SCIENCE CHINA
Earth Sciences

• RESEARCH PAPER •

July 2010 Vol.53 No.7: 1017–1028
doi: 10.1007/s11430-010-0043-x

Spatio-temporal evolution of Beijing 2003 SARS epidemic

CAO ZhiDong[1], ZENG DaJun[1*], ZHENG XiaoLong[1], WANG QuanYi[2], WANG FeiYue[1], WANG JinFeng[3] & WANG XiaoLi[2]

[1] *Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;*
[2] *Beijing Center for Disease Control and Prevention, Beijing 100013, China;*
[3] *State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*

Received May 19, 2009; accepted October 28, 2009; published online May 12, 2010

Studying spatio-temporal evolution of epidemics can uncover important aspects of interaction among people, infectious diseases, and the environment, providing useful insights and modeling support to facilitate public health response and possibly prevention measures. This paper presents an empirical spatio-temporal analysis of epidemiological data concerning 2321 SARS-infected patients in Beijing in 2003. We mapped the SARS morbidity data with the spatial data resolution at the level of street and township. Two smoothing methods, Bayesian adjustment and spatial smoothing, were applied to identify the spatial risks and spatial transmission trends. Furthermore, we explored various spatial patterns and spatio-temporal evolution of Beijing 2003 SARS epidemic using spatial statistics such as Moran's I and LISA. Part of this study is targeted at evaluating the effectiveness of public health control measures implemented during the SARS epidemic. The main findings are as follows. (1) The diffusion speed of SARS in the northwest-southeast direction is weaker than that in northeast-southwest direction. (2) SARS's spread risk is positively spatially associated and the strength of this spatial association has experienced changes from weak to strong and then back to weak during the lifetime of the Beijing SARS epidemic. (3) Two spatial clusters of disease cases are identified: one in the city center and the other in the eastern suburban area. These two clusters followed different evolutionary paths but interacted with each other as well. (4) Although the government missed the opportunity to contain the early outbreak of SARS in March 2003, the response strategies implemented after the mid of April were effective. These response measures not only controlled the growth of the disease cases, but also mitigated the spatial diffusion.

severe acute respiratory syndrome (SARS), Beijing, morbidity rate, spatial analysis, spatio-temporal evolution, control measures

Citation: Cao Z D, Zeng D J, Zheng X L, et al. Spatio-temporal evolution of Beijing 2003 SARS epidemic. Sci China Earth Sci, 2010, 53: 1017−1028, doi: 10.1007/s11430-010-0043-x

Severe acute respiratory syndrome (SARS) is a respiratory disease with a high case-fatality rate, which is thought to be transmitted primarily by close person-to-person contact. Transmission occurs through respiratory droplets produced when an infected person sneezes and/or by touching a surface or object contaminated with infectious droplets. Many natural and social-economic factors contributed to and influenced SARS transmissions. To gain a better understanding of the SARS transmission mechanism and help design and assess related public health response and control measures, public health researchers and practitioners need to carefully and systematically study spatial and temporal evolution of SARS cases, identify risk factors, and evaluate the specific impact of control measures implemented.

The 2003 SARS outbreaks represented one of the most serious public health challenges to China and the world.

*Corresponding author (email: dajun.zeng@ia.ac.cn)

Over the past six years, many scientists around the world have conducted related research in pathology [1], molecular biology [2], clinical medicine [3], epidemiology [4], health economics [5], and emergency management [6]. These studies have produced many important and relevant research results; however, many mysteries of academic interest and practical relevance remain to be further explored.

During the 2003 global SARS crisis, more than half of the SARS infection cases occurred in mainland China. Among Chinese cities, Beijing was hit most severely. SARS infection cases in Beijing accounted for about 30% of the total number of global infections. With an effective treatment or vaccine absent, public health response measures were restricted to case management (e.g., isolating known cases and quarantining close contacts of the patients) and proactive monitoring of identified high-risk groups. As such, from a practical standpoint, it is critical to reveal the spatio-temporal evolutionary patterns of the epidemic and have the capability to quantitatively evaluating the impact of response and control measures taken. Beijing, a major global center of political, economic, and educational and research activities, with a population close to seventeen million, is vulnerable to infectious disease outbreaks such as SARS and influenza A. Our study focuses on 2003 SARS epidemic in Beijing because of both data availability as well as its impact, importance, and representativeness.

Many Chinese researchers have studied Beijing 2003 SARS epidemic from a wide array of perspectives, ranging from descriptive epidemiology [7–9], mathematical modeling for sequential process [10–13], simulation calculation [14] and prediction [15, 16], analysis of impact factors [17, 18] and the impact of SARS epidemic on the national economy [19, 20]. However, based on our literature review, only a handful of studies have touched upon spatio-temporal data analysis, mostly with aggregated data. Wang explored the multidimensional spatial distribution, spatial cluster and spatial correlation based on tens of thousands of close contact incidences [21–24]. Liu et al. [25] distinguished the spatial patterns of the Beijing SARS epidemic between urban and suburban areas. These studies have resulted in many interesting findings. However, the central questions concerning SARS spreading patterns and evaluation of government control measures remain largely unexplored. (1) Close contact incidences cannot substitute for SARS transmission incidences for analysis purposes as the exposure between an infective individual and a susceptible individual needs to be examined through confirmed infection cases. (2) It is crucial to perform spatio-temporal analysis using fine-granularity resolutions such as streets and townships. (3) The existing work has focused on either temporal or spatial analysis, whereas an integrated understanding through space and time is critically needed. (4) The existing work does not support evaluation of public health response and control measures in terms of spatio-temporal evolution of disease cases and disease transmission mechanisms.

In this paper, we report an empirical study of Beijing 2003 SARS epidemiological data from the point of view of spatio-temporal analysis. This works aimes at illustrating the usefulness of spatio-temporal data analysis techniques in public health informatics practice. The empirical findings and policy evaluation work from our study also provide actionable insights that could help prepare for and respond to possible reemergence of the SARS epidemic or other novel epidemics of acute infectious diseases.

The rest of this paper is structured as follows. We begin with a brief introduction to the Beijing 2003 SARS dataset used in our work. Using this dataset, we report an empirical analysis of spatial variation, spatial correlation and clustering patterns, and spatio-temporal evolution of the Beijing 2003 SARS epidemic. We then discuss the inherent SARS transmission mechanisms derived from our spatio-temporal data analysis and evaluate the impact of various government intervention policies during the SARS epidemic.

1 Beijing 2003 SARS dataset

1.1 Survey data

The data used in our study were extracted from a survey of 2444 confirmed SARS patients in Beijing conducted by the Beijing Centers for Disease Control and Prevention, covering the period from March 1, 2003 to May 28, 2003. The questionnaire-based survey covered each patient's detail information, such as gender, age, home address, work address, onset of symptoms, and so on. Figure 1 shows a plot of daily counts of new SARS cases during the Beijing 2003 SARS epidemic. The plot shows clear three stages: slowly increasing during the early stage of the outbreak, sharply increasing during April and rapidly declining during the last period. The epidemic reaches its peak during mid-April.

1.2 Spatio-temporal data elements

Survey data do not support spatio-temporal analysis because it presents time and space information as text format. As such, we need to convert them to spatio-temporal data of GIS format. To analyze the spatial patterns and understand the complexity of spatial transmission risks, we have

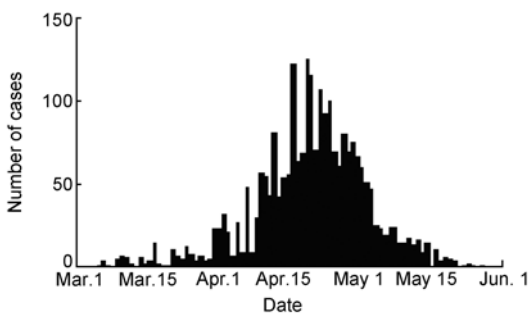

Figure 1 Daily reporting of new SARS cases during the Beijing 2003 SARS epidemic.

adopted home addresses as spatial locations of cases as homes have been widely recognized as one of the highest risk places for SARS transmission. Another reason to choose home address is that it is the most complete data what spatial information we can obtain from the survey data. Figure 2(a) maps the spatial locations of 2321 SARS cases studied. The remaining 123 SARS cases were abandoned because of incomplete, wrong or unavailable spatial information. It is easy to see that most SARS patients aggregated in the urban area.

Table 1 summarizes the rules for spatial location processing. According to these rules, we obtained 2321 available spatial points (out of 2444 cases), where each point corresponds to a SARS case. From Figure 2(a), we see that only a small number of SARS patients lived in outskirts, which represents a potential source of uncertainty to spatio-temporal analysis. Thus for our analysis we focused on these cases located in the urban areas (Figure 2(b)), which covers all regions within 5th Ring Road and includes 2227 cases, accounting for 91.10% of the total number of all reported SARS cases during Beijing 2003 SARS epidemic.

2 Spatial analysis of morbidity

2.1 Morbidity map

The morbidity rate refers to the frequency of new infected cases during a given time interval, which equals the number of new cases occurred during a given time interval divided by the population in the same period.

Figure 3(a) shows the morbidity map connected to the Beijing 2003 SARS epidemic with the spatial data analysis unit set to the street and township level. The area of the smallest spatial unit is about 1.03 km^2 and the largest 163.00 km^2. The number of SARS cases within each unit was obtained by the standard GIS overlay analysis. Overlay analysis combines information from the point layer of SARS cases with the polygon layer of the underlying spatial statistical units to determine whether a given case point belongs to a certain spatial unit. The population data for each unit come from a 1% population sample survey conducted in 2005. The morbidity map shows intuitively that the risks of SARS infection in areas close to the city center are higher than those in the city's peripheral areas. Also, there

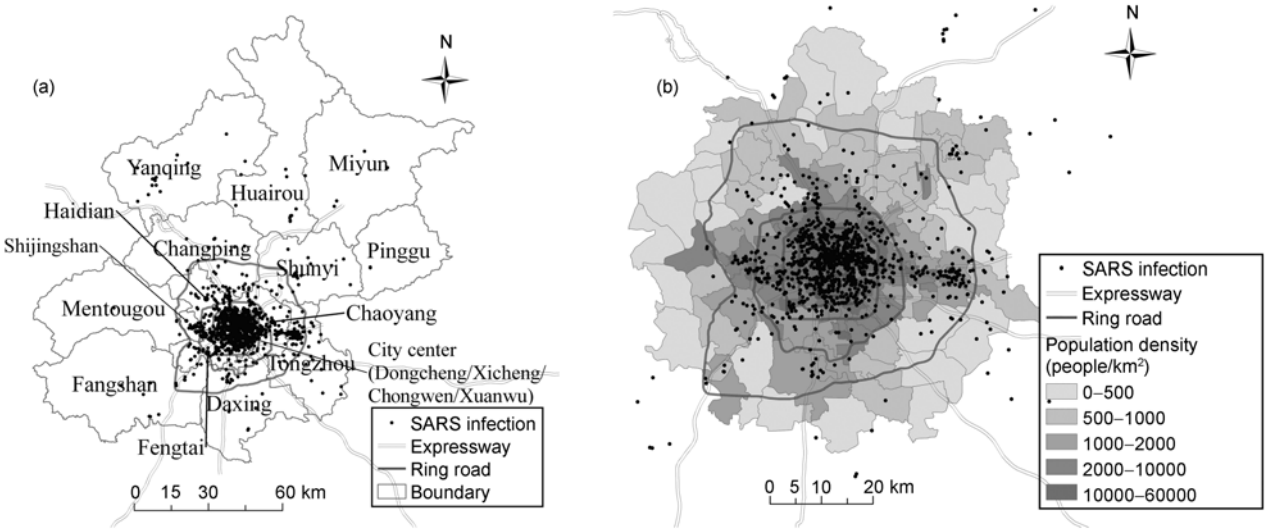

Figure 2 Spatial distribution of all Beijing SARS cases. (a) All 2321 SARS cases based on home address; (b) All 2227 SARS cases located in the urban areas.

Table 1 Spatial location processing [a)]

Is Beijing resident?	Home address	Work address	Spatial location	Resulting statistic	
				Number of cases	Ratio (%)
Yes	√	√	Home address	2158	88.30
	√	×	Home address	0	0
	×	√	Work address	163	6.67
	×	×	Unavailable	0	0
No	√/×	√/×	Unavailable	123	5.03

a) √ stands for complete information; × stands for incomplete, wrong or unavailable spatial information.

might be some relationship between the morbidity rate and these circle roads.

2.2 Bayesian adjustments

Raw morbidity rates might not effectively reflect true risks. The morbidity rates in spatial units with high concentration of population are more reliable indicators due to less randomness. Figure 4(a) illustrates that variations in disease cases are higher in units with smaller population density. However, because of the way these spatial units are partitioned administratively by the government, a unit with higher population density tends to be a smaller area, and vice versa. As a result, the map based on raw morbidity rates can lead to false impression (e.g., some areas of high risk close to the city center could be overlooked).

To solve this problem, we consider the raw, observed morbidity rates as realizations of a non-observable random process and propose to re-estimate morbidity rates using Bayesian adjustments in an effort to uncover true risks with an exposed population.

The average morbidity rate during the Beijing 2003 SARS epidemic was 21.7 infections per 100000 persons over the entire period. As pointed out earlier, in Figure 4(a), we observe that the units with greater population density are associated with rates that are not only closer to the average morbidity, but also with smaller variation. As the size of population decreases, the fluctuation in the raw morbidity rates increases. We also found that the SARS morbidity rates in some units were zeros, which is clearly not a true representation of the potential risk to SARS disease transmission. The Bayesian adjustment technique has been widely applied to solve these problems. Under the Bayesian framework, it is assumed that the rates in neighboring spatial units are correlated, and for the units with low population density, the rates from their neighbors can be used to help establish a more realistic rate.

Figure 3(b) shows the morbidity map after Bayesian adjustments. The Bayesian-adjusted morbidity rates as a function of the raw morbidity rates are plotted in Figure 4(b). An obvious change is that the morbidity rates in the city's peripheral areas with no known SARS cases have been increased. Abnormally high morbidity rates in certain areas have declined. There is little adjustment of the morbidity rates in center areas with high population density. The Bayesian adjustment procedural as applied to SARS morbidity rates is outlined below [26].

Suppose that the real morbidity rate is a random variable θ_i with mean μ_i and variance σ_i^2, in the i-th unit. The observed morbidity rate t_i is a realization of the random variable $N(\mu_i, \sigma_i^2)$. It can be shown that the best Bayesian adjustment is given by a linear combination of the observed rate t_i, and the mean μ_i:

$$\hat{\theta}_i = w_i \cdot t_i + (1 - w_i) \cdot \mu_i, \tag{1}$$

$$w_i = \frac{\sigma_i^2}{\sigma_i^2 + \mu_i / n_i}, \tag{2}$$

where n_i is the population in the i-th unit; the factor w_i is a weight value varying between 0 and 1. The smaller the weight w_i, the greater ratio the mean μ_i, the closer between the Bayesian estimate $\hat{\theta}_i$ and the mean μ_i. Note from eq. (2) that those districts with small population density will have a larger correction.

The only task left as to estimating $\hat{\theta}_i$ is obtaining the estimation of mean μ_i and variance σ_i^2. A simple approach is empirical Bayesian estimation that treats all means and variances as equal for all units, which is based on the hypothesis that the distribution of the random variable θ_i is the same across all areas. Thus we calculate μ and σ^2 directly from the observed data:

$$\mu_i = \hat{\mu} = \sum y_i \Big/ \sum n_i, \tag{3}$$

$$\sigma_i^2 = \frac{\sum n_i (t_i - \hat{\mu})^2}{\sum n_i} - \frac{\hat{\mu}}{\bar{n}}, \tag{4}$$

where the number y_i is the number of SARS infections in the i-th unit; the number $\bar{n}$ is the mean population per unit in Beijing.

2.3 Spatial smoothing

Bayesian adjustments can help alleviate part of the data-related problems. However, additional challenges remain. As shown in Figure 3(a), noises in the raw data have resulted in over-mixing of dark areas (indicative of high risks) and light areas (with low risks), making the task of identifying the overall disease transmission trends hard. Spatial smoothing is a statistical method to solve this problem. The purpose of spatial smoothing is to cope with functional variability that cannot be compensated by spatial normalization and to improve the signal to noise ratio. In other words, smoothing aims to increase statistical power. At the same time, smoothing can lead to reduced sensitivity and loss of spatial details, especially when using a large moving window.

Procedurally, spatial smoothing involves computing the rate in a moving window centered on each spatial unit in turn. The moving window includes the units as well as its neighbors. With respect to the signal to noise ratio, an ideal moving window will be determined based on the specific requirement of the analysis tasks at hand and the target spatial patterns expected to be detected. A moving window determines the neighbors. A traditional way to represent neighborhood relationships is through a $n \times n$ spatial weight

matrix W:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}, \quad (5)$$

where n is the number of spatial statistical units under study; and w_{ij} is the weight of a particular neighborhood relationship. Many methods can be used to define spatial neighborhood relationships. The methods can be divided into two categories: topology-based and distance-based. The former is used to analyze linear or polygonal data with shape and size, whereas the latter is mainly for point data.

Figure 3(c) shows a spatially smoothed morbidity map corresponding to the Beijing 2003 SARS epidemic. We can see a clear spatial trend from this map. The risk of SARS disease transmission in the city center areas is significantly higher than that in the city's peripheral areas. From the center to peripherals, the average morbidity rates gradually decrease, showing a similarly spatial pattern with ring roads. The transmission risk in the east-west direction is strong than that in the north-south direction. We can also identify a local high-risk area in the eastern peripheral area, which might be a special pathway of SASR transmission between the center and the eastern areas.

3 Spatial patterns and spatio-temporal evolutionary patterns

3.1 Quantitative methods to examine spatial association

"Everything is related to everything else, but near things are more related than distant things" [27]. Spatial dependency is one of the essential sources contributing to order, pattern and diversity in nature [28]. Spatial dependency also directly leads to the spatial auto-correlation problem since it violates standard statistical independence assumptions. Population, socio-economic factors, the natural environment, and other spatial factors all have impact on SARS transmission, and spatial locations of SARS cases are clearly not

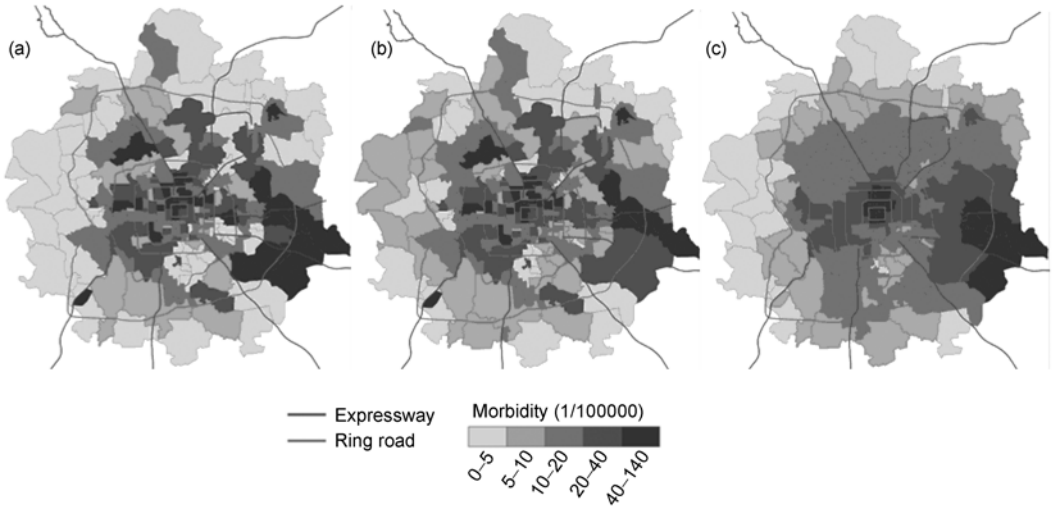

Figure 3 Morbidity map of the Beijing 2003 SARS epidemic. (a) Raw morbidity rates; (b) after Bayesian adjustments; (c) after spatial smoothing.

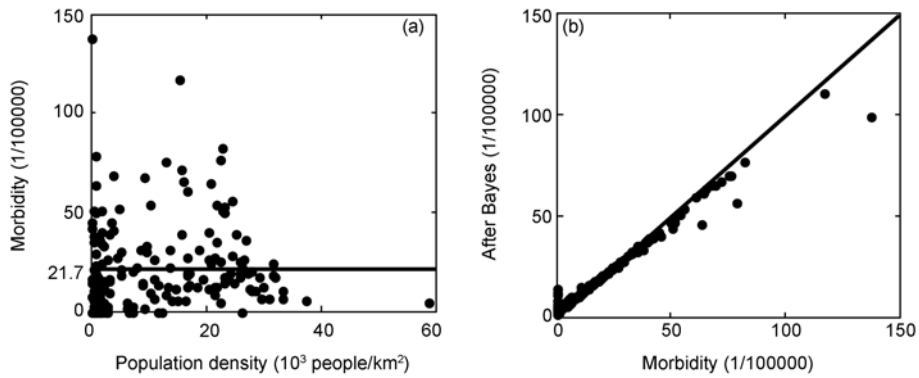

Figure 4 Morbidity rates after Bayesian adjustments. (a) Morbidity rates vs. population density; (b) morbidity rates before and after Bayesian adjustments.

independent. Thus traditional statistical methods cannot be used to analyze spatial patterns. Rather, spatial statistical methods, which are based on the assumption of various forms of spatial dependence, need to be applied.

There are many methods to measure quantitatively spatial association features for geographical objects. Global spatial association indices include Moran's I, Geary's C, Getis's G and Join Count [29]. Local spatial association indices include Local Moran's I, LISA (Local Indicators of Spatial Association) [30] and Getis's G* [31]. In addition, semi-variogram is commonly used to explore spatial dependence [32]. In our study, Moran's I, LISA and semi-variogram were selected to study the spatial association features of the Beijing 2003 SARS epidemic. In this section, we briefly introduce these three methods.

3.1.1 Semi-variogram

Semi-variogram is a function describing the degree of spatial dependence of a spatial random field or stochastic process. It is defined as the expected squared increment of the values between two locations. A variogram consists of two parts: an experimental variogram and a model variogram.

(1) Experimental semi-variogram [32]:

$$\gamma(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}\sum_{|s(x_i)-s(x_j)|\in h\pm\varepsilon}^{N(h)}[z(x_i)-z(x_j)]^2, \tag{6}$$

where $\gamma(h)$ is semi-variogram; $z(x_i)$ is the observation at location $s(x_i)$; $N(h)$ is the number of pairs in the set of pairs of observations i, j such that $|s(x_i)-s(x_j)|\in h\pm\varepsilon$.

(2) Model semi-variogram [32]. A model variogram is a function that models the trend in the experimental variogram. There are three common mathematical functions to define a semi-variogram model, which are listed as follows.

1) The spherical variogram model

$$\gamma(h)=\begin{cases}0, & h=0,\\ C_0+(C-C_0)\left(\dfrac{3h}{2a}-\dfrac{h^3}{2a^3}\right), & 0<h\leqslant a,\\ C, & h>a.\end{cases} \tag{7}$$

2) The Gaussian variogram model

$$\gamma(h)=\begin{cases}0, & h=0,\\ C_0+(C-C_0)(1-\mathrm{e}^{-3h^2/a^2}), & h>0.\end{cases} \tag{8}$$

3) The exponential variogram model

$$\gamma(h)=\begin{cases}0, & h=0,\\ C_0+(C-C_0)(1-\mathrm{e}^{-h/a}), & h>0;\end{cases} \tag{9}$$

where $\gamma(h)$ is semi-variogram; C_0 is the Nugget effect, denoting the randomness and sampling error or short-scale variability, C is sill, the maximum variogram value; a is the range-the distance causing the variogram to reach plateau. The range of the variogram is often close to the average size of physical anomalies in the spatial fluctuation of the z values, and a can have different values in different references, due to the ambiguity in the definition of the range.

3.1.2 Moran's I

Moran's I tests for global spatial autocorrelation. It is a weighed correlation coefficient used to detect departures from spatial randomness. Such departures from randomness can indicate spatial patterns such as clusters. Moran's I may identify other kinds of pattern such as geographic trend. Moran's I is calculated as follow [33]:

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_i-\overline{x})(x_j-\overline{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\cdot\sum_{i=1}^{n}(x_i-\overline{x})^2}, \tag{10}$$

where x_i and x_j are the values or counts at spatial point i and j, $\overline{x}$ is the average value of all the points in the entire region, w_{ij} is the weight of the spatial neighborhood relationship, and n is the population size.

3.1.3 LISA

LISA [30] is a statistic that evaluates the existence of clusters in the spatial arrangement of a given variable. This statistic detects local spatial association and can be used to identify local clusters (i.e., regions where adjacent areas have similar values) or spatial outliers (i.e., areas distinct from their neighbors). The *LISA* statistic is defined as follow:

$$I_i=\frac{n(x_i-\overline{x})\sum_{j-1}^{n}w_{ij}(x_j-\overline{x})}{\sum_{j=1}^{n}(x_j-\overline{x})^2}, \tag{11}$$

where x_i, x_j, and w_{ij} have the same meaning as in Equation (10). Comparing eqs. (10) and (11), we note that the LISA statistics decompose Moran's I into contributions for each location, I_i. The sum of I_i for all observations is proportional to Moran's I.

3.2 Spatial variation of Beijing 2003 SARS epidemic

Noises may drop the power of spatial variation approach. As such, we adopt spatially smoothed morbidity rates to explore spatial variation pattern of the Beijing 2003 SARS epidemic.

A basic assumption behind the semi-variogram-based method is intrinsic stationarity or wide-sense stationarity of the field. This assumption means that the expected z value is an unknown constant and the variance of the difference is the same between any two points that are at the same dis-

tance and direction. As there is a clear directional spatial trend in the morbidity map as shown in Figure 3, this assumption is violated. In our study, we applied a global second-order surface to remove this spatial trend. We have focused on residuals to explore the spatial variation features by the semi-variogram method.

Figure 5 shows the spatial variation of the Beijing 2003 SARS epidemic based on spatially smoothed morbidity rates. The related estimated parameters are summarized in Table 2. Geometric anisotropy, with the semi-variogram reaching the same sill in all directions, but at different ranges, was employed for anisotropy. The longest range, 13.627 km, occurs in northeast-southwest direction, whereas the shortest range, 8.545 km, occurs in northwest-southeast direction. The ratio of the longest range dividing the shortest one is 13.627/8.545 ≈ 1.6, indicating that the spatial risk of SARS transmission is anisotropic. The greater the spatial variability, the slower the speed of spatial diffusion. We conclude that the spatial diffusion of SARS transmission in the northeast-southwest direction is much speeder than that in the northwest-southeast direction. People living in the northeast-southwest suburb areas experienced a more serious threat, which may be the result of more frequency of human activities.

Identification of anisotropy to SARS transmission can be used to guide prevention and control measures. For instance, it would be an effective strategy to allocate more public health resources matching the anisotropy of an elliptical shape shown in Figure 5(b).

3.3 Global spatial association of Beijing 2003 SARS epidemic

The software package GeoDA (https://www.geoda.uiuc.edu/) was used to analyze the global spatial association within the Beijing 2003 SARS data. The Moran's I statistic is 0.2985. In order to check whether the spatial association is statistically significant, we tested it again the Null hypothesis H_0: morbidity rates are randomly distributed spatially. We conducted 999 Monte Carlo simulations with mean $E(I) = -0.0052$, and standard deviation $S_d = 0.0443$. The resulting z-value is 6.74. Based on the P-value (<0.01) we rejected the Null hypothesis H_0 and conclude that the SARS data are positively spatially associated.

We have also tested the effect of four factors, gender, age, onset of symptoms, and exposure record (Table 3). From this table, we make the following observations. (1) The spatial association among women is stronger than that among men; this might be the result of the fact that activities of men are more scatter geographically than those of women. (2) The spatial association among people who have records of exposure is stronger than that of people with no exposure records. Further examination of the data revealed that people being previous exposed are mostly from contacts in families and hospitals. (3) The spatial association is stronger among the older people. This is partially due to the fact that the older people were exposed to SARS mostly through families and hospitals. Figure 6 shows that the ratio between people who have exposure records and those with no exposure records increases with age. (4) The spatial association experienced a weak-strong-weak evolutionary pattern over time, which may be the result of intermingled effect between transmission dynamics and interventions during different periods of SARS transmission in Beijing.

3.4 Spatio-temporal evolution of Beijing 2003 SARS epidemic

In our study, we have used the widely-applied LISA statistic,

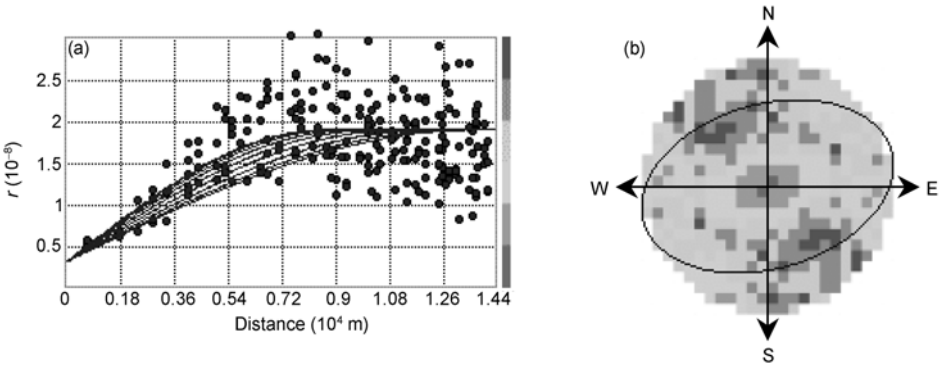

Figure 5 Spatial variation of the 2003 SARS transmission in Beijing. (a) Experimental semi-variogram; (b) anisotropy surface of semi-variogram.

Table 2 Estimated parameters of the semi-variogram function

Range a (km)		Azimuth (°)	Nugget C_0	Sill C	C_0/C
Longest	Shortest				
13.627	8.545	74.70	29.424	188.574	0.156

Table 3 Moran's I statistics computed from the Beijing 2003 SARS data[a)]

Impact factor		Moran's I	999 MC simulations		
			E(I)	S_d	z-score
Gender	Male	0.2231*	–0.0020	0.0448	4.9799
	Female	0.2906*	–0.0025	0.0418	6.9522
Exposure record	yes	0.3406*	–0.0019	0.0405	8.4099
	no	0.1317*	–0.0039	0.0517	2.5474
Age	≤20	0.1291*	–0.0031	0.0446	2.8946
	>20, ≤40	0.1708*	–0.0013	0.0469	3.6418
	>40	0.3667*	–0.0013	0.0394	9.3071
Onset (March 8–May 31)	1st week	0.0158	–0.0037	0.0237	0.6667
	2nd week	0.0731*	–0.0012	0.0252	2.9008
	3rd week	–0.0104	–0.0058	0.0383	–0.2715
	4th week	0.0996*	–0.0055	0.0385	2.5870
	5th week	0.1713*	–0.0088	0.0471	3.6369
	6th week	0.2217*	–0.0036	0.0405	5.4741
	7th week	0.1953*	–0.0041	0.0475	4.1116
	8th week	0.2368*	–0.0013	0.0418	5.6651
	9th week	0.0739	–0.0048	0.0431	1.7146
	10th week	0.1162*	–0.003	0.0426	2.7277
	11th week	0.1997*	–0.0049	0.0416	4.8005
	12th week	0.0463	–0.0039	0.0285	1.6246

a) * refers to a positive spatial association (P<0.01).

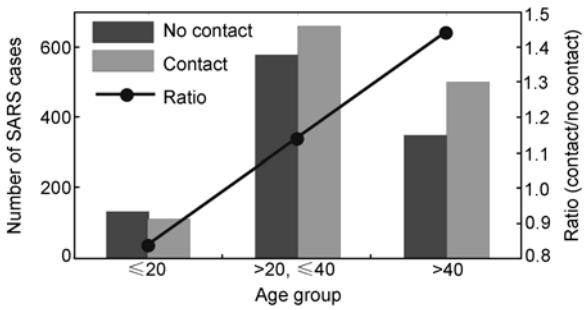

Figure 6 Relationship between age and exposure record.

an indicator of local spatial association, to analyze the morbidity rates of Beijing SARS cases using local clusters and spatial outliers, shown in Figure 7(a). There are four categories of spatial patterns in a LISA map. The high-high and low-low locations (positive local spatial autocorrelation) are typically referred to as spatial clusters, while the high-low and low-high locations (negative local spatial autocorrelation) are termed spatial outliers. A cluster is computed as such when the value at a location (either high or low) is more similar to its neighbors than would be the case under spatial randomness. The high-high locations refers to hot spot areas where the risk of SARS spreading is higher than average where as the low-low locations refers to cool spot areas where the risk of SARS spread is lower than average.

From Figure 7(a), we observe that the high-high locations aggregated in urban areas, covering the northern areas within the 3rd ring road, and the eastern peripheral areas near the Tongzhou district. The low-low locations concentrated in northwest peripheral areas. More specifically, (1) there is a cluster of high-risk areas in the city center, which may be the result of high population density and aggregated hospitals in those areas. (2) The Tongzhou district, located in the far east end of Beijing with low population density and few hospitals receiving SARS patients, also appears to possess high risk. A partial explanation could be economic-social: the cost of living in the Central Business District (CBD) located in the Chaoyang district is expensive, thus many people working in the CBD choose to live in the Tongzhou district, a much affordable area. Also the Beijing-Shenyang Expressway provides a convenient travel for these people. (3) People living in the northwest peripheral areas were better protected, which may be the result of more effective control measures or low intensity and frequency of contacts and interactions among the population in these places. (4) Public health response and control measures against the SARS epidemic seem to be quite effective as the high risk areas of the SARS transmission were confined to a limited area without city-wide outbreaks.

We have also studied the heterogeneity of spatial patterns in terms of three impact factors: gender, age, and exposure record (Figure 7(b)–(h))). Gender has little effect on spatial pattern of SARS transmission (Figure 7(b)–(c)). However, the spatial cluster among people who have exposure records shows a different model with that among people who have no exposure records (Figure 7(d)–(e)). The high-high areas occurred in the center and eastern peripheral areas shown in both Figure 7(d) and 7(e); however, the high-high area occurring in the northwest peripheral areas appeared only in Figure 7(e). Spatial clusters among the people with age<20 or age>40 were more decentralized. We conclude that prevention and control measures were more effective to the people with age between 20 and 40 than the people of age<20 or age>40.

The LISA statistic was applied to study the spatio-temporal evolution of spatial risk of SARS transmission. Figure 8 shows the analysis for each of the 12 weeks covering the entire Beijing SARS outbreak. In Table 4, we summarize our findings concerning observed spatio-temporal patterns, emergency strategies and control measures, and their epidemiological interpretations.

From Table 4, we learn these lessons as follow: (1) SARS disease has the huge potential to destroy our society. When an effective intervention absent, the number SARS-infected increase exponentially, early and effective preparedness and quickly emergency response are crucial to control the SARS epidemic. In the Beijing 2003 SARS epidemic, lack of the early identification of SARS outbreaks leads to miss the opportunity to kill the SARS epidemic in March, 2003. (2) We have identified two spatial clusters: one in the city center and the other in the eastern suburban area. The cluster in the city center occurs in the end of March and dissolves in mid-May, whereas the cluster in the eastern suburban area occurs in mid-April and dissolves in

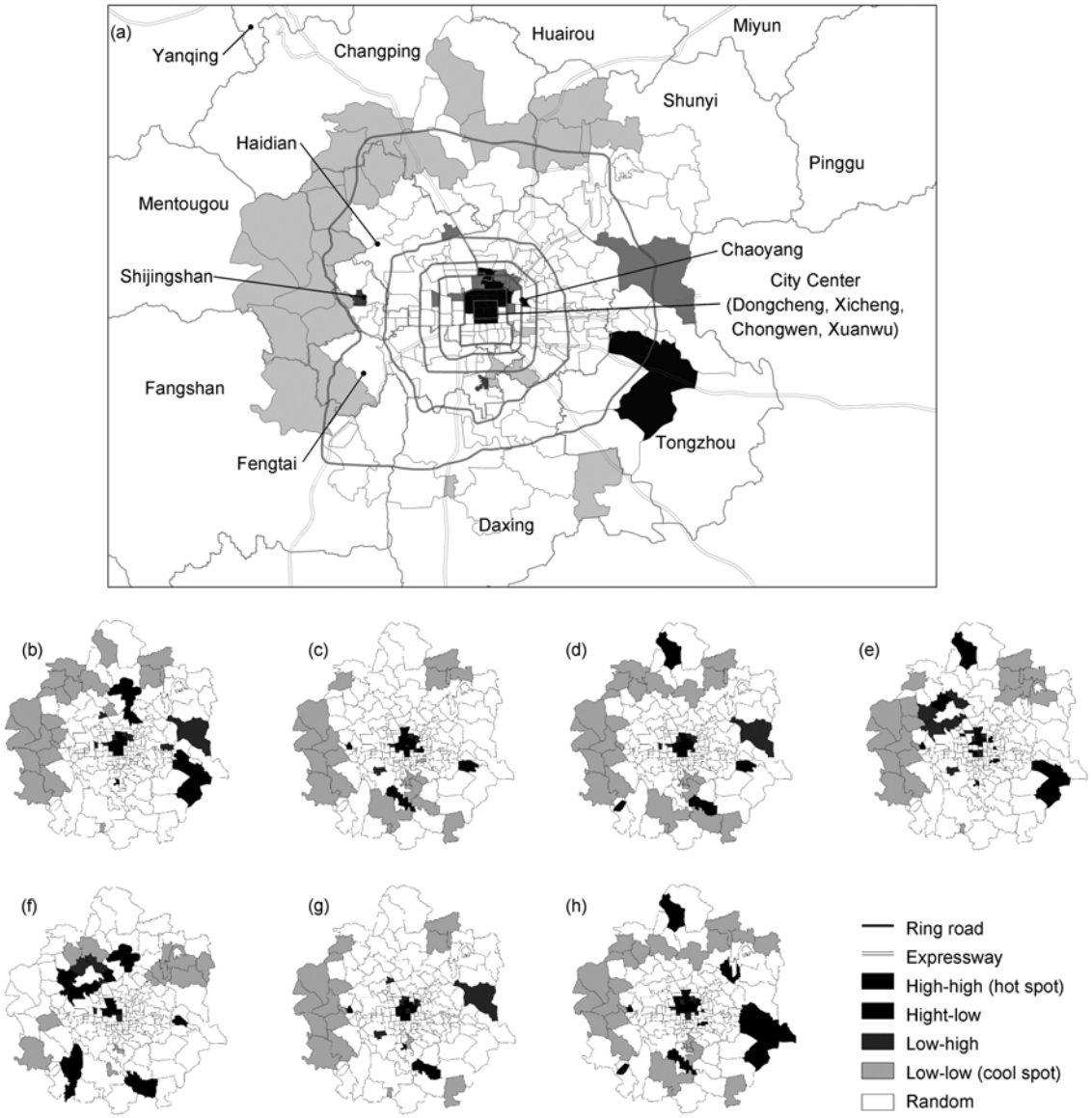

Figure 7 Local spatial association of Beijing 2003 SARS epidemic. (a) All SARS cases; (b) male cases; (c) female cases; (d) cases who have exposure records; (e) cases who have no exposure records; (f) cases of age<20; (g) cases of 20<age⩽40; (h) cases of age>40.

early May, which lasts about one month. (3) Spatial clusters are restricted in limited areas, spatial diffusion of SARS transmission was stopped effectively and timely, especially in the northwest direction. (4) Control measures contained the establishment of Xiaotangshan hospital complemented after April 20 is effective to control the SARS disease. These measures not only cut down the size of Beijing SARS epidemic, but also stop further spatial diffusion.

4 Concluding remarks

In this paper, we explored the spatial patterns and spatio-temporal evolution of the Beijing 2003 SARS epidemic. The relationship between these patterns and various public health response and control measures was studied.

The epidemic of SARS disease in China in 2003 exposed our limit in epidemiology. In public health practice, when an epidemic outbreak occurs, especially for new and emerging diseases, very often action needs to be taken based on sparse knowledge about the pathogen and disease transmission. The public health response mechanisms such as quarantine are often implemented without a proper understanding of their impact. Some measures were complentmented blindly, which resulted in some serious problems, such as high economic and social price. Nowadays, although a large number of researchers have made tremendous progress, but there remains a long distance to reach successes.

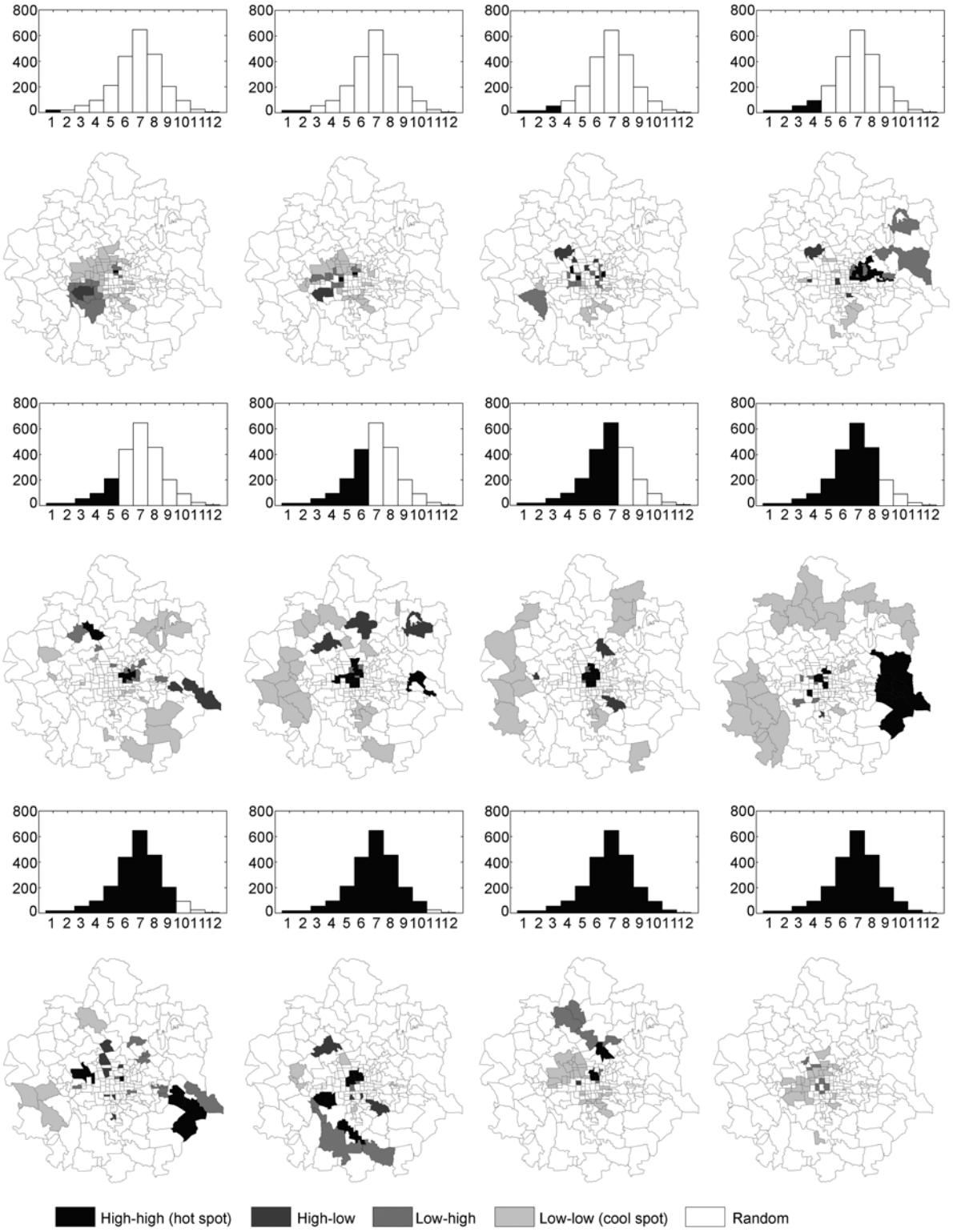

Figure 8 Spatio-temporal evolution of Beijing 2003 SARS epidemic. For each week, a pair of figures are provided. The figure on the top shows the temporal evolution of SARS case counts. The figure on the bottom shows the corresponding spatial pattern.

SARS has pandemic potential, but an effective public health response can curtail this threat. Using the Beijing 2003 SARS epidemic data, this study demonstrates the importance of analyzing spatial patterns and spatio-temporal evolutionary patterns in the context of infectious disease informatics. Our empirical findings provide useful insights that

Table 4 Spatio-temporal evolution of the Beijing SARS epidemic and epidemiological interpretation

Time (week)	Spatio-temporal pattern		Response strategies and control measures	Epidemiological interpretation
	Temporal pattern	Spatial pattern		
1st to 3rd (March 8–28)	Few new cases	Few spatial cluster areas scattered around the city center	SARS information not released to the public; control measures were limited in hospital; because of lack of understanding, many healthcare workers were infected	Early outbreak of SARS occurred in Beijing with the first SARS case imported from Hong Kong. The actual harm caused by SARS was limited but this early stage presented a great (but unfortunately unrealized) opportunity to control the SARS epidemic.
4th to 5th (March 29–April 11)	The number of new cases quickly grew and exponentially accelerated	Rapid growth of spatial cluster areas; diffusion from the city center to the peripheral areas, especially to the eastern areas	WHO issued a warning to China; the then Minister of Health held a press conference announcing that the SARS transmission was under control	Risk of SARS transmission increased quickly. SARS began to spread in the peripheral areas and formed high-risk clusters. However, the spatial diffusion to the northwest suburbs was stopped.
6th to 8th (April 12–May 2)	The number of new cases sharply increased and reached its peak value	Two spatial clusters formed in the city center and the eastern peripheral areas, the former under control but the latter continued to grow	National mobilization of all resources against SARS outbreak; information relate to SARS made open to the general public; SARS hospitals were established with serious infections isolated; travel was subject to control	Major outbreak occurred. Control measures and epidemic information were made accessible to the public. The situation began to improve with lowered case growing speed and no further spatial diffusion.
9th to 12th (May 3–31)	New SARS cases rapidly declined; SARS transmission gradually perished	Spatial clusters in the eastern peripheral areas decayed quickly; the clusters in the city center gradually dissolved	Xiaotangshan hospital was set up quickly to treat all known SARS cases; government continued to restrict and control people's travels and activities	Control measures began to work and the epidemic was prevented effectively. Xiaotangshan hospital ended the SARS transmission in hospitals.

could help better understand disease transmission mechanisms and prepare for and respond to possible reemergence of the SARS epidemic or other novel epidemics of acute infectious diseases.

Over the past two decades, with the development of information technology and its extensive application, mass spatio-temporal data storage, management and dynamic update are no longer a problem. Furthermore, spatial statistical methods have been developed rapidly, which have been used widely in a lot of fields, such as resources, environment, ecology, economic and sociality. Spatial statistical methods have been proved a powerful tool to public health. A new discipline, spatial epidemiology, has formed based on the intersection of epidemiology, spatial statistics and geographic information systems, which is the study of the spatial distribution of disease. At present, researching on spatial epidemiology in China is still in its early stage; this paper can promote its development in public health field and provide a sample of using spatial statistical methods to study the infectious disease epidemics, such as avian flu, hand-foot-mouth disease and Influenza A (H1N1).

This work was supported by National Science Foundation of the United State (Grant Nos. IIS-0839990, IIS-0428241), Department of Homeland Security of the United State (Grant No. 2008-ST-061-BS0002), Important National Science & Technology Specific Projects (Grant Nos. 2009ZX10004-315, 2008ZX10005-013), the Chinese Academy of Sciences (Grants Nos. 2F07C01, 2F08N03), China Postdoctoral Science Fund (Grant No. 20080440559), National High Technology Research and Development Program of China (Grant No. 2006AA010106), and National Natural Science Foundation of China (Grant Nos. 60621001, 40901219, 90924302).

1 Oxford J S, Bossuyt S, Lambkin R. A new infectious disease challenge: Urbani severe acute respiratory syndrome (SARS) associated coronavirus. Immunology, 2003, 109: 326–328
2 Marra M A, Jones S J, Astell C R, et al. The genome sequence of the SARS-associated coronavirus. Science, 2003, 300: 1399–1404
3 Tan Y M, Chow P K H, Soo K C. Severe acute respiratory syndrome—Clinical outcome after inpatient outbreak of SARS in Singapore. Br Med J, 2003, 326: 1394–1394
4 Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science, 2003, 300: 1966–1970
5 Keogh-Brown M R, Smith R D. The economic impact of SARS: How does the reality match the predictions? Health Policy, 2008, 88: 110–120
6 Chen W K, Wu H W, Lin C C, et al. Emergency department response to SARS, Taiwan. Emerging Infect Dis, 2005, 11: 1067–1073
7 Pang X H, Zhu Z H, Xu F J, et al. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. JAMA, 2003, 290: 3215–3221
8 Liang W N, Liu M, Chen Q, et al. Assessment of impacts of public health interventions on the SARS epidemic in Beijing in terms of the intervals between its symptom onset, hospital admission, and notification. Biomed Environ Sci, 2005, 18: 153–158
9 Liang W N, Liu M, Chen Q, et al. Description about onset date,visit date,report date of severe acute respiratory syndromes(SARS) to observe impact of public health interventions (in Chinese). China Public Health, 2004, 20: 129–130
10 Zhang Z B. The outbreak pattern of SARS cases in China as revealed by a mathematical model. Ecol Modell, 2007, 204: 420–426
11 Wang H Y, Rong F, Ke F J, et al. Control dynamics of severe acute respiratory syndrome transmission. Chin Sci Bull, 2003, 48: 2684–2687
12 Cai Q C, Jiang Q W, Xu Q F, et al. To develop a model on severe acute respiratory syndrome epidemics to quantitatively evaluate the effectiveness of intervention measures (in Chinese). Chin J Epidemiol, 2005, 26: 153–158
13 Han W G, Wang J F, Liu X H. Back Analyzing Parameters and Predicting Trend of SAR transmission (in Chinese). Adv Earth Sci, 2004, 19: 925–930

14 ZongY G, Wang L, Qu X L. Monte-Carlo simulation based approach for the study of SARS spatial-temporal scenarios in Beijing (in Chinese). Geogr Res, 2004, 23: 815–824
15 Shi Y L. Stochastic dynamic model of SARS spreading. Chin Sci Bull, 2003,13: 1373–1377
16 Chen Q Z. Application of SIR Model in forecasting and analyzing for SARS (in Chinese). J Beijing Med Univ, 2003, 35: 75–80
17 Wu J, Xu F J, Zhou W G, et al. Risk factors for SARS among persons without known contact with SARS patients, Beijing, China. Emerging Infect Dis, 2004, 10: 210–216
18 Wang Z, Cai D, Li S, et al. On season risk of the prevalence of SARS in China (in Chinese). Geogr Res, 2003, 22: 541–550
19 Li R W K, Leung K W C, Sun F C S, et al. Severe Acute Respiratory Syndrome (SARS) and the GDP. Part I : Epidemiology, virology, pathology and general health issues. Br Dent J, 2004, 197: 77–80
20 Zhu Y B, Ge Q S, Wei X A, et al. A study on the effect of the SARS epidemic on the international tourist arrivals in China (in Chinese). Geogr Res, 2003, 22: 551–559
21 Meng B, Wang J F, Liu J Y, et al. Understanding the spatial diffusion process of severe acute respiratory syndrome in Beijing. Public Health, 2005, 119: 1080–1087
22 Wang J F, Mcmichael A J, Meng B, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. Bull World Health Organ, 2006, 84: 965–968
23 Wang J F, Meng B, Zheng X Y, et al. Analysis on the multi-distribution and the major influencing factors on severe acute respiratory syndrome in Beijing (in Chinese). Chin J Epidemiol, 2005, 26: 164–168
24 Wu J L, Wang J F, Meng B, et al. Spatial association analysis on epidemic of SARS in Beijing, 2003 (in Chinese). J Zhejiang Univ (Agric Life Sci), 2005, 31: 97–101
25 Liu M, Liang W N, Du H, et al. Epidemiological features of severe acute respiratory syndrome in Beijing urban and suburb areas in 2003. Biomed Environ Sci, 2005, 18: 227–232
26 Anselin L, Kim Y W, Syabri I. Web-based analytical tools for the exploration of spatial data. J Geogr Syst, 2004, 6: 197–218
27 Tobler W. A computer movie simulating urban growth in the Detroit region. Econ Geogr, 1970, 46: 234–240
28 Goodchild M F. The application of advanced tecnology in assessing environmental impacts, In: Corwin D L, Loague K. Application of GIS to the Modeling of Non-point Source Pollutants in the Vadose Zone. Soil Sci Soc Am J, 1996, 1–17
29 Haining R P. Spatial Data Analysis: Theory and Practice. Cambridge: Cambridge University, 2003
30 Anselin L. Local indicators of spatial association-lisa. Geogr Anal, 1995, 27:93–115
31 Ord J K, Getis G. Local spatial autocorrelation statistics: Distributional issues and application. Geogr Anal, 1995, 27: 286–306
32 Cressie N. Statistics for Spatial Data. New York: Wiley, 1993
33 Moran P. The interpretation of statistical maps. J R Stat Soc Series B Stat Methodol, 1948, 10: 243–251