ninghe district	2020/1/23	2020/1/26	2020/1/31	0	
...									
Chengdu	73	2	55	Jinniu district	2020/1/26	2020/1/31	2020/2/1	0	
Chengdu	74	2	33	Wuhan city			2020/2/2	1	
Chengdu	75	2	76	Tianfu district			2020/2/2		
Chengdu	76	2	22	Qingbaijiang district	2020/1/31	2020/1/31	2020/2/2	0	relative of one confirmed case
Chengdu	77	1	50	Gaoxin district			2020/2/2		
Chengdu	97	2	47	Wuhou district	2020/2/4	2020/2/4	2020/2/5	0	relative of one case confirmed on Feb 2
Chengdu	98	1	49	Wuhou district	2020/1/27	2020/2/5	2020/2/6	0	
Chengdu	99	2	46	Wuhou district	2020/1/26	2020/2/5	2020/2/6	0	relative of the 98th case

Figure 3

Example of unstructured individual reports and structured databases of Tianjin and Chengdu.

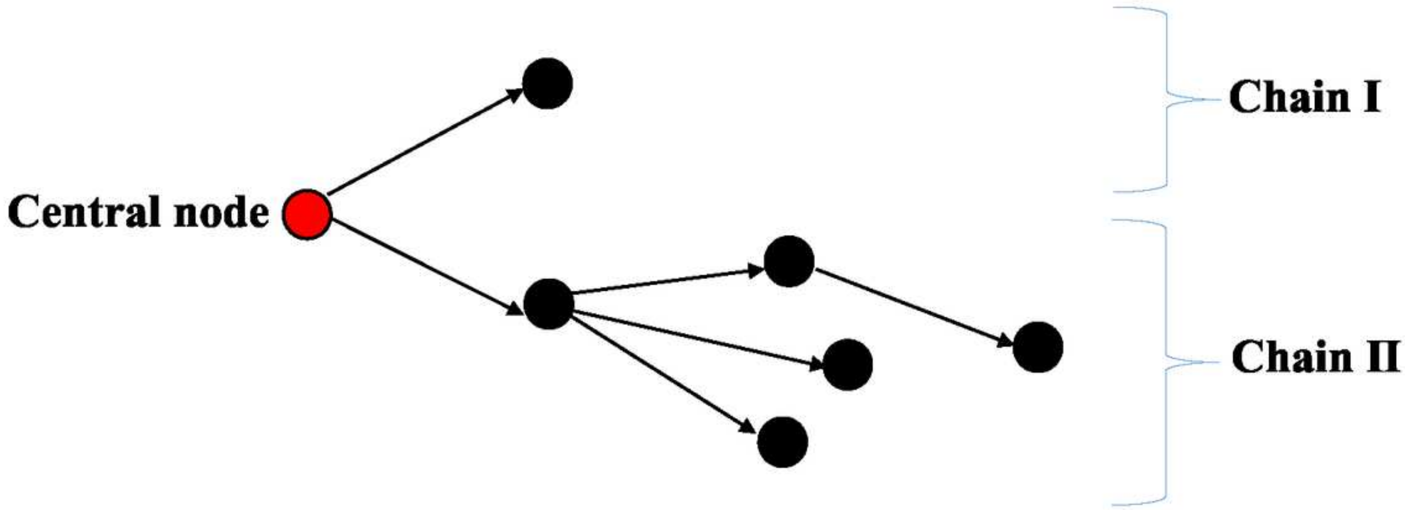


Figure 4

Example of a simplified transmission network with a group of 2 cases as the central node and a total of 2 chains in the network.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.pdf