

Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China

Ze-Liang Chen^{1,2}, Qi Zhang³, Yi Lu⁴, Zhong-Min Guo⁵, Xi Zhang³, Wen-Jun Zhang⁶, Cheng Guo⁷, Cong-Hui Liao¹, Qian-Lin Li¹, Xiao-Hu Han², Jia-Hai Lu¹

¹One Health Center, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong 510080, China;

²Key Laboratory of Livestock Infectious Diseases in Northeast China, Ministry of Education, Shenyang Agricultural University, Shenyang, Liaoning 110866, China;

³College of Animal Science and Veterinary Medicine, Shenyang Agricultural University, Shenyang, Liaoning 110866, China;

⁴Department of Health Law, Policy and Management, School of Public Health, Boston University, Boston, MA 02215, USA;

⁵Animal Experiment Center, Sun Yat-sen University, Guangzhou, Guangdong 510080, China;

⁶Department of Biological Science and Technology, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510080, China;

⁷Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY 10032, USA.

Abstract

Background: The ongoing new coronavirus pneumonia (Corona Virus Disease 2019, COVID-19) outbreak is spreading in China, but it has not yet reached its peak. Five million people emigrated from Wuhan before lockdown, potentially representing a source of virus infection. Determining case distribution and its correlation with population emigration from Wuhan in the early stage of the epidemic is of great importance for early warning and for the prevention of future outbreaks.

Methods: The official case report on the COVID-19 epidemic was collected as of January 30, 2020. Time and location information on COVID-19 cases was extracted and analyzed using ArcGIS and WinBUGS software. Data on population migration from Wuhan city and Hubei province were extracted from Baidu Qianxi, and their correlation with the number of cases was analyzed.

Results: The COVID-19 confirmed and death cases in Hubei province accounted for 59.91% (5806/9692) and 95.77% (204/213) of the total cases in China, respectively. Hot spot provinces included Sichuan and Yunnan, which are adjacent to Hubei. The time risk of Hubei province on the following day was 1.960 times that on the previous day. The number of cases in some cities was relatively low, but the time risk appeared to be continuously rising. The correlation coefficient between the provincial number of cases and emigration from Wuhan was up to 0.943. The lockdown of 17 cities in Hubei province and the implementation of nationwide control measures efficiently prevented an exponential growth in the number of cases.

Conclusions: The population that emigrated from Wuhan was the main infection source in other cities and provinces. Some cities with a low number of cases showed a rapid increase in case load. Owing to the upcoming Spring Festival return wave, understanding the risk trends in different regions is crucial to ensure preparedness at both the individual and organization levels and to prevent new outbreaks.

Keywords: Corona Virus Disease 2019; Temporal; Spatial; Distribution; Outbreak

Introduction

Emerging infectious diseases are a major challenge in the 21st century. In recent years, worldwide outbreaks of Ebola and Middle East Respiratory Syndrome caused great health and economic losses.^[1,2] The ongoing new coronavirus pneumonia (Corona Virus Disease 2019, COVID-19) outbreak is becoming a global public health problem. The COVID-19 outbreak is highly similar to the severe acute respiratory syndrome (SARS) outbreak that occurred in 2003; both outbreaks were caused by new coronaviruses during time periods overlapping with the

Chinese Spring Festival.^[3] On December 31, 2019, the Wuhan Municipal Health Committee reported 27 cases of pneumonia with an unknown cause, and many cases were traced to the Wuhan Southern China Seafood Market, which was subsequently closed on January 1, 2020.^[4] On January 7, 2020, laboratory tests showed that the pathogen causing the previously unexplained pneumonia was a new type of coronavirus; this pneumonia was then officially named COVID-19 by the World Health Organization.^[5,6] The COVID-19 outbreak started in Wuhan and spread rapidly to other provinces and countries.^[7,8] As of January 30, 2020, a total of 34

Access this article online

Quick Response Code:



Website:
www.cmj.org

DOI:
10.1097/CM9.0000000000000782

Ze-Liang Chen, Qi Zhang, and Yi Lu contributed equally to the work.

Correspondence to: Prof. Jia-Hai Lu, One Health Center, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong 510080, China
E-Mail: lujiahai@mail.sysu.edu.cn

Copyright © 2020 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2020;133(9)

Received: 10-02-2020 Edited by: Peng Lyu

provinces and regions in China had reported 9692 cases, and nearly all imported cases were derived from Wuhan in Hubei province.^[9,10]

COVID-19 has been defined as a class B infectious disease but has been managed as a class A infectious disease by the Chinese government. Daily case reports are being released, and any omission or concealment is punishable by law. Currently, the number of cases is still increasing, and the epidemic has not yet reached its peak; however, the situation differs from province to province. Information on the temporal and spatial distributions of cases is important for developing targeted treatment and prevention strategies. Because the return peak of Spring Festival travel is approaching, information on the possible changes in the incidence of COVID-19 in different cities will help in better preparation for disease prevention and management. Therefore, in this study, we investigated the temporal and spatial distributions of the early COVID-19 epidemic to reveal the dynamic changes and trends in reported cases. These results will provide valuable information for disease prevention at both the individual and organization levels.

Methods

Collection of case data

All officially reported confirmed and suspected cases of COVID-19 and related deaths were collected from the official website of health departments or articles citing their reports. Case data were imported into Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and analyzed.

Temporal and spatial distribution and risk analysis

The national and Hubei province shapefiles were used for ArcGIS (Environmental Systems Research Institute, Redlands, WA, USA) analysis. The map was linked to an Excel file containing time and location information. Location data were available for 34 provinces of China and 17 prefecture level cities of Hubei province. The time span was from January 16 to January 30, 2020. The COVID-19 risk analysis was based on the Bayesian space-time model of the WinBUGS (Microsoft Corporation) software.^[11,12] The model was divided into three levels:

(1) Data model

The statistical data on low incidence were assumed to follow a Poisson distribution for the parameters n_i and μ_{it} : $y_{it} \sim \text{Poisson}(n_i \mu_{it})$, where the Hubei province y_{it} was i (1, ..., 17) cities with t (1, ..., 15) days number of cases occurring during the day, and the nationwide y_{it} was the number of cases occurring in t (1, ..., 15) days in i (1, ..., 34) provinces. We assumed that there was no change in the number of people at risk in each city during the study period, such that n_i was the number of people at risk in the town (i), and μ_{it} was the corresponding disease risk in the city (t) per day (i).

(2) Process model

μ_{it} 's logarithmic transformation of disease risk allows the relative risk to be expressed as a linear combination of spatial, temporal, and spatiotemporal

interaction components. The mathematical expression is shown in Equation (1).

$$\log(\mu_{it}) = \alpha + s_i + (b_{0t}^* + v_t) + b_{1i}t^* + \varepsilon_{it} \quad (1)$$

where α is the fixed effect of the overall relative risk in the entire study area within 11 days, and $t^* = t - 5.5$ is the time span relative to the intermediate time point. In this model, the risk of disease is broken down into three parts: spatial change, temporal change, and space-time interaction; s_i is a component of spatial variability, describing the urban disease risk relative to the risk in the entire study region over an 11-day observation period; $b_{0t}^* + v_t$ is the change over time, which represents the overall trend of disease risk in the entire study area relative to that on the medium-term observation day, including the linear trend b_{0t}^* and the time random effect v_t ; b_0 is the time coefficient, representing the time trend in the study area; and $b_{1i}t^*$ allows each city to have different time-varying trends and is part of the spatiotemporal interaction. Relative to b_0 , it represents the trend of local change in each city based on b_0 ; ε_{it} is used to explain local changes that cannot be explained by spatiotemporal random effects.^[13]

(3) Parametric model

According to the Besag York and Mollié (BYM) model,^[14] a spatial structure effect is defined by a prior conditional autoregressive (CAR) structure. In this process, a spatial adjacency weight matrix needs to be defined. If adjacent, the weight $w_{ij} = 1$; otherwise the weight $w_{ij} = 0$, and the special $w_{ij} = 0$. Similarly, b_{1i} is also assumed to follow BYM characteristics. For the time structure effect v_t , a CAR process is used, and the adjacency weight matrix in time is defined. For the over-discrete parameter ε_{it} , according to Gelman, the normal distribution with a mean value of 0 and a variance of $\sigma^2\varepsilon$, is generally assumed and the variance of each parameter obeys Gamma (a, b).^[15] Based on this model, through the spatial component s_i and its posterior probability, high- or low-risk cities (identified based on the average risk [α] in the entire study area) can be identified. By calculating the probability that spatial relative risk $\exp(s_i)$ is greater than 1, regions can be divided into five categories: those with probability >0.8 , 0.6–0.8, 0.4–0.6, 0.2–0.4, and <0.2 are defined as hot spots, secondary hot spots, warm-spots, sub-cold spots, and cold-spots, respectively. Similarly, based on the probability threshold, the differences in these regions can be identified considering the trend over time. Further, based on the probability that $\exp(b_{1i})$ is greater than 1, regions can be divided into five categories: cities with an incidence risk probability greater than 0.8 show a trend for a rapid change in risk relative to the overall change, and those with an incidence risk probability between 0.6 and 0.8 show a trend for a greater change in the incidence risk than the overall change. A value between 0.4 and 0.6 indicates that the change in the occurrence risk is the same as the overall risk change; 0.2 to 0.4, that the trend of change in disease risk is lower than the overall risk change; and less than 0.2, that the trend of change in disease risk is much lower than the overall risk.

Correlation between number of cases and population migration

Population migration data were collected from the Baidu website (<http://qianxi.baidu.com/>). Data on emigration from Wuhan city and Hubei province to other cities and provinces were extracted and edited with Microsoft Excel for Windows (Microsoft Corporation). Emigration intensity was calculated using the migration index multiplied by the migration proportion in the province or city. Correlation analysis was performed using IBM SPSS Statistics software (version 22; International Business Machines Corporation, Armonk, NY, USA). *P* values less than 0.05 were considered statistically significant. Pearson correlation coefficients greater than 0.2 were considered indicators of a positive correlation.

Results

To obtain a general profile of the case distribution, we first analyzed all the available cases during this COVID-19

outbreak.^[16] As shown in Figure 1A, the number of cases remained stable from January 11 to 15, 2020, and the number of new and cumulative cases increased rapidly after January 16. The first death was reported on January 10, and the number of deaths began to increase rapidly from January 17 onwards, with the cumulative number of deaths reaching 213 on January 30 [Figure 1B].^[6] After the nucleic acid assay became available, suspected cases waiting for laboratory confirmation could be diagnosed rapidly.^[17] After January 19, the number of suspected cases increased rapidly, and about 40% to 50% of these suspected cases were then confirmed [Figure 1C]. Before January 19, the number of severe cases remained low, but they increased steadily from January 20 onwards [Figure 1D]. Because Wuhan is the capital city of Hubei province and the virus spread throughout the province quickly, we also analyzed the changes in number of cases in Hubei province. On January 9, 41 cases were first reported, and by January 30, 5806 cases had been reported, accounting for 59.91% (5806/9692) of the total cases in China [Figure 1E]. The cumulative number of

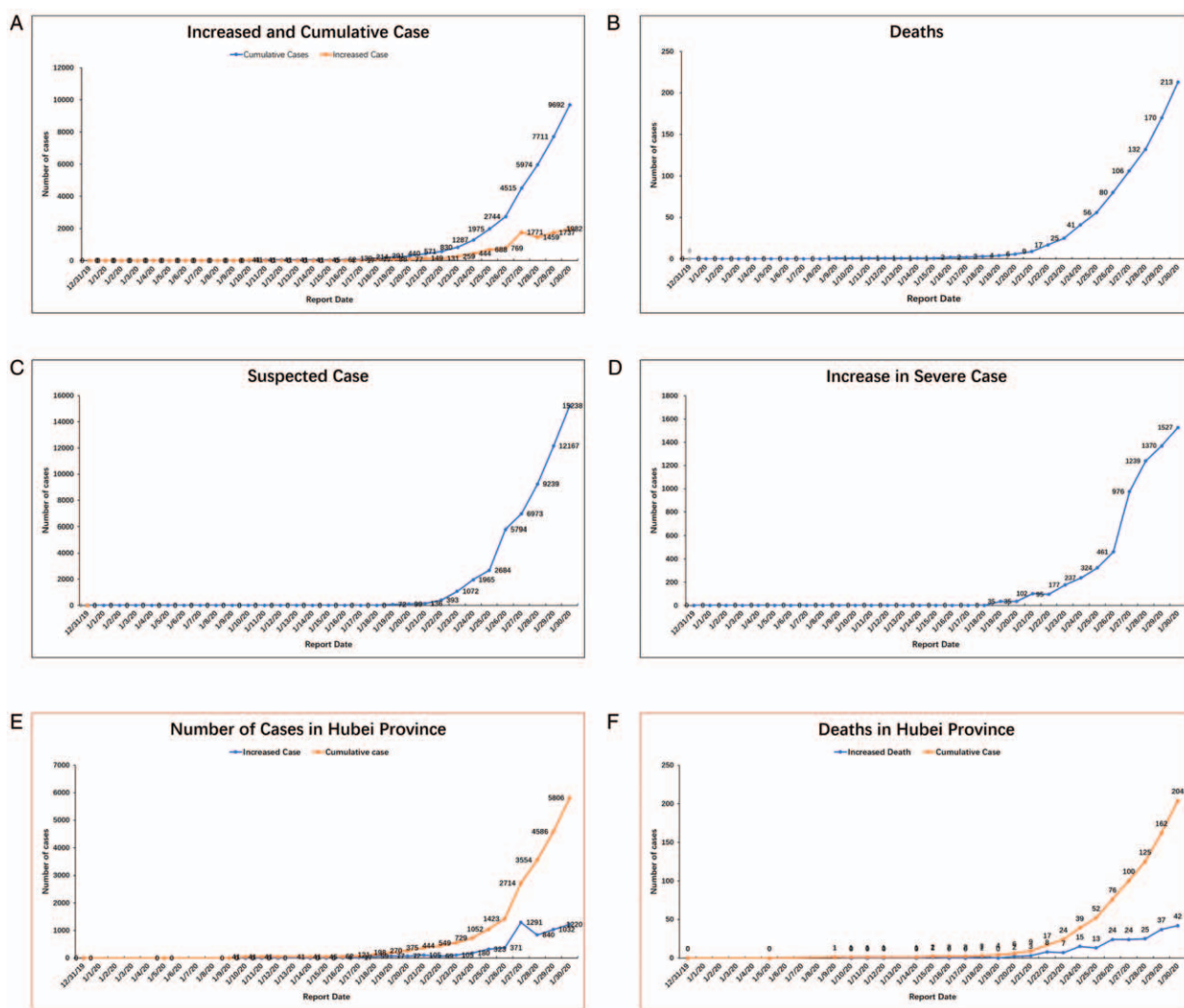


Figure 1: Daily changes of Corona Virus Disease 2019, cases in China. (A) Number of the increased and cumulative cases. (B) Number of death case. (C) Suspected cases. (D) Increase in severe cases. (E) Number of the increased and cumulative cases in Hubei province. (F) Number of the increased and cumulative death case in Hubei province.

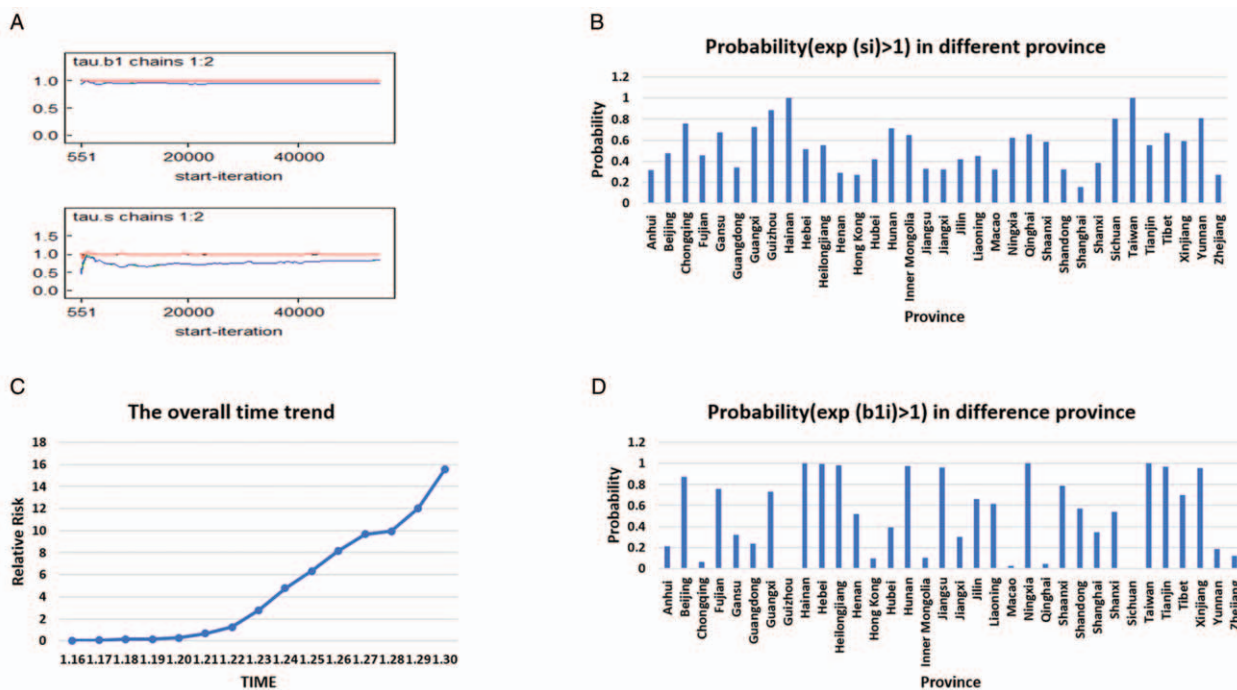


Figure 2: Nationwide distribution of Corona Virus Disease 2019, cases and change in trends across provinces in China. (A) Model convergence analysis. (B) Hot spots and cold spots of case distribution. (C) Overall trendline of relative risk with time. (D) Time risk probability of different provinces.

deaths in Hubei province was 204, accounting for 95.77% (204/213) of the total deaths in China [Figure 1F]. These data indicated that both the incidence and mortality of COVID-19 disease were the highest in Hubei province.^[18]

Before January 16, cases were mainly reported in Hubei province. From January 17 onwards, the outbreak spread to many provinces and the number of cases increased rapidly. Therefore, our spatial and temporal analyses used data from January 17 to 30, 2020. The location of each case was extracted from official reports and mapped onto the national map at the city level using ArcGIS. Of the 362 cities, 307 (84.8%, 307/362) had reported cases. In general, the core outbreak area, Wuhan, and its surrounding cities had the highest number of cases, followed by cities with a high population which are transportation hubs. Spatial distribution was then analyzed with a Bayesian model using WinBUGS. After nearly 100,000 iterations, the model converged successfully. After the model converged, it was iterated another 110,000 times to obtain parameter estimations. Generally, a ratio close to 1 indicates that the two chain iterative sequences are close, and that the model has a good convergence and is stable [Figure 2A]. Using the established model and parameters, hot and cold spots were identified. The results showed that Sichuan, Yunnan, Guizhou, Hainan, and Taiwan were hot spots, and Inner Mongolia, Gansu, Ningxia, Qinghai, Xinjiang, Chongqing, Hunan, and Guangxi were secondary hot spots. Generally, hot spots clustered in the midwest, and cold spots clustered in the southeast [Figure 2B].

The overall temporal trend was calculated using the time risk model ($\exp(b_{0i} * t + v_i)$), which described the general incidence risk according to time between January 16 and

30, 2020. Through the analysis, b_0 was estimated to be 0.4604, that is, the disease risk on the following day was found to be approximately 1.585 times higher than that on the previous day. The relative risk according to time increased steadily from January 20 onwards and the upward trend continued as of January 30 [Figure 2C], indicating that the number of cases nationwide is on the rise. As shown in Figure 2D, Heilongjiang, Hebei, Beijing, Tianjin, Xinjiang, Ningxia, Jiangsu, Hunan, Taiwan, and Hainan showed a faster increase in the number of cases than was observed overall in the country. The increase in the number of cases in Jilin, Liaoning, Shaanxi, Guangxi, and Fujian provinces also occurred relatively fast [Supplementary Table 1, <http://links.lww.com/CM9/A210>]. The increase in other provinces was consistent with or lower than the overall national trend [Figure 2D].

Since Hubei province had the highest number of cases, we analyzed the temporal and spatial distribution in different cities of Hubei province. Wuhan had the highest number of cases, followed by Huanggang and Xiaogan cities. Suizhou, Jingmen, and Xianning were part of the second group with a high number of cases. The spatial convergence analysis had 100,000 iterations [Figure 3A]. Hot spots were identified in the east regions and cold spots were identified in the west regions [Figure 3B]. The overall temporal trend in the change in the number of cases was calculated using the model. The average time trend coefficient b_0 was estimated to be 0.6727, indicating the time risk (occurrence probability in time) on the following day was 1.960 times higher than that on the previous day, suggesting that the daily number of cases in Hubei province is on the rise [Figure 3C]. Xiangyang, Suizhou, Yichang, and Ezhou showed the highest increase rates, and Shiyan, Shennongjia, Xiaogan, and Huangshi showed

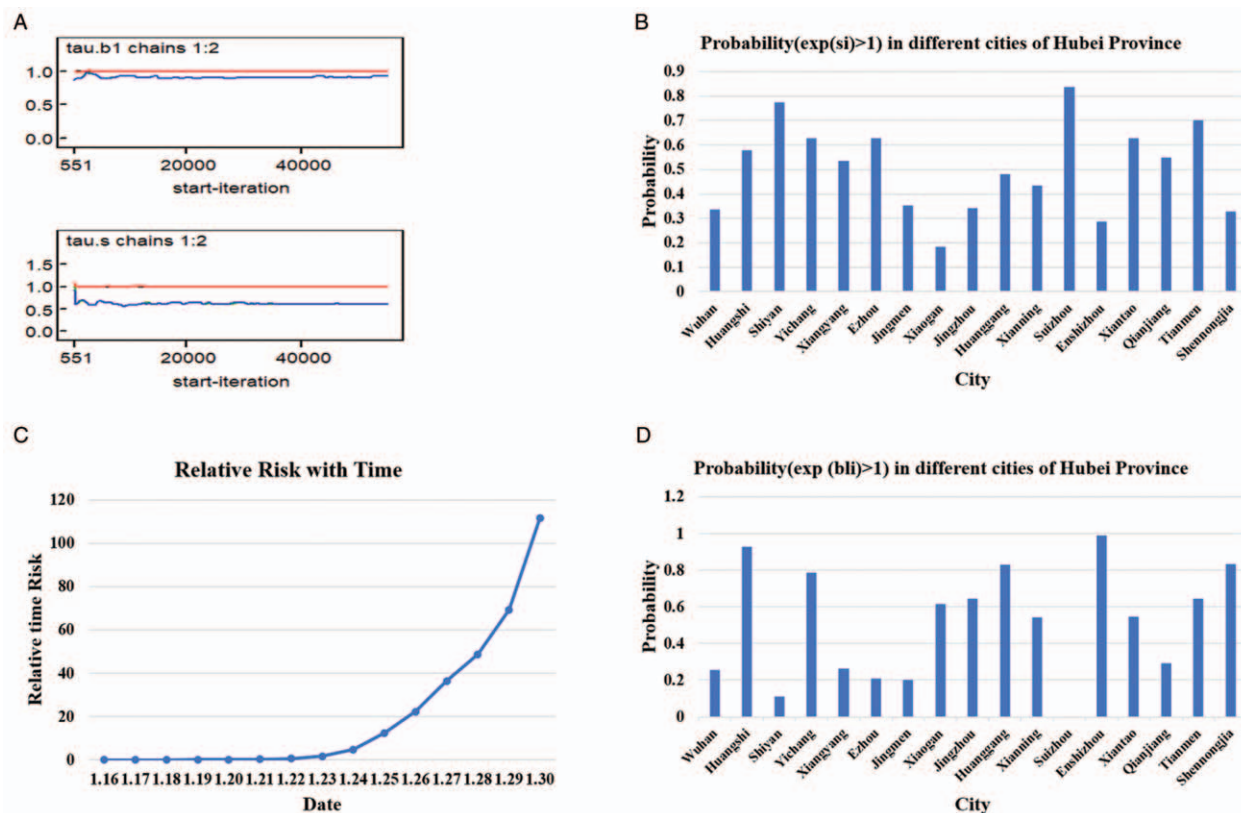


Figure 3: Distribution of Corona Virus Disease 2019, cases and change in trends of cities of Hubei province. (A) Model convergence analysis of case distribution. (B) Hot spots and cold spots of case distribution. (C) Trendline of relative risk with time. (D) Time risk probability of different cities of Hubei province.

relatively high increase rates [Figure 3D]. Other cities had a growth slower than the overall growth in the province [Supplementary Table 2, <http://links.lww.com/CM9/A210>]. The increase rate in Hubei province (1.960) was higher than that in the whole country (1.585), indicating that the rate of increase in Hubei province was significantly higher than that in other provinces in China.

Because the outbreak occurred just before the Spring Festival, large-scale population migration during this period influenced the subsequent epidemic. From January 1 to 23, 2020, the population that migrated out of Wuhan city and Hubei province increased steadily, peaking on January 21 and 22 [Figure 4A]. Wuhan city was under lockdown on January 23, and after that, population migration was greatly inhibited. As observed in 2019, high population migration occurred on January 31; the timely city lockdown prevented a subsequent outbreak burst. We analyzed the migration into and out of Wuhan city and Hubei province. The top targets for emigration included Henan and Hunan provinces [Figure 4B]. More people migrated out of Wuhan than into the city [Figure 4C]. To analyze the correlation between the number of cases and the emigration in Wuhan city and Hubei province, population migration data were collected from Baidu Qianxi. The correlation coefficient between the provincial number of cases and emigration from Hubei province was 0.719 [Figure 4D]. The correlation coefficient between the provincial number of cases and emigration from Wuhan increased to 0.943, with the highest coefficient of 0.996

observed between Wuhan and other cities of Hubei provinces [Figure 4E and 4F; Supplementary Tables 3 and 4, <http://links.lww.com/CM9/A210>]. These data strongly indicated that the number of cases was highly related to population emigration from Wuhan. Although we do not know the exact number of people emigrating from Wuhan, 5 million is an astonishing number, considering that each individual may be a potential virus carrier. If no control measures were implemented, the number of cases would exponentially increase. Of the 5 million emigrants, 74.22% emigrated to other cities of Hubei province [Supplementary Table 3, <http://links.lww.com/CM9/A210>]. Fortunately, 17 cities of Hubei province were under lockdown from January 23 to 26 [Supplementary Table 5, <http://links.lww.com/CM9/A210>]. After the lockdown of Wuhan and other cities of Hubei province, outbreak bursts were prevented, and the number of cases increased steadily but did not show exponential growth.

Because the outbreak duration overlaps with the Spring Festival transport waves, large-scale migration will be a strong determinant of the characteristics of this outbreak. We analyzed the migration in the 3 days before the Spring Festival. The top 50 cities from where emigration occurred before the Spring Festival were mainly located in the south and east of China, with Beijing, Shenzhen, Shanghai, and Guangzhou showing the highest emigration, accounting for over 15% of the migration population [Supplementary Figure 1, <http://links.lww.com/CM9/A211>]. However, cities with high immigration were relatively scattered.

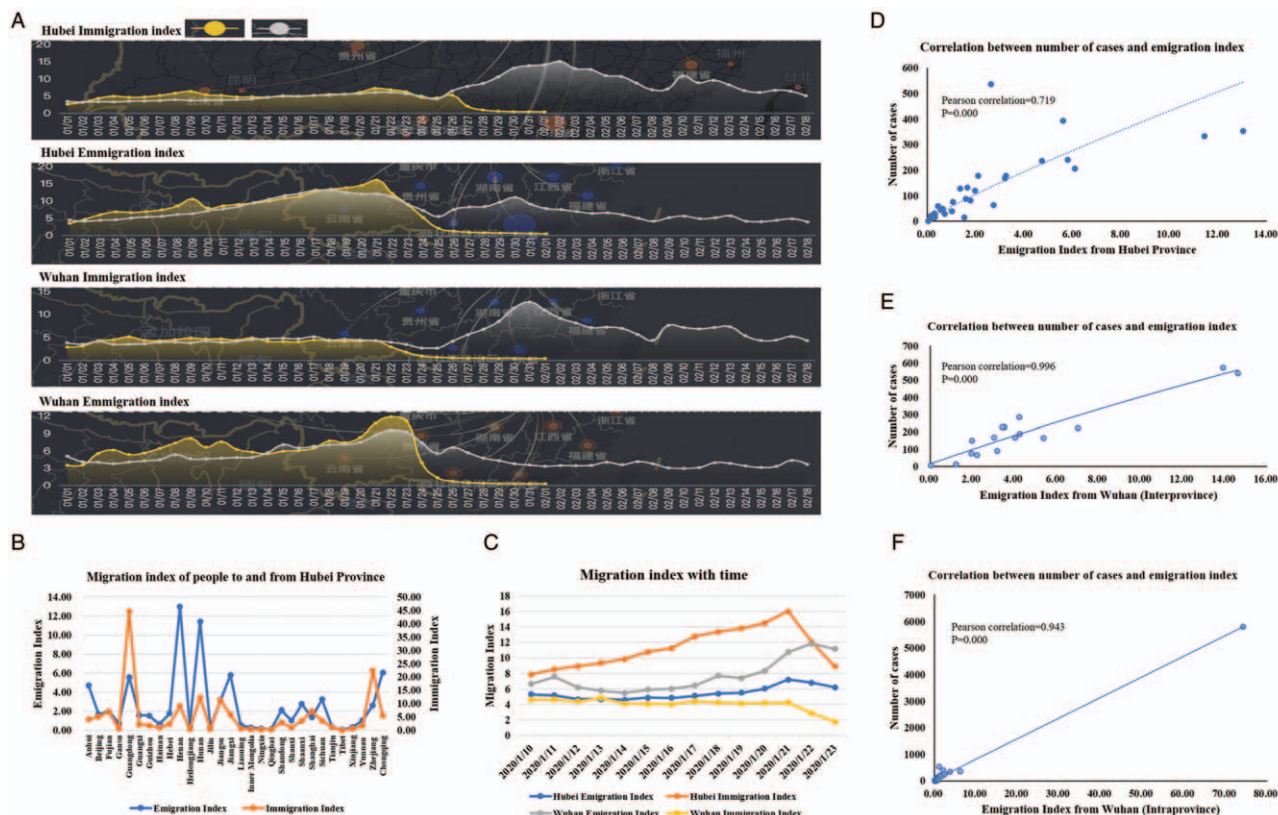


Figure 4: Correlation between migration index and the number of cases. (A) Migration index indicating the movement of people to and from Wuhan city and Hubei province during spring festival (Yellow, 2020; Gray, 2019). (B) Emigration and immigration index of people to and from Hubei province. (C) Emigration and immigration index of people to and from Wuhan city and Hubei province from January 10 to 23, 2020. (D) Correlation between the number of cases and emigration index of people from Hubei province. (E) Correlation between the number of cases and emigration index of people from Wuhan city (inter-province migration). (F) Correlation between the number of cases and emigration index of people from Wuhan city (intra-province migration).

Chongqing experienced the highest immigration, accounting for 1.50% of the total number of immigrants [Figure 4]. As immigrants will be traveling back to work after the Spring Festival, the cities showing high “emigration” may be at a high risk of another wave of new cases owing to the return of the migrants.

Discussion

COVID-19 is causing great public health and economic losses in China. The number of cases has increased rapidly, with over 70% coming from Hubei province.^[16,19] As of January 30, the number of cases has exceeded the total number of cases of the SARS-CoV outbreak.^[20] Until February 15, 2020, the cumulative number of confirmed cases was 70,533, nearly ten times that noted during the SARS outbreak. Prevention and control of the outbreak has required concerted action from the whole population of China. Although all individuals have participated in the campaign against the outbreak, people in areas with a low number of cases assumed that they were safe from the disease. Therefore, awareness of high-risk regions is important for preparing individuals, particularly in regions with low incidence. Further, it must be noted that 5 million persons emigrated from Wuhan to all over the country.^[21] We do not know exactly how many of them are virus carriers, and it is impossible to track and diagnose them all. Evidence from previous cases showed that asymptomatic patients in the incubation period are also infectious,

making it a greater challenge to track virus carriers. Therefore, isolation at home and less contact with others is the most efficient measure to prevent infection and transmission. To reduce transmission, the Spring Festival holiday has been extended from January 31 to February 2. The opening time for all schools and universities has been delayed, and online teaching programs have been launched. Factories have been required to delay resumption or allow work from home.

We analyzed the temporal and spatial distribution of reported cases. In general, the number of cases is still on the rise. For Hubei province, which has the highest number of cases and deaths, the growth trend is relatively stable. Conversely, in other hot spots, the number of cases was not very high, but the growth continued. Hence, these areas should be closely monitored.^[22] It is particularly noteworthy that the cities with the fastest change in temporal risk, such as Chongqing, have large population movements and rapid temporal risk. If they are not strictly monitored, there may be more outbreaks. To prevent disease outbreaks caused by the return travel wave after the Spring Festival, the country has extended the Spring Festival holiday.

Correlation analysis showed that early incidence was closely related to the emigration waves from Wuhan, that is, the higher the migrating population index, the larger was the number of cases. However, with the progress of the

epidemic, migrants are spreading the virus to other people and are becoming an important source of local community transmission. Therefore, it is necessary to strictly implement isolation and related control measures in accordance with the guidelines. Particularly, control measures must be taken to prevent the spread of diseases in communities, which is crucial to prevent a large-scale outbreak.

Very soon, many company staff will return to their workplaces. Because many enterprises in China are labor intensive, with large populations, human-to-human transmission is extremely easy. Therefore, workers need to meet requirements for isolation after returning to the city and use personal protection at work to prevent clustered outbreaks. At present, there have been several reports of employee infections caused by resumption of work; these represent a warning for all enterprises. Super megacities such as Guangzhou, Shenzhen, and Shanghai, which have the largest number of migrant workers, need to be prepared for this.

From February 16, the number of new cases began to decrease, but the epidemic did not stop completely. Therefore, we must act together to stop the spread of the disease. At present, the state has adopted mobility control measures to encourage people to avoid going to public places and wear masks when going out to reduce the risk of human-to-human transmission. We believe that with the joint efforts made by everyone, the number of cases and losses will be kept to a minimum.

Acknowledgements

The authors thank Andre Kiesel for critical revision of this manuscript.

Funding

This work was supported by grants from the National Science and Technology Major Project (No. 2018ZX10101002-001-001), National Key Research and Development Program Projects of China (No. 2017YFD0500305), the State Key Program of National Natural Science of China (No. U1808202), NSFC International (regional) cooperation and exchange program (No. 31961143024), the Key-Area Research and Development Program of Guangdong province (No. 2018B020241002), and the Guangdong Provincial Science and Technology Project (No. 2018B020207013).

Conflicts of interest

None.

References

- Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. *Nature* 2004;430:242–249. doi: 10.1038/nature02759.
- Suwantarat N, Apisarntharak A. Risks to healthcare workers with emerging diseases: lessons from MERS-CoV, Ebola, SARS, and

- avian flu. *Curr Opin Infect Dis* 2015;28:349–361. doi: 10.1097/QCO.0000000000000183.
- Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* 2005;310:676–679. doi: 10.1126/science.1118391.
- Krupovic M, Dolja VV, Koonin EV. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nat Rev Microbiol* 2019;17:449–458. doi: 10.1038/s41579-019-0205-6.
- Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, *et al.* First case of 2019 novel coronavirus in the United States. *N Engl J Med* 2020;382:929–936. doi: 10.1056/NEJMoa2001191.
- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;382:727–733. doi: 10.1056/NEJMoa2001017.
- Carlos WG, Dela Cruz CS, Cao B, Pasnick S, Jamil S. Novel Wuhan (2019-nCoV) coronavirus. *Am J Respir Crit Care Med* 2020;201:7–8. doi: 10.1164/rccm.2014P7.
- Lu H. Drug treatment options for the 2019-new coronavirus (2019-nCoV). *Biosci Trends* 2020. doi: 10.5582/bst.2020.01020 [Epub ahead of print].
- The Lancet. Emerging understandings of 2019-nCoV. *Lancet* 2020;395:311. doi: 10.1016/S0140-6736(20)30186-0.
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020. doi: 10.1056/NEJMoa2001316 [Epub ahead of print].
- Bohning D, Dietz E, Schlattmann P. Space-time mixture modelling of public health data. *Stat Med* 2000;19:2333–2344. doi: 10.1002/1097-0258(20000915/30)19:17/18<2333::aid-sim573>3.0.co;2-q.
- Liao Y, Zhang Y, He L, Wang J, Liu X, Zhang N, *et al.* Temporal and Spatial Analysis of Neural Tube Defects and Detection of Geographical Factors in Shanxi Province, China. *PLoS one* 2016; 11: e0150332. doi: 10.1371/journal.pone.0150332.
- Demirhan H, Kalaylioglu Z. Joint prior distributions for variance parameters in Bayesian analysis of normal hierarchical models. *J Multivar Anal* 2015;135:163–174. doi: 10.1016/j.jmva.2014.12.013.
- Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* 1991;43:1–20. doi: 10.1007/BF00116466.
- Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Perspect* 2004;112:1016–1025. doi: 10.1289/ehp.6740.
- Nishiura H, Jung SM, Linton NM, Kinoshita R, Yang Y, Hayashi K, *et al.* The extent of transmission of novel coronavirus in Wuhan, China, 2020. *J Clin Med* 2020;9. doi: 10.3390/jcm9020330.
- Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol* 2020;92:441–447. doi: 10.1002/jmv.25689.
- Cheng VCC, Wong SC, To KKW, Ho PL, Yuen KY. Preparedness and proactive infection control measures against the emerging Wuhan coronavirus pneumonia in China. *J Hosp Infect* 2020. doi: 10.1016/j.jhin.2020.01.010 [Epub ahead of print].
- Baidu. Real time data report of epidemic situation. Available from: <https://voice.baidu.com/act/newpneumonia/newpneumonia>. [Accessed February 1, 2020].
- Chan PK, Ip M, Ng KC, Rickjason CW, Wu A, Lee N, *et al.* Severe acute respiratory syndrome-associated coronavirus infection. *Emerg Infect Dis* 2003;9:1453–1454. doi: 10.3201/eid0911.030421.
- Hubei Province Press Conference on January 26. Retrieved from <http://www.chinanews.com/shipin/spfts/20200126/2493.shtml>. [Accessed January 27, 2020].
- Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, *et al.* Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int J Infect Dis* 2020;92:214–217. doi: 10.1016/j.ijid.2020.01.050.

How to cite this article: Chen ZL, Zhang Q, Lu Y, Guo ZM, Zhang X, Zhang WJ, Guo C, Liao CH, Li QL, Han XH, Lu JH. Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China. *Chin Med J* 2020;133:1044–1050. doi: 10.1097/CM9.0000000000000782