Understanding the spatial diffusion dynamics of the COVID-19 pandemic in the city system in China

Lijuan Gu [a], Linsheng Yang [a,b], Li Wang [a], Yanan Guo [a], Binggan Wei [a], Hairong Li [a,*]

[a] *Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China*
[b] *College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China*

ARTICLE INFO

ABSTRACT

Investigating the spatial epidemic dynamics of COVID-19 is crucial in understanding the routine of spatial diffusion and in surveillance, prediction, identification and prevention of another potential outbreak. However, previous studies attempting to evaluate these spatial diffusion dynamics are limited. Using city as the research unit and spatial association analysis as the primary strategy, this study explored the changing primary risk factors impacting the spatial spread of COVID-19 across Chinese cities under various diffusion assumptions and throughout the epidemic stage. Moreover, this study investigated the characteristics and geographical distributions of high-risk areas in different epidemic stages. The results empirically indicated rapid intercity diffusion at the early stage and primarily intracity diffusion thereafter. Before countermeasures took effect, proximity, GDP per capita, medical resources, outflows from Wuhan and intercity mobility significantly affected early diffusion. With speedily effective countermeasures, outflows from the epicenter, proximity, and intracity outflows played an important role. At the early stage, high-risk areas were mainly cities adjacent to the epicenter, with higher GDP per capita, or a combination of higher GDP per capita and better medical resources, with more outflow from the epicenter, or more intercity mobility. After countermeasures were effected, cities adjacent to the epicenter, or with more outflow from the epicenter or more intracity mobility became high-risk areas. This study provides an insightful understanding of the spatial diffusion of COVID-19 across cities. The findings are informative for effectively handling the potential recurrence of COVID-19 in various settings.

1. Introduction

The COVID-19 pandemic is impacting all of us. The symptom onset date of the first identified case was on Dec 1, 2019, in Wuhan, China (Huang et al., 2020). Since then, the virus has spread across cities and countries and gradually led to an ongoing worldwide pandemic. Although various efforts have been made to bring the disease under control and many high-income countries have managed to build a robust testing system, implement widespread vaccinations and provide timely and accessible treatments, globally, it continues to wreak havoc (WHO, 2022) and COVID-19 resurgence has been frequently and widely reported (Basov et al., 2022; Fox News, 2022). Vaccines are critical tools in protecting people against diseases caused by many variants (Lopez Bernal et al., 2021) and ending the pandemic. However, vaccines are not 100% protective, and there are vaccine breakthrough infections

(Juthani et al., 2021). The distribution of vaccines is heavily skewed with more than half of the planet being currently unvaccinated (Altmann and Boyton, 2022). Moreover, given the currently rapid transmission worldwide, both the evolution rate and the risk of the emergence of new variants are very high. Because the primary series vaccines show reduced effectiveness against new variants such as Omicron (WHO, 2022), vaccines alone cannot end the pandemic (Zhao et al., 2021). Being well informed about the disease and its transmission process is also vital in solving the pandemic and accelerating global recovery.

The impacts of the COVID-19 pandemic on the health system and economy have been devastating (Gössling et al., 2020). Because a widespread uptake of vaccines and accessible treatments are still underway, a thorough understanding of the linked features is necessary. To date, numerous studies have explored genomic, epidemiological,

clinical and laboratory features, as well as possible treatments and clinical outcomes (Huang et al., 2020). The effectiveness of non-pharmaceutical interventions has been estimated (Tian et al., 2020; Ferguson et al., 2020), and the transmission dynamics have been modeled (He et al., 2020; Kissler et al., 2020; Zhang et al., 2020). Moreover, the role of meteorological features in affecting transmission has received extensive scientific attention (Rahman et al., 2021), and there are some cross-sectional studies working on social environmental factors (Hu et al., 2020; Lee et al., 2021). However, due to a lack of a thorough understanding of the whole transmission process, researchers are still contemplating why and how COVID-19 spreads in colleges (Bahl et al., 2021), across cities (Coşkun et al., 2021), and among countries (Lin et al., 2020).

The values of spatially located data, such as infectious diseases, are very likely to be related to space (Meng et al., 2005). In epidemiology, the spatial pattern of diseases can provide information useful in capturing important facets of their diffusion processes (Kanga et al., 2020). Although spatial analysis is of great help in understanding the route of transmission (Meng et al., 2005), a spatial investigation of COVID-19 is still lagging. Moreover, although the spatial linkages between areas and the spatial patterns of high-risk areas are helpful in surveillance, prediction, identification and prevention (Wang et al., 2006), research on these issues is far from sufficient.

As the first country hit by COVID-19, China reported its first case at the end of December 2019 (Wuhan Municipal Health Commission, 2019), and declared the peak to be over on March 13, 2020 (XinhuaNet, 2020). According to governmental documents (Wuhan Municipal Health Commission, 2020), the symptom onset date of the first confirmed case was December 8, 2019, instead of December 1 as reported in the literature (Huang et al., 2020). The virus spread rapidly at the early stage, during which effective countermeasures were not taken until the lockdown of Wuhan city on January 23, 2020. Subsequently, numerous intervention measures were implemented (Burki, 2020). On February 16, National Health Commission (NHC) spokesman Mi Feng said that controls had started to rein in the virus (Reuters, 2020). On March 12, the peak of this outbreak originating in Wuhan was officially declared to be over (XinhuaNet, 2020). Therefore, going through both a rapid diffusion stage and an effective control process, China provides an ideal case to investigate the spatial diffusion of COVID-19 through an entire epidemic phase. Moreover, spatial scales matter in studying the spatial spread and spatial association of infectious diseases (Wang and Di, 2020; Mu et al., 2020). There are hundreds of cities in China, and they vary in their physical, social, cultural, and economic environments and COVID-19 outcomes. Thus, the city makes a potentially interesting research unit. However, studies on the spatial diffusion dynamics of COVID-19 across cities are rare.

Spatial association is vital in understanding the spatial linkage and diffusion of infectious diseases (Kanga et al., 2020). Using spatial association as the primary analytic strategy, this study explored the spatial diffusion dynamics of COVID-19 across Chinese cities. Various spatial connection assumptions regarding the potential risk factors were considered. Sociodemographic factors, such as population, population density, household size, and socioeconomic factors, such as GDP per capita, urbanization, green space, and medical resources, whose significance in the diffusion of infectious diseases are widely documented (Meng et al., 2005; Kanga et al., 2020; Mu et al., 2020; Lee et al., 2021), were considered. Moreover, given the importance of population movement in shaping the spatiotemporal patterns of epidemics (Balcan et al., 2009), we also included large-scale mobile phone mobility data. Second, to identify the primary epidemic factors at different diffusion stages, the temporal evolution of spatial association was investigated. Finally, to detect high- and low-risk areas under various hypothetical diffusion processes, the geographical distribution of city clusters defined by various factors was presented. The originality of this study concerns an exploration of the spatial dynamics of COVID-19 across cities through an entire phase. Our findings provide a valuable reference in responding to

another potential outbreak of infectious diseases in addition to COVID-19 in various settings.

2. Methodology

2.1. Research unit

A total of 357 mainland cities were included. Among them, 4 were provincial-level municipal cities directly under the administration of the central government, 15 were vice-provincial cities, 318 were prefecture-level cities, and 20 were county-level cities directly under the administration of the province or autonomous region. Their geographical distribution can be seen in Supplementary Figure 1.

2.2. Study period

The study period was January 19, to March 15, 2020, which covered the whole contagion process of local cases. A brief timeline is listed. On January 19, Shenzhen city in Guangdong Province detected its first confirmed case, which was the first reported outside of Wuhan in mainland China. By January 23, a total of 117 cities in 29 of the 31 provinces (or municipal cities) had confirmed cases. By January 29, a total of 288 cities in all 31 provinces (or municipal cities) had reported confirmed cases. By February 13, approximately 90 percent of cities had confirmed cases. With the transmission being brought gradually under control, China started to make dynamic adjustments on February 21 to support a flexible resumption of work and life and downgrade the emergency level. On March 12, the peak of the outbreak was declared to be over. At a press conference on March 16, Mi Feng said that curbing imported cases had become the top priority (National Health Commission, 2020).

2.3. Data sources

Epidemic data. Daily morbidity and newly confirmed cases were used. Morbidity was calculated by dividing accumulated confirmed cases by population. National-level daily aggregated data were acquired from the COVID-19 data repository operated by the Center for System Science and Engineering at Johns Hopkins University (JHU CSSE) (https://github.com/CSSEGISandData/COVID-19). JHU CSSE data aggregates local media and government reports to provide cumulative total cases in near real-time at the provincial level in China (Dong et al., 2020). Except for county-level cities, city-level data were acquired from the China Data Lab (CDL) (https://projects.iq.harvard.edu/chinadatalab). CDL collects data from Ding Xiang Yuan (https://ncov.dxy.cn/ncovh5/view/pneumonia), a professional platform in the medical field providing authoritative public information. Both JHU CSSE and CDL are popular among the research communities in providing reliable officially-reported COVID-19 information (Muhareb and Giacaman, 2020; Liu, 2021). Data for county-level cities were acquired from the official website of the local health commission. Given that imported cases were constantly reported since late February, to ensure accuracy, we manually crosschecked JHU CSSE and CDL data with daily issued official data. Any contradiction was fixed according to official data.

Sociodemographic data. Living in crowded conditions and frequent human contact are major risk factors for infectious diseases (Kanga et al., 2020; Lee et al., 2021; Meng et al., 2005; Tian et al., 2018); thus, population scale, population density and household size were included as potential factors. Because this outbreak covered the Chinese Spring Festival, during which the enormous "floating" population returned to their home of origin to celebrate the Lunar New Year, we used the total number of registered household members as the measure of population. We divided the population by administrative area to generate population density. Household size was the average number of family members per household. Sociodemographic data were collected from the 2019 China City Statistical Yearbook (Department of Urban Surveys, 2019),

the 2019 China County Statistical Yearbook (Department of Rural Surveys, 2019), and the annually posted Statistical Bulletin of National Economic and Social Development from the local government website.

Socioeconomic data. Economically and socially marginalized areas and persons are more vulnerable to infectious diseases (Gardner et al., 2018; Karimi et al., 2021; Mu et al., 2020). Therefore, indicators such as urbanization, GDP per capita and medical resources were included. Moreover, green space was also considered for being closely related to socioeconomic level and opportunities for outdoor activities without the fear of poor ventilation (Allen and Marr, 2020; Lee et al., 2021). Urbanization is measured by the proportion of the urban resident population to the total population. Medical resources included the total number of hospitals, hospital beds, and licensed (assistant) doctors and the number of hospitals per 100 000 persons, hospital beds per 1000 persons, and licensed (assistant) doctors per 1000 persons. Green space was measured by the percentage of green covered area to the completed area. Socioeconomic data were from the 2019 China City Statistical Yearbook, the provincial-level 2019 Statistical Yearbook, the county-level 2019 Statistical Yearbook, and the annually posted Statistical Bulletin of National Economic and Social Development from the local government website.

Migration data. Before Wuhan's lockdown at 10 a.m. on January 23, the number of reported cases was estimated to count merely 14% of the total confirmed cases (Li et al., 2020). According to Xianwang Zhou, Mayor of Wuhan, before the quarantine, more than 5 million people had left (The Guardian, 2020). Moreover, many of the early confirmed cases in other cities were epidemiologically related to Wuhan (Sun et al., 2020). Therefore, the outflows from Wuhan to other cities before lockdown are deemed important. Besides, given the infectious feature of COVID-19, migration data both between and within cities are essential. Because people in other cities were still moving, the intercity movement data are therefore important. Similarly, although under drastic control measures, there were people within cities still on the move, thus the role of intracity movement is also considered.

Migration data were scraped from the Baidu Migration website (https://qianxi.baidu.com/). We collected daily outflows from Wuhan from January 1 to 23, which covered the massive migration period due to the festival (January 10 to 23, 2020) before the lockdown. We also collected each city's inter- and intracity movement data from January 1 to March 31. All the migration data were presented in a movement index with regional and temporal comparisons enabled.

2.4. Methodology

Mapping spatial distribution. There is an intrinsic variance instability in rate data, i.e., the precision of the rate is inversely proportional to the scale of the population at risk (Anselin et al., 2006). To minimize the potentially large standard error estimated from small populations, Empirical Bayes smoothing technique (EB) was applied. EB calculates a weighted average between the raw rate of each unit and their average, with weights proportionally related to the scale of the population (Anselin, 2018). The EB estimate for COVID-19 morbidity in city i is

$$r_i^{EBM} = \omega_i r_i + (1 - \omega_i)\mu$$

$$\omega_i = \frac{\sigma^2}{(\sigma^2 + \mu/N_i)}$$

with r_i and r_i^{EBM} as raw and EB smoothed morbidity, ω_i as weight, which ranges from 0 to 1 and is positively related to population N_i, μ as the overall average of morbidity, and σ^2 as the overall variance. The estimates of μ and σ^2 are given by

$$\mu = \sum_{i=1}^{i=n} C_i \bigg/ \sum_{i=1}^{i=n} N_i$$

$$\sigma^2 = \frac{\sum_{i=1}^{i=n} N_i (r_i - \mu)^2}{\sum_{i=1}^{i=n} N_i} - \frac{\mu}{\sum_{i=1}^{i=n} N_i / n}$$

where n is the total number of cities, and C_i is the number of confirmed cases in city i.

Measuring spatial diffusion. Spatial autocorrelation was used to explore daily spatial patterns. Being arguably the most commonly used indicator of global spatial autocorrelation (Anselin, 2020), Moran's I with various definitions of distance was applied. The formula of Moran's I is

$$I = \frac{\sum_i \sum_j \omega_{ij}(y_i - \bar{y})(y_j - \bar{y}) \big/ \sum_i \sum_j \omega_{ij}}{\sum_i (y_i - \bar{y})^2 / n}$$

with i and j as city indices, ω_{ij} as the distance between city i and j, y_i and y_j as the number of newly confirmed cases (formula for morbidity is illustrated later), and $\bar{y}$ as the daily mean of confirmed cases. Moran's I ranges from -1 to 1, with higher absolute values indicating stronger spatial autocorrelation. Moran's I statistic with pseudo p-value being stably less than 0.05 under different permutations indicated significance.

There are clear city-level differences in population density (see Supplementary Figure 2). For varying population densities, when the variable of interest is a rate, to acquire the correct Moran's I, EB standardization was suggested (Assunção and Reis, 1999). Different from the EB smoothing technique, which computes for a smoothed version, EB standardization turns the crude rate into a transformed standardized random sample with zero mean and unit variance. Therefore, EB standardization minimizes the variance instability problems in rate data. The equation to calculate EB standardized morbidity r^{EBT} is

$$r_i^{EBT} = \frac{r_i - \mu}{\sqrt{\sigma^2 + (\mu/N_i)}}$$

A set of 17 models, namely, 17 different ways to define distance, were used to generate weights and to calculate spatial autocorrelation. The one-order queen contiguity-based weights were applied in Model 1. To further explore the impact of proximity, Model 2 used k-nearest neighbor weights (KNN) (k = 6 to avoid isolates). Model 3 used urbanization. Model 4 used GDP per capita. Population and population density were considered in Models 5 and 6. In Model 7, average household size was used. Model 8 considered the role of green covered areas. The roles of medical resources were investigated in Models 9 through 14. Outflows from Wuhan were considered in Model 15. Specifically, in analyzing spatial diffusion from January 19 to 23, accumulated outflows from January 1 to 18 were used. Since January 24, data from January 1 to 23 were used. Model 16 and Model 17 considered respectively the role of inter- and intracity mobility. A series of daily updated weight matrices were generated. Considering the approximately 4-day incubation period (Guan et al., 2020), the approximately 3-day median time from symptom onset to diagnosis, and the 7-day time lag from infection to report in China since late January (WHO, 2020), we assumed that the reported outcome at Day t was associated with mobility data at Day t-7 or earlier. Therefore, in measuring the spatial diffusion of morbidity at Day t, accumulated data from January 1 to Day t-7 were applied. For newly confirmed cases at Day t, mobility data at Day t-7 were used. This strategy is in accordance with the extant literature (Mu et al., 2020). Given that different definitions of spatial weights may generate different matrices and therefore different results, to verify the robust role of proximity, we compared Moran's I when the contiguity queen and rook, distance band, KNN, inverse distance band and inverse KNN were respectively used.

In investigating the spillover between areas in subjects such as resource utilization efficiency and house price, previous studies argued that it was neither a simple function of spatial propinquity nor a simple function of socioeconomic factors. Instead, a combination of spatial and

economic distance might be more plausible (Sun et al., 2014). Similarly, in analyzing the spatial diffusion of COVID-19, we inferred that large cities might be less remote than their geographical separation would imply because of their smaller socioeconomic distance. Thus, we further used a set of compound weight matrices combining geographical distance with social factors. Referring to the literature (Anselin and Bera, 1998; Fingleton and Le Gallo, 2008), compound matrix ω_{ij}^C was a combination of inverse socioeconomic distance and a negative exponential function of Euclidean distance d_{ij} between cities i and j

$$\omega_{ij}^C = \left[1 \,/\, |x_i - x_j + 1| \right] * \exp\left(- \beta * d_{ij} \right)$$

where x_i and x_j are observations on "meaningful" social characteristics. β is the distance decay rate and is set as 100 following the literature.

To understand the location of clusters or outliers provided by global spatial autocorrelation, a local indicator of spatial association (LISA) was used to acquire the local Moran's I statistic. The equation of the Local Moran statistic is given by

$$I_i = \frac{\sum_j \omega_{ij}(y_i - \bar{y})(y_j - \bar{y}) \big/ \sum_i \sum_j \omega_{ij}}{\sum_i (y_i - \bar{y})^2 \big/ n}$$

Local Moran's I was calculated for models having significant Moran's

I and exhibiting sharp temporal change (from significant to not significant or vice versa). Given that in China, GDP per capita is closely related to many other factors, such as population, urbanization, population density, household size, medical facilities and mobility (see Supplementary Table 1). To understand in-depth the spatial patterns of clusters, we further applied conditional maps to present the distribution of LISA in cities divided by a combination of GDP per capita and certain risk factors. For example, in the conditional LISA map for outflows from Wuhan, we generated a matrix of maps determined by outflows from Wuhan on the horizontal axis and GDP per capita on the vertical axis (both axes were dichotomized by median values). In calculating the LISA for morbidity, EB standardization was applied. Reference distribution approach was used to assess significance.

3. Results

3.1. Descriptive information of COVID-19 and potential risk factors

Cities situated in Central China had the highest value of total confirmed cases, followed by cities situated in Eastern China. In contrast, the majority of cities situated in Western and Northeastern China fell into the 1st or 2nd quartile (Fig. 1(a)). Compared to that of confirmed cases, regional differences in morbidity were more visible,
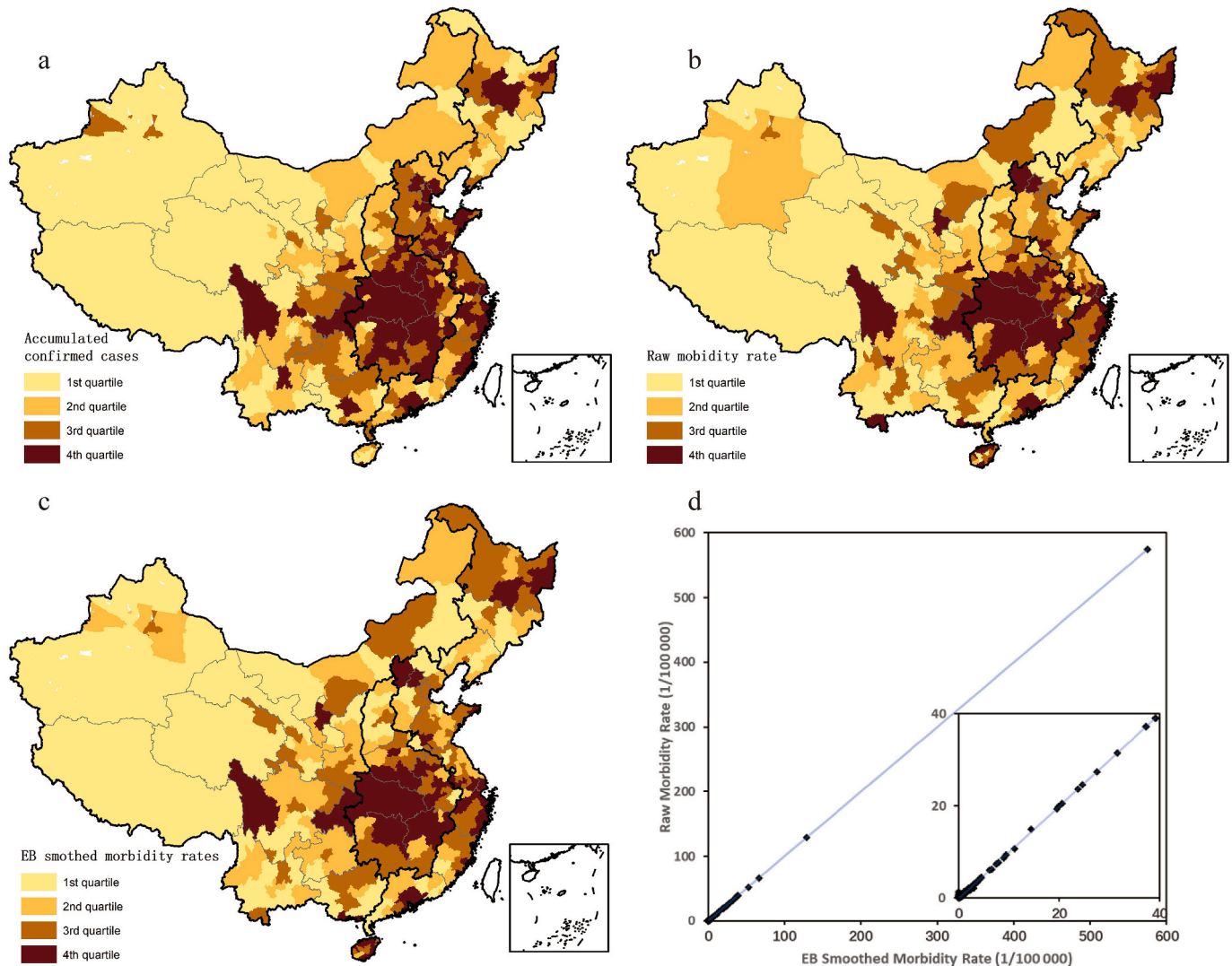

Fig. 1. City-level total confirmed cases and morbidity in mainland China, March 15, 2020. (a) Total confirmed cases. (b) Crude morbidity. (c) EB smoothed morbidity. (d) Scatter plot of crude morbidity vs. EB smoothed morbidity.

with many Central China cities falling into the 4th quartile (Fig. 1(b) and (c)). The EB smoothing technique slightly changed the spatial pattern, lowering rates in some Western cities (Fig. 1(c)). EB smoothing technique mainly smoothed morbidity values less than 20 (1/100 000) (Fig. 1(d)).

This epidemic lasted for approximately 8 weeks. There was a rapid increase in the number of confirmed cases until week 4 and a flattened curve thereafter (Fig. 2(a) and (b)). Fig. 2(c) depicts a rapid increase in cities involved from week 1–2 and a steady increase in morbidity from week 1–4. Combining Fig. 2(b) with 2(d), we can see a rapid increase in cities with newly confirmed cases in week 1 and a steady increase in newly confirmed cases until week 4. Fig. 2(d) also shows a considerable drop in the number of outliers and newly confirmed cases since week 5.

The discrepancies in potential risk factors across cities were visible (Table 1 and Supplementary Figure 2). Fig. 3(a) and (b) indicate a gradual decrease in both inter- and intracity mobility in week 1, compared to two weeks earlier, and a sharpened decrease in both from week 2 to week 5. From week 6, with the resumption of work and production, both inter- and intracity mobility gradually recovered. Geographically, cities situated in Western and Northeastern provinces, such as Tibet, Xinjiang, and Heilongjiang and some cities in Hubei, had a visibly weaker intensity. In comparison, intensity was stronger among municipal cities such as Shanghai and Chongqing, and capital cities such as Changsha and some Eastern coastal cities. Fig. 3(c) and (d) map the accumulated outmigration from Wuhan to other cities. Generally, except for municipal, vice-provincial and some Northeastern cities, which had a higher level of outmigration independent of geographical separation, outflows to other cities decayed with increasing geographical distance.

3.2. The temporal evolution of spatial autocorrelation under various assumptions

Fig. 4(a) shows the temporal changes in Moran's I for morbidity. Only risk factors showing any significance are listed. Generally, Moran's I was significantly different from 0 at week 1 when GDP per capita, number of doctors, hospital beds per 1000 people, and intercity mobility were included. From early week 2, Moran's I was significantly different from 0 when queen contiguity, KNN and outflows from Wuhan were used. In comparison, Moran's I was significantly different from 0 since week 3 when intracity mobility was used. Specifically, when contiguity and KNN distance were applied, Moran's I stayed above 0.4 before February 12 and stayed between 0.2 and 0.3 thereafter. When Wuhan outflows were used, Moran's I stayed higher than 0.6 since January 26. The significant Moran's Is were around 0.04 when GDP per capita, doctors, and hospital beds per 1000 people were used. Moran's I decreased from more than 0.05 to no larger than 0.04 since January 24 when intercity mobility was considered. In contrast, Moran's I using intracity mobility increased to more than 0.05 - after February 11.

Fig. 4(b) maps Moran's I for newly confirmed cases. Since week 1, Moran's I was significantly different from 0 when contiguity, KNN, GDP per capita, hospitals, doctors, beds, outflows from Wuhan, and intercity mobility were used. For contiguity, KNN and Wuhan outflows, the significant Moran's I lasted until week 6. For GDP per capita, doctors, hospital beds and intercity mobility, it lasted until week 2. For hospitals, Moran's I was significant at week 1. For intracity mobility, Moran's I was significantly different from 0 from the end of week 2 until week 7. Specifically, when contiguity and KNN were used, Moran's I was no larger than 0.4. When outflows from Wuhan were used, Moran's I was as large as 0.68. In comparison, for GDP per capita, hospitals, doctors and
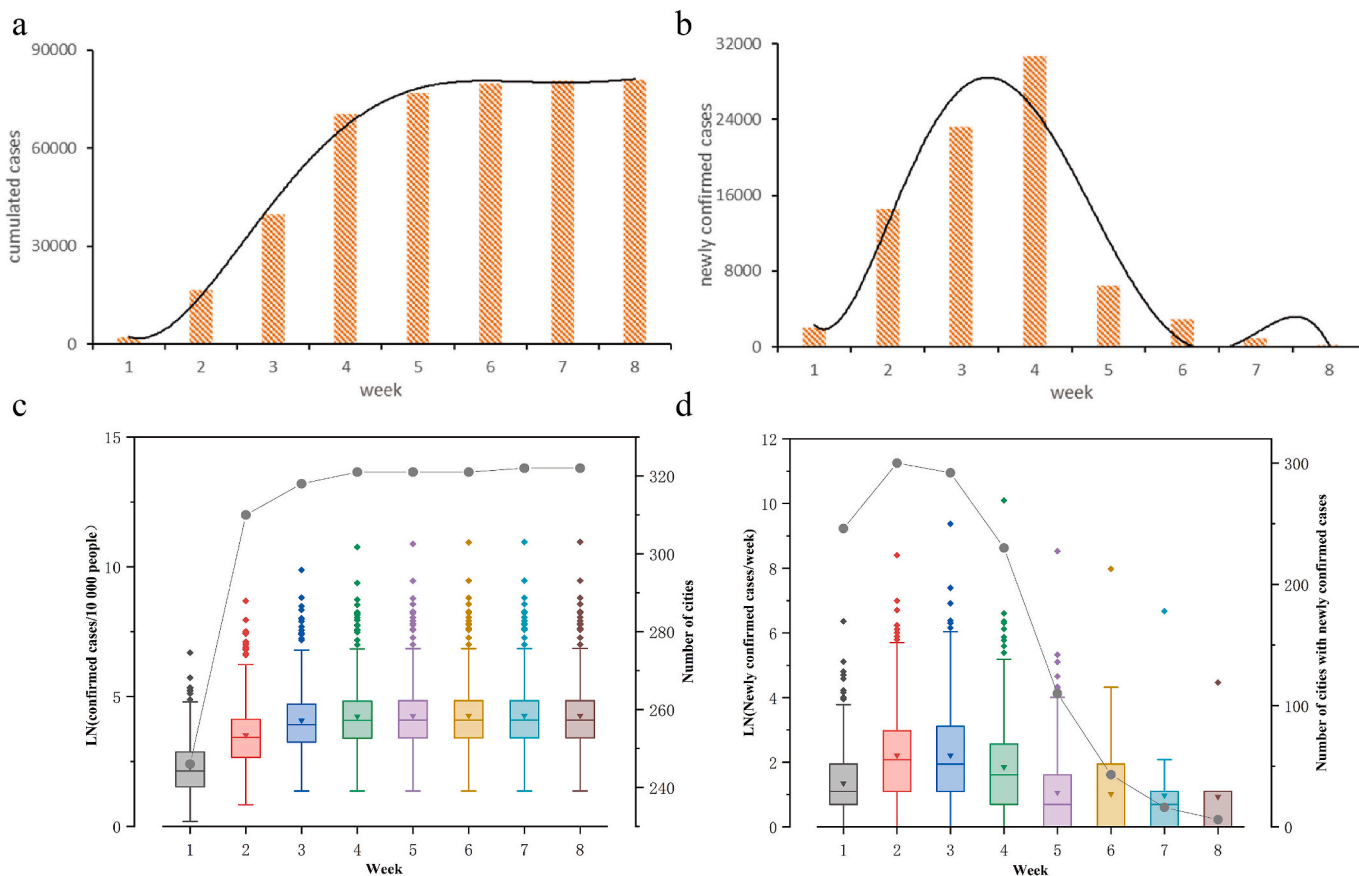

Fig. 2. The diffusion of COVID-19 from January 19 to March 15 by week. (a) Accumulated cases. (b) Newly confirmed cases. (c)Box chart of morbidity (logarithm transformed). (d) Box chart of newly confirmed cases (logarithm transformed).

Table 1
Descriptive statistics of potential risk factors.

| Variables	Observations	Mean	Std. Dev.	Min	Max
Urban (Urbanization, %)	357	55.21	1.47	21.00	99.75
GDP (GDP/capita, yuan)	357	57626.68	35825.93	12094.00	300000.00
Pop (Population, 10 000 people)	357	388.12	328.43	0.25	3397.00
Popden (Population density, people/sq.km.)	357	374.15	353.09	0.36	2578.44
Hsize (Household size, people)	357	3.12	0.51	1.18	4.75
Green (Green covered area as % of completed area, %)	357	39.56	6.40	9.13	64.10
Hospital (Number of hospitals)	357	98.59	104.90	1.00	892.00
Bed (Number of beds)	357	18342.20	19268.44	30.00	162147.00
Doctor (Number of licensed (assistant) doctors)	357	10107.15	11447.51	4.00	109376.00
Ahospital (Number of hospitals/100 000 persons)	357	2.85	1.93	0.04	22.85
Abed (Number of beds/1000 persons)	357	4.67	1.77	1.67	13.68
Adoctor (Number of doctors/1000 persons)	357	2.56	1.16	0.91	8.83
Flow (Migration from wuhan (Jan.1-Jan.23), index)	356	185.48	830.51	0	7939.31
Inter (Accumulated inter-city movement (Jan.1-Mar.8), index)	356	48.29	53.98	0.15	389.58
Intra (Accumulated intra-city movement (Jan.1-Mar.8), index)	356	250.39	44.45	121.12	378.06
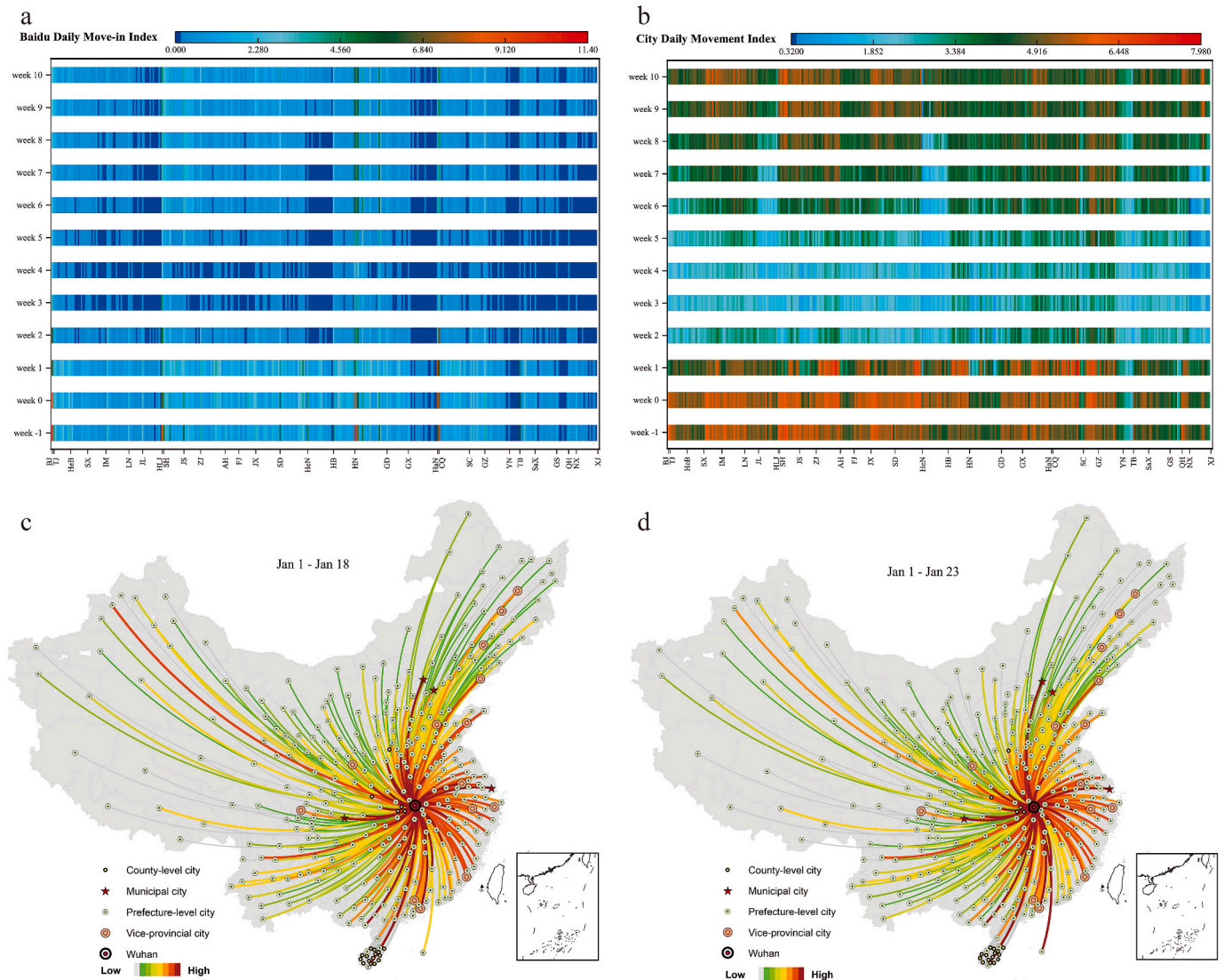
Fig. 3. The distribution of mobility data. (a) Inter-city move-in index by city by province by week. (b) Intra-city mobility by city by province by week. (c) Out-migration from Wuhan (January 1 to 18). (d) Outmigration from Wuhan (Jan 1 to 23). (Note: Movement data two weeks earlier (week 0-Jan 12 to 18, week -1-Jan 5 to 11) and later (week 9-March 16 to 22, week 10-March 23 to 29) were added in Fig. 3(a) and (b) for ease of comparison. BJ-Beijing, TJ-Tianjin, HeB-Hebei, SX-Shanxi, IM-Inner Mongolia, LN-Liaoning, JL-Jilin, HLJ-Heilongjiang, SH-Shanghai, JS-Jiangsu, ZJ-Zhejiang, AH-Anhui, FJ-Fujian, JX-Jiangxi, SD-Shandong, HeN-Henan, HB-Hubei, HN-Hunan, GD-Guangdong, GX-Guangxi, HaN-Hainan, CQ-Chongqing, SC-Sichuan, GZ-Guizhou, YN-Yunnan, TB-Tibet, SaX-Shaanxi, GS-Gansu, QH-Qinghai, NX-Ningxia, XJ-Xinjiang.).
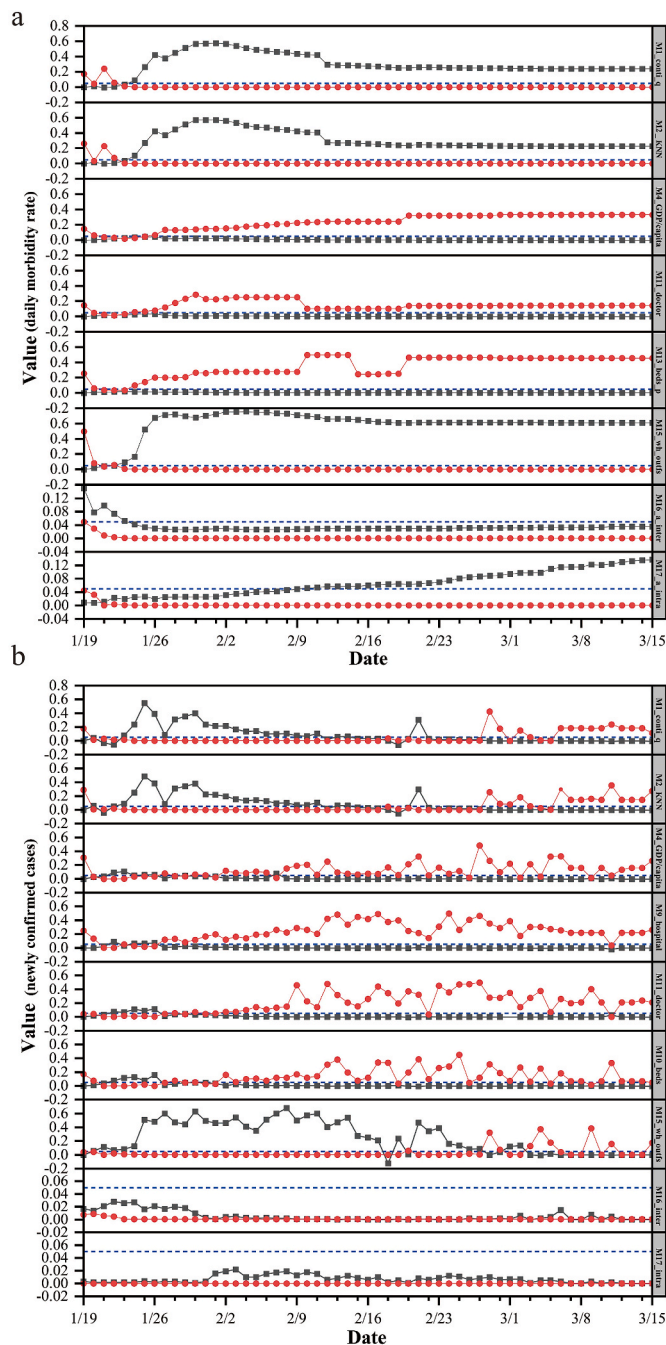
Fig. 4. Temporal change of Moran's I statistics for morbidity and newly confirmed cases under various assumptions. (a) Daily morbidity rate. (b) Daily newly confirmed cases. (Single factor matrix. Black line with squares records Moran's I value, red line with circles records the corresponding p-value, and dark blue dashed line is the 0.05 p-value line). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 5. Temporal change of Moran's I statistics for morbidity and newly confirmed cases under various assumptions. (a) Daily morbidity rate. (b) Daily newly confirmed cases. (Compound factor matrix. Black line with squares records Moran's I value, red line with circles records the corresponding p-value, and dark blue dashed line is the 0.05 p-value line). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

beds, a significant Moran's I did not exceed 0.13. For inter- and intra-city mobility, a significant Moran's I did not exceed 0.03.

Fig. 5 shows the changing Moran's Is with the compound weight matrix. Fig. 5(a) indicates that combining distance with population density, green space, hospital beds per 1000 persons and mobility generated a significant Moran's I for morbidity. It shows similar patterns across factors, with Moran's I being significant since around January 24 and remaining at some 0.8 from Jan 28. Compared with Fig. 4(a), using a compound weight matrix, the index of GDP per capita, hospitals, beds
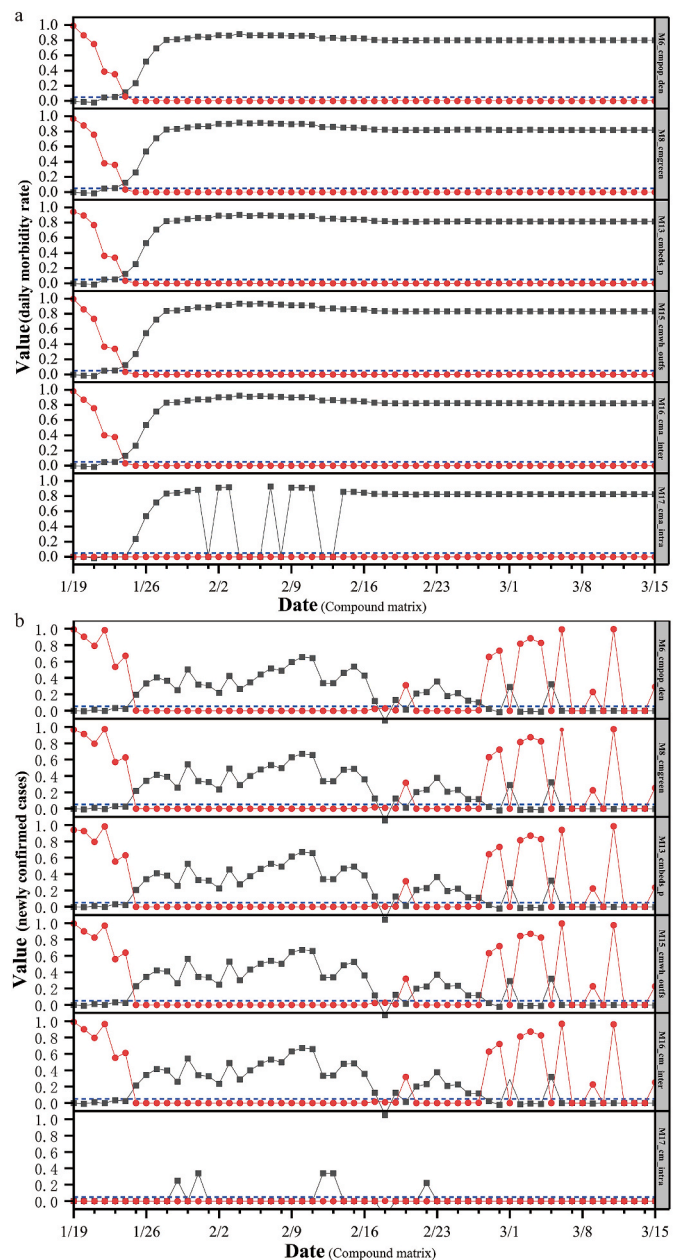
and doctors was no longer significant; the index of population density, green space and hospital beds per 1000 persons - became significant; and the index of mobility stayed significant.

Fig. 5(b) lists the significant Moran's I for newly confirmed cases when a compound matrix between distance and population density, green space, beds per 1000 persons, and mobility were used. Except for intracity mobility, the patterns of other factors were similar, with a significant nonzero Moran's I from the end of week 1 to week 6 (−0.02-0.88). This pattern, although of different scales, was similar to that of contiguity and KNN in Fig. 4(b). Comparing Fig. 5(b) with Fig. 4(b), GDP per capita and doctors were no longer significant, population density and green space became significant, and beds per 1000 persons and

mobility stayed significant. Moran's Is for intracity mobility were significant for only a couple of days.

Except for the lower values when the distance band and inverse distance band were used (which might be caused by their zero threshold to define neighbors), temporal patterns of Moran's I for both morbidity and newly confirmed cases using various spatial weights were similar (Supplementary Figure 3), indicating the robust role of proximity.

3.3. Changing temporal evolution of local spatial autocorrelation under various assumptions

Fig. 6 maps the spatiotemporal conditional clusters for morbidity. Fig. 6(a) depicts the distribution of clusters across cities divided by the median value of GDP per capita when Moran's I turned from not significant in early week 1 into steadily significant from week 2, when queen contiguity was used. The substantial increase in both high-high and low-low clusters is visible in Fig. 6(a2). High-high clusters were mainly situated in Central China, and low-low in Western and Northeastern China. Fig. 6(b) maps clusters defined by GDP per capita in
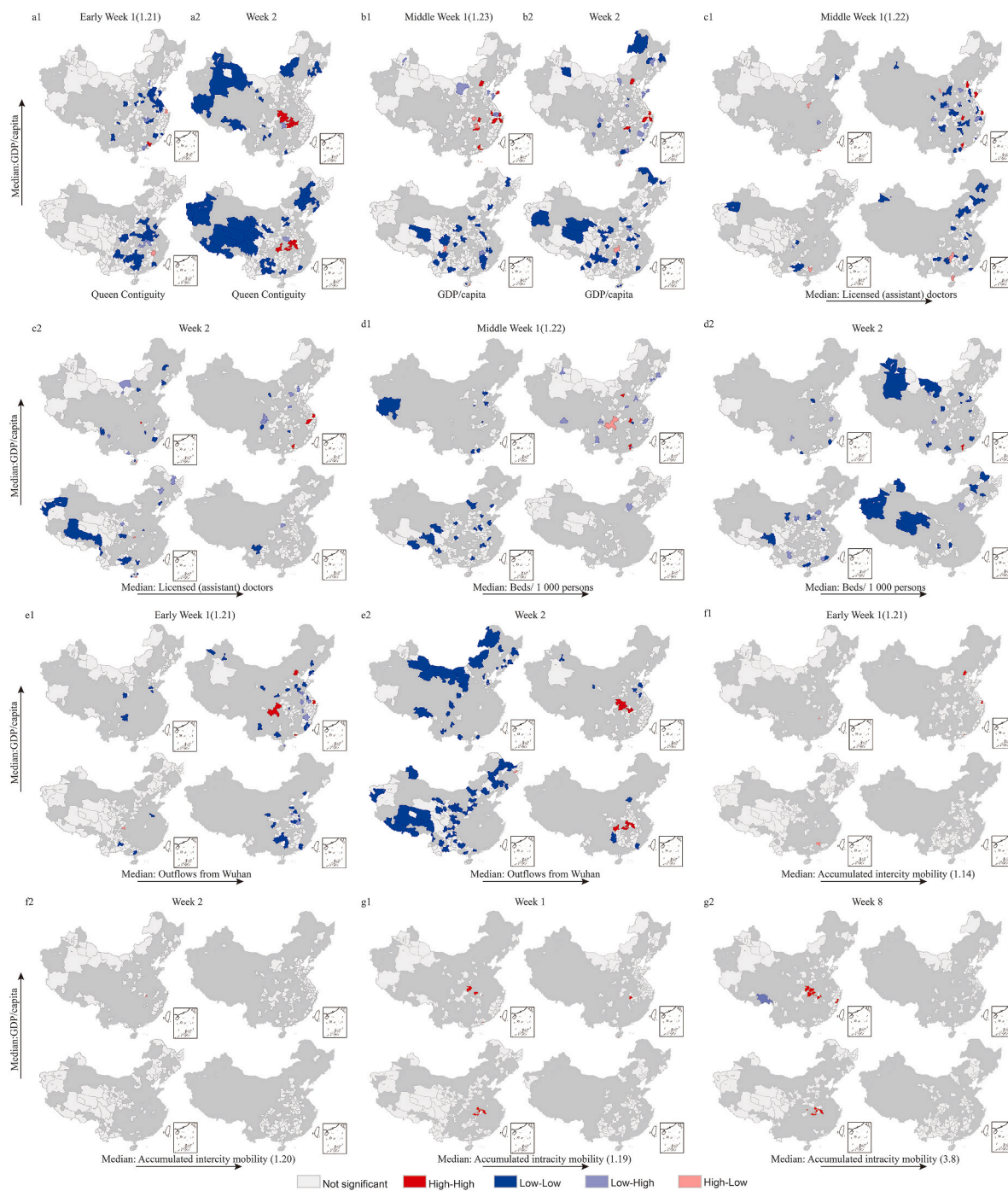

Fig. 6. Clusters of morbidity under various assumptions. (a) Queen based contiguity. (b) GDP/capita. (c) Licensed (assistant) doctors. (d) Beds per 1000 people. (e) Outflows from Wuhan. (f) Accumulated inter-city mobility. (g) Accumulated intra-city mobility. Note: Cities at the right-hand side of the arrow have indicators with median or higher values, cities at the left-hand side of the arrow have indicators with lower than median values.

middle week 1 and week 2, when Moran's I turned into not significant. There was a decrease in high-high clusters in Fig. 6(b2). High-high clusters were mainly cities with higher GDP per capita. Fig. 6(c) maps clusters defined by doctors. In week 1, some high-high and low-low clusters were cities with more licensed (assistant) doctors. In week 2, a few high-high clusters remained in cities with both higher GDP and more doctors. Fig. 6(d) depicts clusters defined by hospital beds per 1000 people from middle week 1 to week 2. A couple of high-high clusters situated in cities with higher GDP and more hospital beds per 1000 people in Fig. 6(d1) were generally no longer significant in Fig. 6 (d2). Fig. 6(e) maps clusters defined by Wuhan outflows. There was a clear increase in both high-high and low-low clusters from early week 1 to week 2. High-high clusters were mainly cities in Central China with more outflows. Fig. 6(f) maps clusters defined by accumulated intercity mobility. A few high-high clusters in higher GDP and intercity mobility cities (Fig. 6(f1)) no longer existed in week 2 (Fig. 6(f2)). Clusters defined by accumulated intracity mobility are mapped in Fig. 6(g). Although Moran's I became significant, some high-high clusters were found in cities with lower intracity mobility in both weeks.

Fig. 7 depicts clusters of newly confirmed cases. Fig. 7(a) maps clusters in weeks 1 and 6 using queen-based contiguity. The substantial decrease in clusters in week 6 is visible. High-high clusters were mainly in Central China. Fig. 7(b) maps clusters in weeks 1 and 2 using GDP per capita. In week 1, there were some high-high clusters situated primarily in higher GDP cities. In week 2, there left a few high-high clusters mainly in higher GDP cities. Fig. 7(c) maps clusters with hospitals used in weeks 1 and 2. High-high clusters were mainly cities with higher GDP and more hospitals, and low-low clusters were mainly cities with fewer hospitals in week 1 (Fig. 7(c1)). The number of clusters shrank considerably in week 2. Fig. 7(d) shows the distribution of clusters in weeks 1 and 3 with doctors being used. High-high clusters were mainly cities with more doctors and higher GDP (Fig. 7(d1)), and the number of high-high clusters dropped considerably in week 3 (Fig. 7(d2)). Fig. 7(e) maps clusters in weeks 1 and 3 with hospital beds. High-high clusters were mainly cities with higher GDP and more beds, and low-low clusters were mainly cities with fewer beds. Fig. 7(f) maps clusters in weeks 1 and 6 when Wuhan outflows were used. High-high clusters were mainly cities with more outflows, and low-low clusters were mainly cities with fewer. Fig. 7(g) depicts clusters with intercity mobility. High-higher clusters were mainly cities with higher GDP and intercity mobility (Fig. 7(g1)). It is interesting to note that high-high clusters were mainly cities with higher intracity mobility in week 1 when intracity mobility was used (Fig. 7(h1)). In week 3, when Moran's I turned significant, high-high clusters were mainly cities with lower intracity mobility (Fig. 7(h2)).

Clusters and conditional clusters with a compound weight matrix are not presented in detail because of word limits and similar patterns across factors, with cities from Central China being high-high clusters (see Supplementary Figures 4 to 7 for details).

4. Discussion

Although investigating spatial epidemic dynamics is crucial in understanding the routine of spatial diffusion and in surveillance, prediction, identification and prevention of another outbreak (Wang et al., 2006), studies on the spatial analysis of COVID-19 transmission are still lagging. Using spatial association as the primary strategy, this study examined the spatial linkage of COVID-19 under various diffusion assumptions. Moreover, the geographical distributions and characteristics of city clusters in various epidemic stages were also explored.

We found significant Moran's Is for morbidity and newly confirmed cases in week 1 (to January 26) and/or 2 (to February 2) when GDP per capita and indicators of medical resources were risk factors. Given the gradual establishment of countermeasures from January 23 to February 2 and their lagged effects (Wang et al., 2020), this indicates the importance of GDP per capita and medical resources in diffusion before countermeasures take effect. Compared to morbidity, the spatial

association of newly confirmed cases seems to be more sensitive with a longer period and higher values. These differences are understandable, given that as a direct measure of population, changes in newly confirmed cases are more flexible and sensitive. GDP per capita as a factor at an early stage is understandable, since development is normally correlated with timely confirmation (Gardner et al., 2018). The important role of medical resources at an early stage echoes the findings of Meng et al. (2005) and Kanga et al. (2020), which argued the importance of doctors and hospitals in impacting the spread of SARS in Beijing and the spread of COVID-19 across provinces at an early stage. Before effective countermeasures were implemented, the exponential growth of infections and social panic often led to overloaded hospitals, and numerous hospital-acquired infections (Zhou, 2020), making hospitals the primary place of spreading and doctors a high-risk group (Wang et al., 2020).

We found a significant Moran's I for morbidity since early week 1 and a significant Moran's I for newly confirmed cases since January 20, when contiguity, distance and Wuhan outflows were used. This finding suggests the key role of these factors during the whole diffusion process. Moran's I valued higher when outflows were used, compared to that of contiguity and distance, implying its dominant role. This finding echoes previous research (Li et al., 2020), which argued that before Wuhan's lockdown, the number of reported confirmed cases counted merely 14% of the total confirmed cases; and that the majority of infections were transmitted through this group of people. Moreover, this finding also echoes previous studies on environmental factors of COVID-19 (Liu et al., 2021; Mu et al., 2020), which argued that variations between cities are mainly driven by Wuhan outflows. Our finding expands the prior work of Kanga et al. (2020), Liu et al. (2021) and Mu et al. (2020) by analyzing the whole epidemic period and by finding that the role of Wuhan outflows decreased since week 4 and was no longer significant since week 6. In line with previous research (Han et al., 2021; Kanga et al., 2020; Meng et al., 2005), the significant role of contiguity and distance in the study indicated the importance of proximity in epidemic diffusion. Our findings also extend prior research by indicating that, although the role of proximity is always important to morbidity, for newly confirmed cases under effective countermeasures, it may soon attenuate and gradually no longer be significant.

We found a rapid increase in cities with reported cases before week 1 (January 26) and a steadily rapid rise in both morbidity and newly confirmed cases from weeks 2–4. We also found a significant role of intercity mobility until the middle of week 2 (around January 30). Combining the role of intracity mobility, which was significant since late in week 2, as well as the roughly 7-day lag from infection to diagnosis in China since late January (WHO, 2020), we inferred a primary intercity diffusion (transfer transmission) stage before January 23, and a primary intracity diffusion (local or community transmission) stage later. Given the implementation of countermeasures since January 23 and the time lag, the significant role of intercity mobility until late January is understandable. This finding is in general agreement with that of Mu et al. (2020), which declared a major transfer diffusion before Chinese New Year (January 24) and a subsequent local diffusion thereafter. Given that previous studies on mobility and COVID-19 are limited and focused mainly on accumulated data on Wuhan outflows, this finding broadens our understanding of the spatial diffusion of COVID-19 by empirically illustrating its transmission stage with various categories of real-time mobility data.

Our compound weight matrix combining spatial distance with social distance indicated partially different findings from that of the single weight matrix. Although the significant role of hospital beds per 1000 persons and mobility remained, the role of population density and green space became significant, and GDP per capita, doctors and hospitals were no longer significant. Searching for literature, previous studies comparing the results of spatial analyses with respectively the compound weight matrix and single weight matrix are rare. On reflection, we speculate that the geographical pattern of potential risk factors may
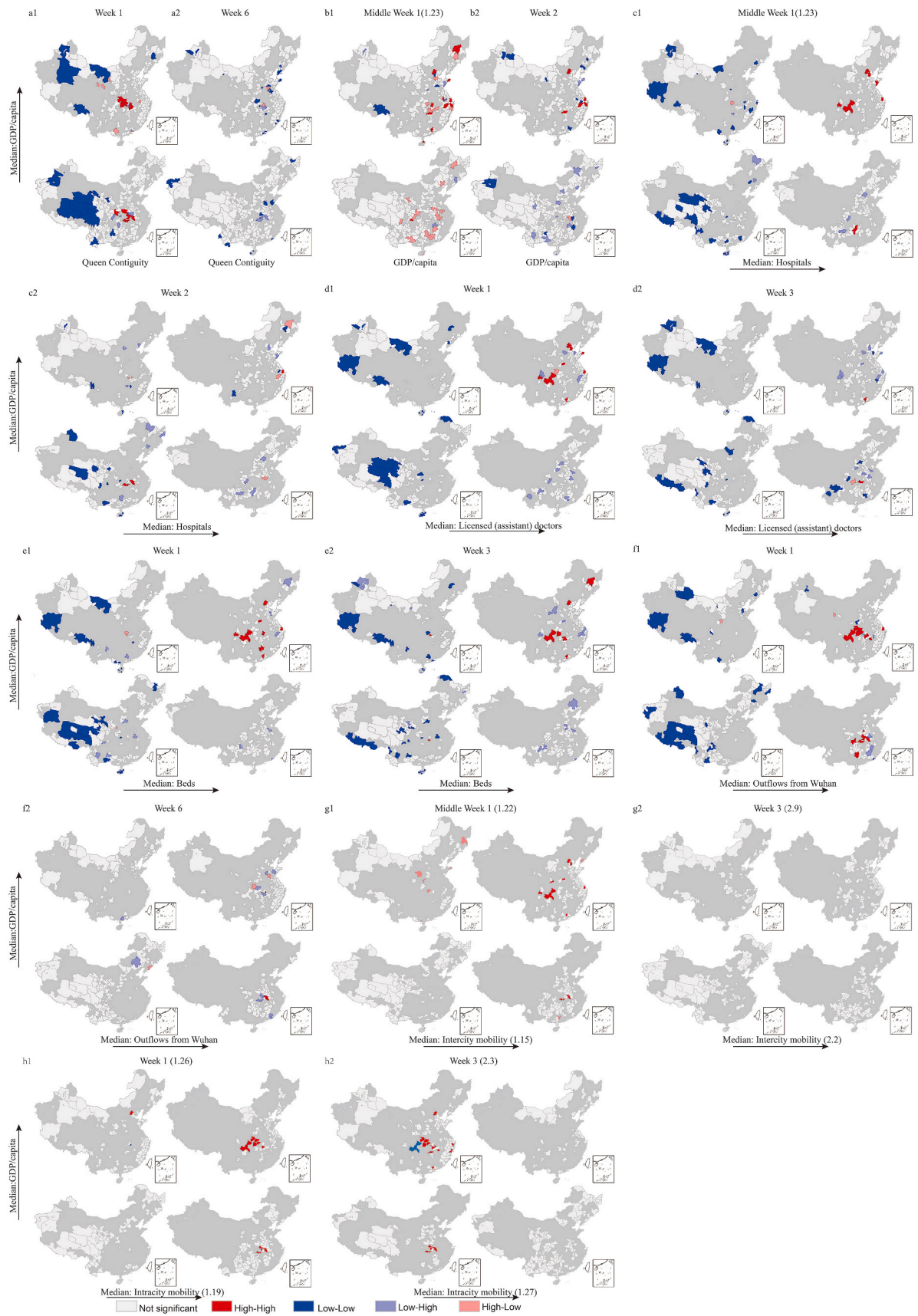
Fig. 7. Clusters of newly confirmed cases under various assumptions. (a) Queen_based contiguity. (b) GDP/capita. (c) Hospitals. (d) Licensed (assistant) doctors. (e) Beds. (f) Outflows from Wuhan. (g) Inter-city mobility. (h) Intra-city mobility. Note: Cities at the right-hand side of the arrow have indicators with median or higher values, cities at the left-hand side of the arrow have indicators with lower than median values.

provide some reasonable explanations. In both Figs. 6 and 7, there was a scattered distribution of clusters with high-high clusters surrounded by other types when GDP per capita, hospitals and doctors were used. After combining these factors with spatial distance, the weights of those formally scattered high-value neighbors with long spatial distance decreased, and the weights of their nearby low-value neighbors increased. Therefore, with the compound weight matrix, the clusters disappeared and the roles of GDP per capita, hospitals and doctors were no longer significant. In comparison, the distribution of clusters when mobility and beds per 1000 persons were used were more clustered. Hence, the relative weights of neighbors should not change much with the consideration of spatial distance, and their significant role stayed. The significant role of population density and green space in the compound matrix could be explained similarly. Our results indicated that although a compound matrix could highlight the importance of spatial lagging by improving the value of Moran's I and facilitate the next-step analysis (such as spatial regression models adopted in Fingleton and Le Gallo (2008) and Sun et al. (2014)), it might also suppress the impact of some factors that were significant in a single matrix.

We failed to find any significant role of population and household size, which is similar to Mu et al. (2020). Our findings are different from COVID-19 studies in the US (Hu et al., 2020; Lee et al., 2021) and previous studies on other infectious diseases (Dalziel et al., 2018; Meng et al., 2005; Tian et al., 2018), which underscores the important role of household size and population scale. On reflection, we infer that the instant intervention measures taken in China have minimized the role of these traditional factors. Searching for literature, the substantial impact of China's suppression measures, such as the lockdown of Wuhan, strict movement restrictions, and mitigation measures, such as social distancing, in reducing case importations and local transmission is widely supported (Ferguson et al., 2020; Zhang et al., 2020). Different from previous studies that argued for the significant role of population density and green space in explaining COVID-19 variations (Han et al., 2021; Lee et al., 2021), population density and green space were significant risk factors only when combined with spatial distance. We believe that spatial distance might play a moderating role in influencing the effect of population density and green space. More in-depth studies are required to compare single and compound weight matrices in spatial analysis.

We found a change in the geographical distribution of high-risk areas. At the early stage and before interventions took effect, high-risk areas were mainly cities adjacent to the epicenter, or with higher GDP per capita, or with a combination of higher GDP per capita and more medical resources, or more outflows from the epicenter, or more intercity mobilities. After intervention measures were effected, cities adjacent to the epicenter, cities with more outflows from the epicenter and cities with more intracity mobility were high-risk areas. Throughout the whole epidemic period, cities adjacent to Wuhan and cities with more outflows from Wuhan, which were mainly cities situated in Central China, were high-risk areas. This indicates Wuhan to be the only epicenter. By further exploring both the geographical distribution and characteristics of high-risk cities throughout the whole process, our findings have expanded current research cross-sectionally focusing on factors of transmission in a given period.

Although understanding the spatial diffusion process of infectious diseases is vital in supporting decision-making pertaining to epidemic control, previous studies attempting to evaluate the spatial spread of COVID-19 during a whole transmission period are limited. As the first country detected with confirmed cases, China has managed to control the epidemic rapidly and effectively (Burki, 2020) and thus provides a good example to study the spatial diffusion of COVID-19. Using city as the study unit, this study has managed to explore the temporal change in primary risk factors influencing the spatial diffusion of COVID-19 under various diffusion assumptions and throughout the whole epidemic period. Moreover, this study has investigated both the characteristics and geographical distributions of high-risk areas in different epidemic stages, which enables more targeted intervention measures with the evolution of the epidemic process.

Some limitations need to be mentioned. First, because various characteristics of a country, such as the speed of response, the health system, the epidemic curve, the coordination between government sectors, and civil compliance with regulations, may impact the effectiveness of countermeasures (Burki, 2020), the spatial process and spatial association of this outbreak across the Chinese city system may not generalize to other settings. Therefore, the findings of this study merely provide a reference, or the possibility of a potential pattern, under the promise that there is timely implementation of both suppression and mitigation countermeasures at an early stage. Second, although spatial association analyses enable us to understand the temporal evolution of spatial diffusion under various assumptions, neither the potentially lagged effect of previously confirmed cases nor interactions between possible risk factors were acquired. An innovative methodology combining spatiotemporal analysis and prospective and interactive investigation is therefore urgently needed. Third, similar to the strategy that many extant studies adopted (Liu et al., 2021; Mu et al., 2020), officially reported data are the most available data source we can use in analyzing the spatiotemporal diffusion of COVID-19 and potential factors. However, because of limited case detection capacities at the early stage (Niehus et al., 2020), and the fact that asymptomatic cases and patients with very mild symptoms might not be identified (Baud et al., 2020), we are aware that official data closely pertain to clinically apparent and confirmed cases. Therefore, as indicated by the literature (Imai et al., 2020), an underestimation of prevalence is highly possible. Finally, as an ecological study at the city level, although our application of Baidu microphone-based mobility data that aggregated individual travel data could provide very useful information in understanding the role of mobility in disease diffusion, there is substantial heterogeneity in individual travel behaviors across socioeconomic groups and possibly also disparities in disease exposure and transmission rates (Brough et al., 2021). Combining our ecological study with future individualistic studies linking COVID-19 outcomes to individual travel behavior will provide a more insightful understanding of disease diffusion.

5. Conclusion

This study seeks to arrive at an in-depth understanding of the spatial diffusion pattern of COVID-19 across Chinese cities and throughout the entire epidemic period. We found a rapid intercity diffusion process at the early stage and a primarily intracity diffusion process thereafter. Before countermeasures took effect, GDP per capita, medical resources, and intercity mobility significantly impacted the early diffusion process. With speedily effective countermeasures, intracity mobility played an important role. The roles of proximity and outflows from Wuhan were important throughout the entire stage. At the early stage, high-risk areas were mainly cities adjacent to the epicenter, or with higher GDP per capita, or a combination of higher GDP per capita and better medical resources, or more outflows from the epicenter, or more intercity mobility. After intervention measures took effect, cities adjacent to the epicenter, cities having more outflows from the epicenter or cities with more intracity mobility were high-risk areas. This study provides valuable insights into understanding the spatial diffusion patterns of COVID-19 across the city system. The findings are informative in effectively handling the recurrence of COVID-19 in China and abroad.

Credit authorship contribution

Lijuan Gu: Conceptualization, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **Linsheng Yang:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Writing - review & editing. **Li Wang:** Formal analysis, Methodology, Software. **Yanan Guo:** Resources, Writing - review& editing. **Binggan Wei:** Resources, Writing - review &editing.

Hairong Li: Resources, Writing - review &editing.

Ethical approval

This is an analysis of public data so ethics approval is not required.

Declaration of competing interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2022.114988.

References

Allen, J.G., Marr, L.C., 2020. Recognizing and controlling airborne transmission of SARS-CoV-2 in indoor environments. Indoor Air 30, 557–558.

Altmann, D.M., Boyton, R.J., 2022. COVID-19 vaccination: the road ahead. Science 375 (80), 1127–1132. https://doi.org/10.1126/science.abn1755.

Anselin, L., 2018. In: Maps for Rates or Proportions. https://geodacenter.github.io/workbook/3b_rates/lab3b.html#empirical-bayes-eb-smoothed-rate-map.

Anselin, L., 2020. In: Global Spatial Autocorrelation: Visualizing Spatial Autocorrelation. https://geodacenter.github.io/workbook/5a_global_auto/lab5a.html.

Anselin, L., Bera, A.K., 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Handbook of Applied Economic Statistics. CRC Press, Boca Raton.

Anselin, L., Lozano, N., Koschinsky, J., 2006. Rate Transformations and Smoothing.

Assunção, R.M., Reis, E.A., 1999. A new proposal to adjust Moran'sI for population density. Stat. Med. 18, 2147–2162.

Bahl, R., Eikmeier, N., Fraser, A., et al., 2021. Modeling COVID-19 spread in small colleges. PLoS One 16.

Balcan, D., Colizza, V., Gonçalves, B., et al., 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc. Natl. Acad. Sci. U.S.A. 106, 21484–21489.

Basov, V., 2022. In: St Barbara Reports COVID-19 Outbreak at Simberi, Withdraws Annual Guidance. https://www.kitco.com/news/2022-02-18/St-Barbara-reports-COVID-19-outbreak-at-Simberi-withdraws-annual-guidance.html.

Baud, D., Qi, X., Nielsen-Saines, K., et al., 2020. Real estimates of mortality following COVID-19 infection. Lancet Infect. Dis. 20, 773.

Brough, R., Freedman, M., Phillips, D., 2021. Understanding socioeconomic disparities in travel behavior during the COVID-19 pandemic. J. Reg. Sci. 61, 753–774.

Burki, T., 2020. China's successful control of COVID-19. Lancet Infect. Dis. 20, 1240–1241.

Coşkun, H., Yıldırım, N., Gündüz, S., 2021. The spread of COVID-19 virus through population density and wind in Turkey cities. Sci. Total Environ. 751.

Dalziel, B.D., Kissler, S., Gog, J.R., et al., 2018. Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. Science (80-.) 362, 75–79.

Department of Rural Surveys, 2019. China Statistical Yearbook (County-Level). China Statistics Press, Beijing, China.

Department of Urban Surveys, 2019. China City Statistical Yearbook (2019). China Statistics Press, Beijing, China.

Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis.

Ferguson, N., Laydon, D., Nedjati-Gilani, G., et al., 2020. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imp. Coll. COVID-19 Response Team.

Fingleton, B., Le Gallo, J., 2008. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties. Pap. Reg. Sci. 87, 319–339.

Fox News, 2022. In: Shanghai Extends COVID-19 Lockdown, Outbreak Remains "Extremely Grim". https://www.foxnews.com/world/shanghai-extends-covid-19-lockdown-outbreak-extremely-grim.

Gardner, L.M., Bóta, A., Gangavarapu, K., et al., 2018. Inferring the risk factors behind the geographical spread and transmission of Zika in the Americas. PLoS Neglected Trop. Dis. 12.

Gössling, S., Scott, D., Hall, C.M., 2020. Pandemics, tourism and global change: a rapid assessment of COVID-19. J. Sustain. Tourism 1–20.

Guan, W., Ni, Z., Hu, Yu, et al., 2020. Clinical characteristics of coronavirus disease 2019 in China. N. Engl. J. Med. 382, 1708–1720.

The Guardian, 2020. China Coronavirus: Mayor of Wuhan Admits Mistakes.

Han, Y., Yang, L., Jia, K., et al., 2021. Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. Sci. Total Environ. 761.

He, J., Chen, G., Jiang, Y., et al., 2020. Comparative infection modeling and control of COVID-19 transmission patterns in China, South Korea, Italy and Iran. Sci. Total Environ. 747.

Hu, M., Roberts, J.D., Azevedo, G.P., et al., 2020. The role of built and social environmental factors in Covid-19 transmission: a look at America's capital city. Sustain. Cities Soc. 65.

Huang, C., Wang, Y., Li, X., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395, 497–506.

Imai, N., Dorigatti, I., Cori, A., et al., 2020. Estimating the Potential Total Number of Novel Coronavirus (2019-nCoV) Cases in Wuhan City. China, London.

Juthani, P.V., Gupta, A., Borges, K.A., et al., 2021. Hospitalisation among vaccine breakthrough COVID-19 infections. Lancet Infect. Dis. 21, 1485–1486.

Kanga, D., Choia, H., Kimb, J., et al., 2020. Spatial epidemic dynamics of the COVID-19 outbreak in China. Int. J. Infect. Dis. 96–102.

Karimi, S.E., Ahmadi, S., Soleimanvandiazar, N., 2021. Inequities as a social determinant of health: responsibility in paying attention to the poor and vulnerable at risk of covid-19. J. Public health Res. 10, 1–2.

Kissler, S., Tedijanto, C., Goldstein, E., et al., 2020. Projecting the transmission dynamics of SARS-CoV-2 through the post-pandemic period. Science 80.

Lee, W., Kim, Honghyok, et al., 2021. Urban environments and COVID-19 in three Eastern states of the United States. Sci. Total Environ. 779.

Li, Q., Guan, X., Wu, P., et al., 2020. Early transmission dynamics in wuhan, China, of novel coronavirus–infected pneumonia. N. Engl. J. Med. 382, 1199–1207.

Li, R., Pei, S., Chen, B., et al., 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368 (80), 489–493.

Lin, Y., Liu, C., Chiu, Y., 2020. Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries. Brain Behav. Immun.

Liu, C., 2021. Chinese public's support for covid-19 surveillance in relation to the west. Surveill. Soc. 19, 89–93.

Liu, M., Liu, M., Li, Z., et al., 2021. The spatial clustering analysis of COVID-19 and its associated factors in mainland China at the prefecture level. Sci. Total Environ. 777.

Lopez Bernal, J., Andrews, N., Gower, C., et al., 2021. Effectiveness of covid-19 vaccines against the B.1.617.2 (delta) variant. N. Engl. J. Med. 385, 585–594.

Meng, B., Wang, J., Liu, J., et al., 2005. Understanding the spatial diffusion process of severe acute respiratory syndrome in Beijing. Publ. Health 119, 1080–1087.

Mu, X., Yeh, A.G.O., Zhang, X., 2020. The interplay of spatial spread of COVID-19 and human mobility in the urban system of China during the Chinese New Year. Environ. Plan. B Urban Anal. City Sci.

Muhareb, R., Giacaman, R., 2020. Tracking COVID-19 responsibly. Lancet.

National Health Commission, 2020. In: The Preventation of Imported Cases Has Become the Priority of the Epidemic Control in China (In Chinese). http://www.gov.cn/xinwen/gwylflkjz59/wzsl.htm.

Niehus, R., De Salazar, P.M., Taylor, A.R., et al., 2020. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. Lancet Infect. Dis. 20, 803–808.

Rahman, M., Islam, M., Shimanto, M.H., et al., 2021. A global analysis on the effect of temperature, socio-economic and environmental factors on the spread and mortality rate of the COVID-19 pandemic. Environ. Dev. Sustain. 23, 9352–9366.

Reuters, 2020. In: China Health Commission Official Says Controls Starting to Rein in Virus. https://www.reuters.com/article/idUSKBN20A05I.

Sun, C., Zhao, L., Zou, W., et al., 2014. Water resource utilization efficiency and spatial spillover effects in China. J. Geogr. Sci. 24, 771–788.

Sun, K., Chen, J., Viboud, C., 2020. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. Lancet Digit. Heal. 2, e201–e208.

Tian, H., Hu, S., Cazelles, B., et al., 2018. Urbanization prolongs hantavirus epidemics in cities. Proc. Natl. Acad. Sci. U.S.A. 115, 4707–4712.

Tian, H., Liu, Y., Li, Y., et al., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 368 (80), 638–642.

Wang, Y., Di, Q., 2020. Modifiable areal unit problem and environmental factors of COVID-19 outbreak. Sci. Total Environ. 740.

Wang, J., McMichael, A., Meng, B., et al., 2006. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. Bull. World Health Organ. 84, 965–968.

Wang, H., Wang, Z., Dong, Y., et al., 2020. Phase-adjusted estimation of the number of coronavirus disease 2019 cases in wuhan, China. Cell Discov 6.

WHO, 2020. In: Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19.

WHO, 2022. COVID-19 Weekly Epidemiological Update.

Wuhan Municipal Health commission, 2019. Urgent Notice on Treatment of Patients with Pneumonia of Unknown Cause.

Wuhan Municipal Health commission, 2020. Information Circular on Viral Pneumonia of Unknown Cause.

XinhuaNet, 2020. China Says its COVID-19 Peak Is over. http://www.xinhuanet.com/english/2020-03/12/c_138870547.htm.

Zhang, J., Litvinova, M., Wang, W., et al., 2020. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive

and modelling study. Lancet Infect. Dis. 20, 793–802. https://doi.org/10.1016/S1473-3099(20)30230-9.

Zhao, S., Sha, T., Wu, C.-I., et al., 2021. In: Will the Large-Scale Vaccination Succeed in Containing the COVID-19 Pandemic and How Soon?. https://ssrn.com/abstract=3826205.

Zhou, M., 2020. In: COVID-19: Why Is Medical System in Metropolises So Vulnerable?. http://www.china.org.cn/opinion/2020-04-21/content_75957964.htm.