Geographical information systems-based spatial analysis and implications for syphilis interventions in Jiangsu province, People's Republic of China

Yue-Jia Cheng[1], Jessie Norris[2], Chang-Jun Bao[3], Qi Liang[3], Jian-Li Hu[3], Ying Wu[3], Fen-Yang Tang[3], Wen-Dong Liu[3], Ke-Qin Ding[1], Yang Zhao[1], Zhi-Hang Peng[1], Rong-Bin Yu[1], Hua Wang[3], Hong-Bing Shen[1], Feng Chen[1]

[1]Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, 210029, People's Republic of China; [2]National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, People's Republic of China; [3]Jiangsu Province Center for Disease Control and Prevention, Nanjing 210009, People's Republic of China

Abstract. Spatial distribution rules and risk factors for syphilis were studied in Jiangsu province, People's Republic of China during 2005 and 2009. Trend surface analysis, spatial autocorrelation analysis and spatio-temporal clustering were applied with the incidence rates of the various counties in the province to determine spatial distribution rules and risk factors. Syphilis was found to be most severe in the southern region of the province where many counties could be shown to be hotspots with positive autocorrelation. Clusters were detected in the south-western region of Jiangsu with the county-level city of Yixing as the centre. Temperature, distance from railways and highways, and the normalised difference vegetation index were determined as supporting variables with regard to the transmission of the disease by both univariate and multivariate spatial correlation analyses. Interventions, including health education and awareness campaigns, should be strengthened throughout the province targeting the south-western areas, especially the clusters and hotspots detected in order to improve the situation.

Keywords: syphilis, geographical information systems, trend surface analysis, spatial autocorrelation analysis, spatial correlation analysis, People's Republic of China.

Introduction

Before the establishment of the People's Republic of China (P. R. China) in 1949, syphilis was widespread. After a series of intervention measures such as mass screenings and free treatment programmes were carried out by the government in the 1950s, P. R. China was gradually relieved of the pervasiveness of the disease (Chen, 2008). In 1978, however, reformation and the new open policy in the country not only resulted in economic development, but also led to an increase in the prevalence of syphilis and other sexually transmitted diseases (STDs) (Wu et al., 2010) that began in the coastal areas and later spread inland (Gong et al., 2000). Migration, social changes, increasing numbers of high-risk groups such as female sex workers (FSWs),

intravenous drug users and men who have sex with men (MSM) have contributed to the spread of syphilis in P. R. China (Lin et al., 2006; Li et al., 2009; Guo et al., 2010). The incidence rate of syphilis nationwide in 2005 was over 28 times as high as that in 1993 (Chen et al., 2007) and there were already 305,762 cases recorded in 2009 according to the Chinese Center for Disease Control and Prevention (CDC). Jiangsu, a highly developed coastal province, is currently experiencing a resurgence of the disease with incidence rates having increased continuously; from 12.0 (per 100,000 residents) in 2005 to 32.1 in 2009 according to the statistical data from the Jiangsu provincial CDC.

Studies both at home and abroad have focused on syphilis' herd susceptibility, vulnerable groups and aetiology (Mak et al., 2004; Martin et al., 2009). Pregnant women infected with *Treponema pallidum* (TP) are prone to premature birth, still birth or low infant birth weight (Zhang et al., 2007). Syphilis also has implications for other STDs such as human immunodeficiency virus (HIV), which in turn can facilitate reinfection with syphilis (Liu, 2000; Chen and Zhang, 2001). One study found that STDs such as syphilis showed spatial variability (Law et al., 2004). This study was initiated

Corresponding author:
Zhi-Hang Peng
Department of Epidemiology and Biostatistics
School of Public Health, Nanjing Medical University
140 Hanzhong Road, Nanjing 210029
People's Republic of China
Tel. +86 25 8686 2755; Fax +86 25 8686 7613
E-mail: zhihangpeng@yahoo.com.cn

to discover the geographical variation of syphilis in Jiangsu province, seeking clusters and hotspots. It aims to explore risk factors contributing to the transmission of syphilis in the province to provide strategies for future syphilis control and prevention.

Materials and methods

Study area

Jiangsu province is located in the eastern part of P. R. China between longitudes 116°21-121°54 E and latitudes 30°46-35°08 N. Due to its East Asian monsoon climate, the annual mean temperature varies between 13.6 °C and 16.1 °C. The population was 76,772,923 people in 2009 (data from the Chinese National Statistics Department). The province is geographically divided into the southern, northern and middle regions. The southern prefectures include Suzhou, Wuxi, Changzhou and Zhenjiang and the northern Xuzhou, Lianyungang, Suqian, Huai'an and Yancheng. The prefectures of Nanjing, Yangzhou, Taizhou and Nantong form the middle region.

Data collected

The cases included new syphilis cases recorded between 2005 and 2009 (no asymptomatic carriers or latent cases were included). The total number of cases were composed of primary syphilis (infection within the last three months), secondary syphilis (infection within the last two years) and tertiary syphilis cases (infection prior to the last two years). The incidence data for the approximately 106 counties were obtained from the Jiangsu provincial CDC. Demographic data from every county in Jiangsu province were collected from the Chinese National Statistics Department. Incidence rates (cases per 100,000 residents) were calculated using the number of cases and demographic data. The counties were the primary sampling units.

Mapping and statistical methods

ArcGIS software, version 9.3 (ESRI, Redlands, CA, USA) was used to map and analyse the cases and incidence rates of syphilis by county. A trend surface analysis was applied to study changes in the distribution of syphilis around the province. Spatial autocorrelation analysis and spatio-temporal clustering analysis denoted clusters and hotspots. Spatial correlation analysis was also used to identify risk factors associated with the infection.

Trend surface analysis

Trend surface analysis, which describes spatial structure or spatial changes, was used in an attempt to highlight scattered observations or geographical anomalies as done by Zhou (2009). Here, spatial data were divided into three parts: regional trends that change as locations move, local anomalies that indicate sudden changes in data and random errors that comprise sampling and measurement errors. Polynomial functions were applied in interpolating the incidence rates at the unknown nodes of the projection. Longitudes (denoted as X) and latitudes (denoted as Y) were independent variables, while the incidence rates of counties (denoted as Z) were dependent variables. In order to reduce the expectations of local anomalies and random errors to zero and minimizing squared deviations, ordinary least square (OLS) methods were used to calculate the coefficients of the polynomial functions as described by He et al. (2008). A trend surface map was produced with the x-axis pointing east and the y-axis north.

Spatial autocorrelation analysis

Spatial autocorrelation reflects the correlations of one region with its adjacent regions with regard to the same attributes, indicating a spatial independency on the data of the nodes. In autocorrelation theory, a short distance increases the likelihood of relationships among the attributes. Attributes of similar high or low values result in positive autocorrelation by definition, while opposing high and low attribute values result in negative autocorrelation (Liu et al., 2008).

General spatial autocorrelation analysis considers Jiangsu province as one entity assuming that the counties within the province do not differ from each other. The aim of this analysis, however, was to detect whether the epidemic was aggregated at the provincial level, thus determining whether Jiangsu province has clusters of syphilis. As general spatial autocorrelation cannot locate the clusters, general Moran's I was applied as used by He et al. (2008):

$$I = \frac{n}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}} \times \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum\limits_{i=1}^{n}(x_i - \bar{x})}$$

(Eq. 1)

where n is the number of spatial units (the counties), x_i and x_j the observations from unit i to unit j about the phenomenon x (in this study x = the incidence rate and $\bar{x}$ = the average incidence rate of all counties by

that year) and w_{ij}(d) the adjacent weight matrix from distance. If unit i is adjacent to unit j then w_{ij}(d) is 1, otherwise it is 0. The general Moran's I was then compared with the critical value of Z. If p-value <0.05, then Moran's $I \neq 0$, meaning that the area included a syphilis cluster.

Given that the syphilis epidemic may vary geographically by source of infection, mode of transmission, population characteristics and behavioural distinctions, provincial counties may not always exhibit the same features as neighbouring counties. Since the results of general spatial autocorrelation, therefore, could mask the autocorrelation of some counties, local spatial autocorrelation was used to examine whether there were regional patterns instead of considering the province as a whole. Local Moran's I (LMi) (Eq. 2) and local Getis Ord G index as by He et al. (2008) (G_i) (Eq. 3) were calculated using the following equations:

$$I_i = \frac{n^2}{\sum\limits_i \sum\limits_j w_{ij}} \times \frac{(x_i - \bar{x}) \sum\limits_j w_{ij} (x_j - \bar{x})}{\sum\limits_j (x_j - \bar{x})^2} \qquad \text{(Eq. 2)}$$

$$G_i(\text{d}) = \frac{\sum\limits_{i=1}^{n} \sum\limits_{j=1, j\neq i}^{n} w_{ij}(\text{d})x_j}{\sum\limits_{i=1}^{n} \sum\limits_{j=1, j\neq i}^{n} x_j} \qquad \text{(Eq. 3)}$$

LMi classified the counties into three patterns, positive autocorrelation, negative autocorrelation and dispersed distributed counties, and while it could determine whether there was a cluster and how it was aggregated, the local G index (G_i) could indicate the hotspots or cold spots among the positive autocorrelations. Moran's I ranges from -1 to 1 and the G index ranges from 0 to 2. Positive autocorrelation or hotspot occurs when Moran's I >0 or G_i >1. Dispersal distributed counties held LMi and G_i closer to 0.

For the local correlation analysis, the full ranges of LMi and Gi were classified into five levels based on the incidence rates. Adapted from the optimum partition method as used by Peng et al. (2010), the partition successfully achieves the minimum differences between samples at the same level of incidence rates, and between samples at different levels.

Spatio-temporal cluster analysis

Analysis of spatio-temporal clusters of attributes was applied to avoid the bias caused by using spatial or temporal cluster statistics individually. It relies on moving, circular windows with geographical bases for location and heights for time to scan the study areas. The resulting "cylinders" change their shapes (diame-ters and heights) while moving, and calculations for anomalies on incidence rates both inside and outside the scanning window are automatically performed including the expected incidence rates of the window. The results were inspected through log likelihood ratio (LLR) methods (Kulldorff et al., 1998). The "cylinder" with the minimum number of anomalies and an incidence rate above the expected was defined as indicating the most likely cluster.

Spatial correlation analysis

The goal was to detect potential relationships or associations between different variables (Wang, 2006). As climate and environmental factors play an important role in influencing public health and thus have wide applications for epidemic studies, the authors selected six risk factors for the spatial correlation analysis: the annual mean temperature, the normalised difference vegetation index (NDVI) for spring, summer, autumn and winter and the distribution of lakes, rivers, railways and highways (Chen et al., 2006; Jagai et al., 2009; Shakoor et al., 2010; Dambach et al., 2012).

The spatial correlation analyses in this study involved univariate analysis and a regression model that combines all underlying factors. Univariate analysis was conducted to describe the relationships between the incidence rates and each potential risk factor independently. Then Poisson regression was applied in multivariate analysis by assuming the incidence rates had a Poisson distribution. Multivariate regression model was established concerning risk factors by a series of selections through all six risk factors. The selections were based on knowledge of epidemiology and biostatistics throughout the regression. According to the degree-of-fit results, factors were not included until P ≤0.05 with a high correlation coefficient R. In both univariate and multivariate regression analysis, coefficients of determination (R^2) and F value were calculated to measure the goodness of fit between the observed data or regression results. Based on these results, the annual mean temperature and hibernal NDVI were finally excluded in the regression process, while the distribution of railways, highways, rivers and lakes were included.

Results

All primary, secondary and tertiary syphilis cases showed increasing trends from 2005 to 2009. Primary syphilis cases increased generally, while secondary
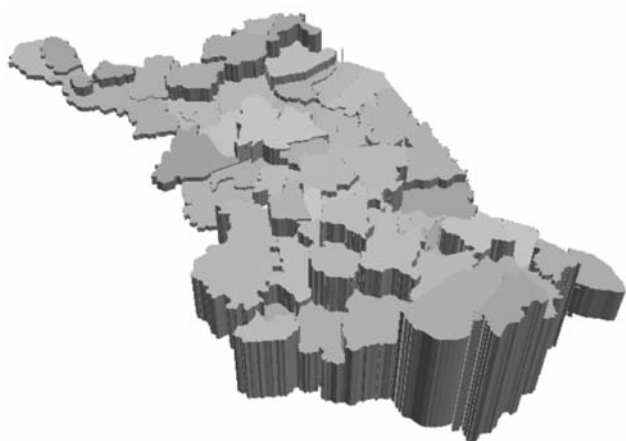
Fig. 1. Incidence map for syphilis incidence rates of Jiangsu province in 2009.

syphilis decreased between 2007 and 2008 but rose afterwards. The total annual syphilis incidence rates (cases per 100,000 residents) in Jiangsu province from 2005 to 2009 were 12.0, 16.2, 21.5, 27.8 and 32.1, respectively. In most counties, the incidence rates of 2009 accounted for the greatest proportion in all five years.

Geographically, the syphilis incidence rates were found to be more severe in the southern regions, though a few northern counties also showed high rates. In the incidence map for 2009 (Fig. 1) each polygon represents a county, the height of which represents the incidence, i.e. greater height indicates higher incidence. The risk map illustrates that the southern counties and the south-western areas were at higher risk for syphilis. Indeed, the primary, secondary and tertiary syphilis incidence rates were all elevated in these counties. In addition, a few regions in the west, north-east and north-west also threatened to increase the transmission of syphilis through their relatively high incidence rates.

Trend surface analysis

For the trend surface analysis, several locations were selected on the maps to disclose the changes of syphilis incidence rates in different directions. "Locations"

represent the perspective views of the trend surface maps and with the perspective located at 0°, the epidemic trends curves show a west to east or a north-to-south direction. Other locations represent changes in the north-west to south-east direction or in the north-east to south-west direction, e.g. east by north (41 degrees). The x-axis represents longitudinal increase and the y-axis latitudinal. The height of the z-axis represents the incidence rate.

In general, the syphilis incidence rates for the years 2005-2009 increased in a western to eastern direction and decreased from south to north (Figs. 2a-e). Moving from west to east the incidence rates increased in an exponentially in 2005 and 2007 but only linearly in 2006, 2008 and 2009. The detailed result of the trend surface analysis on the incidence rates is listed in Table 1. While the incidence of primary syphilis increased from west to east (the rates increased more quickly in 2006 than before but slowed again in 2009 (Fig. 2f), the secondary syphilis incidence rates in the same geographical direction accelerated continuously since the start of the study in 2005. The acceleration of secondary syphilis incidence in the north-south direction peaked in 2006 (Fig. 2g). With respect to tertiary syphilis, the trends showed only smooth changes and in all locations and has remained at a low incident level (Fig. 2h).

In order to further confirm the degrees of fit, and the results for the trend surface analysis, coefficients of determination R^2, F-values and p-values were calculated for all axes (Table 2). According to the evaluation, the curves fitted the results well, and changes detected were statistically significant.

Spatial autocorrelation analysis

The general spatial autocorrelation analysis showed that the syphilis incidence rates were aggregated at Jiangsu province. The Moran's *I* for general spatial autocorrelation from 2005 to 2009 was 0.235, 0.220, 0.249, 0.278 and 0.273, respectively, all statistically significant at P ≤0.001.

The local spatial autocorrelation analysis, conduct-

Table 1. Trend surface analysis of total syphilis incidence and the geography of Jiangsu province during the time of the study.

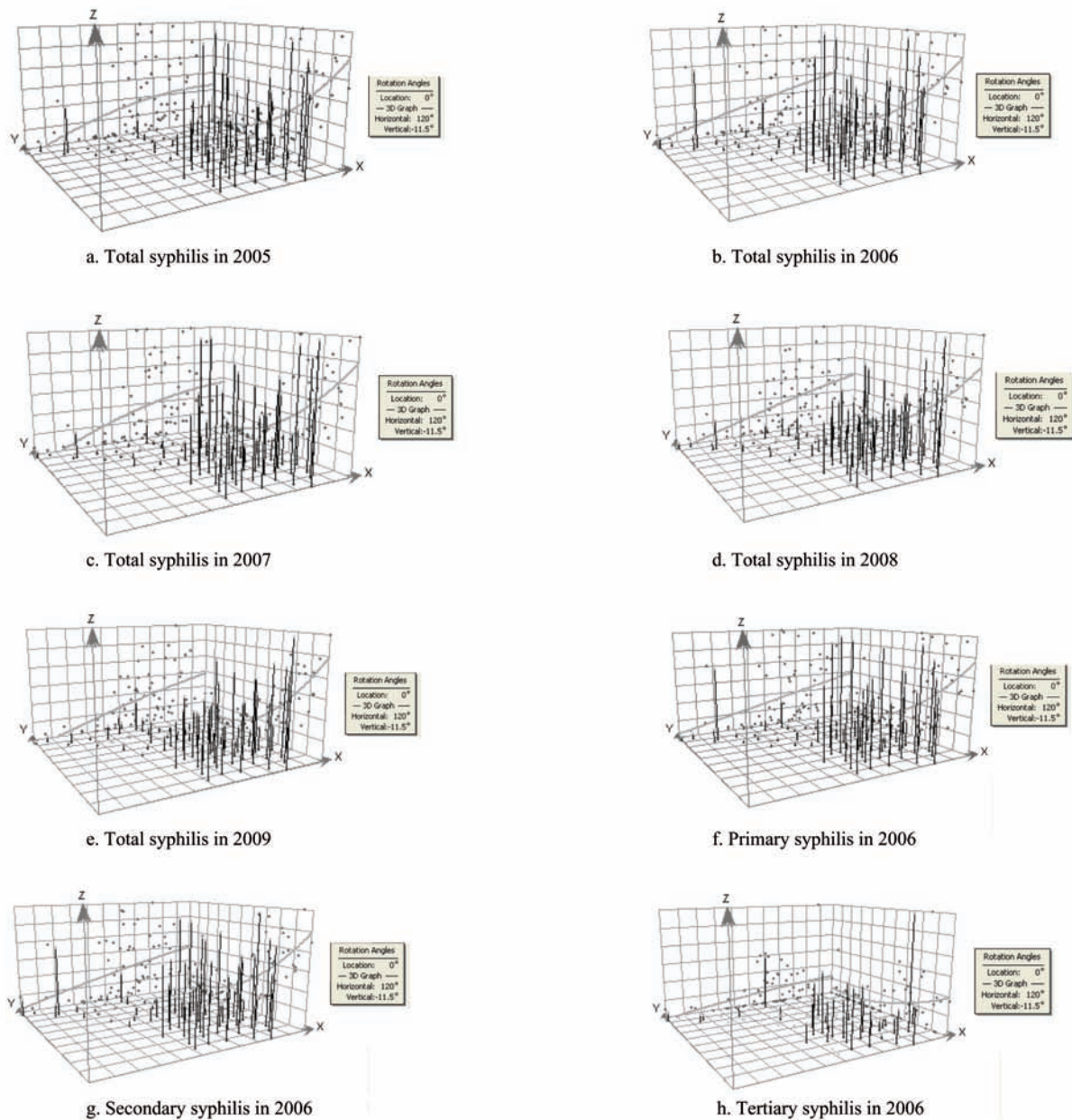| Years	West to East	North to South	North-west to South-east	North-east to South-west
2005	Logarithmic increase	Exponential increase	Linear increase	Exponential increase
2006	Linear increase	Exponential increase	Linear increase	Linear increase
2007	Logarithmic increase	Exponential increase	Increase	Linear increase
2008	Linear increase	Exponential increase	Exponential increase	Linear increase
2009	Linear increase	Increase	Exponential increase	Linear increase

a. Total syphilis in 2005

b. Total syphilis in 2006

c. Total syphilis in 2007

d. Total syphilis in 2008

e. Total syphilis in 2009

f. Primary syphilis in 2006

g. Secondary syphilis in 2006

h. Tertiary syphilis in 2006

Fig. 2. Trend surface graphs for syphilis.

Table 2. Trend surface analysis fitting degrees.

Year	Axis	R^2	F	P-value
2005	X	0.09	6.71	0.012
	Y	0.55	39.16	<0.001
2006	X	0.12	8.95	0.004
	Y	0.51	33.14	<0.001
2007	X	0.25	21.32	<0.001
	Y	0.65	120.89	<0.001
2008	X	0.28	24.63	<0.001
	Y	0.68	68.15	<0.001
2009	X	0.30	28.34	<0.001
	Y	0.73	85.36	<0.001

ed to further analyse the clusters within the province, indicated that the clusters are both hotspots for syphilis and positively autocorrelated (Table 3). The risk maps (Fig. 3) for local spatial autocorrelation analysis divided the counties into five levels according to the LMi or the Gi. The deeper colour in the LMi risk maps indicates a higher index and thus a higher aggregated level of syphilis incidence. Likewise, a deeper colour in the Gi risk maps indicates a greater likelihood of being a syphilis hotspot. Syphilis was particularly severe in these counties and the clusters verified as "high-high", defined as locations with high

Table 3. Clusters for local spatial autocorrelation analysis.

Year	County	Incidence	LMi	LMi Z*	LMi P**	Gi z*	Gi P**
2005	Suzhou	25.534	<0.001	9.991	<0.001	2.354	0.019
	Zhenjiang	26.547	0.001	9.585	<0.001	2.485	0.013
	Yangzhou	24.222	<0.001	6.058	<0.001	2.184	0.029
	Yangzhong	24.490	<0.001	6.684	<0.001	2.218	0.027
2006	Suzhou	27.004	<0.001	9.419	<0.001	2.318	0.020
	Changzhou	24.935	<0.001	8.151	<0.001	2.061	0.039
	Zhenjiang	26.457	<0.001	7.696	<0.001	2.250	0.024
	Yangzhou	27.305	<0.001	6.236	<0.001	2.356	0.018
2007	Suzhou	42.247	<0.001	10.666	<0.001	2.223	0.026
	Wujiang	43.478	<0.001	10.252	<0.001	2.318	0.020
	Yangzhou	43.160	<0.001	7.696	<0.001	2.294	0.022
	Yizheng	43.086	<0.001	7.338	<0.001	2.288	0.022
2008	Suzhou	37.411	0.001	12.645	<0.001	2.332	0.020
	Wujiang	43.701	0.001	13.753	<0.001	2.933	0.003
	Jintan	34.426	0.001	9.141	<0.001	2.047	0.041
	Danyang	34.969	0.001	8.886	<0.001	2.099	0.036
2009	Suzhou	59.501	0.001	12.844	<0.001	2.181	0.029
	Wujiang	73.390	0.001	15.078	<0.001	2.963	0.003
	Danyang	59.730	0.001	8.193	<0.001	2.194	0.028

*Score; **value.

incidence rates surrounded by counties that also have high incident levels. These counties were also found to be hotspots by local Getis Ord G analysis. Among the clusters, Suzhou was demonstrated as a hotspot every year for all five years of study, Yangzhou in 2005, 2006 and 2007 and Wujiang from 2007 to 2009. Clusters were only found in Zhenjiang in 2005 and 2006, Danyang in 2008 and 2009, while other counties, such as Yangzhong, Changzhou, Yizheng and Jintan, appeared only once each in the list of autocorrelations.

Spatio-temporal cluster analysis

According to our analysis, the distribution of syphilis in Jiangsu province was not dispersed. A space-time cluster was found in the south-western part of the province in 2008 and 2009. The cluster was made up of 27 counties belonging to the south-western region, with an average incidence rate of 31.2 per 100,000 residents. The center of the cluster was at coordinates 31.375° N and 119.78° E in Yixing county. The diameter for the cycloid base of the scanning
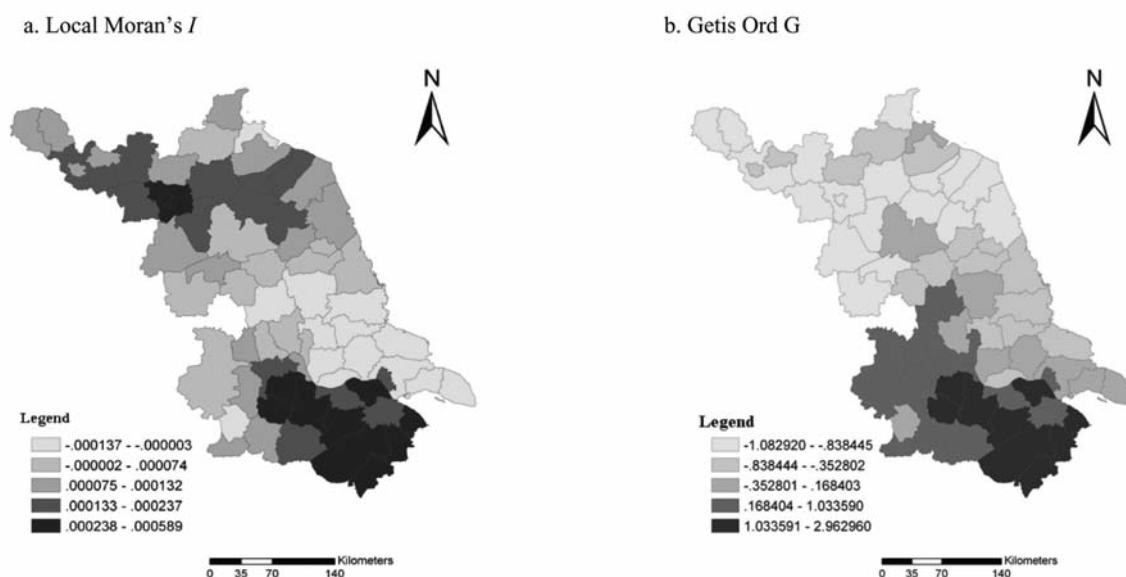

Fig. 3. Maps for local autocorrelation analysis in 2009.

window was 261.52 km. The 27 counties included in the scanning window were Jiangdu, Yangzhou, Yizheng, Nanjing, Taizhou, Yangzhong, JuRong, Tanshui, Gaochun, Liyang, Jintan, Danyang, Taixing, Rugao, Jingjiang, Changzhou, Yixing, Wuxi, Jiangyin, Zhangjiagang, Nantong, Changshu, Suzhou, Taicang, Kunshan, Wujiang and Zhenjiang. The incidence rate within the range of the scanning window was 27.5 cases per 100,000 residents, with an expected incident rate of 12.1 cases per 100,000 residents. The LLR was 7,332.72 (P <0.001).

Spatial correlation analysis

Annual mean temperature data, distance from lakes and rivers, railways/highways and NDVI for spring, summer, autumn and winter were extracted from the database and compared to the syphilis incidence rates. The layout charts are shown in Fig. 4.

From 2005 to 2009, the maximum annual mean temperature in Jiangsu Province was 22.6 °C and the minimum was 14.5 °C (Fig. 4a). The annual mean temperature was found to be positively correlated with syphilis incidence rates. Higher temperatures were associated with higher primary and secondary syphilis rates, with R^2 values of 0.296 (F = 27.33, P <0.001) and 0.310 (F = 29.18, P <0.001), respectively.

Concerning the distances of each county centroid from railways and highways (Fig. 4b-c), the incidence rates correlation coefficients turned out to be 0.953 (F = 1019.50, P <0.001) and 1 (F = 3519.25, P <0.001), respectively, indicating a close relationship between spatial communication and the syphilis epidemic. With respect to the presence of railways and highways, the rates of all forms of syphilis (primary, secondary and tertiary) showed a high correlation. However, this relationship varied. The incidence rates for all kinds of syphilis increased with distance peak-
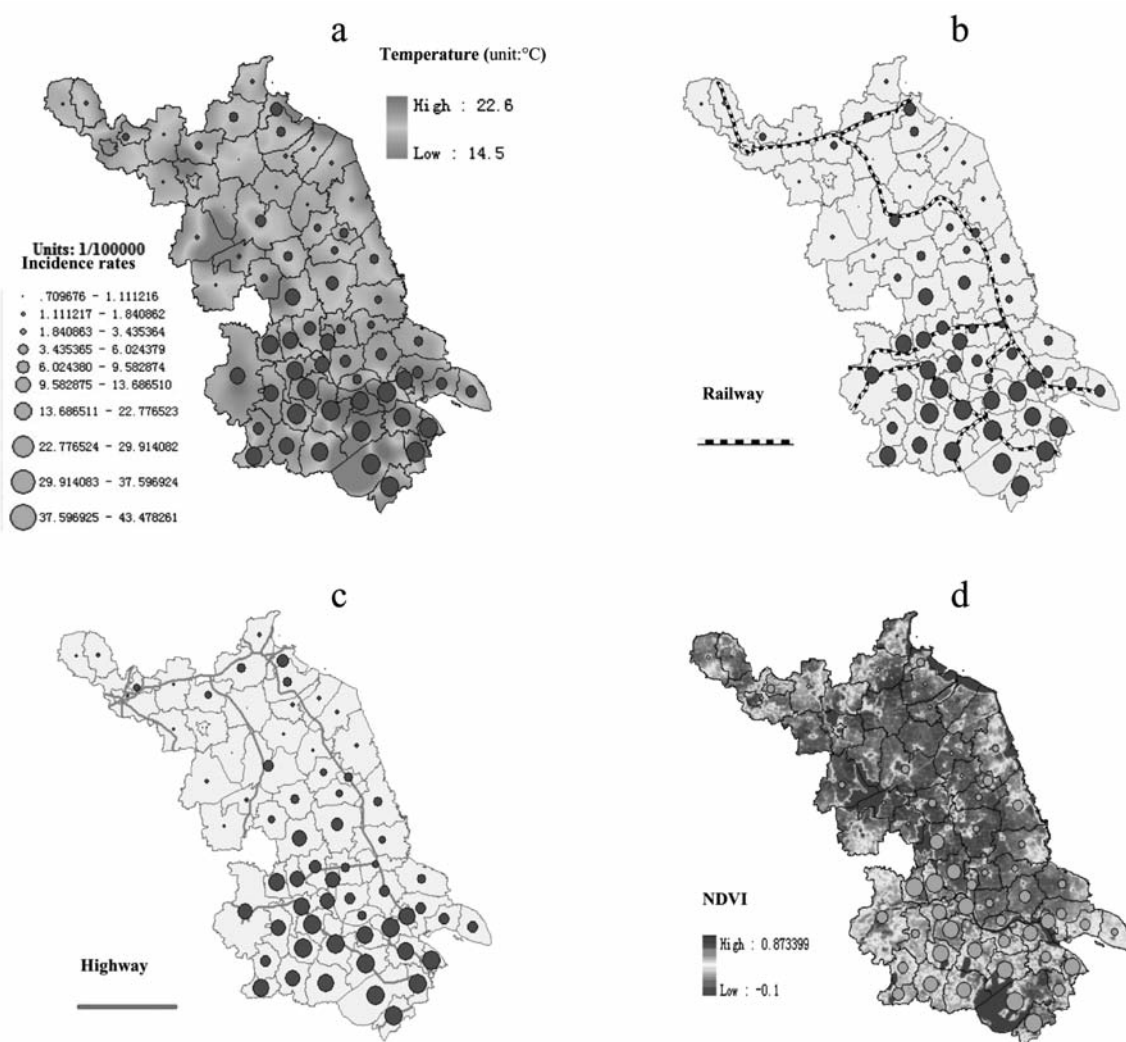

Fig. 4. Layout charts for syphilis based on risk factors.

Table 4. Parameter estimates for Poisson regression analysis.

Factor	IRR*	95% CI	P-value
Highway	1.15	1.11~1.19	<0.001
Railway	1.07	1.06~1.09	<0.001
River	1.01	1.01~1.02	<0.001
Lake	0.80	0.77~0.83	<0.001

*Incidence rate ratio

ing at about 15 km from the highways, while the situation was the opposite with regard to the railways, i.e. primary syphilis decreased with distance up to 5 km, beyond which incidence rates changed without showing a trend. Secondary and tertiary syphilis rates were negatively correlated with all distances from the railways investigated.

NDVI was used to represent the vegetation coverage during the four seasons (Fig. 4d). Syphilis was not found to be correlated with NDVI in most of the seasons but the winter season had a slight negative correlation with primary and secondary syphilis incidence rates. There was no correlation between tertiary syphilis and the NDVI. Correlation coefficients were calculated to be 0.545 (F = 77.80, P <0.001) and 0.444 (F = 51.90, P <0.001), respectively. No obvious correlation was observed with regard to the location of rivers and lakes.

Based on the multivariate regression analysis and Poisson regression, the model was established as:

$$y = 4.378 + 2.81{\times}10^{-5}\ e^{4.29}x_1 + 5.67{\times}10^{-4}\ x_2 + 0.0103x_3 - 3.91{\times}10^{-4}\ x_3^2 - 1.39\ \mathrm{Ln}\ x_4$$

where y represents total syphilis incidence rate (cases per 100,000 residents), x_1 and x_2 are distances (km) from the nearest railways and highways and x_3 and x_4 stand for distances (km) from the nearest rivers and lakes, respectively. The regression model is significant at P = 0.001 level (F = 1980.42) with R of 0.923 (R^2 = 0.852). The incidence rate ratios (IRR) for highway, railway, river and lake were 1.15, 1.07, 1.01 and 0.80, respectively. The Poisson regression results are presented in Table 4.

Discussion

Unlike traditional analytical methods, risk maps and trend surface graphs are presented in three dimensions with coordinates. Geographical information systems (GIS) made it possible to locate the spatial or temporal clusters and hotspots for syphilis and detect potential risk factors contributing to syphilis transmission.

The trend surface analysis was consistent with the risk maps, indicating that the epidemic was more severe in the south, especially the south-western counties. General spatial autocorrelation indicated that syphilis incidence was aggregated at the provincial level, and local spatial autocorrelation analysis found that the clusters comprised nine counties, among which Suzhou was a hotspot every year. In addition, spatio-temporal cluster analysis detected a cluster in 2008 and 2009 at the south-western region with Yixing county at its center.

The demonstrated incidence rates from 2005 to 2009 illustrate that syphilis continued to spread around Jiangsu province within the whole study period increasing the burden for control and prevention. The geographical variation of incidence rates of primary and secondary syphilis in contrast to the rates of tertiary syphilis, which was more evenly spread, illustrates that the former, rather than the latter, was the cause of the disproportionate distribution.

According to the Moran's *I*, the general aggregation level became increasingly higher between 2005 and 2009, indicating that there might be more clusters or, alternatively, that the clusters were more significant in the aggregation. All local autocorrelations detected were found at the middle or southern regions of the province. After 2008, and consistent with the clusters that were found in the spatio-temporal cluster analysis, Wujiang replaced Yangzhou, which illustrates that the clusters are moving southwards. The seven clusters detected in the earlier years were all located inside the range of the scanning window with Suzhou, Jintan and Danyang located adjacent to the center of the spatio-temporal cluster. In addition, the syphilis cases were more highly aggregated during 2008 and 2009. In recent years, the economic prosperity of the south-western counties attracted large numbers of migrants resulting in high population densities and increased transmission of all kinds of STDs.

High temperature was found to be slightly associated with higher syphilis incidence rates. Not surprisingly, seasons with higher temperature influence the sexual behaviour resulting in a higher possibility of getting infected with *T. pallidum* (Liang, 2007). Also, a significant negative correlation was found between NDVI in the winter and the syphilis rates. Winter conditions make transportation difficult, which affects commercial sexual behaviour, while the more frequent application of antibiotics for respiratory infections may limit syphilis transmission (Wright and Judson, 1978).

Migrant workers are believed to be a high-risk group for syphilis transmission due to commercial sex,

illegal blood sale and drug smuggling. The spatial correlation analysis illustrates that a positive correlation exist between syphilis incidence and transportation. Both railways and highways play a role in bringing in migrants presumably facilitating syphilis transmission. It is also well know that changes in the social environment and decreased social constraint when away from the family increase the risk for commercial sexual encounters. However, railways seemed to have a stronger influence on the surrounding syphilis rates (within a range of 5 km) than the highways. There are many small hotels near railway stations where commercial sex businesses are known to thrive, and more stations are found along railways than highways (Tucker et al., 2010b), which could explain this. In addition, the flow of travelers is greater on the railways than on the highways as people generally prefer the railway for long-distance travel. Under these circumstances, migrant workers are more likely to aggregate near the railway stations than near the highways.

The Possion regression analysis developed a more complex model. It detected that water resources (rivers and lakes) and transportation factors (railways and highways) could have a shared influence on syphilis, while no correlation was observed with rivers or lakes in univariate analysis. The model provided information that water resources may not influence syphilis transmission directly but could have reduced syphilis transmission due to the vicinity of railways and highways.

Among the potential risk factors for clusters and hotspots, seven autocorrelations were detected by local spatial autocorrelation analysis. However, clusters were only detected once in four of the counties (Yangzhong, Changzhou, Yizheng and Jintan). Since the spatial autocorrelation analysis was limited to spatial boundaries only and did not consider temporal factors, these clusters could have occured by chance. Further observations about the syphilis incidence rates in these areas are needed to ensure whether these areas really are at risk. NDVI in winter was calculated assuming a negative correlation with the syphilis incidence rates. However, it was not possible to detect whether the vegetation levels actively worked to reduce syphilis transmission. More high-risk behaviour that might increase syphilis transmission was not included in the study due to limited data and the difficulty of conducting the statistical analyses as indicated by Fuente (2010).

Conclusions

Despite relatively low incidence levels of syphilis in P. R. China, this STD should not be ignored. Since

migrant populations might return north on holidays, syphilis could spread to counties with currently low infection rates. Attention should be paid to the prevalence of syphilis throughout Jiangsu province targeting areas with clusters. The range of free screening and treatment programmes should be expanded, especially in the rural areas. Transmission routes and the use of prophylactic measures should be broadcast regularly.

Awareness campaigns, advocating changes in sexual behaviour, should be strengthened in areas with heavy migrant populations. Health education should be given to local residents to raise their awareness of self-protective behaviour and eliminate discrimination towards those who are already infected as mentioned by Tucker et al. (2010a). Free condoms and educational leaflets on syphilis could be distributed in railway and bus stations.

The quality of the syphilis reporting system needs to improve to avoid duplication, misreporting and under-reporting. Spatial epidemiological studies on syphilis, currently limited, should be broadened to allow interventions based on convincing spatial epidemiological studies such as cluster detection.

References

Chen XS, 2008. Epidemic updates and control strategies of syphilis in China. Int J Epidemiol Infect Dis 35, 73-77.

Chen XS, Yin YP, Gong XD, Liang GJ, Zhang WY, Poumerol G, Shi MQ, Wu SQ, Zhang GC, 2006. Prevalence of sexually transmitted infections among long-distance truck drivers in Tongling, China. Int J STD AIDS 17, 304-308.

Chen YF, Zheng DC, 2001. Analysis on the epidemiology of syphilis in Guangdong province from 1995-1999 and the strategy for syphilis control. Chin J Epidemiol 22, 191-193.

Chen ZQ, Zhang GC, Gong XD, Lin C, Gao X, Liang GJ, Yue XL, Chen XS, Cohen MS, 2007. Syphilis in China: results of a national surveillance program. Lancet 369, 132-138.

Dambach P, Machault V, Lacaux JP, Vignolles C, Sié A, Sauerborn R, 2012. Utilization of combined remote sensing techniques to detect environmental variables influencing malaria vector densities in rural West Africa. Int J Health Geogr 11, 8-20.

Fuente MJ, 2010. Syphilis, a resurgent disease. Actas

Dermosifliogr 101, 817-819.

Gong XD, Jiang WH, Wang QP, Zhang JY, 2000. Epidemiological analysis of syphilis in China from 1979 to 1998. Chin J Public Health 16, 1020-1022.

Guo W, Qu SQ, Ding ZW, Yan RX, Li DM, Wang L, Wang L, 2010. Situations and trends of HIV and syphilis infections among drug users in China, 1995-2009. Chin J Epidemiol 31, 666-669.

He ZG, Han SM, Cui DY, Zhang ZJ, 2008. Discussion on statistic of spatial autocorrelation analysis. Chin J Schisto Control 20, 315-317.

Jagai JS, Castronovo DA, Monchak J, Naumova EN, 2009. Seasonality of cryptosporidiosis: a meta-analysis approach. Environ Res 109, 465-478.

Kulldorff M, Athas W, Feuree EJ, Miller BA, Key CR, 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am J Public Health 88, 1377-1380.

Law DCG, Serre ML, Christakos G, Leone PA, Miller WC, 2004. Spatial analysis and mapping of sexually transmitted diseases to optimise intervention and prevention strategies. Sex Transm Infect 80, 294-299.

Li CM, Xiao Y, Liu JB, Ding XB, Jia YJ, 2009. HIV and syphilis infections among men who have sex with men in Chongqing municipality, China. Chin J Epidemiol 30, 126-130.

Liang MM, 2007. Study of the seasonal distribution of syphilis in Baoshan district with the circular distribution method. Occup Health 23, 1011-1012.

Lin CC, Gao X, Chen XS, Chen Q, Cohen MS, 2006. China's syphilis epidemic: a systematic review of seroprevalence studies. Sex Transm Dis 33, 726-736.

Liu H, 2000. Analysis on the epidemiology of syphilis from 1994 to 1998 in Beijing. Chin J Epidemiol 21, 85-87.

Liu XN, Huang F, Wang P, 2008. Spatial analysis theories and methods on GIS. Beijing: Science Press, 189-194.

Mak DB, Johnson GH, Plant AJ, 2004. A syphilis outbreak in remote Australia: epidemiology and strategies for control. Epidemiol Infect 132, 805-812.

Martin IE, Gu W, Yang Y, Tsang RS, 2009. Macrolide resistance and molecular types of *Treponema pallidum* causing primary syphilis in Shanghai, China. Clin Infect Dis 49, 515-521.

Peng ZH, Bao CJ, Zhao Y, Yi HG, Xia LT, Yu H, Hong BS, Chen F, 2010. Weighted Markov chains for forecasting and analysis in incidence of infectious diseases in Jiangsu province, China. J Biomed Res 24, 207-214.

Shakoor H, Madeline O, Tom K, 2010. Health effects of hot weather: from awareness of risk factors to effective health protection. Lancet 375, 856-863.

Tucker JD, Chen XS, Peeling RW, 2010a. Syphilis and social upheaval in China. N Engl J Med 362, 1658-1660.

Tucker JD, Hawkes SJ, Yin YP, Peeling RW, Choen MS, Chen XS, 2010b. Scaling up syphilis testing in China: implementation beyond the clinic. Bull World Health Organ 88, 452-457.

Wang F, 2006. Quantitative methods and applications in GIS. London, UK. CRC Press, 265 pp.

Wright RA, Judson FN, 1978. Relative and seasonal incidences of the sexually transmitted diseases. Brit J Vener Dis 54, 433-440.

Wu J, Huang JY, Xu DR, Lu CY, Deng XQ, Zhou XL, 2010. Infection status and risk factors of HIV, HBV, HCV, and syphilis among drug users in Guangdong, China - a cross-sectional study. BMC Public Health 10, 657-664.

Zhang RL, Chen LP, Chen QY, Wang XY, Zhang LP, Xiu XY, 2007. Epidemiological study on 38418 syphilis cases during pregnancy. Chin J Epidemiol 28, 749-752.

Zhou XN, 2009. Spatial epidemiology. Beijing: Science Press, 241-242.