



ARTICLE



<https://doi.org/10.1057/s41599-023-02071-5>

OPEN

The effect of official intervention on reducing the use of potentially discriminatory language during the COVID-19 pandemic in China

Yiwei Jiang^{1,2}, Hsin-Che Wu³  [✉] & Yihang Zuo⁴

This study used the Baidu Index to analyze searches using discriminatory language toward the source of the outbreak during the initial stages of the COVID-19 pandemic. The study found a strong correlation between official intervention and a reduction in searches using discriminatory language. Following the official naming of the virus by the government, the general public understood that searches using discriminatory language should be avoided, and search behavior dramatically changed. Meanwhile, this study also found that official intervention also overcame the effects of local socioeconomic factors. Regions with higher levels of socioeconomic development and higher population density had a stronger discriminatory language in internet searches, while the opposite was the case for regions with lower CPI and unemployment rates. However, after the official intervention, these factors no longer had a significant effect. It is recommended that officials worldwide provide early official media guidance during infectious disease outbreaks to reduce the use of discriminatory language and ensure the delivery of transparent and effective information to the public.

¹ Department of Public Administration at School of Economics and Management, Huzhou University, 313000 Huzhou, China. ² School of International Relations & Public Affairs, Fudan University, 200433 Shanghai, China. ³ School of Political Science and Public Administration, Shandong University, 266237 Qingdao, China. ⁴ School of Social and Political Science, The University of Edinburgh, Edinburgh EH8 9LD, UK. [✉]email: hcw@email.sdu.edu.cn

Introduction

The sudden emergence of COVID-19 at the end of 2019 caused people to rethink the importance of public health crisis management. This includes establishing emergency response organizations, improving epidemic awareness, evaluating the effectiveness of institutional arrangements, enhancing scientific decision-making and implementation, and addressing psychological crises during the pandemic (Nuzzo et al., 2020). However, one aspect is frequently overlooked: eliminating the “social stigma directed toward the source of the outbreak.”

Social stigma, prejudice, and punitive remarks can destroy all goodwill, potentially diminishing the desire to maintain transparency in epidemic regions (Smallman, 2015; WHO, 2020; Chapelan, 2021). History has recorded such hostility countless times. The pandemic that swept through Europe from 1918 to 1919 became known as the Spanish flu; the prion disease affecting cattle that began in the 1980s is widely referred to as British mad cow disease; Ebola, which impacted all of West Africa from 2014 to 2015, is named after a river in the Democratic Republic of the Congo. Other examples include Mexican influenza (H1N1 influenza) and Zika virus, named after the Zika Forest in Uganda, where it was first discovered. The stigmatization of the source of a disease outbreak often undermines disease control, incites hostility toward infected people or regions, and leads to local protectionism fueled by fear of the disease (Abdelhafiz & Alorabi, 2020; UNICEF, 2020; UNICEF Sudan, 2021; Rewerska-Juško & Rejdak, 2022; Yousefi et al., 2022). Social stigma can greatly hinder the effectiveness of efforts to fight pandemics. As such, appropriate government intervention becomes critical to alleviate or eradicate this issue.

World Health Organization (WHO) indicates that “Governments, citizens, media, key influencers, and communities have an important role to play in preventing and stopping stigma surrounding people from China and Asia in general” (UNICEF, 2020). This article focuses on official intervention to utilize the behavioral data of hundreds of millions of internet users in the Baidu Index to examine changes in the use of potentially discriminatory language, like “Wuhan pneumonia”, in internet searches following the official naming of the virus. By analyzing the usage patterns of the keywords “Wuhan pneumonia” and “novel coronavirus pneumonia,” this study aims to shed light on the impact of official intervention on the use of discriminatory language when searching the internet.

The analysis in this article has three main focuses. The first is to suggest the importance of government intervention in reducing social stigma. The second is to examine the searching behavior of potential discriminatory language over time. Through the time dimensions, we can observe the change in the searching behavior before and after official intervention. This paper also uses panel data at the provincial level to analyze socioeconomic factors that would limit the effect of official intervention to searching behavior of potential discriminatory language. And finally, coping strategies for reducing social stigma directed toward the source of the outbreak.

Social stigma and official intervention

An individual can be stigmatized due to various factors, including race, culture, gender, intelligence, and health, leading others to perceive them as different (Goffman, 1963). Social stigma in the field of infectious disease prevention and control, as defined by the World Health Organization (WHO) and related United Nations organizations, refers to the labeling, stereotyping, or isolation of individuals affected by a disease or regions experiencing an outbreak. This stigma extends not only to those directly affected by the disease but also to individuals residing in areas

with infectious disease outbreaks, resulting in discrimination and isolation (Abdelhafiz & Alorabi, 2020; WHO, 2020; UNICEF, 2020; CDC USA, 2021; UNICEF Sudan, 2021).

On 12 February 2020, the WHO officially named the disease caused by the new coronavirus as COVID-19, which stands for Coronavirus Disease 2019. WHO Director-General Tedros Adhanom Ghebreyesus (2020) emphasized that the chosen name deliberately avoided any reference to a person, place, or animal associated with the virus to prevent stigmatization. In a tweet, Tedros Adhanom Ghebreyesus stated, “We now have a name for the disease caused by the novel coronavirus: COVID-19. Having a name matters to prevent the use of other names that can be inaccurate or stigmatizing” (Tedros, 2020). This tweet marked the first instance where the significance of virus naming was conveyed to the global public, and it was also the first time the public has become aware of the issue of social stigma related to the naming of infectious diseases.

When society stigmatizes a disease, those who are infected are likely to feel reluctant to confront the illness and may avoid seeking treatment to evade negative labeling (Abdelhafiz & Alorabi, 2020; Chew et al., 2021; UNICEF Sudan, 2021). The consequences of this phenomenon were similar to the stigma faced by African residents due to AIDS (Shilts, 2011). Sontag also points out that patients with certain diseases, such as cancer and AIDS, were deliberately isolated and attacked, resulting in them feeling guilty. Due to the public’s lack of understanding of the disease, societal symbols and myths can prevent patients from effectively confronting the illness, ultimately leading to higher mortality rates (Sontag, 1978; 1997). In China, the influence of ancient Chinese philosophies and attitudes contribute to regional prejudices and social stigma toward ethnic groups in southwest China (Zhang, 2005).

In addition to diseases such as AIDS or cancer, the stigma associated with infectious diseases is widespread, posing a serious barrier to disease prevention and treatment. Once a disease name is established, it is very difficult to alter or remove it. Historically, this has happened repeatedly, particularly in contexts where empathy, respect for others, and human rights are paramount. For instance, the term “Spanish flu” fueled irrational fears and stigma (Hoppe, 2018). The term “Ebola” has led to survivors of Ebola in West Africa facing long-term humiliation and stigmatization within their communities (Overholt et al., 2018). Furthermore, the term Mexican influenza (H1N1) influenced some Americans’ perceptions of Latinos in the United States, adversely affecting their social welfare (McCauley et al., 2013).

Following the outbreak of COVID-19 in early 2020, the use of the term “Wuhan pneumonia,” including in internet searches, associated the city of Wuhan with the disease. Unknowns and the fear of these unknowns are the key drivers of social stigma and discrimination. (UNICEF, 2020; UNICEF Sudan, 2021) In addition, due to the initial lack of official and public understanding of COVID-19 as well as its symptoms, such as rapid infection and relatively high mortality, the disease quickly became a marker. As a result, individuals infected with the disease and residents of Wuhan not only had to deal with the physical pain of the disease but also with societal pressure. This stigma then extended to groups associated with those afflicted by COVID-19, typically starting with those closest to the patients and then spreading to those in the same geographical area. This stigma can lead to a reluctance to undergo COVID-19 testing, posing challenges for disease control authorities in their epidemic prevention efforts (Chew et al., 2021).

Furthermore, the highly contagious nature of the COVID-19 virus means that any stigmatization of an area or an infected person, which may lead to avoidance of testing, hampers the

efforts of health organizations in various countries to contain the virus's spread (Althoff et al., 2021).

In order to prevent social stigma, the WHO (2015) has stated that newly identified human diseases should be assigned socially acceptable names that are non-offensive to individuals, regions, or countries and avoid reference to animals. Naming that prioritizes convenience and inherent stereotypes may inadvertently promote stigma and symbolic associations, leading to public attitudes of alienation and hostility towards stigmatized groups, which could hinder the effectiveness of public health measures during an epidemic. The WHO has stated that government intervention is effective in reducing social stigma during pandemics (UNICEF, 2020). Scholars have also pointed out that official intervention has a positive effect on reducing social stigma during the COVID-19 epidemic. (Abdelhafiz & Alorabi, 2020; Huda et al., 2020). Moreover, both the British and Australian governments have made specific commitments and implemented practical policies to reduce social stigma during pandemics (UK Parliament, 2015; Watkins, 2020).

The initial COVID-19 outbreak occurred suddenly in China in early 2020. In January and February, cities were successively placed under lockdown. Moreover, the government and ordinary people were unaware of the nature of the COVID-19 virus. This situation was consistent with the WHO's definition of social stigma due to disease. (UNICEF, 2020; UNICEF Sudan, 2021) This article uses the changes in the Baidu index to explore changes in people's search behavior before and after the Chinese government intervention in order to understand the effect of official intervention.

Measurement of searching potential discriminatory language

The COVID-19 pandemic provided material to understand the use of potential discriminatory language among ordinary people in China. Internet search behavior is an important source of online big data (Liu and Xu, 2015). The Baidu Index reflects the subjective preferences of users as well as the influence of the macro-social environment when searching online. This paper utilized the Baidu Index, a public big data analytics tool derived from the Baidu search engine and Baidu news service. The index represents the query volume for specific keywords within a certain area, thus reflecting the popularity of these terms among users as well as the search preferences and tendencies of users at a specific moment. (Liu and Liao, 2021).

In the early days of the COVID-19 outbreak in January 2020, due to the limited understanding of the virus, the sudden lockdown of Wuhan and then followed by more and more cities and provinces, and the high death toll of infected individuals, fear of the disease was understandably a common reaction when people searched for information about the virus online. According to the World Health Organization (WHO) and UNICEF (2021), this fear often extended to infected individuals and even the city where the outbreak originated.

Discriminatory behavior concerning the disease's origin can manifest itself in the naming, discussion, and searching of the disease or its point of origin. While search behavior cannot wholly encapsulate all behavioral traits, it constitutes an integral part of an individual's response to the disease and then form tendencies to the disease. Especially in China, search behavior can boost the formation of trending topics, such as Weibo trending topics, which are ranked according to users' search click-through rate. This dynamic can inadvertently fuel discrimination related to the place of origin, continuously reinforcing prejudiced perceptions of the disease and its geographical source. In the early days of the outbreak, the demand for searches related to "Wuhan pneumonia" came after Wuhan was locked down and the Chinese

government officially disclosed the disease. As the WHO and UNICEF (2021) showed, this unprecedented lockdown of a major city in China exerted a significant psychological impact on the populace, thereby shaping inherent biases about the disease and the city of Wuhan when searching.

However, this article also believes that the causal relationship between those conducting online searches and discrimination cannot be directly deduced from internet search behavior alone. Consequently, the primary objective of this study is to investigate the manifestations of search behavior using potentially socially discriminatory language in the initial stages of the pandemic and the effect of official intervention, employing the Baidu Index as our primary research tool.

This paper assessed the proclivity towards discrimination in relation to the source of an outbreak as follows: Internet users who used the term "Wuhan pneumonia" in Baidu searches were believed to have discriminatory views toward the source of the outbreak. Conversely, the use of "novel coronavirus pneumonia," a neutral term, does not display the same tendency toward social stigma. In order to explore the fluctuations in social stigmatization across different time periods and analyze the influence of macro social-environmental factors on these changes, this paper selected a three-week period from January 19 to February 8 as our study period. Landmark events during this period were selected as significant moments and categorized into distinct periods, as presented in Table 1.

Data sources and research methods

Search engine data: the Baidu Index. The Baidu search engine is the most influential internet search tool in China, with an average daily volume of more than seven billion searches (Meng and Zhao, 2019; Fang, et al., 2021). This paper used the most recent Baidu Index (Baidu Index, 2019), which is similar to Google Trends, to extract relevant search data. The Baidu Index serves as a data analysis platform based on the behavioral data of hundreds of millions of internet users, making it a crucial statistical analysis tool for both the modern internet era and the big data age. Its personal computer (PC) trends capture data from June 2006 to the present, while its mobile trends capture data from January 2011 to the present.

The Baidu search index used in this paper is based on the search volume of regular Internet users on Baidu. This study analyzed and calculated the weighted sum of the search frequencies of each keyword in Baidu web searches.

The Baidu Index has compiled an information database reflecting the degree of interest, evolving trends, regional distribution, and search preferences of internet users regarding social issues across different periods (Meng and Zhao, 2019). Depending on the search source, the search index is divided into a PC search index and a mobile search index. The Baidu Index data used in this article combines the PC and mobile indices. The Baidu Index also has some ancillary features, such as statistics on audience demographics, primarily including two types of social attributes: age distribution and gender ratio. Additionally, it includes a statistical measure (media index) for the most frequently appearing media articles following a search.

Given the above method of measuring the social stigma toward the source of the outbreak as well as notable differences in the population base and number of internet users in each province, this paper calculated internet searches using potentially discriminatory language toward the outbreak's source in different regions by dividing the search volume (PC and mobile) of "Wuhan pneumonia" by the aggregate search volume of "Wuhan pneumonia" and "novel coronavirus pneumonia" on each day. Table 2 presents the descriptive statistics of the discriminatory views toward the source of the outbreak.

Table 1 Explanation of the measurement of potentially discriminatory language toward the source of the outbreak in internet searches.

Variable name	Explanation of time period selection
Potentially discriminatory language in internet searches in the initial stage of the outbreak	A lockdown was imposed on Wuhan at 10:00 a.m. on 23 January 2020, and search volume soared.
Potentially discriminatory language in internet searches before official guidance on the use of discriminatory language	No guidance for social stigma toward the source of the outbreak. On 29 January 2020, the government began to issue guidance on discriminatory language. CCTV stated, "Regional emphasis (due to infectious diseases) cannot be turned into regional social stigma." China's National Health Commission emphasized, "Our common enemy is the disease, not the people of Wuhan." Official guidance on social stigma toward the source of the outbreak began on January 29, 2020, in the second week of the COVID-19 pandemic.
Potentially discriminatory language in internet searches after official guidance on the use of discriminatory language	The search volume for "novel coronavirus pneumonia" overtook "Wuhan pneumonia" on 30 January 2020, the day after the new official guidance was issued.
Potentially discriminatory language in internet searches following the official naming of the disease	On 8 February 2020, the Joint Prevention and Control Mechanism of the State Council temporarily named pneumonia caused by the new coronavirus "novel coronavirus pneumonia" (NCP). After the release of the official name of the new coronavirus, all media were obliged to follow the new name.
Trends in potentially discriminatory language in internet searches (3-week average)	The 21-day mean of the index of social stigma toward the source of the outbreak

Table 2 Descriptive statistical analysis of potentially discriminatory language toward the source of the outbreak in internet searches.

Variable name	Mean	Standard deviation	Min	Max
Potentially discriminatory language in internet searches in the initial stage of the outbreak	0.7312	0.0615	0.5770	0.8473
Potentially discriminatory language in internet searches before official guidance on the use of discriminatory language	0.6677	0.0885	0.4427	0.8257
Potentially discriminatory language in internet searches after official guidance on the use of discriminatory language	0.2337	0.0382	0.1602	0.3353
Potentially discriminatory language in internet searches following the official naming of the disease	0.1715	0.0274	0.1316	0.2448
Trends in potentially discriminatory language in internet searches (3-week average)	0.4360	0.0448	0.3411	0.5489

The above data was calculated by dividing the search volume (PC and mobile) of "Wuhan pneumonia" by the aggregate search volume of "Wuhan pneumonia" and "novel coronavirus pneumonia" on each day. Source: Baidu Index.

Collection of data on macro-influencing factors. In order to examine the impact of macro-social environmental factors on discriminatory views toward the source of disease, this paper collected recent economic and social development variables. The data on population density, GDP, CPI, registered unemployment rate, and the ratio of fiscal expenditures were sourced from the 2018 statistics of the National Bureau of Statistics.

Land area data were drawn from the Ministry of Civil Affairs' national administrative division information query platform. Internet penetration rate data were obtained from Wangsu's 2018 China Internet Development Report and calculated by dividing the number of internet users in each province by the population of that province.

Temporal and regional analysis and media effects

This section examines the changes in social stigma toward the source of the outbreak across different time periods, focusing on regional and temporal distribution, media effects, and the media's role in promoting or curbing discriminatory views.

Regional and temporal analysis of potentially discriminatory language toward the source of the outbreak in internet searches. This paper examined the evolution of social stigma toward the source of the outbreak over a 21-day period. During this period, the index score gradually decreased from a peak of 0.933

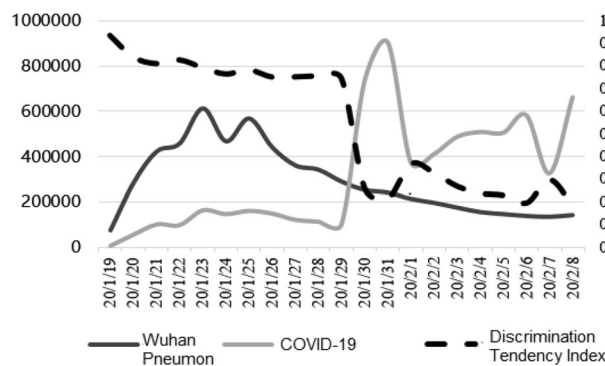


Fig. 1 Searching trend of potential discriminatory language. Searching potential discriminatory language toward the source of the outbreak and keyword search trends on the Baidu Index.

to 0.179 (as shown by the dotted line in Fig. 1). Combining the search trends of the Baidu index keywords "Wuhan pneumonia" (solid black line in Fig. 1) and "novel coronavirus pneumonia" (solid gray line in Fig. 1), we can identify several key points over time:

1. After the lockdown of Wuhan on January 23, the search volume for "Wuhan pneumonia" peaked, with a value of 613,238, while the search volume for "novel coronavirus pneumonia" at

Table 3 Regional ranking of searches using potentially discriminatory language toward the source of the outbreak in different periods.

Potentially discriminatory language in the initial stage of the outbreak		Potentially discriminatory language in internet searches before official guidance on the use of discriminatory language		Potentially discriminatory language in internet searches after official guidance on the use of discriminatory language		Potentially discriminatory language in internet searches following the official naming of the disease		Trends in potentially discriminatory language in internet searches (3-week average)			
1	Shanghai	0.8473	1	Shanghai	0.8257	1	Shanghai	0.2448	1	Shanghai	0.5489
2	Guangdong	0.8194	2	Beijing	0.7892	2	Shanghai	0.2102	2	Beijing	0.5113
3	Beijing	0.8115	3	Guangdong	0.7690	3	Tianjin	0.2086	3	Chongqing	0.4903
4	Jiangsu	0.7824	4	Jiangsu	0.7489	4	Chongqing	0.2057	4	Hunan	0.4796
5	Tianjin	0.7796	5	Chongqing	0.7352	5	Qinghai	0.2017	5	Hainan	0.4766
6	Henan	0.7725	6	Zhejiang	0.7290	6	Hunan	0.2000	6	Guangdong	0.4753
7	Sichuan	0.7692	7	Tianjin	0.7256	7	Jiangsu	0.1934	7	Tianjin	0.4736
8	Zhejiang	0.7665	8	Anhui	0.7209	8	Guangdong	0.1728	8	Hainan	0.4697
9	Anhui	0.7618	9	Henan	0.7189	9	Jilin	0.1708	9	Jiangsu	0.4641
10	Jiangxi	0.7540	10	Hunan	0.7152	10	Hainan	0.2500	10	Anhui	0.4626

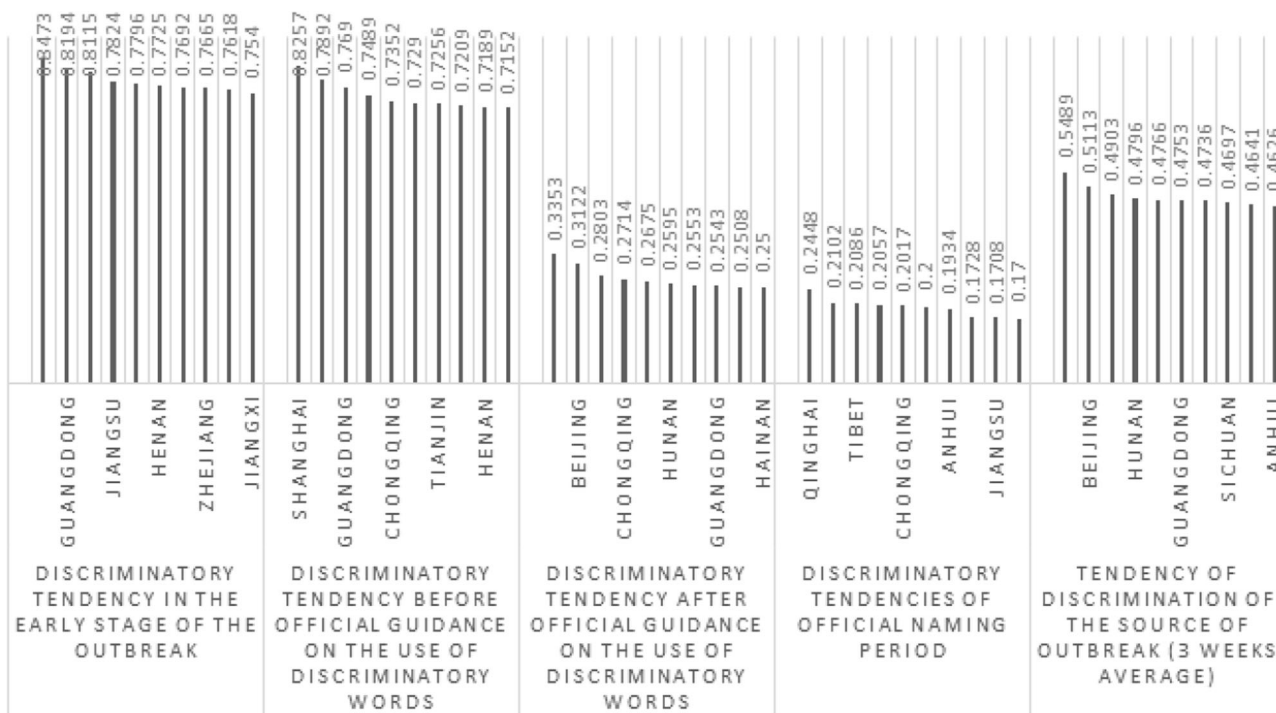


Fig. 2 Searches using potentially discriminatory language in different provinces and periods. Searches using potentially discriminatory language toward the source of the outbreak in different provinces in different periods.

the same point was 162,438, and the discrimination index was 0.7906.

2. Between January 29 and January 30, the search volume for “novel coronavirus pneumonia” overtook that of “Wuhan pneumonia.” The discrimination index also dropped from 0.7422 on January 29 to 0.2574 on January 30, indicating that the official intervention in the early days of the pandemic significantly affected people’s behavior.

3. The search volume of “novel coronavirus pneumonia” was at a high level after January 30. After the Joint Prevention and Control Mechanism of the State Council officially named the disease on February 8, the search volume for “novel coronavirus pneumonia” surpassed the search volume for “Wuhan pneumonia” by a large margin. The discrimination index also continued to decline.

Timeline analysis revealed the sudden shift from January 28 to January 29 shows that the government’s initial intervention

significantly affected behavioral choices. We found that the while the official guidance given to the media played a major role during this period, the shift in views was also affected by macro-social factors in each region.

We observed regional differences in searching potential discriminatory language toward the source of the outbreak at different time points. However, there were some common patterns across the study period (see Table 3 and Fig. 2):

1. Shanghai consistently ranked in the top five in searches using potentially discriminatory language: It was fourth in the initial stage, third before official guidance was issued, fourth after official guidance was issued, and fifth following the official naming of the disease. It was also fourth in the 3-week average.

2. Guangdong and Beijing also ranked highly in the initial stage, but unlike Shanghai, they showed a significant fall after the official guidance was issued and were no longer in the top ten when the disease was named.

Table 4 The most frequently accessed information from searches for “Wuhan pneumonia”.

Date	Most accessed news titles	Publisher	Search index
1.23	People Involved in Wuhan Pneumonia: More Than a Dozen Merchants Selling Game in the Huanan Seafood Market	China Business Journal	613,328
	Severely Ill Patients With Wuhan Pneumonia, Hospital Beds Hard to Find	Sina News	
	Regarding Wuhan Pneumonia, These Rumors Have Been Officially Refuted	Tencent news	
1.25	Overseas Cases of Wuhan Pneumonia Increase, and Many Countries Accelerate the Development of Vaccines	Caijing Magazine	568,762
	Cedar Trust Rushed to Help the Wuhan Pneumonia Epidemic, Donating 10 Million Yuan and the First Batch of Urgently Needed Materials	Banyuetan, Xinhua News	
	Good News For Wuhan Pneumonia! The First Virus Strain Has Been Successfully Isolated	Sina Weibo	
1.27	5794 Suspected Cases of Wuhan Pneumonia Nationwide	Sina Finance	361,933
	Use AI to Help the Development of a Wuhan Pneumonia Vaccine!	Sohu Technology	
	U.S. Stock Futures Fall More Than 300 Points in a Day as Wuhan Pneumonia Cases Increase	Sina Finance	
1.29	Cnlight Inc Sends More Than 3000 Disinfection Lamps on New Year’s Eve to Fight the Wuhan Pneumonia Epidemic	South CN plus	290,570
	Recovery from Wuhan Pneumonia: Lung Function Stopped For a Time	The Beijing News	
	Mayday Donate 5 Million Yuan to Assist In the Response to the Wuhan Pneumonia Epidemic	Taiwan, CN	
2.4	Questions Raised by Investors: Does the Company Assume Social Responsibility and Actively Participate in the Development of Apis for Wuhan Pneumonia?	Sina Finance	156,977
	Questions Raised by Investors: How Much Will the Spread of Wuhan Pneumonia During the Chinese New Year Impact the Company’s Factory Production?	Sina Finance	
	How Much Money Did Car Manufacturers Donate During the Wuhan Pneumonia Epidemic?	ifeng.com	

Source: top three news articles on each date according to Baidu Media Index.

3. Qinghai and Tibet showed the opposite trend to Guangdong and Beijing, as searches using potentially discriminatory language in these areas actually increased after the official guidance was issued. Therefore, the significant decline in discrimination tendencies in other provinces was not observed in these western inland provinces.

This article also indicates two aspects of regional differences:

First, it is often assumed that those in positions of dominance—whether economically, politically, or culturally—are more likely to succumb to regional social stigma. (Connerton, 1989; Halbwachs, 1992; Harvey and Bourhis, 2012; Kuppens et al., 2018) This paper’s analysis of regional differences confirms that this phenomenon does exist, albeit with variations. For example, in Guangdong and Beijing, searches using potentially discriminatory language were obvious in the initial stage of the outbreak. However, searches using potentially discriminatory language showed a sharp decline following the issuing of official guidance. Therefore, the previous assumption that those in dominant positions succumb to regional social stigma may only be partially true. Changes in search behavior may show different trends over time which are closely related to factors such as the local internet penetration rate and exposure to official news media. Although those in a dominant position are the most prone to using potentially discriminatory language when searching online, they also have the greatest potential for change.

Second, the opposite patterns occurred in Qinghai and Tibet, due to the combination of a lack of exposure to official news media and low internet penetration rates, which resulted in a sense of political alienation. Therefore, the issuing of official guidance failed to significantly reduce discrimination toward the source of the outbreak.

Media effects on searches using potentially discriminatory language toward the source of the outbreak. Internet searches conducted by the general public reflect their personal inclinations. However, the news content they encounter often serves to amplify this search behavior, potentially deepening their perceptions and reinforcing their viewpoints, thereby creating a cumulative effect. Therefore, by using the most frequently accessed information from these searches, this study analyzed the media landscape

during the 3-week period from 23 January to 8 February 2020 (spanning 3 weeks). Supplementary research materials for this study were also provided by the Baidu Media Index.

Tables 4 and 5 present the top three news stories and their search index at a number of key time points using data from the Baidu Media Index. The types of media outlets were as follows. Fifteen news articles with “Wuhan pneumonia” in the headline were provided by portal websites, including Sina, NetEase, Sohu, and Tencent, accounting for 53% of the total. Only BanYueTan and the Beijing News were sponsored by the government or under the direct control of the Central Committee of the Communist Party of China (CPC), accounting for 13.3%.

After February 8, there were no news headlines that used the term “Wuhan pneumonia” in any newspapers or publications under the direct control of the CPC Central Committee. Of the 18 news articles with “novel coronavirus pneumonia” in the main title of the news, there were 10 articles from party newspapers and publications directly under the CPC Central Committee or media sponsored by the Chinese government, including five articles from CCTV News, one article from People.cn, and one article from the central government. There was also one article on the internet, one article on China Economic Net, one article on Huanqiu.com, and one article on Beijing News. These accounted for 55.6% of the total. Among these news headlines, the term “novel coronavirus pneumonia” was used the most by the National Health Commission, where it appeared five times.

Some intuitive feelings can be generated from the headlines and statistics of these reports. China’s central media (including CPC and government media) were more cautious when dealing with discourse related to local social stigma, whether this was in the initial stage of the outbreak or later on. However, the speed of information dissemination and catchy titles were more important for web portals, producing a less cautious approach.

The above Results demonstrate how the government and official media intervened to limit the use of potentially socially discriminatory language. This action has resulted in an awareness among the public that such language might exacerbate social stigma. Following government intervention, the “correct” terminology pertaining to

Table 5 The frequently accessed information from searches for “novel coronavirus pneumonia”.

Date	Most searched news items	Publisher	Search index
1.21	High-Level Expert Group of the National Health Commission answers reporters' questions on novel coronavirus pneumonia coronavirus pneumonia Hong Kong finds 117 patients suspected of being infected with novel coronavirus pneumonia Fifteen doctors and nurses in Wuhan diagnosed with novel coronavirus pneumonia	CCTV News ifeng.com People.cn	100,167
1.23	National Health Commission announces 17 deaths from novel coronavirus pneumonia Bai yansong refutes the rumor: smoking and drinking cannot prevent novel coronavirus pneumonia WHO reconvenes on novel coronavirus pneumonia outbreak	Sina Technology CCTV News CCTV News	162,438
1.25	One Patient With Novel Coronavirus Pneumonia Was Discharged From Hospital in Beijing Today 1287 Confirmed Cases of Novel Coronavirus Pneumonia Nationwide, 41 Deaths Australia's First Confirmed Case of Novel Coronavirus Pneumonia	/ The Beijing News Huaxia.com	159,828
1.31	Henan Requires That Infection With Novel Coronavirus Pneumonia Due to the Performance of Duties Is Promptly Recognized as a Work-Related Injury National Health Commission: Patients Who Have Been Cured of Novel Coronavirus Pneumonia Are Still Likely to Be Re-Infected Experts Interpret the Fatality Rate of Novel Coronavirus Pneumonia	CCTV News Sina Technology CCTV News	902,431
2.4	National Health Commission: The Vast Majority of Patients With Novel Coronavirus Pneumonia Have Mild Symptoms National Health Commission: 80% of Deaths From Novel Coronavirus Pneumonia in the Country Are Elderly People Over 60 Years Old Forward to Others! Eight Recent Explanations about the Novel Coronavirus Pneumonia	163 (NetEase) News Huanqiu.com Gov.CN	506,917
2.6	28 Days of Her Fight Against the Novel Coronavirus Pneumonia Ctrip Corporate Travel Continues to Act on Prevention and Guidance for Novel Coronavirus Pneumonia From 6:00 p.m. on February 5 to 12:00 p.m. on the 6th, Tianjin Added 9 New Confirmed Cases of New Coronavirus Pneumonia, With a Total of 78 Confirmed Cases	China Economic Network ifeng.com ifeng.com	580,704

Source: top three news articles on each date according to Baidu Media Index.

COVID-19 has been officially confirmed, significantly influencing people’s search behavior.

Regression analysis of provincial data

This paper found that there are regional differences in official intervention in the above section. In order to understand whether the effect of official intervention on the use of discriminatory language may be influenced by regional or socioeconomic factors, a regression analysis is required.

The above analysis has defined the source of social stigma in four different periods as the dependent variable. Previous research has provided some pointers on how to effectively analyze the effect of socioeconomic influencing factors on the dependent variable.

Demographic and geographical factors, such as economic development levels and media orientation factors, should be the most important variables. Some scholars have proposed that the fundamental influencing factors of regional social stigma are differences in the natural environment, cultural and environmental factors, and individual environmental perception (Chen and Xie, 2017; Ran et al., 2021). Other scholars have also emphasized the guiding role of mass media, even beyond socioeconomic factors (Cattaneo, 2014). Therefore, this section combines these variables to analyze the source of discrimination against people in different regions, using the provincial cross-sectional data of four time periods to carry out regression analysis to investigate the effect of official intervention on the use of potentially discriminatory language in searches.

Definition of variables. This paper used cross-sectional data and panel data for analysis. The dependent variable was the use of potentially discriminating language toward the source of the outbreak in internet searches in each province across several time periods, including the initial stage of the outbreak, the period

before official guidance on the use of discriminatory language was issued, the period after official guidance on the use of discriminatory language was issued, and following the official naming of the disease. Official intervention variables are Official guidance * Average number of official social media posts and Official guidance which are put in the last model (The final column of Table 7).

The regional variables included the population and geographical factors of each province (land area, population density), the level of economic development (GDP, the proportion of fiscal expenditure), livelihood factors (CPI, registered unemployment rate), the education level (years of education per capita), and media effect factors (internet penetration rate, average number of official social media posts). Table 6 presents the descriptive statistical analysis of variables.

Regression findings of provincial data. In order to conduct a more effective analysis, this paper first conducted regression analysis on the cross-sectional data on searches using potentially discriminatory language toward the source of the outbreak across four different times.

The advantage of cross-sectional data is that it is based on a one-dimensional data set composed of different regions (objects) at a single point in time. This approach highlights the spatial differences of the social phenomenon at a specific point in time and shows the uniqueness of each sample. This paper analyzed the reasons for different patterns in social stigma across 30 provinces in the time section.

The analysis results presented in Table 7 indicate that the model proposed in this paper demonstrates sufficient explanatory power before and after the initial stage of the outbreak, with an adjusted R square reaching 90%. During this period, many factors, including population, geography, economic development, livelihood, cultural literacy, and media effects, had a significant impact on social stigma.

Table 6 Descriptive statistical analysis of variables.

Variable name	Variable description	Mean	Standard deviation	Min	Max
Land area	Land area of each province (10,000 sq km)	31.23132	38.83751	0.6339	164.0016
Population density	Population density by province (10,000 people/sq km)	0.0465676	0.0716272	0.0002881	0.3823947
GDP	Regional GDP (100 million yuan)	29178.03	24242.48	1477.63	97277.77
Proportion of fiscal expenditure	Proportion of local fiscal expenditure to GDP	0.2960293	0.2204662	0.1258956	1.333676
CPI	Consumer Price Index (previous year=100)	102.08	0.2998856	101.5	102.5
Registered unemployment rate	Urban registered unemployment rate (%)	3.123333	0.5769616	1.4	4
Years of education per capita	Average years of education of the population	9.206466	1.149103	5.721324	12.67539
Internet penetration rate	Proportion of internet users to population in each province	0.5573333	0.082208	0.43	0.75
Average number of official social media posts	Average number of official Weibo posts in each province	56.98444	65.42586	3.780488	341.2941

The internet penetration rate was calculated by dividing the number of internet users in each province by the population in each province. The average post value of social media posts was calculated by dividing the number of official Weibo posts in each province by the number of administrative levels above the county level in each province.
 Data sources: data on population density, GDP, CPI, registered unemployment rate, and proportion of fiscal expenditure are from 2018 data from the National Bureau of Statistics. The data on land area is from the Ministry of Civil Affairs National Administrative Division Information Inquiry Platform. The data on internet penetration is from Wangsu's 2018 China Internet Development Report. The data on official information channels are from the 43rd Statistical Report on Internet Development in China.

In the initial stage of the outbreak, discriminatory search behavior was positively affected by population density, level of economic development, and per capita years of education. This implies that these factors significantly increased discriminatory search behavior across each province. Factors related to livelihoods, such as the consumer price index (CPI) and registered unemployment rate, along with those related to media effects, such as the internet penetration rate and the average number of official social media posts, contributed to reducing searches using potentially discriminatory language in each province.

The effect of GDP on searches using potentially discriminatory language in each province was linked to the local internet penetration rate, indicating that internet penetration strengthens the effect of low economic development on social stigma. At the same time, the *P*-value of the regression equation for this time section, as determined by the *F*-test was 0.000, indicating the model's overall statistical significance. The heteroscedasticity test yielded a *P*-value of 0.6415, exceeding 0.05, indicating that the model is not subject to heteroscedasticity problems. Hence, this model is particularly suitable for analyzing factors influencing the public's social stigma toward the source of the outbreak before official guidance was issued. However, the model's explanatory power was poor in the periods after the official guidance was issued and after the disease was named.

This paper initially posits that due to the official guidance, searches using potentially discriminatory language declined rapidly, and the data structure changed, rendering the original model inapplicable. By the time the official name was established, the *P*-value obtained by the *F*-test of the model was 0.3282, no longer suggesting a linear relationship.

In order to better explain the failure of the model in the second half of the timeline, this paper separated the data from the four aforementioned periods, calculated the decline of searches using potentially discriminatory language toward the source of the outbreak between adjacent periods, and reintegrated them with the independent variables from the original model to form panel data. The analysis of this data focused on whether official directives led to a rapid decline in potentially discriminatory search behavior toward the source of the outbreak and what factors worked together with the official intervention.

In this paper, the Hausman test, LM statistic test, and *F*-test were performed on the fixed-effects model, random-effects model, and mixed regression analysis model in the panel data, respectively. Based on the previous discussion, this paper concluded that the key factor was the official intervention in guiding public opinion. Therefore, January 29, the day when the official intervention began, was set as the boundary, and a dummy variable measuring the effect of the official intervention was set up in order to construct a more suitable model.

The final column of Table 7 shows that the official intervention had a positive and significant impact on reducing social stigma toward the source of the outbreak. When the government began issuing guidance, the number of official social media posts in each province played an important positive role in the decline of potentially discriminatory search behavior.

Socioeconomic factors, such as income or education, had significant and positive effects on searches using potentially discriminatory language during the initial stage of the outbreak and before the release of official guidance on the use of discriminatory language. However, after the release of official guidance on discriminatory language, people with higher incomes and higher education were no longer more likely to use potentially discriminatory language when searching online. In China, official propaganda has a considerable influence on people's cognition (Wu et al., 2021; Piao & Wu, 2023). Therefore,

Table 7 Regression results of factors influencing searches using potentially discriminatory language toward the source of the outbreak.

	Searches using potentially discriminatory language toward the source of the outbreak			Reduction in searches using potentially discriminatory language toward the source of the outbreak before and after the issuing of official guidance on the use of discriminatory language
	Initial stage of the outbreak	Period before official guidance on the use of discriminatory language	Period after official guidance on the use of discriminatory language	
Population geography				
Land area	-0.000783*** (-5.65)	-0.000862*** (-4.07)	-0.000433** (-2.32)	-0.000106 (-0.67)
Population density	0.177* (2.01)	0.218 (1.62)	0.165 (1.39)	-0.0956 (-1.44)
Economic development				
lnGDP	0.0704*** (7.05)	0.0724*** (4.74)	-0.000891 (-0.07)	0.0328* (1.86)
Proportion of fiscal expenditure	0.203*** (3.96)	0.147* (1.88)	0.0729 (1.06)	0.0902 (1.40)
Livelihood				
CPI	-0.0515*** (-3.42)	-0.0282 (-1.23)	0.00751 (0.37)	-0.00973 (-0.50)
Registered unemployment rate	-0.0275*** (-3.21)	-0.0226* (-1.73)	0.00435 (0.38)	-0.000258 (-0.03)
Education level				
Years of education per capita	0.0429*** (4.99)	0.0477*** (3.63)	0.00985 (0.85)	0.0268** (2.42)
Media effect				
Internet penetration rate	-0.0165** (-2.40)	-0.00777 (-0.74)	0.00359 (0.39)	-0.00723 (-0.81)
*lnGDP	-0.000387*** (-2.98)	-0.000398* (-2.01)	0.0000802 (0.46)	-0.000665*** (-3.28)
Average number of official social media posts	-	-	-	0.000703** (2.60)
Official guidance	-	-	-	0.323*** (10.25)
Constant	5.050*** (3.27)	2.504 (1.06)	-0.668 (-0.32)	0.572 (0.30)
N	30	30	30	90
R ²	0.9315	0.9229	0.6782	0.5085

t statistics in parentheses.
*p < 0.1, **p < 0.05, ***p < 0.01.

after the official guidance was released, these people may soon have become aware of which language to avoid from official propaganda and adjusted their search behavior accordingly.

Overall, the panel model explained 50% of the reduction in social stigma toward the source of the outbreak during this short period, and the *P*-value of the *F*-test was 0.0000, indicating a strong model.

Discussion

The analysis in this paper highlights several key points that could provide an important reference for understanding the emergence and evolution of discriminatory tendencies toward the source of an outbreak.

As for the social stigma over time, official intervention influences discriminatory tendencies toward the source of an outbreak, especially in today's internet age. Whether in the early stages of the outbreak or later time, the government can play a role in increasing public awareness of potentially discriminatory language and reducing the number of searches using such language. Meanwhile, if a region has high internet penetration and a large number of official social media posts, it can play a role in suppressing discriminatory tendencies toward the source of an outbreak in its early stages. During the period of state regulation of discriminatory public opinion, these two factors played a complementary role in reducing discriminatory tendencies toward the source of the outbreak. However, this effect could not be useful in larger geographical area, whereas areas with high population density lean the other way.

The official intervention also diminish effects of socioeconomic factors. Considering the region factors, economic development significantly increases people's discriminatory tendencies. It is worth noting that the data also revealed that areas with high unemployment rates or CPI had relatively low levels of discrimination. These findings in the initial stage of the outbreak may be related to the experimental findings of scholars in social psychology: highly educated individuals and wealthier people are often more discriminatory in their attitudes (Harvey and Bourhis, 2012; Kuppens, et al., 2018). As such, the relatively disadvantaged people in society, who "can understand what it is like to be stigmatized and form a circle of interconnection" (Goffman, 1963), are more aware of the harm of discrimination. Having often been victims of discrimination themselves, their inclination to discriminate could be lower. However, above effects were reversed after the official intervention, which demonstrate the importance of official attitudes and actions in preventing disease stigma.

Conclusion

In the face of the COVID-19 pandemic, effectively utilizing big data for public opinion monitoring and behavior prediction to provide data analysis and policy recommendations for crisis response has become an important task in the academic community. This paper first examined the literature and evolution of the concept of discriminatory tendencies toward the source of the outbreak, and the importance of government intervention in increasing the public's awareness of social stigma.

The study's findings highlight the effectiveness of the official intervention and the official naming of the virus as "COVID-19." The sudden change observed between January 28 and January 29 shows that the official intervention significantly affected people's behavioral choices. Shanghai exhibited the highest discriminatory tendencies of the provincial-level units in the study. Guangdong and Beijing also ranked highly in the initial period, but their ranking fell sharply after official guidance was issued, and they dropped out of the top ten after the disease was officially named. In terms of official media effects, official party media was

more cautious when dealing with discourse related to local social stigma. On the other hand, portal websites prioritized information dissemination speed and headline attractiveness.

Although official public opinion guidance helped to reduce discriminatory language in internet searches, the panel data analysis also revealed some limitations in its effectiveness. Factors such as the CPI and registered unemployment rate, internet penetration rate, and the number of official social media posts helped reduce potential discriminatory searching behavior. The effect of GDP on the discriminatory tendencies correlated with each province's internet penetration rate. Low internet penetration rates provinces, such as Qinghai, and Tibet, which could result in low exposure to official news media. After officially issuing of public opinion guidance, their potential discriminatory searching behavior slightly heightened rather than subsided.

Based on the regression analysis and big data description, this paper proposes the following suggestions to address social stigma toward the source of an outbreak:

First, official intervention is effective in reducing discriminatory language in internet searches. Our analysis showed that the number of official social media posts in various regions significantly and positively reduces potentially discriminatory language searching. As such, when governments attempt to diminish discriminatory tendencies, official guidance plays a crucial role and is a necessary task for officials to enhance epidemic prevention effectiveness.

Second, news media should adhere to the World Health Organization's rules when naming diseases, especially epidemic diseases, to avoid linking the disease to a specific geographical location, population, or animal. Failing to do so can lead to negative outcomes, including discrimination against those in the outbreak location and potential concealment of infections, hindering epidemic prevention efforts.

Third, countries should ensure open and transparent communication of disease control information. By promoting knowledge of infectious diseases and ensuring transparency in epidemic prevention information, the general public can eliminate their fears and understand diseases more objectively.

Data availability

Materials were collected from the Baidu Index (similar to Google Trends), which is publicly available at <https://index.baidu.com/v2/index.html/>.

Received: 6 January 2023; Accepted: 5 September 2023;

Published online: 14 September 2023

References

- Abdelhafiz AS, Alorabi M (2020) Social stigma: the hidden threat of COVID-19. *Front Public Health* 8:429
- Althoff K, McGinty B & Stuart E (2021) COVID-19 and Stigma: Why shame and blame won't help fight the pandemic, and what we should be focusing on instead. Johns Hopkins Bloomberg School of Public Health. <https://publichealth.jhu.edu/2021/covid-19-and-stigma>
- Baidu Index (2019) Baidu Index. <https://index.baidu.com/v2/index.html/>
- Cattaneo C (2014) Which factors explain the rising ethnic heterogeneity in Italy?: An empirical analysis at province level. *Fondazione Eni Enrico Mattei (FEEM)* 56(6):34–50
- CDC, USA (2021) Reducing Stigma. <https://www.cdc.gov/mentalhealth/stress-coping/reduce-stigma/index.html>
- Chapelan A (2021) Swallowing the red pill: the coronavirus pandemic and the political imaginary of stigmatized knowledge in the discourse of the far-right. *J Transatl Stud* 19:282–312. <https://doi.org/10.1057/s42738-021-00073-2>
- Chen D, Xie H (2017) The geographical thinking on regional discrimination. *Chinese J. Popul. Resour. Environ.* 27(201):295–298. (Chinese)

- Chew CC, Lim XJ, Chang CT et al. (2021) Experiences of social stigma among patients tested positive for COVID-19 and their family members: a qualitative study. *BMC Public Health* 21:1623. <https://doi.org/10.1186/s12889-021-11679-8>
- Connerton P. (1989) *How societies remember*. Cambridge University Press
- Fang J, Zhang X, Tong Y, Xia Y, Liu H, Wu K (2021) Baidu index and COVID-19 epidemic forecast: evidence from China. *Front Public Health* 9:685141
- Goffman, E. (1963) *Englewood Cliffs*. Prentice-Hall, New Jersey
- Halbwachs, M (1992). *On collective memory*. University of Chicago Press
- Harvey S-P, Bourhis RY (2012) Discrimination in wealth and power intergroup structures. *Group Process Intergr Relat* 15(1):21–38. <https://doi.org/10.1177/1368430211408364>
- Hoppe T (2018) “Spanish flu”: when infectious disease names blur origins and stigmatize those infected. *Am J Public Health* 108(11):1462–1464
- Huda MN, Islam R, Qureshi MO, Pillai S, Hossain SZ (2020). Rumours and social stigma as barriers to the prevention of coronavirus disease (COVID-19): what solutions to consider? *Global Biosecurity*. <https://researchdirect.westernsydney.edu.au/islandora/object/uws:62765/datastream/PDF/view>
- Kuppens T, Spears R, Manstead AS, Spruyt B, Easterbrook MJ (2018) Educationism and the irony of meritocracy: negative attitudes of higher educated people towards the less educated. *J. Exp. Soc. Psychol* 76:429–447
- Liu T, Xu X (2015) Can internet search behavior help to forecast the macro economy. *Econ Res J* 12:68–83. (Chinese)
- Liu Y, Liao W (2021) Spatial characteristics of the tourism flows in China: a study based on the Baidu index. *ISPRS Int. J. Geoinf* 10(6):378. <https://doi.org/10.3390/ijgi10060378>
- McCaughey M, Minsky S, Viswanath K (2013) The H1N1 pandemic: media frames, stigmatization and coping. *BMC Public Health* 13:1116. <https://doi.org/10.1186/1471-2458-13-1116>
- Meng T, Zhao J (2019) Internet search behavior and public attention in the era of big data: A dynamic analysis based on Baidu Index from 2011 to 2017. *Xuehai* 3:41–48. (Chinese)
- Nuzzo J, Bouey J, Klain R (2020) The Wuhan Coronavirus: Assessing the Outbreak, the Response, and Regional Implications. Subcommittee Hearing. <https://foreignaffairs.house.gov/hearing/the-wuhan-coronavirus-assessing-the-outbreak-the-response-and-regional-implications/>
- Overholt L, Wohl DA, Fischer 2nd WA, Westreich D, Tozay S, Reeves E, Pewu K, Adjasso D, Hoover D, Merenbloom C, Johnson H, Williams G, Conneh T, Diggs J, Buller A, McMillian D, Hawks D, Dube K, Brown J (2018) Stigma and Ebola survivorship in Liberia: results from a longitudinal cohort study. *PLoS ONE* 13(11):e0206595. <https://doi.org/10.1371/journal.pone.0206595>
- Piao L, Wu H-C (2023) The Effect of the Chinese Government’s Political Propaganda and Individual Characteristics of Anti-U.S. Sentiment. *Asian Surv* 63(3):381–406
- Ran MS, Hall BJ, Su TT et al. (2021) Stigma of mental illness and cultural factors in Pacific Rim region: a systematic review. *BMC Psychiatry* 21:8. <https://doi.org/10.1186/s12888-020-02991-5>
- Rewerska-Juško M, Rejdak K (2022) Social stigma of patients suffering from COVID-19: challenges for health care system. *Healthcare* 10(2):292. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8872526/>
- Shilts, R (2011) *And the band played on: Politics, people, and the AIDS epidemic*. Souvenir Press
- Smallman, Shawn C (2015) Whom do you trust? Doubt and conspiracy theories in the 2009 influenza pandemic. *J. Int. Stud* 6(2). <https://digitalcommons.lindenwood.edu/jigs/vol6/iss2/1>
- Sontag S (1978) *Illness as Metaphor*. Farrar, Straus & Giroux, New York
- Sontag S (1997) AIDS and its metaphors. *The disability studies reader*, 232–240
- Tedros Adhanom Ghebreyesus (2020) Tedros’s twitter. <https://twitter.com/drtedros/status/1227297754499764230?lang=zh-Hant>
- UK Parliament (2015) R Mental health stigma. Key issue, UK Parliament. <https://www.parliament.uk/business/publications/research/key-issues-parliament-2015/social-change/mental-health-stigma/>
- UNICEF (2020) Social stigma associated with the coronavirus disease (COVID-19). <https://www.unicef.org/documents/social-stigma-associated-coronavirus-disease-covid-19>
- UNICEF Sudan (2021) COVID-19 & stigma: How to prevent and address social stigma in your community. <https://www.unicef.org/sudan/covid-19-stigma-how-prevent-and-address-social-stigma-your-community>
- Watkins, H (2020) Reducing stigma and discrimination needs urgent government action. Newsroom, The University of Melbourne. <https://www.unimelb.edu.au/newsroom/news/2020/october/reducing-stigma-and-discrimination-needs-urgent-government-action>
- World Health Organization (2015) *World Health Organization Best Practices for the Naming of New Human Infectious Diseases*. World Health Organization https://apps.who.int/iris/bitstream/handle/10665/163636/WHO_HSE_FOS_15.1_eng.pdf
- World Health Organization (2020) *Social Stigma associated with COVID-19*. <https://www.who.int/docs/default-source/coronaviruse/covid19-stigma-guide.pdf>
- Wu H-C, Weatherall M, Huang K-P (2021) Propagating “democracy” in China: a two-way communication explanation. *J. Contemp. China* 30(130):596–612. <https://doi.org/10.1080/10670564.2020.1852736>
- Yousefi, Z, Abbasian, M, Sullman, MJ, & Azizi-Zeinalhajlou, A (2022) COVID-19 stigmatization: consequences and solutions. *archives of clinical infectious diseases*, 17(6). <https://brieflands.com/articles/archcid-120644.html>
- Zhang W (2005) Regional prejudice and ethnic discrimination: a cultural interpretation of miasma and miasma in ancient China. *Ethno-National Studies* 3:68–77. (Chinese)

Acknowledgements

The research is supported by 2023 Zhejiang Province Philosophy and Social Science Planning Special Project (No: 039); The National Social Science Fund of China (No: 22BSH002); The Humanities and Social Sciences Research Youth Fund Project of the Ministry of Education of China (No: 21YJC840006).

Author contributions

Y.J. and H.W. were responsible for writing the full text. Y.Z. analyzed, modeled, and visualized the data. W.H. was responsible for revising the article. Y.J. and H.W. were responsible for final approval of the version to be published. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Hsin-Che Wu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023