Chinese Public Attention to COVID-19 Epidemic: Based on Social Media

Yuxin Zhao, Sixiang Cheng, Xiaoyan Yu, Huilan Xu

Table of Contents

Chinese Public Attention to COVID-19 Epidemic: Based on Social Media

Yuxin ZhaoMD, ; Sixiang ChengPhD, ; Xiaoyan YuMD, ; Huilan XuPhD,

Corresponding Author:
Huilan XuPhD,
Phone: +8613637480958
Email: xhl_csu@163.com

Abstract

Background: Since the new coronavirus epidemic in China in December 2019, information and discussions about COVID-19 have spread rapidly on the Internet and have quickly become the focus of worldwide attention, especially on social media.

Objective: This study aims to investigate and analyze the public's attention to COVID-19-related events in China at the beginning of the COVID-19 epidemic in China (December 31, 2019, to February 20, 2020) through the Sina Microblog hot search list.

Methods: We collected topics related to the COVID-19 epidemic on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020 and described the trend of public attention on COVID-19 epidemic-related topics. ROST CM6.0 (ROST Content Mining System Version 6.0) was used to analyze the collected text for word segmentation, word frequency, and sentiment analysis. We further described the hot topic keywords and sentiment trends of public attention. We used VOSviewer to implement a visual cluster analysis of hot keywords and build a social network of public opinion content.

Results: The study has four main findings. First, we analyzed the changing trend of the public's attention to the COVID-19 epidemic, which can be divided into three stages. Second, the hot topic keywords of public attention at each stage are slightly different. In addition, the emotional tendency of the public toward the COVID-19 epidemic-related hot topics has changed from negative to neutral, with negative emotions weakening and positive emotions increasing as a whole. Finally, we divided the COVID-19 topics with the most public concern into five categories: (1) new COVID-19 epidemics and their impact; (2) frontline reporting of the epidemic and prevention and control measures; (3) expert interpretation and discussion on the source of infection; (4) medical services on the frontline of the epidemic; and (5) focus on the global epidemic and the search for suspected cases.

Conclusions: This is the first study of public attention on the COVID-19 epidemic using a Chinese social media platform (i.e., Sina Microblog). Our study found that social media (e.g., Sina Microblog) can be used to measure public attention to public health emergencies. During the epidemic of the novel coronavirus, a large amount of information about the COVID-19 epidemic was disseminated on Sina Microblog and received widespread public attention. We have learned about the hotspots of public concern regarding the COVID-19 epidemic. These findings can help the government and health departments better communicate with the public on health and translate public health needs into practice to create targeted measures to prevent and control the spread of COVID-19.

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
    Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
    Only make the preprint title and abstract visible.
    No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
    Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
    Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Yuxin Zhao, MD[1], Huilan Xu, PhD[1*]

Affiliations: Department of Social Medicine and Health Management, Xiangya School of Public Health, Central South University, China

Correspondence to: Huilan Xu, PhD

Department of Social Medicine and Health Management, Xiangya School of Public Health, Central South University, Changsha, Hunan, 410078, China

E-mail: xhl_csu@163.com

Tel.: 86-13637480958

Chinese Public Attention to COVID-19 Epidemic: Based on Social Media

Abstract

Background: Since the new coronavirus epidemic in China in December 2019, information and discussions about COVID-19 have spread rapidly on the Internet and have quickly become the focus of worldwide attention, especially on social media.

Objective: This study aims to investigate and analyze the public's attention to COVID-19-related events in China at the beginning of the COVID-19 epidemic in China (December 31, 2019, to February 20, 2020) through the Sina Microblog hot search list.

Methods: We collected topics related to the COVID-19 epidemic on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020 and described the trend of public attention on COVID-19 epidemic-related topics. ROST CM6.0 (ROST Content Mining System Version 6.0) was used to analyze the collected text for word segmentation, word frequency, and sentiment analysis. We further described the hot topic keywords and sentiment trends of public attention. We used VOSviewer to implement a visual cluster analysis of hot keywords and build a social network of public opinion content.

Results: The study has four main findings. First, we analyzed the changing trend of the public's attention to the COVID-19 epidemic, which can be divided into three stages. Second, the hot topic keywords of public attention at each stage are slightly different. In addition, the emotional tendency of the public toward the COVID-19 epidemic-related hot topics has changed from negative to neutral, with negative emotions weakening and positive emotions increasing as a whole. Finally, we divided the COVID-19 topics with the most public concern into five categories: (1) new COVID-19 epidemics and their impact; (2) frontline reporting of the epidemic and prevention and control measures; (3) expert interpretation and discussion on the source of infection; (4) medical services on the frontline of the epidemic; and (5) focus on the global epidemic and the search for suspected cases.

Conclusions: This is the first study of public attention on the COVID-19 epidemic using a Chinese social media platform (i.e., Sina Microblog). Our study found that social media (e.g., Sina Microblog) can be used to measure public attention to public health emergencies. During the epidemic of the novel coronavirus, a large amount of information about the COVID-19 epidemic was disseminated on Sina Microblog and received widespread public attention. We have learned about the hotspots of public concern regarding the COVID-19 epidemic. These

findings can help the government and health departments better communicate with the public on health and translate public health needs into practice to create targeted measures to prevent and control the spread of COVID-19.

Keywords: COVID-19; Sina Microblog; Public attention; Social media; China; Public health emergency

Introduction

Pneumonia caused by the novel coronavirus (COVID-19) is a new infectious disease that is mainly transmitted by respiratory droplets and contact and is generally infectious to human beings[1]. On January 11, 2020, after pathogenic nucleic acid testing (NAT), China reported 41 cases of pneumonia infected with the novel coronavirus (SARS-CoV-2) for the first time[2], which was the world's first reported human infection with the novel coronavirus. On January 30, 2020, the World Health Organization listed the novel coronavirus epidemic as a Public Health Emergency of International Concern (PHEIC) [3]. As of February 20, 2020, a total of 75,465 confirmed cases and 2,236 deaths have been reported in mainland China [4]. The novel coronavirus has caused great challenges and threats to public health in China and has quickly become the focus of worldwide attention. Information and discussions about COVID-19 have spread rapidly online, especially on social media.

To fight against COVID-19 and promote the prevention and control of the epidemic, researchers have recently made efforts in various aspects; their research involves epidemiological research [5,6,7], diagnostic methods of the novel coronavirus [8,9,10,11], clinical characteristics of the disease [12,13,14,15,16], characteristics of disease transmission [17,18,19], development of candidate therapies [20,21,22], the identification of animal hosts [23,24,25,26], etc. However, there has been no research conducted on the public's attention to COVID-19. Since public participation is required to prevent and control the epidemic spread of infectious diseases, it is extremely important to learn about the public's attention to COVID-19 during the current epidemic. Such knowledge is of great significance when guiding people to respond appropriately to the epidemic and helping them learn how to cope with the sudden infectious disease of COVID-19, and it also supports social stability [27].

Social media has developed rapidly in recent years. Increasing numbers of public health departments and individuals are using social media platforms to communicate and share information during public health emergencies. Social media has become an important channel for promoting risk communication during the crisis[28,29]. The use of social media to measure public attention has also been gradually applied to research on infectious diseases, such as H7N9[27,30,31], Ebola[28,32,33, 34,35,36], Zika virus[29, 37,38], MERS-CoV[39], and Dengue fever[40].

Sina Microblog ("Weibo" for short) is one of the most popular social media platforms in China and is equivalent to Twitter within China. As of the fourth quarter of 2018, the number of monthly active users had reached 462 million, and approximately 200 million people are using Sina Microblog every day[41]. The Sina Microblog hot search list is the ranking of the most followed and hottest information on Sina Microblog and is the most popular functional module in Sina Microblog applications[42, 43]. This ranking is sorted according to the search volume of the hot topics that users searched for within a certain period. The higher the search volume is, the higher the ranking is, which directly reflects the public attention and attitude toward the topic.

This paper studies public attention given to COVID-19 on Sina Microblog by searching for and

analyzing topics related to the COVID-19 epidemic on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020. This is the first time that a Chinese social media platform (i.e., Sina Microblog) has been used to study the public attention of COVID-19 [44]. In this study, we describe the trend of public attention given to topics related to the COVID-19 epidemic and the hot topic keywords of public concern, we analyze the emotional tendencies and trends of hot topics related to the COVID-19 epidemic, and we conduct a visual cluster analysis of the hot topic content. This approach is used to obtain timely access to public responses so that the government and the health department can better communicate with the public on health issues and take appropriate measures to prevent and control the epidemic [35, 39].

Methods

Research Overview

The research process mainly includes five steps: (1) collect topics related to the COVID-19 epidemic on the Sina Microblog hot search list, (2) segment the collected text into words, (3) describe the Sina Microblog search trend about the COVID-19 epidemic; (4) evaluate public opinion through word frequency and sentiment analysis; and (5) construct a social network of public opinion through the subject analysis of public opinion content. Each step is described in detail below.

Data Collection

We obtained information on the COVID-19 infection in mainland China from the National Health Committee of the People's Republic of China (PRC NHC) [45]. The Wuhan Municipal Health Committee first reported viral pneumonia of unknown causes on December 31, 2019 [46]. This study collected topics on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020, using Weibo Search Rank[47] from ENLIGHTENT[48], and selected topics and search volume related to the COVID-19 epidemic. A total of 4,056 topics related to the COVID-19 epidemic were on the hot search list, and 3,234 remained after excluding duplicate topics; this group was used as the data basis for further processing, analysis, and discussion.

Data Processing

The topics related to the COVID-19 epidemic on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020, were summarized and classified daily in chronological order. We used the Chinese words segmentation function in ROST CM6.0 (ROST Content Mining System Version 6.0) to segment the content of hot-search topics on Sina Microblog into words [49]. After word segmentation, the text was processed by merging synonyms and deleting nonsense words to provide a basis for subsequent research.

Data Analysis

Trend Analysis

We plotted the number of topics related to the COVID-19 epidemic on the Sina Microblog hot search list and the cumulative search volume by date to explore the public's attention given to the COVID-19 epidemic over time. Pearson correlation analysis was used to explore the relationship between the number of hot search topics related to the COVID-19 epidemic and the cumulative search volume.

Word frequency and sentiment analysis

After Chinese word segmentation and invalid word filtering, we used ROST CM6.0 software to perform word frequency statistics and sentiment analysis on the Sina Microblog hot topics related to the COVID-19 epidemic. We calculated the frequency of keywords appearing in the hot search topics on Sina Microblog, explored the sentiment tendency of each hot search topic to the COVID-19 epidemic, and calculated the emotional score.

Social network analysis and visualization

The high-frequency keywords and their frequencies were extracted from the Sina Microblog hot topic text after word segmentation and invalid word filtering. We used VOSviewer software to create a visual knowledge map of keyword co-occurrence analysis and cluster analysis, using the keyword co-occurrence frequency as the weight[50].

Results

Search trend of COVID-19 epidemic on Sina Microblog hot search list

Figure 1 lists the number of COVID-19 epidemic-related topics and the cumulative search volume on the Sina Microblog hot search list from December 31, 2019, to February 20, 2020. Pearson correlation analysis shows that the number of topics related to the COVID-19 epidemic is positively correlated with the cumulative search volume of the topic per day (Pearson correlation coefficient = 0.767, P = .000). In other words, the more topics are listed, the greater the cumulative search volume related to the COVID-19 epidemic per day. We can see that the public's attention to the COVID-19 epidemic on Sina Microblog can be divided into three stages. Stage A (December 31, 2019, to January 18, 2020) has low and unstable public attention, which represents the incubation period. Stage B (January 19 to 26, 2020) has a concentrated increase in public attention, which represents the epidemic period. Stage C (January 27 to February 20, 2020) demonstrates continued public attention to the epidemic, representing a widespread period.
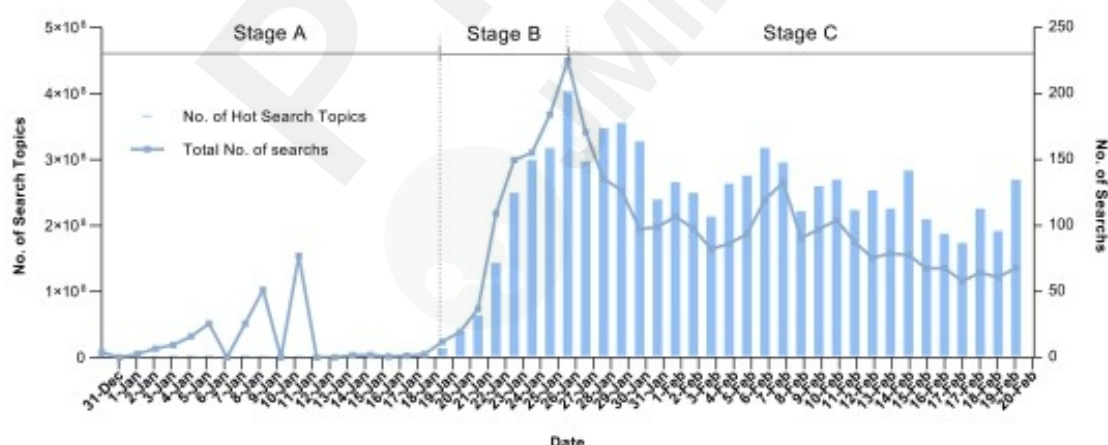

Figure 1. From December 31, 2019, to February 2, 2020, the number of topics and cumulative search volume of the COVID-19 epidemic on the Sina Microblog hot search list.

Frequency and distribution of hot search keywords related to the COVID-19 epidemic on Sina Microblog

The top 15 keywords and their frequencies for the three stages of the public's attention to the COVID-19 epidemic are shown in Table 3. We can find that "Wuhan", "case" and "pneumonia" always appear in three periods as hot keywords, and the remaining keywords in the different periods are slightly different. In stage A, "unknown cause" and "new coronavirus" are the main keywords, indicating that in the initial stage of the epidemic, viral pneumonia had just been discovered, and the cause was unknown. After the pathogen was initially identified as a new coronavirus, the public began to search for information on new coronaviruses to learn relevant knowledge. Compared with stage A, stage B has new keywords such as "new", "mask", "first case", and "first-level response". The reason for this is that the COVID-19 epidemic had spread across the country during this stage. The first cases had appeared successively throughout the country, and the number of confirmed cases was increasing. The outbreak of the epidemic made the people and the government aware of the importance of prevention. People began to buy masks, and governments at all levels initiated first-level responses to major public health emergencies. Compared with the previous two stages, the main keywords appearing in stage C are "discharged", "national", "materials" and "Huoshenshan Hospital". At this stage, the public's attention had shifted to material donation and medical service assistance in the key epidemic areas in Wuhan. Moreover, the epidemic had spread throughout the country, and the public was more concerned about the rehabilitation of patients.

Table 1. Top 15 keywords about the COVID-19 epidemic on the Sina Microblog hot search list in the three periods.

| Stage A | | Stage B | | Stage C | |
Dec 31, 2019, to Jan 18, 2020		Jan 19 to Jan 26, 2020		Jan 27 to Feb 20, 2020	
Key Word	Freq	Key Word	Freq	Key Word	Freq
Wuhan	17	COVID-19	141	Wuhan	316
pneumonia	14	Wuhan	124	COVID-19	283
unknown cause	9	confirmed diagnosis	87	case	212
novel coronavirus	6	case	84	confirmed diagnosis	212
patient	5	add	52	add	199
case	5	masks	37	epidemic	190
Thailand	4	epidemic	31	Hubei	159
leave hospital	2	Hubei	29	masks	149
death	2	startup	25	patient	149
person-to-person spread	2	first case	22	hospital	98

add	2	pneumonia	20	leave hospital	90
Japan	2	Beijing	20	nationwide	83
COVID-19	2	patient	19	Beijing	20
epidemic disease	2	first-level response	18	goods and materials	59
eliminate	2	novel coronavirus	17	Huoshenshan hospital	59

[a] Abbreviations: Freq, Frequency; COVID-19, coronavirus disease - 2019.

Sentiment analysis of the hot search topics related to the COVID-19 epidemic on the Sina Microblog

Afer segmentation, we imported the Sina Microblog hot topic text related to the COVID-19 epidemic into the ROST6.0 sentiment analysis tool and obtained the sentiment and proportion of the three stages of public attention given to the topics related to the COVID-19 epidemic on Sina Microblog (Figure 2). Emotions are classified as positive, negative, and neutral emotions. Based on this, positive and negative emotions are subdivided into three categories: general, moderate, and high. Neutral emotions are not subdivided. We found that the sentiment of the hot topics of the COVID-19 epidemic on Sina Microblog in stage A tended to be negative, accounting for 29.17%, of which high and moderate negative emotions accounted for 4.17% and 16.67%, respectively, and positive emotions accounted for the lowest proportion at 8.33%. In stage B, the majority of the hot search topics were neutral; however, 22.85% of the hot search topics showed negative emotions, with 0.31% showing high negative and 3.60% showing moderate negative, while positive emotions accounted for 12.52%. In stage C, 15.44% of the hot search topics showed positive emotions, of which moderate positive emotions accounted for 2.26% and highly positive emotions accounted for 0.31%; in addition, 24.33% of the hot search topics showed negative emotions. Based on the comprehensive analysis of the three-stage emotional tendencies, the public's negative emotions toward the COVID-19 epidemic were weakened as a whole, and their positive emotions were generally enhanced.
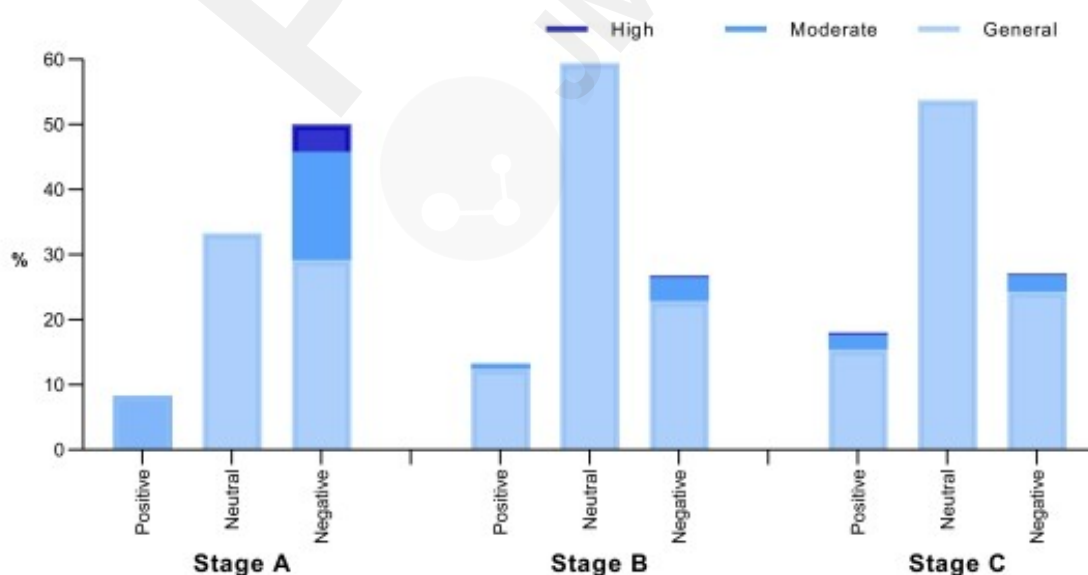
Figure 2. The sentiment statistics of the Sina Microblog hot search topics related to the COVID-19 epidemic.

Figure 3 shows the trend of the proportion of the daily emotional tendencies of the hot search topics related to the COVID-19 epidemic from December 31, 2019, to February 20, 2020. We can see that the three kinds of emotions are relatively unstable before January 9, 2020. From January 9 to January 20, 2020, negative emotions accounted for the largest proportion, followed by neutral emotions, and positive emotions accounted for the smallest proportion.

Compared with before January 20, 2020, the positive emotions of the hot search topics related to the COVID-19 epidemic after January 20, 2020 are generally on the rise, the negative emotions decline on the whole, and the emotions tend to be stable, as seen in Figure 3. This outcome shows that as the COVID-19 epidemic began to spread throughout the country after January 20, 2020, the public eased their concerns and fears caused by their uncertainty toward and ignorance of the epidemic and responded to the COVID-19 epidemic with a more objective attitude.
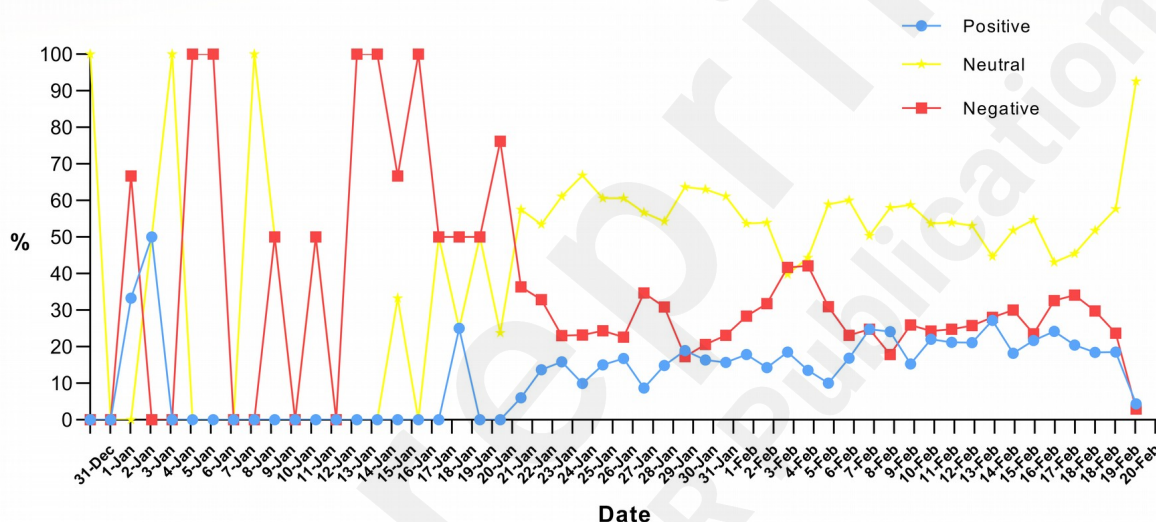
Figure 3. The sentiment trend of hot search topics related to the COVID-19 epidemic from December 31, 2019, to February 2, 2020.

Social semantic network analysis of hot search topics related to the COVID-19 epidemic on Sina Microblog

To explore the themes reflected by the related topics of the COVID-19 epidemic on the Sina Microblog hot search list, this study used VOSviewer to generate clusters and co-occurrence networks of topic keywords. The results are shown in Figure 4. In the figure, the larger the size of the nodes and the font, the greater the weight of the keyword is, and the more it is in the core position. The connection between the nodes indicates that the keywords on the two nodes have appeared together. The thicker the connection is, the higher the frequency of co-occurrence and the closer the connection between the nodes. Nodes with the same color belong to the same cluster.
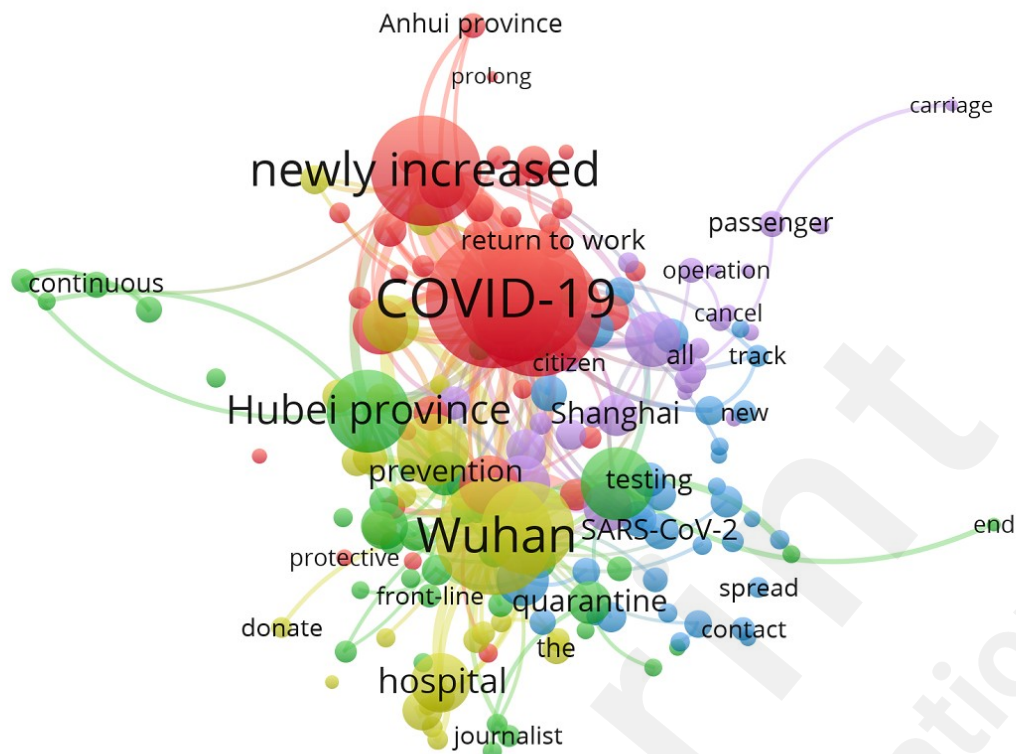
Figure 4. Social network of high-frequency keywords in the Sina Microblog hot search topics related to the COVID-19 epidemic.

According to the network visualization graph constructed in Figure 4, we can see that the keyword "COVID-19" is at the core node position, and the two nodes "Wuhan" and "add" are next. The core topic of public concern about the COVID-19 epidemic is COVID-19 itself, and the public was extremely concerned about the status of the epidemic in Wuhan and the new cases. We can divide the high-frequency keywords of topics into the following five clusters:

Cluster 1 (Red Cluster): Discussion on the new cases, the outbreak of COVID-19 across the country and the impact of the epidemic on the resumption of school and work. The keywords included are "case", " newly increased ", "appear", "Hubei Province", "Beijing", etc. (Figure 5).

Cluster 2 (Green Cluster): Search for news reports on the frontline of the epidemic and related measures for prevention and control. The keywords include "Wuhan", "front-line", "inpatient", "quarantine", "living at home", etc. (Figure 6).

Cluster 3 (Blue Cluster): Search for interpretations of the epidemic situation and prevention and control knowledge of experts and relevant health departments, as well as the discussion on the source of infection. The keywords in this cluster include "academician", "WHO", "face mask", "epidemic prevention", "symptom", "virus", and "novel coronavirus". (Figure 7).

Cluster 4 (Yellow Cluster): Search for frontline medical services such as frontline hospital construction and medical team support. This cluster includes the keywords "patient", "hospital", "Leishenshan", "Huoshenshan", "first batch", etc. (Figure 8).

Cluster 5 (Purple Cluster): Search for the global spread of the disease and searching online for

fellow passengers with confirmed cases. This cluster contains the keywords "passenger", "carriage", "cruise", "suspend", "infection", "Japan", etc. (Figure 9).
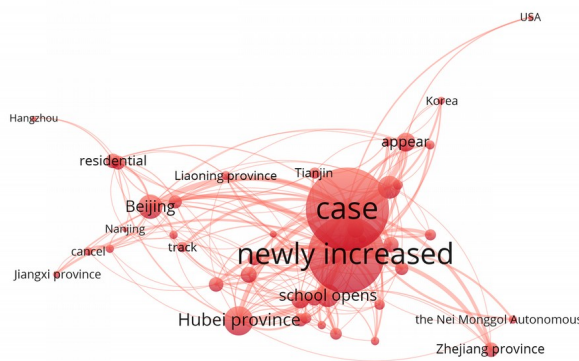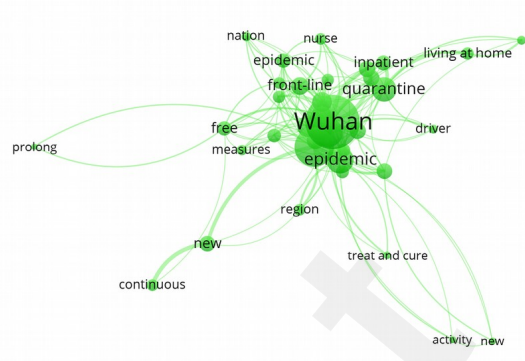

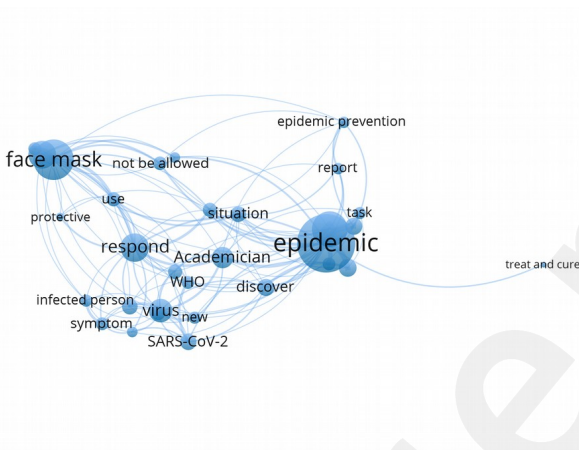Figure 5. Cluster 1   Figure 6. Cluster 2
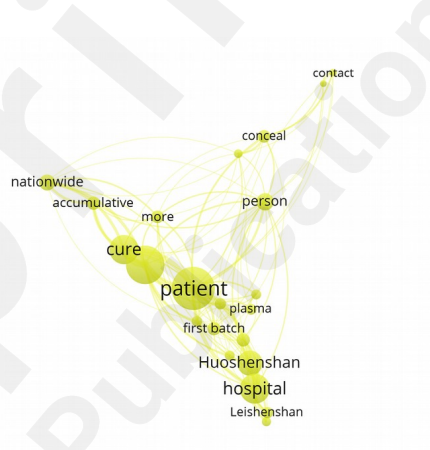

Figure 7. Cluster 3   Figure 8. Cluster 4

Figure 9. Cluster 5

Discussion

Principal Results

In recent years, as people increasingly seek health information on social media, social media

has played an increasingly important role in public health emergencies[29,35]. However, there have been no relevant studies that have incorporated social media analysis into the public response to the COVID-19 epidemic. We used the Sina Microblog hot search list to analyze the public's attention to the COVID-19 epidemic from December 31, 2019, to February 20, 2020, which was 52 days after the first disclosure of COVID-19 by the Chinese health department. There are four main findings from this study. First, we analyzed the changing trend of public attention given to the COVID-19 epidemic, which can be divided into three stages. Second, the hot topic keywords of the public's attention at each stage are slightly different. In addition, the public's emotional tendency toward the hot search topics related to the COVID-19 epidemic changed from negative to neutral across the study period. As a whole, negative emotions weakened, and positive emotions increased. Finally, we divided the topics of public concern about the COVID-19 epidemic into five categories through social semantic network analysis. This study analyzes the public's response to and attention given to the COVID-19 epidemic, which will help public health professionals monitor public response, identify public needs as early as possible, make timely public health prevention and control measures, and disseminate knowledge to citizens in a targeted manner to better respond to the current COVID-19 epidemic.

From December 31, 2019, to February 20, 2020, the public's attention given to the COVID-19 epidemic on Sina Microblog can be divided into three stages. In the beginning, there was little public attention paid to the epidemic, and then the concentration of attention increased. Next, the public's attention given to the epidemic generally declined, but continued attention was still paid. In the first stage (December 31, 2019, to January 18, 2020), the Chinese public paid less attention, as pneumonia was only reported in Wuhan, China. On January 19, 2020, suspected cases appeared in Shanghai and Shenzhen, and body temperatures began to be measured at the Wuhan Airport and Railway Station. People began to notice the severity of the epidemic, and the level of attention paid to the COVID-19 epidemic began to increase until January 26, 2020. We can see that when the epidemic began to spread across the country, the public responded quickly to news of the COVID-19 outbreak on Sina Microblog[39]. This is because COVID-19 is a new infectious disease, which means that no effective treatment has been found and no corresponding vaccine has been developed. When the outbreak began, the public was eager to search the relevant knowledge and information online to meet their own protection needs[27,30]. In the third stage (January 27 to February 20, 2020), although the number of cases of COVID-19 was still increasing, the number of Sina Microblog topics and search volume was decreasing. The reason for this may be because the epidemic information was relatively saturated, and it was difficult for people to acquire more new knowledge through a Sina Microblog search[27]. It may also be that people are no longer paying close attention to the epidemic dynamic to gain a sense of security as time goes on, and the public's consciousness tends to be rational[51].

The keywords for the three different stages of the public's attention to the COVID-19 epidemic are slightly different. From the hot search keywords of each stage, we found that the public did not know much about the virus and its causes in the first stage. The main keywords of the search to seek relevant knowledge were "unknown cause" and "novel coronavirus". During the second stage, the epidemic began to spread throughout the country, the first cases were reported in different areas successively, and the number of confirmed cases continued to increase, thereby making the public and government aware of the importance of prevention. During the third stage, the epidemic spread widely throughout the country; Wuhan was the most seriously affected area, and the city was locked down. The public's attention was mainly shifted to material donation and medical assistance in Wuhan. Compared with a qualitative research, it is more accurate to obtain the hot spots of public attention through keyword

frequency analysis, as this process has higher accuracy and credibility in the research of hot spots and their development trends[40]. Based on our high-frequency keyword analysis of the Sina Microblog topics related to the COVID-19 epidemic, we can obtain information about the concerns and opinions of Sina Microblog users at different stages[52]. Studies have found that people's interest in infectious diseases on social media is linked to the latest news and major events. Studies have also shown that people will pay attention to and search for disease-related words as the spread of infectious diseases changes [30].

According to the sentiment tendencies of the hot search topics on the COVID-19 epidemic on Sina Microblog, the first stage of emotion was negative, and the second and third stages were neutral. On the whole, negative emotions weakened, and positive emotions increased. Previous studies have pointed out that there is also an important relationship between emotions and content on social media[53]. The content analysis of social networks can identify people's attitudes or reactions to specific health hazard events[30,39]. In the first stage, there was less public attention given to the Sina Microblog hot search list. Most of the topics related to the COVID-19 epidemic were about the notification of pneumonia and the virus, and the emotions tended to be negative. At that time, the public had a strong demand for information on public health emergencies such as infectious diseases. When the information demand could not be fully satisfied, the users' emotions were negative[40]. As the epidemic progressed to the second and third stages, the public sentiment tended to be neutral because increasingly more news was reported at this stage, and objective events became the mainstream information on the Sina Microblog hot search list. The public reduced their previous levels of worries and fears about the epidemic; their negative feelings weakened, and their positive emotions increased. More hot search topics mentioned information about prevention or protection, which is conducive to public health communication and promotion.

We divided the COVID-19 topics with the highest levels of public concern into five categories: (1) new COVID-19 epidemics and their impact, (2) frontline reporting of the epidemic and prevention and control measures, (3) expert interpretation and discussion on the source of infection, (4) medical services on the frontline of the epidemic, and (5) focus on the global epidemic and look for suspected cases. From the search of the subject content, we can see that during the outbreak of the novel coronavirus, the public, the news media, and the health department all actively used Sina Microblog as a platform for disseminating COVID-19-related information[29], indicating that Sina Microblog is a communication channel for both individuals or organizations to publicize COVID-19 symptoms, preventive measures and related policies[37]. In addition, as a real-time and extensive online platform, Sina Microblog provides a channel for information dissemination. For example, in this emergency outbreak, the public has made good use of the platform to find fellow passengers with confirmed cases, thereby playing an important role in preventing and controlling the disease transmission.

Limitations

There are some limitations to this study. First, our study was limited to the period 52 days after COVID-19 was first disclosed by the Chinese health department; thus, the situation after February 20, 2020, was not included in this study. Second, the data source for this study is relatively narrow. We regard Sina Microblog as the only social media platform in China, excluding other popular social media data sources, such as WeChat and ByteDance, and the study is limited to the publicly available data on the Sina Microblog hot search list. In addition, due to the lack of detailed information about the users who contributed to the search volume of the Sina Microblog hot search list, we could not describe the social demographic information of

the Sina Microblog users, and we failed to obtain the geospatial distribution of Sina Microblog active users; thus, we could not calculate the average attention of the public in different regions of China.

Conclusions

Our study found that social media platforms (i.e., Sina Microblog) can be used to measure public attention given to public health emergencies. Our study shows that a large amount of information about the COVID-19 epidemic was disseminated and received widespread public attention on Sina Microblog during the novel coronavirus epidemic. We have learned about the hotspots of public concern regarding the COVID-19 epidemic on Sina Microblog. These findings can help the government and health departments better communicate with the public about public health and then translate public health needs into practice to create targeted measures to prevent and control the spread of COVID-19.

Acknowledgments

Not applicable.

Conflicts of Interest

None declared.

Abbreviations

COVID-19: Corona Virus Disease 2019

References

1. 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□[in Chinese]. *Chinese Journal of Epidemiology*. 2020(02):139-144. [doi:10.3760/cma.j.issn.0254-6450.2020.02.002]

2. Wuhan Municipal Health Committee. Report of Wuhan Municipal Health Committee on viral pneumonia of unknown cause [in Chinese]URL: http://wjw.wuhan.gov.cn/front/web/showDetail/2020011109035 [accessed 2020-02-21]

3. World Health Organization. Novel Coronavirus(2019-nCoV) Situation Report - 11[in Chinese]URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf?sfvrsn=de7c0f7_4 [accessed 2020-02-21]

4. National Health Commission of the people's Republic of China. Pneumonia situation in novel coronavirus infection on January 21st[in Chinese]URL: http://www.nhc.gov.cn/xcs/yqtb/202001/930c021cdd1f46dc832fc27e0cc465c8.shtml [accessed 2020-02-21]

5. Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health.* [doi:10.1016/S2589-7500(20)30026-1]

6. Xu B, Kraemer MUG, Xu B, et al. Open access epidemiological data from the COVID-19

outbreak. *The Lancet Infectious Diseases.* [doi:10.1016/S1473-3099(20)30119-5]

7. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. *New England Journal of Medicine.* 2020. [doi:10.1056/NEJMp2002125]

8. Mohamed E-T, Haim H. B, Jinzhao S. *A Single and Two-Stage, Closed-Tube, Molecular Test for the 2019 Novel Coronavirus (COVID-19) at Home, Clinic, and Points of Entry.* 2020. [doi:10.26434/chemrxiv.11860137.v1]

9. Li Q, Ding X, Xia G, et al. A simple laboratory parameter facilitates early identification of COVID-19 patients. *medRxiv.* 2020:2020.2002.2013.20022830. [doi:10.1101/2020.02.13.20022830]

10. To KK-W, Tsang OT-Y, Yip CC-Y, et al. Consistent Detection of 2019 Novel Coronavirus in Saliva. *Clinical Infectious Diseases.* 2020. [doi:10.1093/cid/ciaa149]

11. Yang Y, Yang M, Shen C, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv.* 2020:2020.2002.2011.20021493. [doi:10.1101/2020.02.11.20021493]

12. Pan X, Chen D, Xia Y, et al. Asymptomatic cases in a family cluster with SARS-CoV-2 infection. *The Lancet Infectious Diseases.* [doi:10.1016/S1473-3099(20)30114-6]

13. Xu X-W, Wu X-X, Jiang X-G, et al. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series. *BMJ.* 2020;368:m606. [doi:10.1136/bmj.m606]

14. Silverstein WK, Stroud L, Cleghorn GE, Leis JA. First imported case of 2019 novel coronavirus in Canada, presenting as mild pneumonia. *The Lancet.* 2020;395(10225):734. [doi:10.1016/S0140-6736(20)30370-6]

15. Pan F, Ye T, Sun P, et al. Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19) Pneumonia. *Radiology.* 2020:200370. [doi:10.1148/radiol.2020200370]

16. Fang Y, Zhang H, Xu Y, Xie J, Pang P, Ji W. CT Manifestations of Two Cases of 2019 Novel Coronavirus (2019-nCoV) Pneumonia. *Radiology.* 2020:200280. [doi:10.1148/radiol.2020200280]

17. Liang W, Guan W, Chen R, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *The Lancet Oncology.* 2020;21(3):335-337. [doi:10.1016/S1470-2045(20)30096-6]

18. Chen J. Pathogenicity and transmissibility of 2019-nCoV—A quick overview and comparison with other emerging viruses. *Microbes and Infection.* 2020. [doi:10.1016/j.micinf.2020.01.004]

19. Mizumoto K, Kagaya K, Chowell G. Early epidemiological assessment of the transmission potential and virulence of 2019 Novel Coronavirus in Wuhan City: China, 2019-2020. *medRxiv.* 2020:2020.2002.2012.20022434. [doi:10.1101/2020.02.12.20022434]

20. Micholas S, Jeremy C. S. *Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface.* 2020. [doi:10.26434/chemrxiv.11871402.v4]

21. Anh-Tien T, Francesco G, Michael H, Fuqiang B, Artem C. *Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds.* 2020. [doi:10.26434/chemrxiv.11860077.v1]

22. Rimanshee A, Amit D, Vishal P, Mukesh K. *Potential inhibitors against papain-like protease of novel coronavirus (SARS-CoV-2) from FDA approved drugs.* 2020. [doi:10.26434/chemrxiv.11860011.v2]

23. Liu P, Jiang J-Z, Wan X-F, et al. Are pangolins the intermediate host of the 2019 novel coronavirus (2019-nCoV)? *bioRxiv.* 2020:2020.2002.2018.954628. [doi:10.1101/2020.02.18.954628]

24. Xiao K, Zhai J, Feng Y, et al. Isolation and Characterization of 2019-nCoV-like Coronavirus from Malayan Pangolins. *bioRxiv.* 2020:2020.2002.2017.951335. [doi:10.1101/2020.02.17.951335]

25. Lam TT-Y, Shum MH-H, Zhu H-C, et al. Identification of 2019-nCoV related coronaviruses in Malayan pangolins in southern China. *bioRxiv.* 2020:2020.2002.2013.945485. [doi:10.1101/2020.02.13.945485]

26. Wong MC, Javornik Cregeen SJ, Ajami NJ, Petrosino JF. Evidence of recombination in coronaviruses implicating pangolin origins of nCoV-2019. *bioRxiv.* 2020:2020.2002.2007.939207. [doi:10.1101/2020.02.07.939207]

27. Gu H, Chen B, Zhu H, et al. Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *J Med Internet Res.* 2014;16(1):e20-e20. [doi:10.2196/jmir.2911] [Medline: 24440770]

28. Househ M. Communicating Ebola through social media and electronic news media outlets: A cross-sectional study. *Health Informatics J.* 2016;22(3):470-478. [doi:10.1177/1460458214568037] [Medline:25656678]

29. Gui X, Wang Y, Kou Y, et al. Understanding the Patterns of Health Information Dissemination on Social Media during the Zika Outbreak. *AMIA Annu Symp Proc.* 2018;2017:820-829. [PMID: 29854148] [Medline:29854148]

30. Chen Y, Zhang Y, Xu Z, Wang X, Lu J, Hu W. Avian Influenza A (H7N9) and related Internet search query data in China. *Sci Rep.* 2019;9(1):10434-10434. [doi: 10.1038/s41598-019-46898-y] [Medline: 31320681]

31. Xie T, Yang Z, Yang S, Wu N, Li L. Correlation between reported human infection with avian influenza A H7N9 virus and cyber user awareness: what can we learn from digital epidemiology? *Int J Infect Dis.* 2014;22:1-3. [doi: 10.1016/j.ijid.2013.11.013] [Medline: 24576842]

32. Fung IC-H, Fu K-W, Chan C-H, et al. Social Media's Initial Reaction to Information and Misinformation on Ebola, August 2014: Facts and Rumors. *Public Health Rep.* 2016;131(3):461-473. [doi: 10.1177/003335491613100312] [Medline:27252566]

33. Fung IC-H, Duke CH, Finch KC, et al. Ebola virus disease and social media: A systematic review. *Am J Infect Control.* 2016;44(12):1660-1671. [doi: 10.1016/j.ajic.2016.05.011] [Medline:27425009]

34. Wong R, Harris JK. Geospatial Distribution of Local Health Department Tweets and Online Searches About Ebola During the 2014 Ebola Outbreak. *Disaster Med Public Health Prep.* 2018;12(3):287-290. [doi: 10.1017/dmp.2017.69] [Medline:28835312]

35. Feng S, Hossain L, Crawford JW, Bossomaier T. Quantifying Network Dynamics and Information Flow Across Chinese Social Media During the African Ebola Outbreak. *Disaster Med Public Health Prep.* 2018;12(1):26-37. [doi: 10.1017/dmp.2017.29] [Medline:28760166]

36. Seltzer EK, Jean NS, Kramer-Golinkoff E, Asch DA, Merchant RM. The content of social media's shared images about Ebola: a retrospective study. *Public Health.* 2015;129(9):1273-1277. [doi: 10.1016/j.puhe.2015.07.025] [Medline:26285825]

37. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. *Public Health.* 2017;150:170-175. [doi:10.1016/j.puhe.2017.07.015] [Medline:28806618]

38. Fu K-W, Liang H, Saroha N, Tse ZTH, Ip P, Fung IC-H. How people react to Zika virus outbreaks on Twitter? A computational content analysis. *Am J Infect Control.* 2016;44(12):1700-1702. [doi: 10.1016/j.ajic.2016.04.253] [Medline:27566874]

39. Fung IC-H, Fu K-W, Ying Y, et al. Chinese social media reaction to the MERS-CoV and

avian influenza A(H7N9) outbreaks. *Infect Dis Poverty.* 2013;2(1):31-31. [doi: 10.1186/2049-9957-2-31] [Medline:24359669]

40. Wang X-Y, Hao Y-H, Wu Q-H, et al. Visual analysis of official microblogging network public opinion on Dengue events[in Chinese]. *Chinese Preventive Medicine.* 2018;19(06):401-406. [doi: 10.16506/j.1009-6639.2018.06.001]

41. Sina microblog data center. 2018 Sina microblog user development report[in Chinese]URL:https://data.weibo.com/report/reportDetail?id=433 [accessed 2020-02-21]

42. Chen Y. Analysis of the Social Functions of Microblog Hot Search in the Perspective of Development Communication[in Chinese]. *Research on Transmission Competence.* 2019;3(04):76.

43. Zhang Y-H. Analysis of the Reasonability of Microblog Hot Search List in the Perspective of Communication[in Chinese]. *Journal of News Research.* 2018;9(16):29-31.

44. Hu G, Han X, Zhou H, Liu Y. Public Perception on Healthcare Services: Evidence from Social Media Platforms in China. *International journal of environmental research and public health.* 2019;16(7):1273. [doi: 10.3390/ijerph16071273] [Medline:30974729]

45. National Health Commission of the people's Republic of China. Efforts to Prevent and Control Epidemic Situation of COVID-19[in Chinese]URL:http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml [accessed 2020-02-21]

46. Wuhan Municipal Health Committee. Report of Wuhan Municipal Health Committee on the current situation of pneumonia in our city[in Chinese]URL: http://wjw.wuhan.gov.cn/front/web/showDetail/2019123108989 [accessed 2020-02-21]

47. Weibo Search Rank [in Chinese] URL: https://www.enlightent.cn/research/rank/weiboSearchRank [accessed 2020-02-21]

48. ENLIGHTENT[in Chinese]URL: https://www.enlightent.cn/ [accessed 2020-02-21]

49. Fang Q. Research on the Impact of ROST Content Mining System on Content Analysis[in Chinese]. *Course Education Research.* 2014(01):234-235.

50. VOSviewer - Visualizing scientific landscapes. URL: https://www.vosviewer.com/ [accessed 2020-02-21]

51. Liu B, Wang Z, Qi X, Zhang X, Chen H. Assessing cyber-user awareness of an emerging infectious disease: evidence from human infections with avian influenza A H7N9 in Zhejiang, China. *Int J Infect Dis.* 2015;40:34-36. [doi: 10.1016/j.ijid.2015.09.017] [Medline:26432409]

52. Lazard AJ, Wilcox GB, Tuttle HM, Glowacki EM, Pikowski J. Public reactions to e-cigarette regulations on Twitter: a text mining analysis. *Tob Control.* 2017;26(e2):e112-e116. [doi: 10.1136/tobaccocontrol-2016-053295] [Medline:28341768]

53. Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC. Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter. *J Med Internet Res.* 2016;18(12):e318-e318. [doi: 10.2196/jmir.6670] [Medline:27919863]
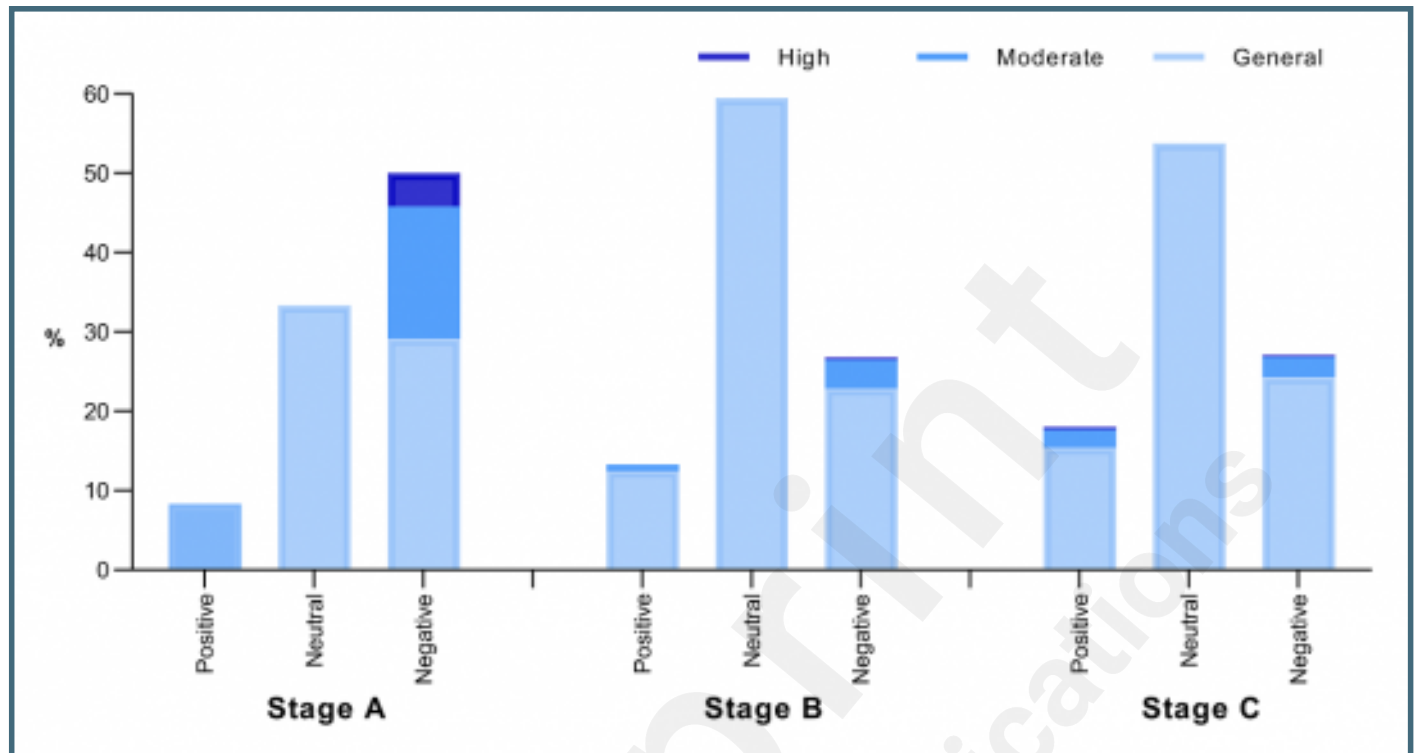
Supplementary Files

Figures

The number of topics and the cumulative search volume of the coronavirus disease epidemic on the Sina Microblog hot search list.
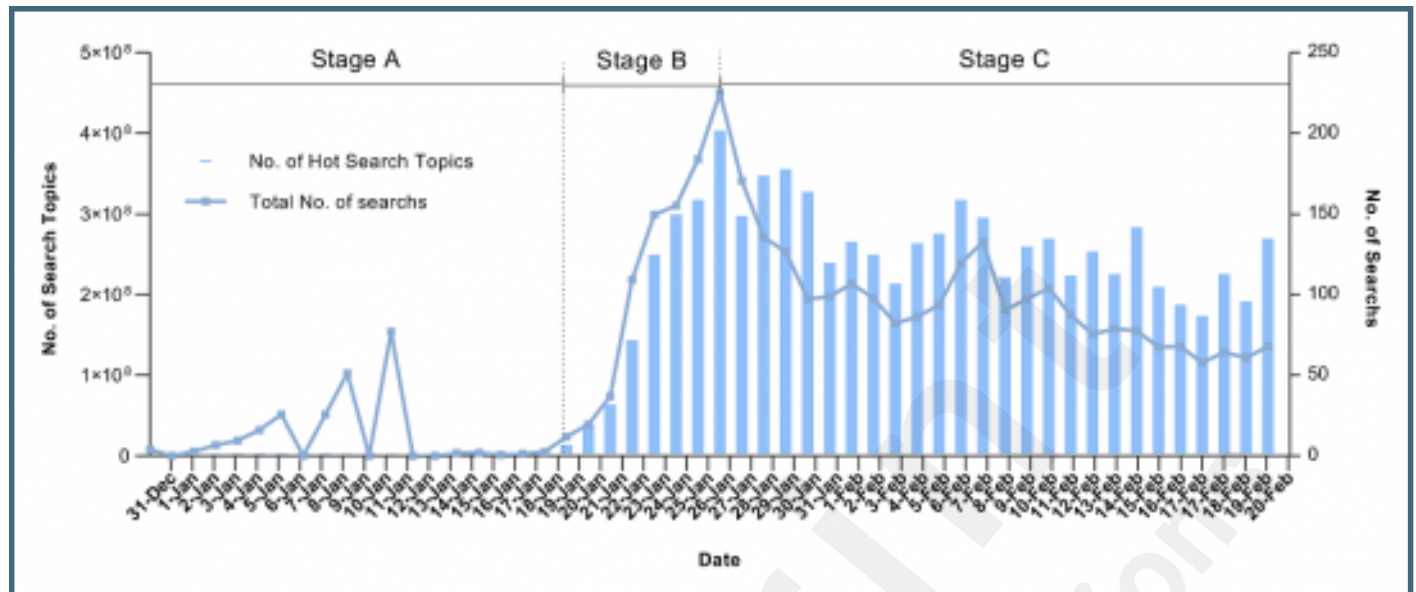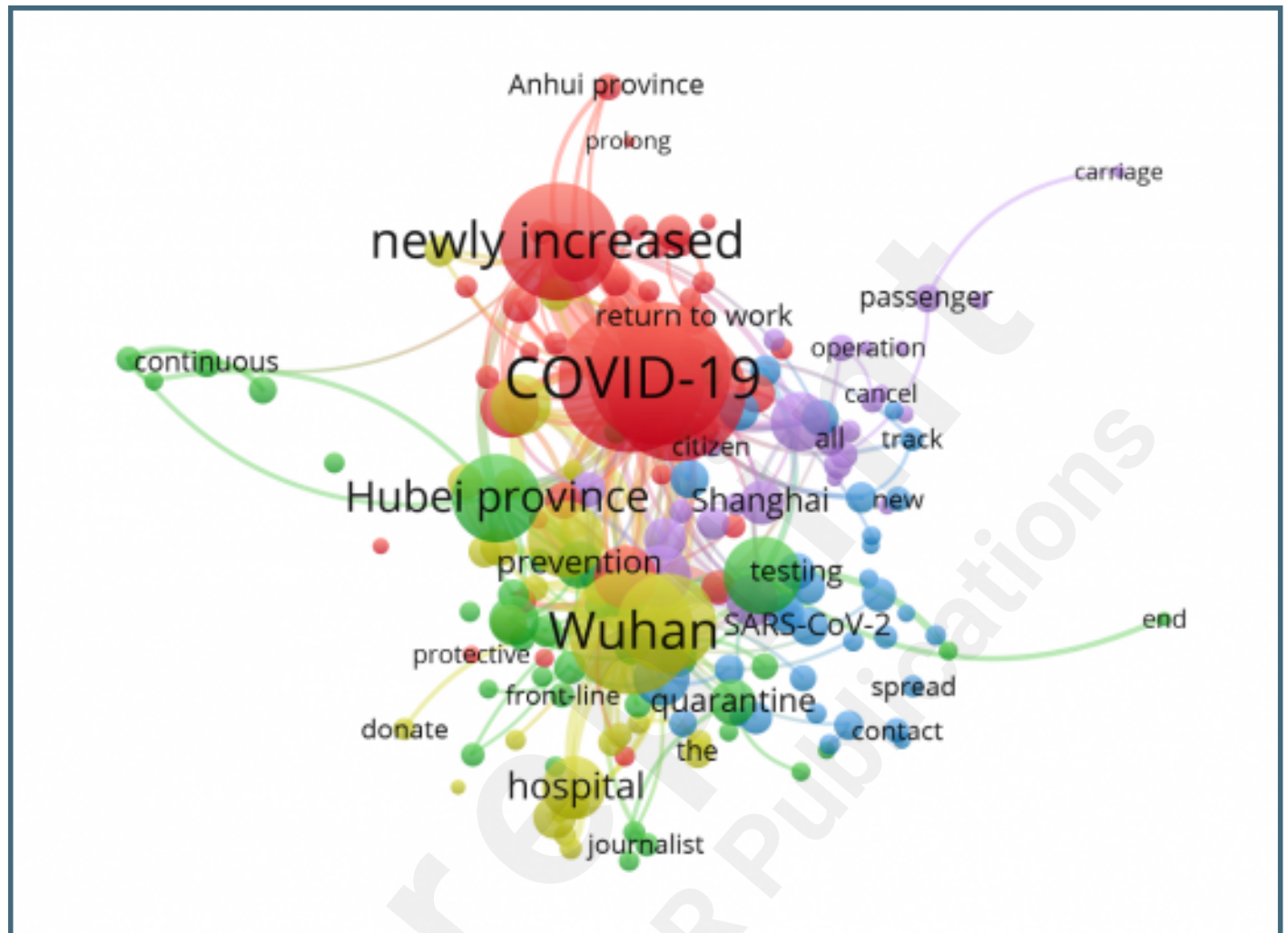
The sentiment statistics of the Sina Microblog hot search topics related to the coronavirus disease epidemic.

The sentiment trend of hot search topics related to the coronavirus disease epidemic from December 31, 2019, to February 20, 2020.
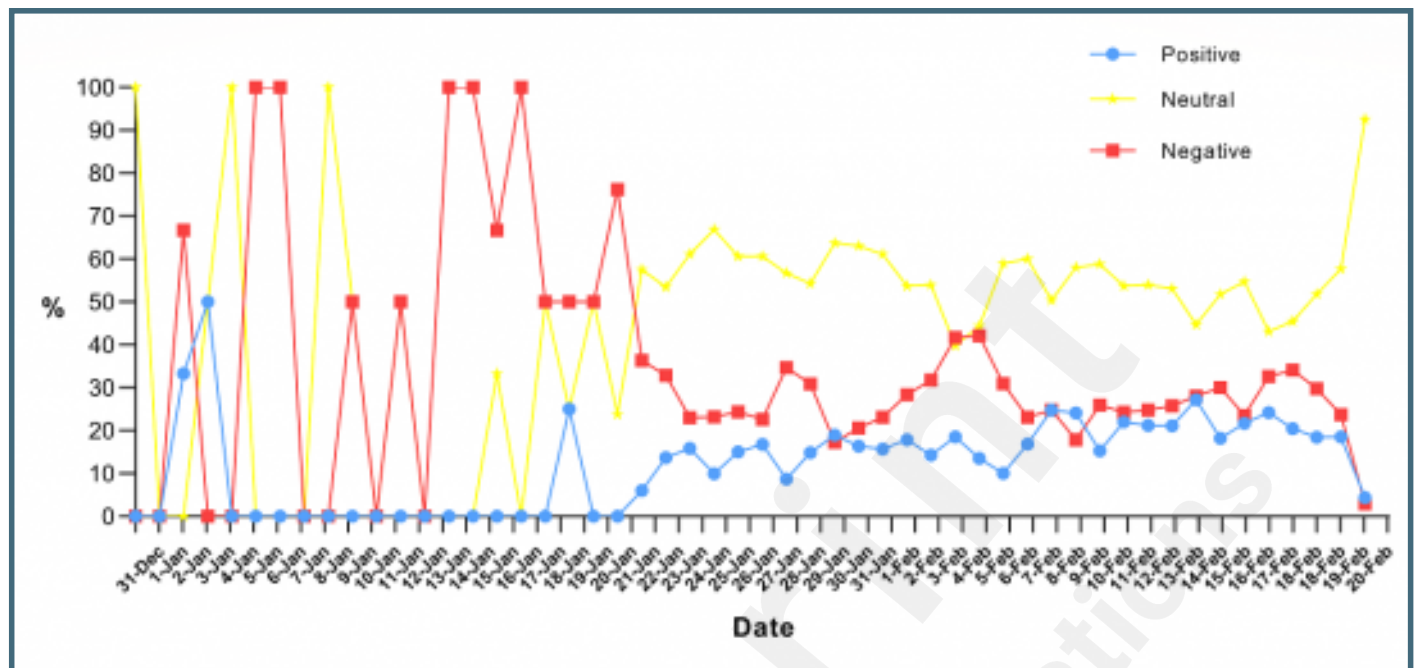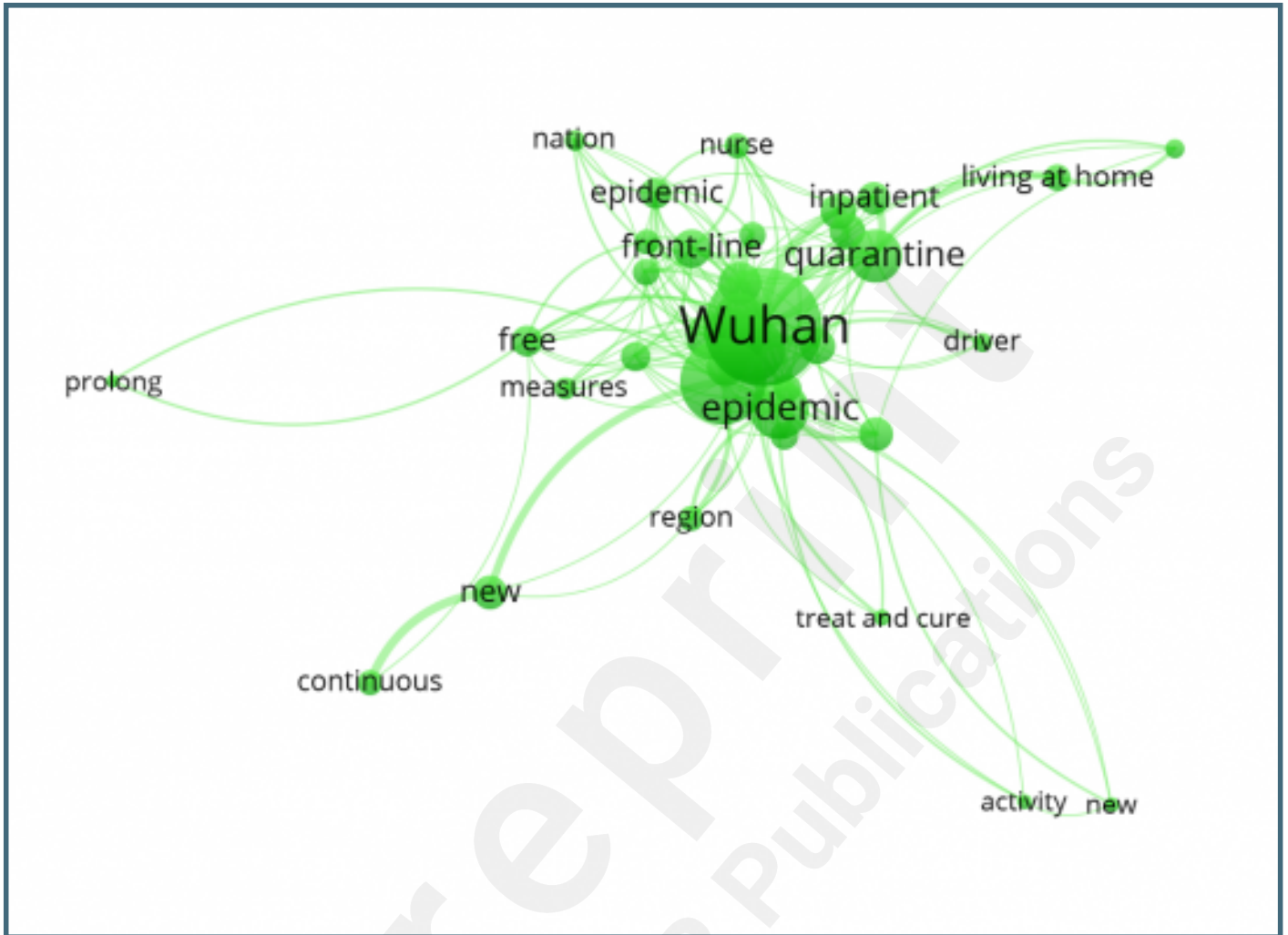
Social network of high-frequency keywords in the Sina Microblog hot search topics related to the COVID-19 epidemic.
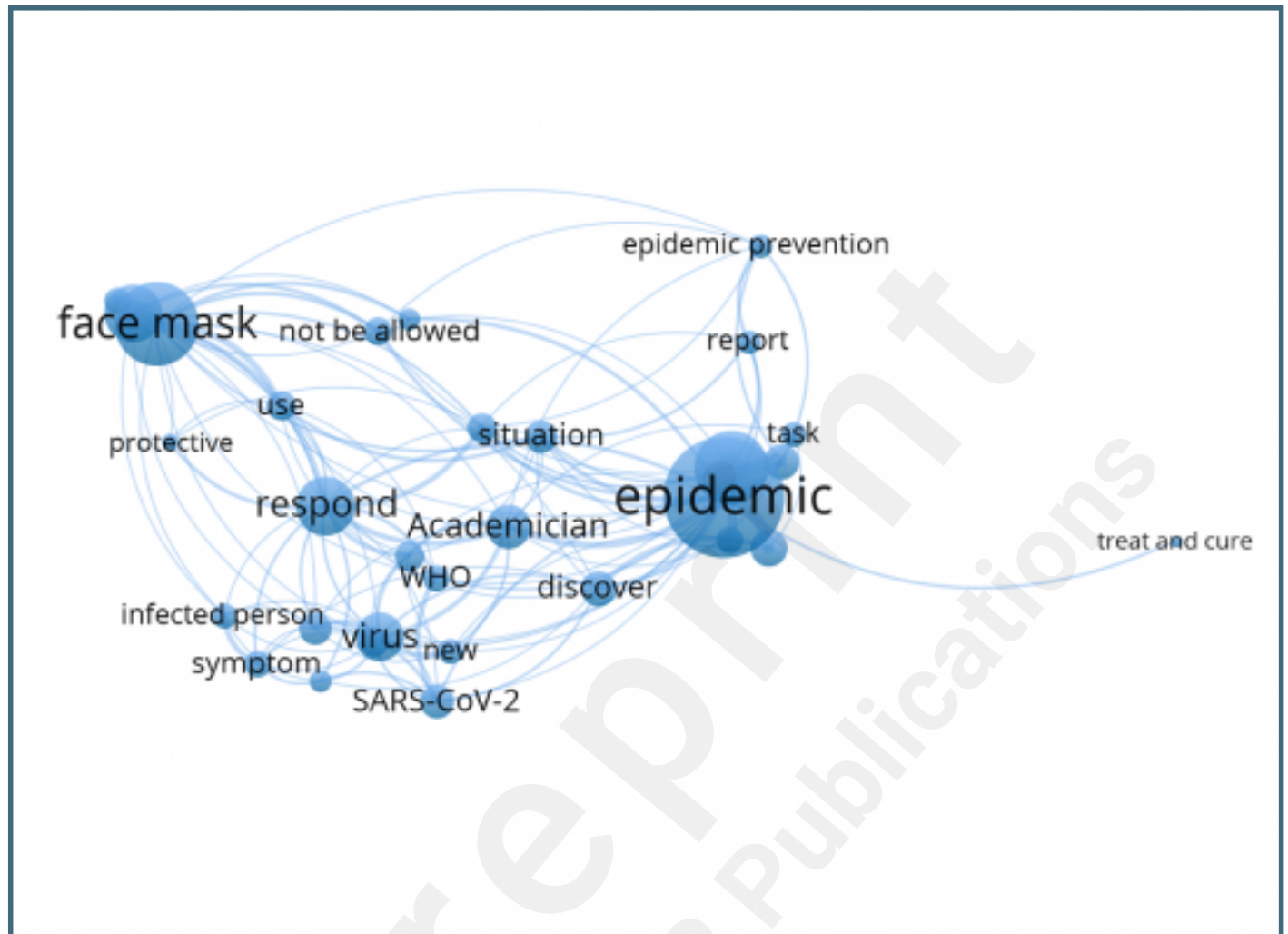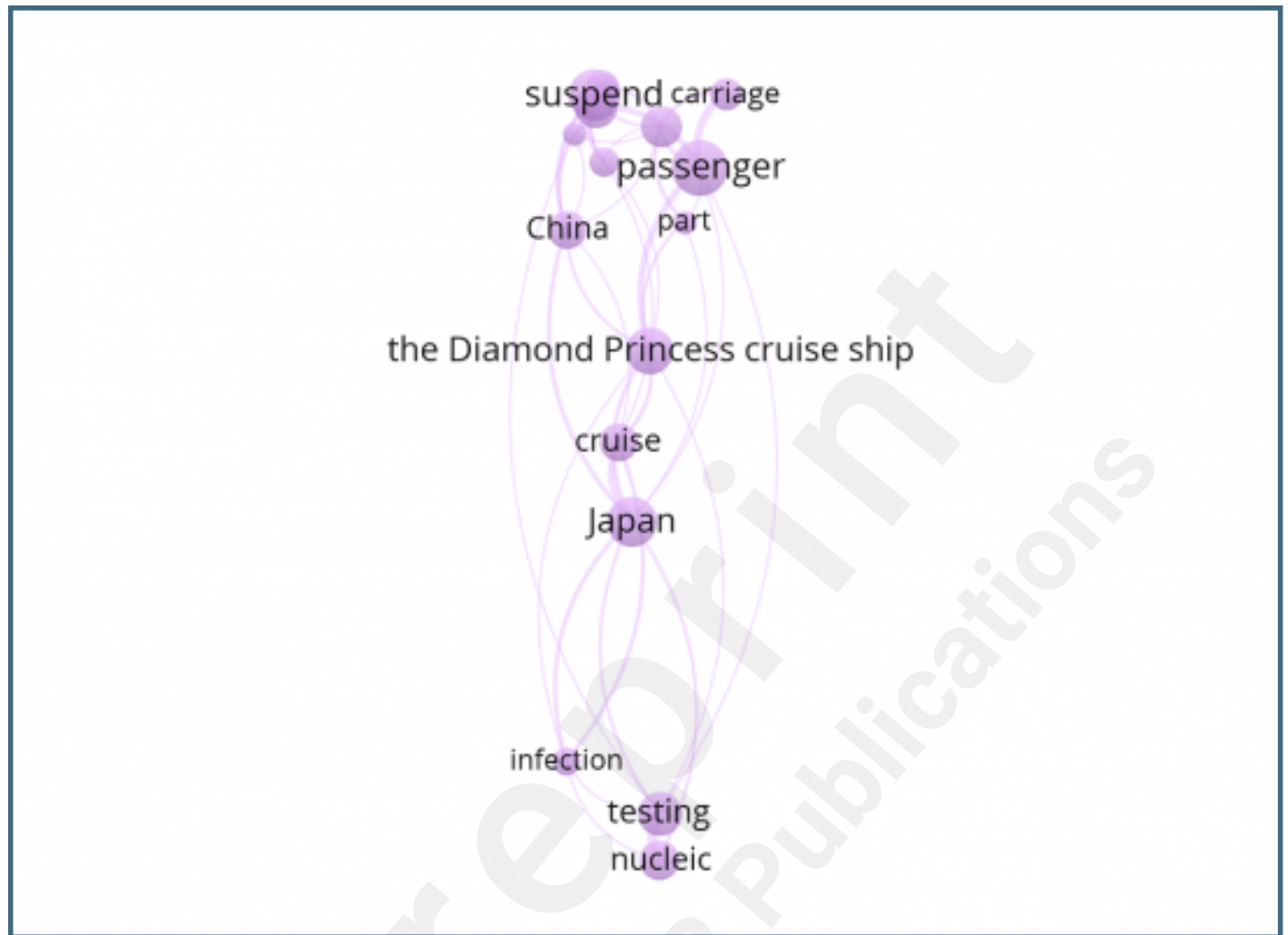
Cluster 1.
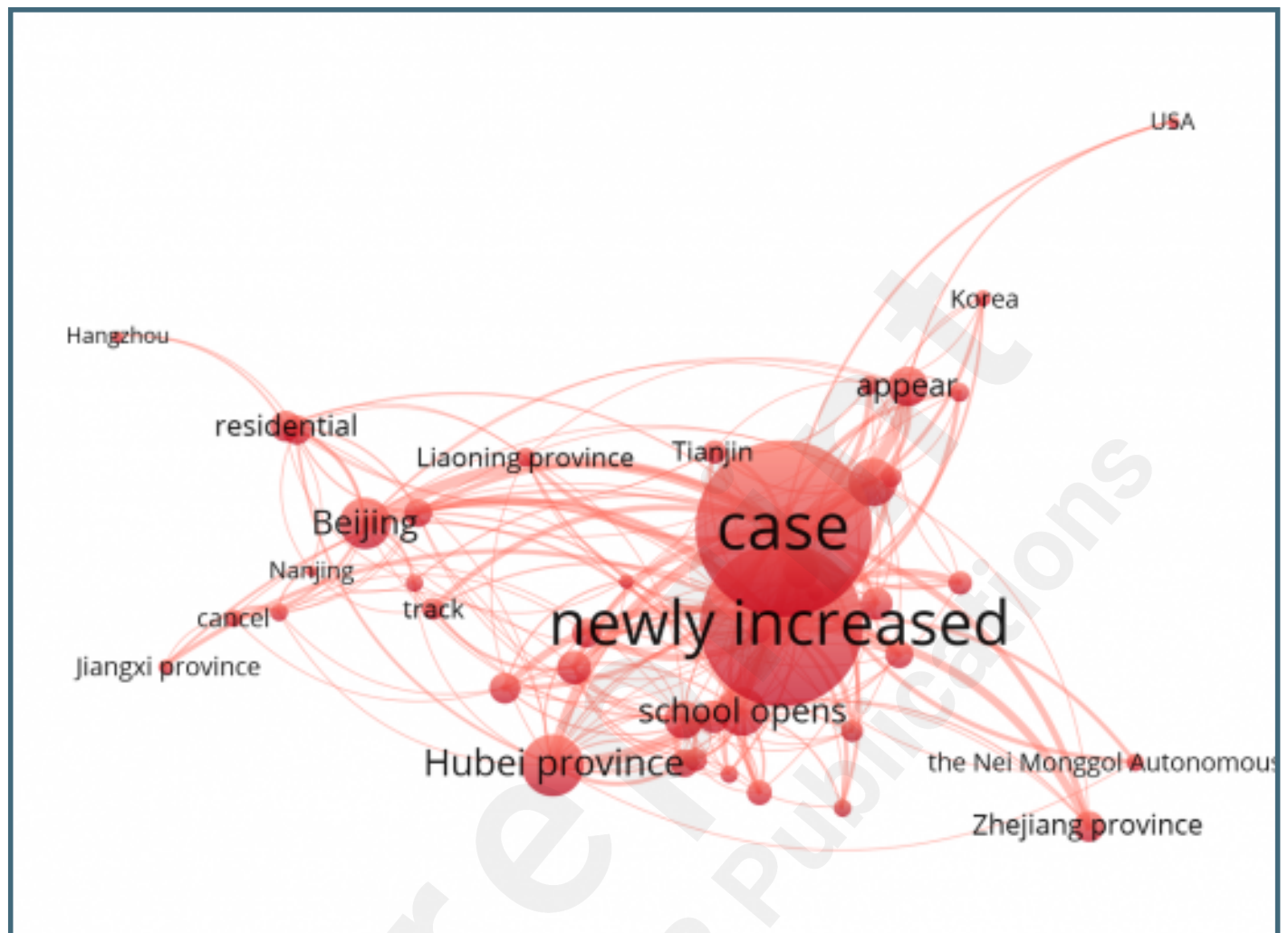
Cluster 2.

Cluster 3.

Cluster 4.

Cluster 5.

TOC/Feature image for homepages

Chinese public searches for COVID-19 on Sina Weibo.