Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Spatiotemporal characteristics and factor analysis of SARS-CoV-2 infections among healthcare workers in Wuhan, China

P. Wang [a], H. Ren [a], X. Zhu [a, d, e, *], X. Fu [a], H. Liu [c], T. Hu [b, **]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China
[b] Center for Geographic Analysis, Harvard University, Cambridge, MA, USA
[c] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China
[d] Collaborative Innovation Center of Geospatial Technology, Wuhan, China
[e] Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, Wuhan University, Wuhan, China

A R T I C L E   I N F O

S U M M A R Y

Background: Studying the spatiotemporal distribution of SARS-CoV-2 infections among healthcare workers (HCWs) can aid in protecting them from exposure.
Aim: To describe the spatiotemporal distributions of SARS-CoV-2 infections among HCWs in Wuhan, China.
Methods: In this study, an open-source dataset of HCW diagnoses was provided. A geographical detector technique was then used to investigate the impacts of hospital level, type, distance from the infection source, and other external indicators of HCW infections.
Findings: The number of daily HCW infections over time in Wuhan followed a log-normal distribution, with its mean observed on January 23rd, 2020, and a standard deviation of 10.8 days. The implementation of high-impact measures, such as the lockdown of the city, may have increased the probability of HCW infections in the short term, especially for those in the outer ring of Wuhan. The infection of HCWs in Wuhan exhibited clear spatial heterogeneity. The number of HCW infections was higher in the central city and lower in the outer city.
Conclusion: HCW infections displayed significant spatial autocorrelation and dependence. Factor analysis revealed that hospital level and type had an even greater impact on HCW infections; third-class and general hospitals closer to infection sources were correlated with especially high risks of infection.
© 2021 The Healthcare Infection Society. Published by Elsevier Ltd. All rights reserved.

* Corresponding author. Address: State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, China. Tel.: 86-13808691386.
** Corresponding author. Address: Center for Geographic Analysis, Harvard University, Cambridge, MA, USA. Tel.: 1-3309908943.
E-mail addresses: xinyanzhu@whu.edu.cn (X. Zhu), taohu@g.harvard.edu (T. Hu).

Introduction

At the end of 2019, SARS-CoV-2 was discovered and it spread rapidly around the world. By October of 2020, more than 200 countries had been infected, and there were more than 30 million confirmed cases and more than 900,000 deaths attributed to the virus [1]. At present, the management of and
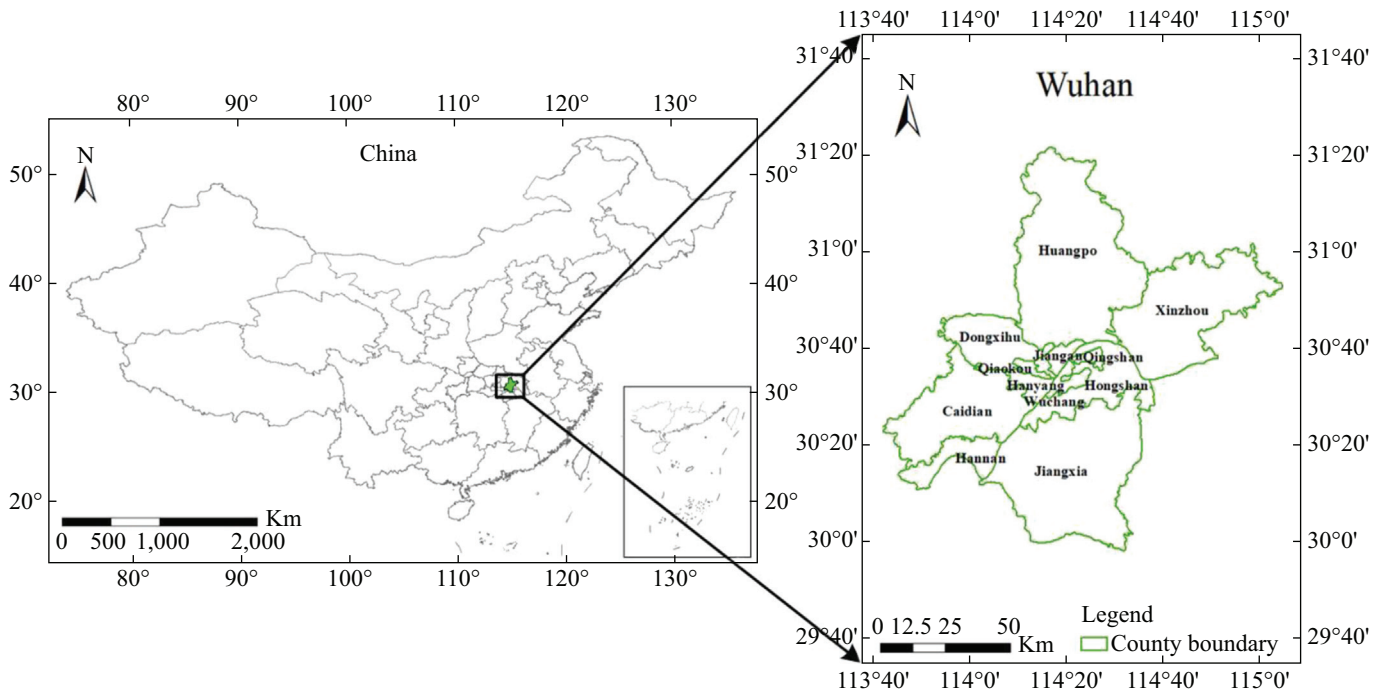
Figure 1. Map of the study area.

response to SARS-CoV-2 are common challenges facing all of humanity [2−4,31,32]. During the fight against the epidemic, healthcare workers (HCWs) have always been on the front line. While helping patients combat the virus, HCWs are also exposing themselves to high levels of the virus at close range [5,6]. Thus, how to better protect HCWs is a key issue at both the city and national levels.

Since the outbreak of SARS-CoV-2 began, different degrees of HCW infections have occurred in many countries and regions [7,8]. For example, the number of HCW infections in China accounts for ∼4% of all infections nationwide [9,10], whereas those in Italy account for ∼10%. Studies have shown that the infection rate of HCWs is significantly higher than that of non-HCWs [11]. In order to deal with HCW infections, researchers in China and elsewhere have explored the causes of HCW infections, revealed the transmission routes leading to HCW infections, and formulated allocation schemes to address resource limitations and protective measures [12−16].

Spatiotemporal analysis involves the use of statistics, geographical, and time-series data to study the patterns and mechanisms underlying a given phenomenon over time and space [17,18]. In the field of public health, even the introduction of more simplified analytical methods to evaluate spatial and temporal relationships could be beneficial [19]. Thus, researchers used geographic information systems to understand the spatiotemporal patterns of viral transmission among reported cases in the early stages of the epidemic [20−24]. However, due to the lack of data on HCW diagnoses, few existing HCW infection-related studies have been able to reveal the spatiotemporal characteristics of HCW infections and the external environmental factors influencing infections from a geographic perspective.

Therefore, we aimed to study the spatial distributions of HCW infections in Wuhan using partitioning statistics, distribution fitting, and spatial autocorrelation techniques [25,26]. A geographical detector was then used to explore the impacts of external factors, such as hospital level, type, and distance from the infection source, on HCW infections [27].

Methods

Study area and data sources

Wuhan is not only the first city in China where the SARS-CoV-2 was discovered, but also the city with the most serious infections among HCWs in China. Since the outbreak began, more than 3000 HCWs have been infected in Wuhan, accounting for 82.5% of infections nationwide. Wuhan is in Central China and is composed of a central city and an outer city (Figure 1). The central city includes the districts of Jiangan, Jianghan, Qiaokou, Hanyang, Wuchang, Qingshan, and Hongshan, while the outer city includes Huangpi, Xinzhou, Dongxihu, Jiangxia, Caidian, and Hannan.

The data used in this study can mainly be divided into two types: (i) hospital data from Wuhan and (ii) confirmed data on HCW infections in Wuhan.

Hospital data from Wuhan were mainly taken from the Health Commission of Hubei Province (http://wjw.hubei.gov.cn/). Web crawler technology was used to search for medical institutions in Wuhan and the returned data were parsed to obtain hospital information. After parsing, a total of 285 pieces of hospital information were obtained. Each record contained the unique identity, name, rank, type, latitude, and longitude of the hospital. Based on the hospital grading system of China, the ranks mainly include first-, second-, and third-class hospitals [28]. The number of beds in first-class hospitals is generally ≤100, the number of beds in second-class hospitals is generally 101−500, and the number of beds in third-class hospitals is

generally ≥501. Hospital types include community, specialized, and general. Community hospitals refer to medical institutions that provide basic medical services, general hospitals refer to medical institutions that have no restrictions on the scope of disease treatment, and specialist hospitals refer to medical institutions that have certain restrictions on the scope of disease treatment. To study the impact of the distance between a hospital and the infection source on HCW infections, we further calculated the straight-line distance between each hospital and the South China Seafood Market [29]; the calculated results are represented by the 'distance' throughout this work.

Data on HCW diagnoses in Wuhan were mainly derived from the Chinese Red Cross Foundation (https://www.crcf.org.cn/), which distributes relief funds to every confirmed healthcare worker [30]. As of September 11th, 2020, 83 groups of HCWs had received foundation assistance. We used crawler technology to obtain 3703 publications that elucidated the conditions of HCWs that suffered from SARS-CoV-2, including 3655 from Hubei Province and 3058 from Wuhan City. Each record contains the province, city, and hospital name of confirmed HCW infections. In addition, the 'hospital name' field of confirmed HCW infections is linked with the same field in the hospital data.

Methodological framework

The methodological framework of this study is shown in Figure 2. First, two types of HCW infection inventories, based on the crawled HCW infection data, were constructed: the Grid-Level Healthcare Worker Infection Inventory and the Hospital-Level Healthcare Worker Infection Inventory. The infection inventory records daily infection number for HCWs at different levels. Using the Grid-Level Healthcare Worker Infection Inventory, the spatiotemporal distributions and spatial autocorrelation of HCW infections were explored. Finally, based on the Hospital-Level Healthcare Worker Infection Inventory, a geographical detector technique was used to explore the impacts of hospital level, type, distance from the infection source (i.e. South China Seafood Market), and other environmental indicators on HCW infections. The details of the methodology framework are described in Appendices A.1−A.5.

Results and discussion

Temporal characteristics of HCW infections

To assess the temporal characteristics of HCW infections, the statistical distributions and changes in HCW infections over time were analysed. Supplementary Figure S2 shows the results fitted to normal, log-normal, and gamma distributions for HCW infections in Wuhan. The mean square error (MSE) was used to evaluate the fit of each distribution. The results showed that a log-normal distribution was the most suitable to describe the evolution of HCW infections in Wuhan. Supplementary Table S2 shows the statistical characteristics of HCW infections, for which 95% confidence intervals were calculated via bootstrap resampling. Based on the log-normal distribution, the mean
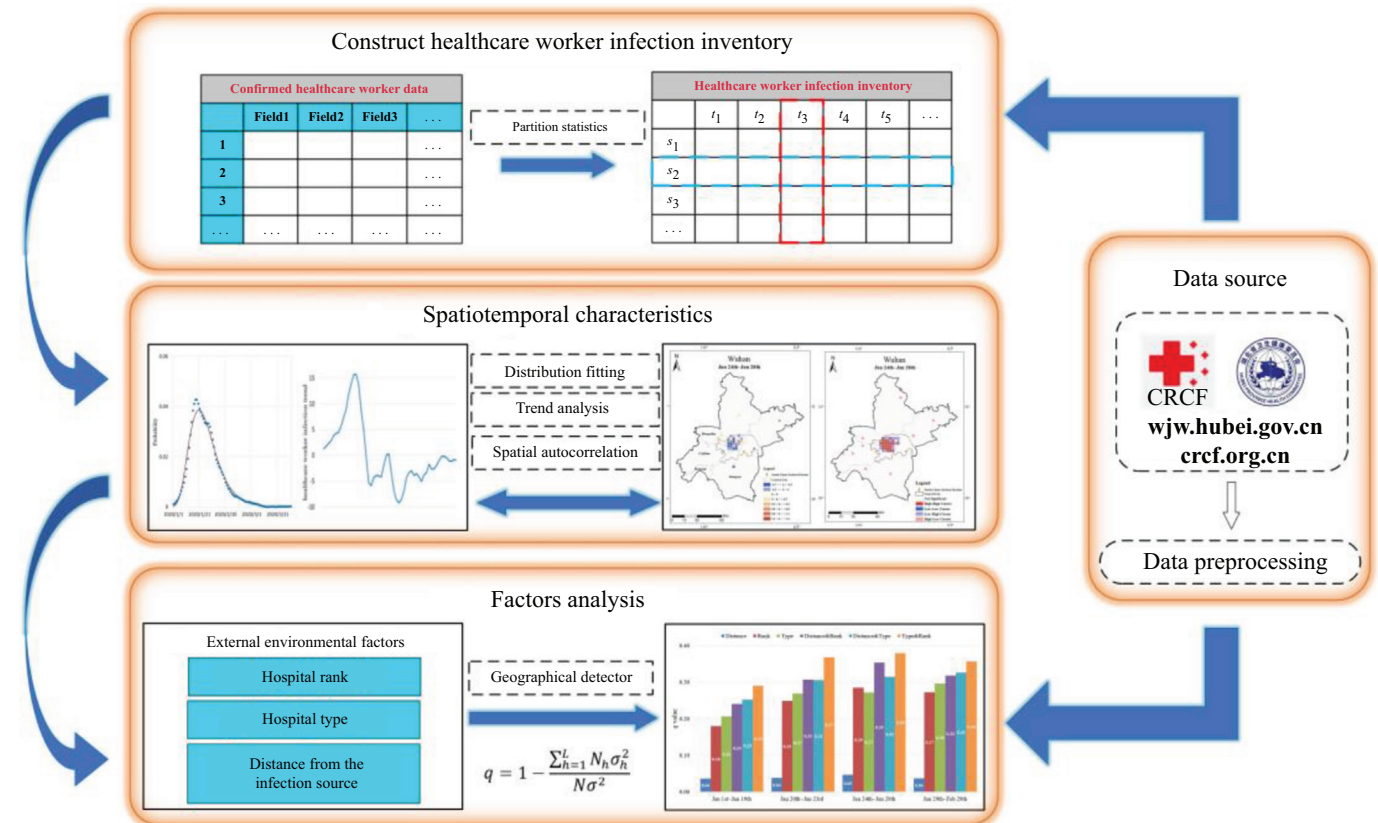

Figure 2. Research framework of spatiotemporal characteristics and correlation analyses for healthcare worker infections.

day of infections in Wuhan was estimated to be January 23rd, 2020, according to the data on daily cases of HCW infections, and the estimated standard deviation was 11.8 days.

The changes in the daily infection number of HCWs with time are shown in Supplementary Figure S3. Overall, HCW infections in Wuhan mainly occurred between January 1st and February 29th, 2020. Within that period, infections rose from January 1st to 19th, whereas from January 19th to February 29th, they declined. Moreover, from January 23rd to 28th, the downward trend of HCW infections gradually decreased, and the infection rate curve showed an upward trend again between January 25th and 28th; thereafter the HCW infections fell again. There are two main reasons for this change. First, the lockdown measures largely contained the spread of the SARS-CoV-2, but in a short period of time it led to a large number of symptomatic patients visiting the hospitals, which increased the intensity of work for medical staff, leading to increased infection rates among them. Second, as of January 28th, there were >6000 HCWs in Hubei, which greatly eased the workload, thus decreasing the rates of infection. Supplementary Figure S1b clearly shows the changes in the trend of HCW infections over time. The decreasing trend of HCW infections began to slow on January 21st. This time node is highly consistent with that determined by human-to-human transmission.

In general, the temporal characteristics of HCW infections were highly correlated with the times of release of news or policies. In particular, the release of isolation measures (e.g. lockdowns) may exacerbate HCW infections in the short term. Therefore, before the implementation of such policies, it is particularly important to strengthen the measures in place to protect HCWs from infection.

Spatial characteristics of HCW infections

As shown in Supplementary Figure S2, the infection of HCWs can be roughly divided into four stages: (1) January 1st to 19th, (2) January 20th to 23rd, (3) January 24th to 28th, and (4) January 29th to February 29th. Here, we show the spatial distributions and trends, and analyse the spatial autocorrelations during these four stages.

Spatial distributions

The spatial distributions of HCW infections are shown in Supplementary Figure S4. The spatial distribution of HCW infections during the four stages showed the same trend, in that the rates of infection in the central urban area were higher than those in the outer urban area. This indicates that infections among HCWs in Wuhan were extremely unbalanced and spatially heterogeneous.

The areas with severe HCW infections were mainly located near the South China Seafood Market in the central urban area, as this market was a key source of the SARS-CoV-2 epidemic in Wuhan. However, there were slight differences in the spatial distributions of HCW infections between the four stages. Compared with the distribution of HCW infections from January 1st to 19th, the rate of HCW infections in the outer city of Wuhan from January 20th to February 29th gradually increased.

Spatial trends

The spatial trends of HCW infections are shown in Supplementary Figure S5. From January 1st to 19th, the overall rate of HCW infection in Wuhan increased rapidly, with the fastest growth rate in the areas near the South China Seafood Market. From January 20th to 23rd, HCW infections exhibited a downward trend for the first time, among which the rate of decline was relatively fast in Wuchang and Jiangan. However, across Wuhan, the infection of HCWs continued to increase. From January 24th to 28th, the rate of HCW infections in Wuhan showed a declining trend overall, while the rates of infection in Dongxihu, Caidian, Hannan, and Jiangxia showed an upward trend again. Combined with Supplementary Figure S2, these results demonstrate that intense measures, such as the lockdown of Wuhan, may suddenly increase the workload of HCWs in mild and non-infected areas, and thus increase the probability of HCW infections in outer city areas. From January 28th to February 29th, the overall trend of HCW infection decreased, but there were still some areas with slight increases in rates of infection near the South China Seafood Market, indicating that HCWs near the infection source were always more susceptible to infection.

Spatial autocorrelations

The Global Moran's I was used to explore the global spatial dependency of the four stages of HCW infections (Supplementary Table S3). In each study interval, there was a significant spatial dependence of HCW infections, and, with increases in time, this dependence gradually increased. From January 24th to 28th, the spatial dependency was the strongest, where the Global Moran's $I = 0.475$, $Z > 2.58$, and $P < 0.001$, indicating that HCW infections in Wuhan were significantly and positively correlated at the grid-level. In other words, grids with high numbers of infections among HCWs had higher numbers of HCW infections across the grid, whereas grids with low numbers of HCW infections had lower numbers of HCW infections across the grid.

The Local Moran's I was used to explore the local spatial dependency of the four stages of HCW infections (Supplementary Figure S6). Three clustering patterns were observed in the four stages, namely high–high, high–low, and low–high abnormal clustering. High–high clustering occurred in the infection hot spot area, which was mainly concentrated around the South China Seafood Market in the central city. High-low clustering was mainly distributed near the high–high clustering, indicating that HCW infections were essentially controlled within a certain range. Abnormal low–high clustering was mainly distributed in the districts of the outer city in a discrete manner, and the low–high areas gradually increased over time, indicating that some areas of the outer city also had a high incidence of HCW infections. Combined, the Global and Local Moran's I values reveal that HCW infections exhibited regional agglomeration.

Analysis of factors related to HCW infections

In this study, geographic detectors were used to explore the impacts of external environmental factors, such as hospital level, type, and distance from the infection source, on HCW infections in the four stages previously identified (Supplementary Figure S7). From a single-factor perspective, hospital level and type had greater impacts on HCW infections, with q-values mainly falling between 0.18 and 0.35, indicating that hospital type and rank could explain 18–35% of HCW infections. Meanwhile, distance had a lesser impact on HCW

infections, with q-values ranging from 0.03 to 0.04, indicating that distance could only explain 3–4% of HCW infections. To further verify the validity of the q-value, Supplementary Table S4 showed the hypothesis test results of each variable. The P-values of the three variables were all <0.05, that is, the three variables could explain the phenomenon of HCW infections. With respect to the interactions between two factors, hospital level and hospital type worked together to explain HCW infections to the greatest extent. At the same time, whereas the distance had a small impact on HCW infections, the combination of distance and either hospital level or type could cause a significant increase in the q-values, which indicates that, combined, these factors could explain HCW infections to a great degree. Additionally, with increases in time, the explanatory power of hospital level or type on HCW infections gradually increased, while the explanatory power of distance changed little.

To further explore the influences of hospital level, type, and distance from the infection source on HCW infections, the percentages of daily HCW infections to total HCW infections for each factor were computed (Supplementary Table S5). HCWs in third-class and general hospitals were more likely to be infected. Whereas distance had little explanatory power, HCWs who were close to the infection source were still more vulnerable to infection. As shown in Supplementary Figure S6, HCWs at third-class and general hospitals closer to the infection source were at the greatest risk of infection.

As the first line of defence in the fight against the SARS-CoV-2 epidemic, HCWs help patients combat the virus, while exposing themselves to high concentrations of the virus. It is important to study the spatiotemporal distributions and factors influencing SARS-CoV-2 infections among HCWs to better protect them. Existing studies related to HCW infections have largely emphasized infection rates and protective measures, whereas the spatiotemporal patterns and related external factors related to HCW infections have been unclear. To fill this knowledge gap, we first studied the spatiotemporal distributions of infections of HCWs in Wuhan using statistical partitioning, curve-fitting, and spatial autocorrelation. The results showed that the lowest MSE was obtained using a log-normal distribution, indicating that this distribution was the most suitable to describe the evolution of HCW infections over time. The mean date of infection was January 23rd, with a standard deviation of 10.8 days. The implementation of substantial measures, such as the lockdown of Wuhan, may have contributed to short-term increases in the probability of infection among HCWs, especially in outer city areas. This indicates that it is especially important to strengthen the protective measures for HCWs before such policies are implemented. The rate at which HCWs were infected in Wuhan displayed apparent spatial heterogeneity. The number of HCW infections was higher in the central city and lower in the outer city and the likelihood of infection exhibited significant spatial autocorrelation and spatial dependency.

To better understand the temporal and spatial distributions of HCW infections, geographical detector techniques were used to explore the impacts of external environmental factors, such as hospital level, type, and distance from the infection source on HCW infections. It was found that hospital level and type significantly impacted the probability of HCW infections, among which third-class and general hospitals carried the greater risks of infection. This conclusion is similar to that of

Zheng *et al.*, which independently validates the accuracy of our results to a certain extent [11]. Moreover, although the distance from the infection source had little impact on HCW infections, the combined effect of distance and hospital level or type was shown to increase the explanatory power of HCW infections, that is, third-class and general hospitals closer to the infection source had the greatest risks of HCW infection. Therefore, investing more medical supplies into hospitals with higher risks of HCW infection is vital.

This study had four primary limitations. First, the selection of influencing factors excluded factors that are known to contribute to the incidence of HCW infections, such as access to sufficient medical supplies, sufficient knowledge of protective protocols, and patient contact, among others. However, due to the difficulty of data acquisition, only three external factors — hospital rank, type, and distance from the infection source — were considered. Second, the coverage of HCW infection data was narrow. Only data on HCW infections in China were used to explore related spatiotemporal characteristics and influencing factors; data from other countries were not included, which may limit the generalizability of our results. Third, the infection data for HCWs was not comprehensive. As the information on HCW infections is published on the official website of the Chinese Red Cross Foundation in batches, additional information may continue to be published in the future that could not be included here. Finally, our study used the number of HCW infections instead of the infection rate, ignoring the heterogeneity of HCWs in Wuhan. In light of these limitations, future studies should focus on collecting additional domestic and foreign HCW infection data to more accurately and comprehensively analyse the spatiotemporal characteristics and factors influencing SARS-CoV-2 infections among HCWs.

Conflict of interest statement
None declared.

Data availability statement
Data supporting the findings of this study are available in 'figshare.com' (https://doi.org/10.6084/m9.figshare.13372985).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhin.2021.02.002.

References

[1] World Health Organization. Coronavirus disease (COVID-19). 2020. September 27th. Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200928-weekly-epi-update.pdf [last accessed February 2021].

[2] Fang Z, Yi F, Wu K, Lai K, Sun X, Zhong N, et al. Clinical characteristics of coronavirus pneumonia 2019 (COVID-19): an updated systematic review. medRxiv p. 2020.03.07.20032573, 2020. doi: 10.1101/2020.03.07.20032573.

[3] Guan W-j, Ni Z-y, Hu Y, Liang W-h, Ou C-q, He J-x, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382:1708–20.

[4] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395(10223):497–506.

[5] Wańkowicz P, Szylińska A, Rotter I. Assessment of mental health factors among health professionals depending on their contact with COVID-19 patients. Int J Environ Res Publ Health 2020;17:5849.

[6] Xiang Y-T, Jin Y, Wang Y, Zhang Q, Zhang L, Cheung T. Tribute to health workers in China: a group of respectable population during the outbreak of the COVID-19. Int J Biol Sci 2020;16:1739.

[7] Jones NK, Rivett L, Sparkes D, Forrest S, Sridhar S, Young J, et al. Effective control of SARS-CoV-2 transmission between healthcare workers during a period of diminished community prevalence of COVID-19. Elife 2020;9:e59391.

[8] Rivett L, Sridhar S, Sparkes D, Routledge M, Jones N, Forrest S, et al. Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission. Elife 2020;9:e58728.

[9] Gao W, Sanna M, Tsai M-K, Wen CP. Geo-temporal distribution of 1,688 Chinese healthcare workers infected with COVID-19 in severe conditions – a secondary data analysis. PLoS One 2020;15:e0233255.

[10] Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. Chin J Epidemiol 2020;41:145–51.

[11] Zheng L, Wang X, Zhou C, Liu Q, Li S, Sun Q, et al. Analysis of the infection status of healthcare workers in Wuhan during the COVID-19 outbreak: a cross-sectional study. Clin Infect Dis 2020;71:2109–13.

[12] Wang J, Zhou M, Liu F. Reasons for healthcare workers becoming infected with novel coronavirus disease 2019 (COVID-19) in China. J Hosp Infect 2020;105:100–1.

[13] Ferioli M, Cisternino C, Leo V, Pisani L, Palange P, Nava S. Protecting healthcare workers from SARS-CoV-2 infection: practical indications. Eur Resp Rev 2020;29(155).

[14] Siow WT, Liew MF, Shrestha BR, Muchtar F, See KC. Managing COVID-19 in resource-limited settings: critical care considerations. Berlin: Springer; 2020.

[15] Tan L. Preventing the transmission of COVID-19 amongst healthcare workers. J Hosp Infect 2020;105:364–5.

[16] Tan L, Kovoor JG, Williamson P, Tivey DR, Babidge WJ, Collinson TG, et al. Personal protective equipment and evidence-based advice for surgical departments during COVID-19. Aust NZ J Surg 2020;90:1566–72.

[17] Yu IT, Wong TW, Chiu YL, Lee N, Li Y. Temporal–spatial analysis of severe acute respiratory syndrome among hospital inpatients. Clin Infect Dis 2005;40:1237–43.

[18] Vernaz N, Sax H, Pittet D, Bonnabry P, Schrenzel J, Harbarth S. Temporal effects of antibiotic use and hand rub consumption on the incidence of MRSA and Clostridium difficile. J Antimicrob Chemother 2008;62:601–7.

[19] Davis G, Sevdalis N, Drumright L. Spatial and temporal analyses to investigate infectious disease transmission within healthcare settings. J Hosp Infect 2014;86:227–43.

[20] She B, Hu T, Zhu X, Bao S. Bridging open source tools and Geo-portals for interactive spatial data analytics. Geo-spatial Inform Sci 2019;22:185–92.

[21] Zhu X, Hu T, Ye X, Guo W, Huang L, Xiong H, et al. Development and implementation of a dynamic and 4D GIS based on semantic location model. Geo-spatial Inform Sci 2019;22:193–213.

[22] Wang SY, Liu Y, Struthers J, Lian M. Spatiotemporal characteristics of COVID-19 epidemic in the United States. Clin Infect Dis 2020 Jul 8:ciaa934. https://doi.org/10.1093/cid/ciaa934 [online ahead of print].

[23] Jia J, Lu X, Yuan Y, Xu G, Jia J, Christakis N. Population flow drives spatio-temporal distribution of COVID-19 in China. Nature 2020;582(7812):1–11.

[24] Liu Q, Sha D, Liu W, Houser P, Zhang L, Hou R, et al. Spatio-temporal patterns of COVID-19 impact on human activities and environment in mainland China using nighttime light and air quality data. Remote Sensing 2020;12:1576.

[25] Moran PAP. Notes on continuous stochastic phenomena. Biometrika 1950;37:17–23.

[26] Anselin L. Local indicators of spatial association – LISA. Geogr Anal 1995;27:93–115.

[27] Wang JF, Li XH, Christakos G, Liao YL, Zhang T, Gu X, et al. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. Int J Geogr Inform Sci 2010;24:107–27.

[28] National Bureau of Statistics, China. 2017 national economic industry classification (GB/T 4754-2017). 2017. Available at: http://www.stats.gov.cn/statsinfo/auto2073/201710/P020181204372998240694.pdf [last accessed February 2021].

[29] Kang Y, Xu S. Comprehensive overview of COVID-19 based on current evidence. Dermatol Ther 2020;33:e13525.

[30] Chinese Red Cross Foundation. Management rules of Chinese Red Cross Foundation ByteDance healthcare workers humanitarian relief fund (revised edition). 2020. July 10th. Available at: https://new.crcf.org.cn/article/19718 [last accessed February 2021].

[31] Hu T, Guan WW, Zhu X, Shao Y, Liu L, Du J, et al. Building an open resources repository for COVID-19 research. Data Inform Manage 2020;4(3):130–47.

[32] Hu T, Yue H, Wang C, She B, Ye X, Liu R, et al. Racial segregation, testing site access, and COVID-19 incidence rate in Massachusetts, USA. Int J Environ Res Public Health 2020;17(24):9528.