Preplanned Studies

Social Network Analysis of a Norovirus Outbreak at a Primary School — Zhuhai City, Guangdong Province, China, 2023

Xiling Yin[1]; Songjian Xiao[1]; Xuebao Zhang[1]; Feng Ruan[1,#]

Summary

What is already known about this topic?

The investigations and analyses limited to epidemiological characteristics are insufficient to analyze the spread patterns of norovirus outbreaks in schools.

What is added by this report?

Norovirus outbreaks in primary schools are a dynamic process that spreads through social networks. The use of a social network analysis method to measure and identify key nodes for simulating control evolution was proven effective.

What are the implications for public health practice?

Infected students exhibit priority connection characteristics at different developmental stages in the network topology. Identifying and deliberately targeting key nodes could destroy network connectivity and help reduce the spread of the outbreak.

Norovirus is currently considered the leading cause of acute gastroenteritis worldwide (*1*). Clustering among children in schools can easily lead to norovirus outbreaks (*2*). Rapid and effective control of such outbreaks in schools remains a significant public health challenge (*3*). Norovirus spread should be regarded as a diffusion behavior in social networks and as a dynamic, continuously evolving process. Social network analysis, as a highly flexible research method, can recognize the inherent complexity of individuals' connections. By simulating the transmission dynamics of social networks based on a norovirus outbreak in a primary school, 12 key nodes were identified during the spread and recession periods. This approach demonstrated that targeted control strategies, which transcend inherent epidemiological approaches, are effective and provide insights for emergency management.

Using data from a norovirus outbreak reported in a Zhuhai City primary school (*4*), the epidemic lifecycle — from emergence to spread and decline — was analyzed. A 63×63 transmission matrix was constructed for all nodes. Network analysis and characteristic parameter calculation were performed using UCINET (version 6.528, Analytic Technologies, USA). NetLogo (version 3D 6.1.1, Northwestern's Center for Connected Learning and Computer-Based Modeling, USA) was used to model and simulate the effects of interventions during different epidemic periods. Methods for social network analysis and modeling are described in Supplementary Material (available at https://weekly.chinacdc.cn/).

The norovirus outbreak lasted 23 days, with 63 total cases reported across 6 grades and 18 classes. The class attack rate ranged from 1.96% (1/51) to 42.9% (21/49), and the overall attack rate was 4.2% (63/1,500). The outbreak initially affected 21 cases (Cases 1–21) in 3 classes over 2 days. This initial outbreak was primarily concentrated in Grade 3, Class 3, involving 19 cases, including the index case who experienced vomiting in the classroom on October 13, 2020. The incidence curve indicated a point-source exposure pattern. On October 14, classes for Grade 3, Class 3 were suspended. However, the epidemic continued with a spreading period involving 28 cases (Cases 22–49) across 13 classes over 9 days, suggesting a person-to-person transmission model. On-campus housing was suspended until the evening of October 23. On October 24, a recession period began, involving 14 cases (Cases 50–63) in 10 classes. This period lasted 12 days, and the epidemic curve showed a tailing pattern (Figure 1).

A rapid point-source outbreak occurred in Grade 3, Class 3, with 10 (Cases 2, 5, 7, 11–15, 19, and 20) of the 21 reported cases involved in mixed-class hosting at the school on October 13 and 14, 2020. This activity resulted in the direct infection of 19 students in other classes. The outbreak spread further to the classes of these students and through off-campus hosting centers. Case 23, which was hosted at hosting centers B and C at midday and in the evening, may have caused internal transmission at these two centers. None of the reported cases in Grade 3, Class 3, were cared for at hosting centers A and C. These two centers had longer

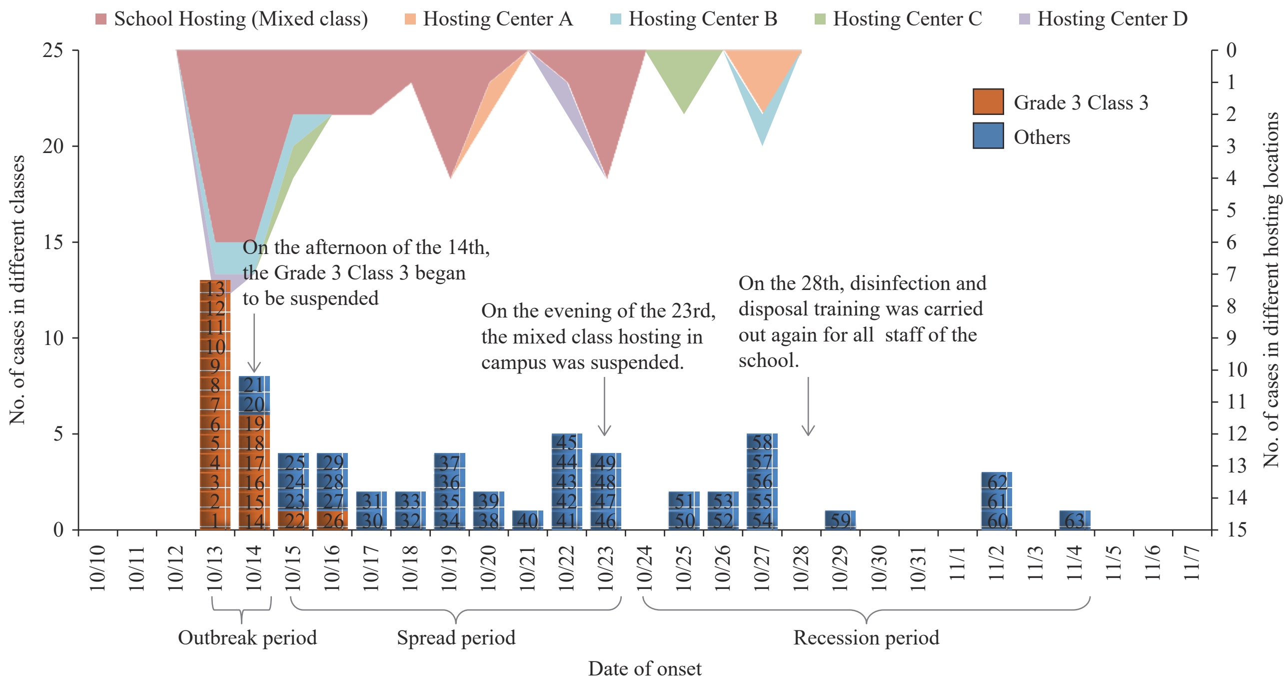

FIGURE 1. The onset date in different classes and hosting locations and periods of development of a norovirus outbreak at a primary school in Zhuhai City, Guangdong Province, China, 2023.

transmission chains. Additionally, Case 9 in Grade 3, Class 3, and Case 43 in Grade 6, Class 2, were both at hosting center D; however, their 8-day onset time interval suggests latent infections within the network as possible sources or bridges of infection. Notably, Case 61 in Grade 6, Class 3, had no clear social relationships within the network (Figure 2).

During the outbreak period, the network was the tightest (density=0.871), with 10 main key nodes (cases 2, 5, 7, 11–15, 19, and 20), which were included in 17 core nodes, as shown by the core-periphery analysis results. The average distance between nodes in the outbreak period was 1.085, meaning any two nodes could communicate through an average of 1.1 intermediate nodes. During the spread period, network tightness decreased (density=0.495), and cases 9, 43, 23, 3, 16, 2/5/7/11–15/19/20 had greater betweenness centralization, meaning they were in a relatively central position in the network. During the recession period, the network was relatively loose (density=0.099). Five core nodes (41, 46–49) were in the core matrix, and cases 48 and 59 were the key nodes as the spread of vomiting according to social network topology. The degree and closeness centralization indicators of the whole network both point to key nodes 2, 5, 7, 11–15, 19, and 20. The average distance was 2.292, and 29 of the 63 cases were in the core matrix (Table 1, Figure 3).

The curve generated with 10 randomly controlled nodes (decay step length=1,083) was similar to the curve generated without controlling any nodes (1,091). Controlling 10 key nodes (2, 5, 7, 11–15, 19, and 20) and eliminating the factors of mixed-class hosting at school significantly reduced the step length of the curve (819). Eliminating the key nodes (48,59) are controlling the spread of vomiting during the recession period further reduced the step length (680) (Figure 4).

DISCUSSION

The network density of the norovirus outbreak gradually decreased from 0.871 to 0.099. This change was likely due to the concentration of cases within a single class during the outbreak period. As the outbreak progressed to the spread period, more classes were involved due to mixed-class activities at school and off-campus hosting. During the recession period, sporadic cases arose, likely due to the untimely isolation of infected individuals and improper handling of vomitus. Previous studies have identified risk factors for norovirus outbreaks in schools, including vomiting on campus and case activity in public areas (*5*). In this outbreak, mixed-class hosting during the spread period and the failure to standardize the handling of vomitus during the recession period were key factors

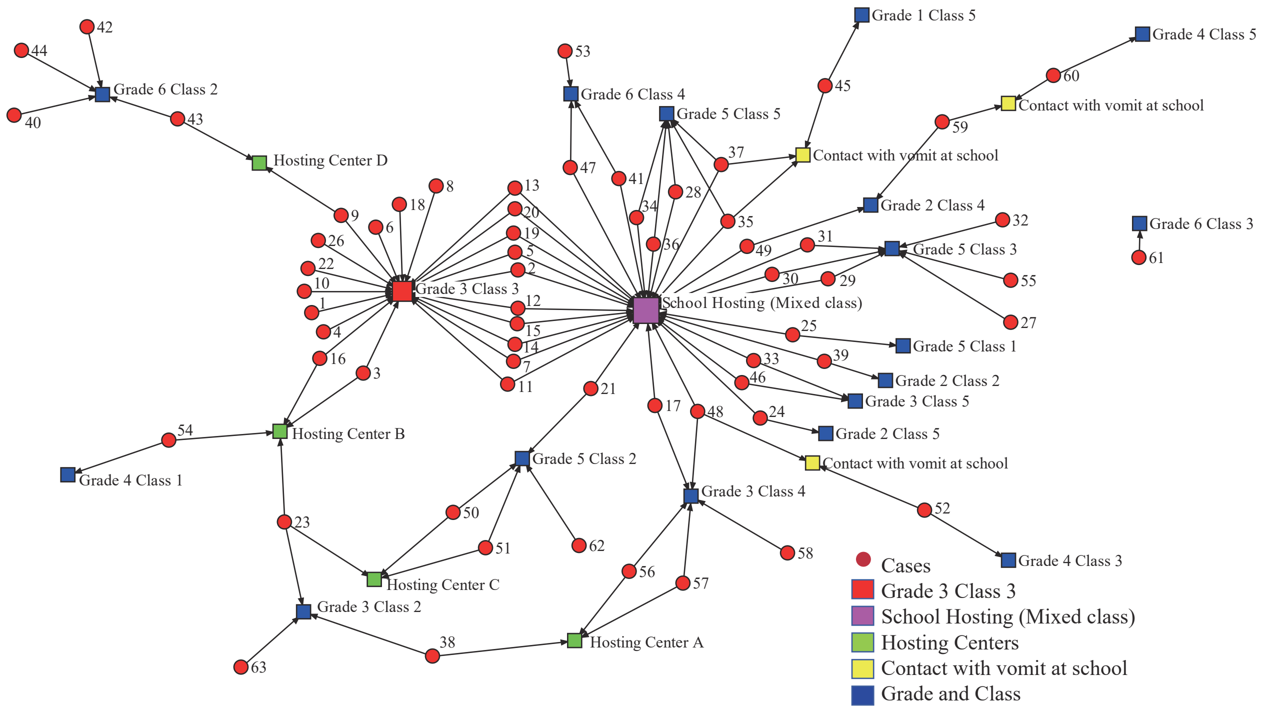

FIGURE 2. The social network topology of 63 cases of a norovirus outbreak at a primary school in Zhuhai City, Guangdong Province, China, 2023.

TABLE 1. The social network characteristic parameters and key nodes of different periods of development of a norovirus outbreak at a primary school in Zhuhai City, Guangdong Province, China, 2023.

Period	Density	Degree centralization		Betweenness centralization	Closeness centralization		Average distance	Core-periphery
		Out	In		Out	In		
Outbreak	0.871	0.083	0.135	0.005	0.502	0.890	1.085	17 core nodes: 2–9, 11–16, 18–20
		Rank 1–10: 2, 5, 7, 11, 12, 13, 14, 15, 19, 20						
Spread	0.495	0.303	0.324	0.139	0.274	0.579	1.793	27 core nodes: 2, 5, 7, 11–15, 19–21, 24, 25, 28–31, 33–37, 39, 41, 46, 47, 49
		Rank 1–10: 2, 5, 7, 11, 12, 13, 14, 15, 19, 20		Rank 1–10: 9, 43, 23, 3, 16, 2/5/7/11–15/19/20	Rank 1–10: 2, 5, 7, 11, 12, 13, 14, 15, 19, 20			
Recession	0.099	0.229	0.324	0.119	0.666	0.664	2.292	5 core nodes: 41, 46–49
		Rank 1–10: 48, 47, 41, 49, 53, 46, 56, 42, 58, 57		Rank 1–7: 48, 53, 49, 42, 41, 47, 59; Other node: 0	Rank 1–10: 52, 53, 41, 47, 48, 49, 46, 60, 56/57/58	Rank 1–10: 42, 43, 44, 48, 49, 41, 47, 46, 56/57/58		
Whole network	0.314	0.304	0.320	0.085	0.125	0.306	2.100	29 core nodes: 2, 5, 7, 11–15, 17, 19–21, 24, 25, 28–31, 33–37, 39, 41, 46–49
		Rank 1–10: 2, 5, 7, 11, 12, 13, 14, 15, 19, 20		Rank 1–10: 9, 43, 23, 48, 3, 16, 17, 49, 21, 29/31/30	Rank 1–10: 2, 5, 7, 11, 12, 13, 14, 15, 19, 20			

contributing to the outbreak and the tailing of the epidemic curve. These factors created critical nodes in the transmission network with priority connection characteristics, increasing the likelihood of disease spreading at each stage. Based on theoretical parameters and operational feasibility, we selected 10 key nodes in the transmission network as primary targets for spread control and 2 key nodes for regression control. Modeling simulations demonstrated that implementing corresponding control measures could effectively reduce the extent and spread of the norovirus.

Social network studies have contributed significantly to understanding the occurrence and development of infectious diseases in recent years (*6*). Studies have shown that immunizing or isolating a small number of nodes can effectively control infectious disease outbreaks (*7*–*8*). This study analyzed the social

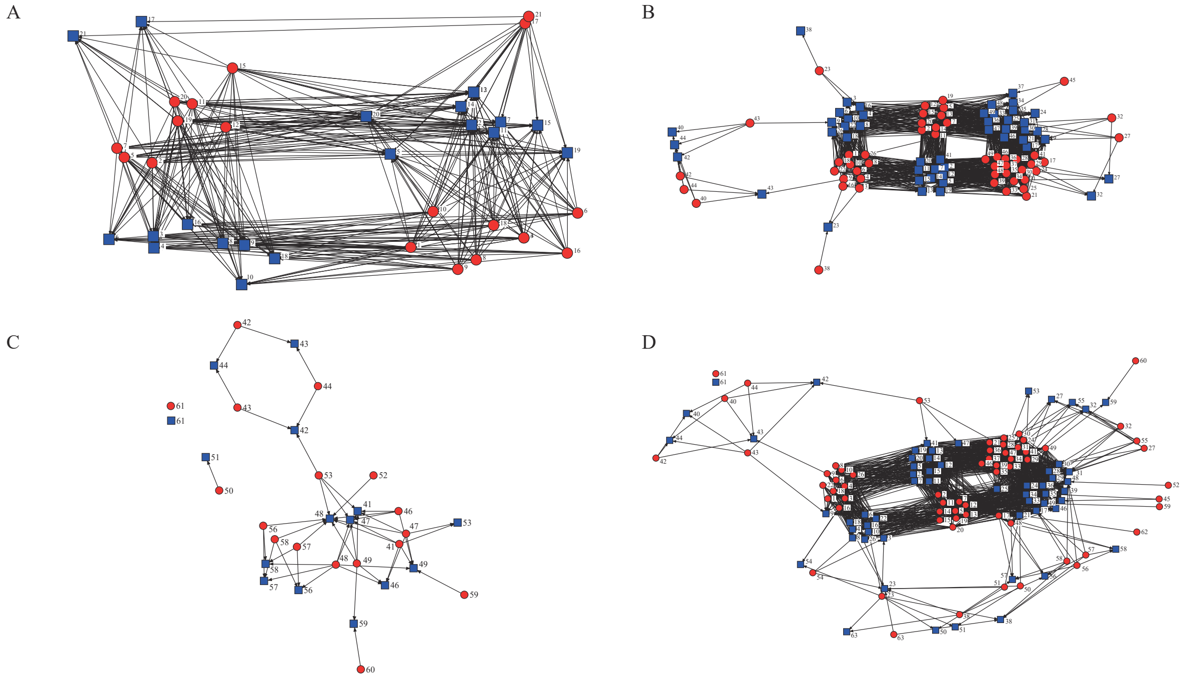

FIGURE 3. The social network matrix of different periods of development of a norovirus outbreak at a primary school in Zhuhai City, Guangdong Province, China, 2023. (A) Outbreak; (B) Spread; (C) Recession; (D) Whole network.

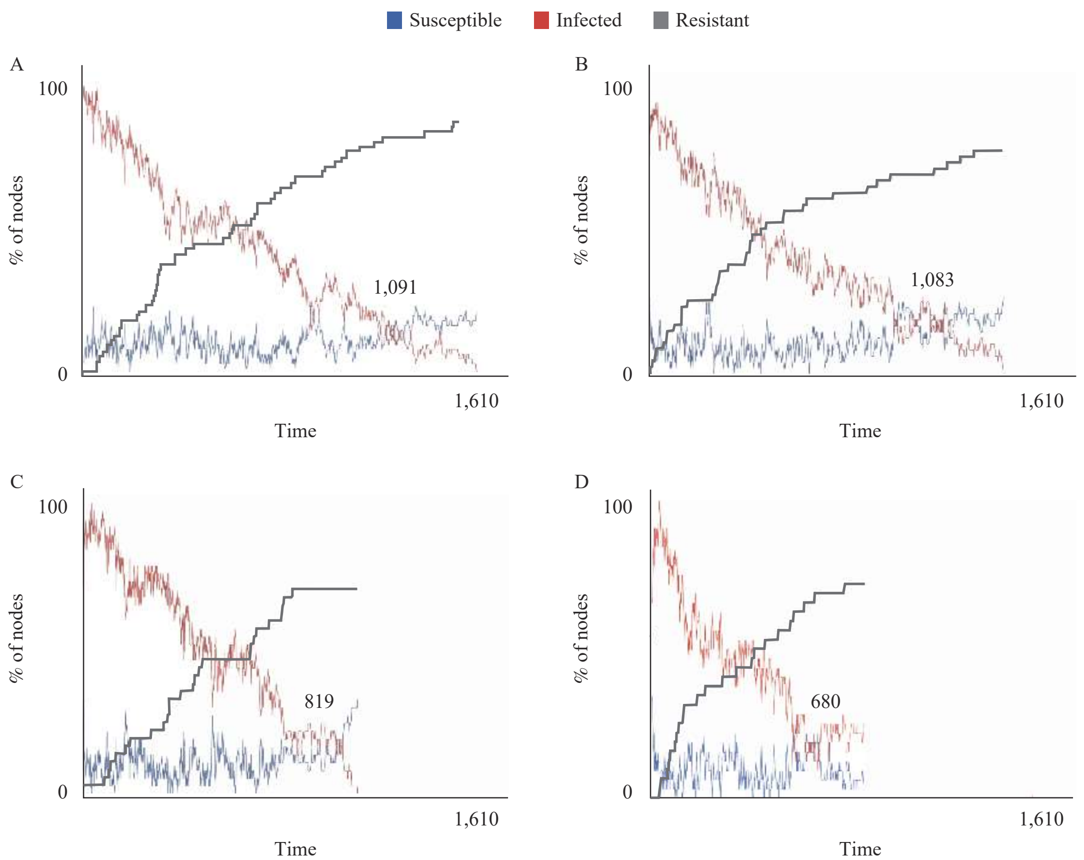

FIGURE 4. The modeling network status of different situations of a norovirus outbreak at a primary school in Zhuhai City, Guangdong Province, China, 2023. (A) Actual situation; (B) Ten random controlled nodes; (C) During the spread period, 10 key nodes (cases 2, 5, 7, 11–15, 19, and 20) were controlled; (D) Control of 10 key nodes and 2 key nodes (cases 48 and 59) in the recession period.

network spread patterns and relationships of 63 cases to verify the effectiveness of this approach in identifying network structures, clarifying core members, and improving decision-making. Using social network analysis, researchers can feasibly identify core nodes and key relationships in infectious disease transmission to accurately and quickly interrupt transmission pathways and prevent large-scale spread or new outbreaks by bridging crowds.

Case 61, who lacked clearly defined social relationships, appeared peripherally in the network. Additionally, despite exhibiting social relationships, some nodes within the network had symptom onset intervals exceeding the incubation period. Notably, studies have reported that latent norovirus infections can account for up to 17.6% of outbreak cases (*9*). Therefore, latent infection within the network may represent a potential source of transmission.

This survey did not include collecting and testing of samples from healthy students or school staff for norovirus nucleic acid, precluding assessment of the proportion of latent infections. Factors such as inadequate vomit disposal practices during the initial stages, non-standard disinfection protocols, and cross-contamination from shared mop use between classes may confound the assessment of network relationships.

CDC physicians might consider incorporating social network analysis when managing norovirus outbreaks in schools. Understanding transmission patterns to identify and control key nodes could help to terminate outbreaks expeditiously.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the Medical Science and Technology Research Fund Project of Guangdong Province (A2021021).

doi: 10.46234/ccdcw2024.232

Corresponding author: Feng Ruan, jkzxrf@zhuhai.gov.cn.

[1] Department of Health Emergency Management, Center for Disease Control and Prevention of Zhuhai City, Zhuhai City, Guangdong Province, China.

Submitted: April 03, 2024; Accepted: October 24, 2024

REFERENCES

1. Tan MD, Tian Y, Zhang DT, Wang QY, Gao ZY. Aerosol transmission of norovirus. Viruses 2024;16(1):151. https://doi.org/10.3390/v16010151.
2. Jin M, Wu SY, Kong XY, Xie HP, Fu JG, He YQ, et al. Norovirus outbreak surveillance, China, 2016-2018. Emerg Infect Dis 2020;26(3):437 – 45. https://doi.org/10.3201/eid2603.191183.
3. Liu J. Whole network approach-a practical guide to UCINET. 3nd ed. Shanghai: Shanghai People's Publishing House 2019. (In Chinese). https://weread.qq.com/web/bookDetail/b37321b0811e3e505g01678.
4. Xiao SJ, Yin XL, Lin XT, Liu DM, Wu YF, Ruan F. Investigation and analysis of an outbreak of norovirus infection in a School. Henan J Prev Med 2022;33(10):768 – 71. https://doi.org/10.13515/j.cnki.hnjpm.1006-8414.2022.10.010.
5. Liu BW, Wang Y, Tan MD, Boran E, Zhang DX, Yan HQ, et al. Risk factors for norovirus outbreaks in schools and kindergartens-Beijing municipality, China, July 2017–June 2022. China CDC Wkly 2024;6(33):841-5. https://weekly.chinacdc.cn/en/article/doi/10.46234/ccdcw2024.181.
6. Liu MX, He SS, Sun YZ, Liu MX, He SS, Sun YZ. The impact of media converge on complex networks on disease transmission. Math Biosci Eng 2019;16(6):6335 – 49. https://doi.org/10.3934/mbe.2019316.
7. Morone F, Makse HA. Influence maximization in complex networks through optimal percolation. Nature 2015;524(7563):65 – 8. https://doi.org/10.1038/nature14604.
8. Guo WF, Zhang SW, Shi QQ, Zhang CM, Zeng T, Chen LN. A novel algorithm for finding optimal driver nodes to target control complex networks and its applications for drug targets identification. BMC Genomics 2018;19(Suppl 1):924. https://doi.org/10.1186/s12864-017-4332-z.
9. Wang J, Ji ZH, Zhang SB, Yang ZR, Sun XQ, Zhang H. Asymptomatic norovirus infection during outbreaks in China: a systematic review and meta-analysis. J Med Virol 2024;96(1):e29393. https://doi.org/10.1002/jmv.29393.

SUPPLEMENTARY MATERIAL

Social network analysis is a research method used to visualize and analyze relationships and connections between entities or individuals within a network. It has emerged as a key technique in modern sociology (*1*).

Relationship Matrix

Social network analysis characterizes networked structures in terms of nodes (individual actors, people, or things within the network) and the ties, edges, or links (relationships or interactions) that connect them. These relationships or interactions can then be coded and converted into data suitable for network analysis (*2*).

The cases reported were used as the nodes and the relationships (links) between cases were identified in our study. It contained the following situations: ① in the same class, ② co-participation in school hosting (mixed class), ③ co-participation in the same off-campus hosting, or ④ contact with vomit (confirmed by the on-site investigation). The first index case reported was used as the initial node. A 63×63 transmission relationship matrix between each node (case) was established.

Visualizing Network

The social network topology was a visualization that represented the nodes and links of a network as a series of connected points. In our study, information was used to visually identify network factors. The cases that serve as bridges between different groups were shown. Nodes were taken in different shapes, colors, or sizes to reflect different characteristics of the cases (Figure 2 in the main text).

Procedure for using the UCINET 6.528 (*3*): Click on the Visualize→NetDraw→File→Open→UCINET dataset→Network

Characteristic Parameters

The characteristic parameters of the network from the initial outbreak through its spread and eventual decline were analyzed in our study (Table 1 and Figure 3 in the main text).

Density

Density is the percentage of possible links that are present in the network. More dense networks are characterized by having more connections between nodes, with 0 representing a network with no connections, and 1 representing a network where all nodes are connected to each other.

Procedure for using the UCINET 6.528: Click on the Network→Cohesion→Density

Centralization

There are three types of centralization (degree, betweenness, and closeness centralization), each corresponding to a different aspect of connectivity and centrality.

Degree Centralization

Degree centralization represents the number of links that a node has. It can be used to identify the most connected cases in the network. In-degree is the number of in-coming links, or the number of predecessor nodes; out-degree is the number of out-going links, or the number of successor nodes. For degree centralization, higher values mean that the node is more central. These cases are considered popular or active and they often have a strong influence within the network due to their higher values of degree centralization.

Procedure for using the UCINET 6.528: Click on the Network→Centrality and Power→Degree

Betweenness Centralization

Betweenness centralization is a centrality measure based on the number of paths of nodes between other nodes. It measures the extent to which a point is located in the "middle" of the other "point pairs" in the network. It can be used to identify these cases as having a unique position where they connect different parts of the network, facilitating or controlling the flow of information between others.The higher the values of betweenness centralization, the higher the ability of cases as mediators in the network.

Procedure for using the UCINET 6.528: Click on the Network→Centrality and Power→Betweenness

Closeness Centralization

A measure of how quickly a node can reach every other node in the network via the shortest paths. These nodes can disseminate information or exert influence quickly due to their close proximity to all other nodes. The higher the values of closeness centralization, the more important the cases were in the network.

Procedure for using the UCINET 6.528: Click on the Network→Centrality and Power→Closeness

Average Distance

The average distance is the average length of all shortest paths between all pairs of connected vertices in the corresponding network. It can be used to measure the efficiency of the information flow within the network which means that any two nodes can communicate with each other at an average distance between intermediate nodes.

Procedure for using the UCINET 6.528: Click on the Network→Cohesion→Multiple cohesion measures →Average distance

Core-periphery

Core-periphery analysis can identify which nodes are in the tightly connected core region and which nodes are in the loosely distributed periphery in the whole network.

Procedure for using the UCINET 6.528: Click on the Network→Core/Periphery→Categorical

Modeling Network Status

The modeling network parameters were set as follows: average node degree=10, initial outbreak size=21, and virus spread chance=4.2%.

Procedure for using the NetLogo 6.1.1: Click on the File→Model library→Networks→Virus on a network→Setup parameters→Go

REFERENCES

1. Broda MD, Granger K, Chow J, Ross E. Using social network analysis in applied psychological research: a tutorial. Psychol Methods 2023;28(4):791 – 805. https://doi.org/10.1037/met0000451.
2. Tabassum S, Pereira F S F, Fernandes S, Gama J. Social network analysis: an overview. WIREs: Data Min Knowl Discovery 2018;8(5):e1256. https://doi.org/10.1002/widm.1256.
3. Liu J. Lectures on whole network approach-a practical guide to UCINET. 2nd ed. Shanghai: Shanghai People's Publishing House 2014. http://find.nlc.cn/search/showDocDetails?docId=-7227913879254502304&dataSource=ucs01&query=%E6%95%B4%E4%BD%93%E7%BD%91%E5%88%86%E6%9E%90. (In Chinese).