The spatial distribution pattern of human immunodeficiency virus/acquired immune deficiency syndrome in China

Ying Wang,1 Yongli Yang,1 Xuezhong Shi,1,2 Saicai Mao,1 Nian Shi,3 Xiaoqing Hui1

1School of Public Health, Zhengzhou University, Zhengzhou; 2Editorial Department, Zhengzhou University, Zhengzhou, China; 3Cardiovascular Department of Internal Medicine, University College London Medical School, London, UK

Abstract

Human immunodeficiency virus (HIV) infection and the acquired immune deficiency syndrome (AIDS) exhibit variable patterns among the provinces of China. Knowledge of the geographical distribution of the HIV/AIDS epidemic is needed for the prevention and control of AIDS. Thus, the cumulative number of reported cases of HIV/AIDS from the period 1985-2013, and the incidence rate of AIDS in 2013 were determined. Spatial autocorrelation analysis and hotspot analysis were conducted using ArcGIS10.2 to explore the spatial distribution of the HIV/AIDS epidemic. Both the thematic map and the global spatial autocorrelation Moran’s I statistics revealed a clustered distribution of the spatial pattern. A local spatial autocorrelation analysis indicated hotspots of AIDS incidence rate that were confined to the provinces of Guangxi, Yunnan and Sichuan. The hotspots encompassed Guangxi and Yunnan, while Henan Province displayed a negative autocorrelation with more variable numbers that included neighbouring regions. The Getis-Ord G* statistics identified 6 hotspots and 8 coldspots for the incidence of AIDS, and 7 hotspots and 1 coldspot for the cumulative number of reported cases of HIV/AIDS. The spatial distribution pattern of the HIV/AIDS epidemic in China is clustered, demonstrating hotspots located in the Southwest. Specific interventions targeting provinces with severe HIV/AIDS epidemic are urgently needed.

Introduction

Since the discovery of human immunodeficiency virus (HIV) in the early 1980s and the disease caused by the infection, i.e. the acquired human immunodeficiency syndrome (AIDS), has posed a serious threat to human health. The HIV/AIDS epidemic has become a serious public health problem because of its rapid worldwide spread (Tanser et al., 2009). Studies investigating HIV/AIDS in China and abroad have shown that the epidemic varies geographically (Al-Ahmadi and Al-Zahrani, 2013). Based on available figures from 2012, sub-Saharan Africa has the most serious epidemic of this virus in the world, with roughly 25 million people living with HIV or AIDS, while the Middle East and North Africa (MENA) region has one of the lowest rates prevalence, with an estimated 260,000 people living with HIV (UNAIDS, 2013). Even within a country, the epidemic can vary widely (Li et al., 2014b). For example, in China, the cumulatively reported number of HIV infections for the period 1985 to 2013 in the provinces of Yunnan, Guangxi, Sichuan and Henan exceeded 50,000, while fewer than 1000 HIV infections were reported in Qinghai Province and the Autonomous Region of Ningxia that are situated near each other but separated by Gansu Province in the centre-northwest of the country. This finding demonstrated the strong geographical variation of the epidemic. The HIV infectiousness governs its occurrence, but this is also related to the local environment, social economy and culture among other factors. Because of this, the development and prevalence of AIDS is a spatial phenomenon with interacting diffuse characteristics. For example, individuals infected with HIV are more likely to develop tuberculosis (TB) resulting in increasing numbers of TB-HIV co-infections, and AIDS can be transmitted from high-risk groups to
the general population (Silva and Torres, 2015). To understand these developments, the relevant data need to be based on the spatial properties of the infection. Compared with traditional statistical methods, spatial statistical methods have certain advantages. For example, spatial statistics can be used to analyse correlations in space and the clustering analysis can be used to describe the occurrence of HIV infection in specified areas (Qian et al., 2014).

Geographic information systems (GIS) represent an important tool for spatial analysis that plays a vital role in strengthening the whole process of epidemiological surveillance, information management and analysis (Kandwal et al., 2009). It can effectively manage spatial data, describe the geographical distribution and variation of diseases and analyse spatial and temporal trends (Tang et al., 2014). Advances in GIS technology provide new opportunities for epidemiologists to study associations between demographics and the spatial distribution of disease (Vanmeulebrouk et al., 2008). Some scholars have applied GIS technology to explore temporal and spatial differences in infectious diseases such as malaria and severe acute respiratory syndrome, among other infections (Shiode et al., 2015). However, studies on the spatial distribution of HIV infection are limited. Peng et al. (2011) described the geographical characteristics of HIVAIDS in 125 counties of Yunnan Province based on spatial statistics and using data from the HIVAIDS sentinel surveillance system of the Chinese Center for Disease Control and Prevention (China CDC) from 1991 to 2009. The results indicate that the epidemic was most severe in the far west and northeastern regions of the Yunnan Province. Zhang et al. (2015) analysed the spatial-temporal clustering of the HIVAIDS epidemic in Chongqing, a large city in central-north China that previously belonged to Sichuan Province but now constitutes a separate province of its own. They use data from the annual reports of Chongqing Municipal Center for Disease Control from 2006 to 2012. The results revealed epidemic hotspots distributed in 15 mid-western counties. Since the spatial distribution and potential HIVAIDS clusters in the provincial level of the mainland of China has not been clarified, an improved understanding of the distribution pattern of the epidemic is needed.

Because the incubation period of AIDS is long (average estimated period of 7-10 years), it is difficult to acquire the incidence rate of HIV. On the other hand, we have a good source of data on AIDS incidence as HIV carriers are followed up every 6 months. In the present study, province-level AIDS incidence rates and the cumulative number of reported cases of HIV/AIDS in mainland China was used to examine the spatial distribution and hotspots of the HIV/AIDS epidemic. Our goal was to provide a more comprehensive understanding of the epidemic and its geographical distribution to provide a rational basis for the development of a regional policy investigating AIDS prevention and control (Xiao et al., 2013).

Materials and Methods

Data sources

We obtained the data for cumulative number of reported cases of HIV/AIDS in mainland China for the period 1985 to 2013 from the surveillance system of the China CDC, also called HIV network direct reporting information system, established by the Ministry of Health based on the disease prevention and control information system platform to improve the quality and relevance of the report. Data for the incidence of AIDS in 2013 was obtained from the China Public Health Statistics Yearbook (2014).

Statistical approach

In order to create a thematic map, data for the incidence rates of AIDS and the cumulative number of reported cases of HIV/AIDS in the 31 provinces were loaded into China’s provincial boundary administrative division vector diagram. The data were classified by the natural breaks method, which identifies the partitions capable of achieving the minimum difference between the specimens in the same segment and the maximum difference between specimens in the various segments (Bivand, 2013).

General spatial autocorrelation

Moran’s autocorrelation coefficient ($I$) was used to measure the correlation among neighbouring observations and the levels of spatial clustering among neighbouring districts (Zulu et al., 2014). It indicates the direction and degree of a single variable that is strictly attributable to a relatively close position and can therefore be used to explore overall spatial patterns (Osei and Duker, 2008). The general Moran’s $I$ was calculated as follows:

$$I = \frac{n}{\sum_{i=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}, \quad i \neq j$$

where $n$ is the number of study areas, $w_{ij}(d)$ represents the adjacent weight matrix from the distance and $x_i$ refers to the incidence rate of AIDS (or the cumulative number of reported cases of HIV/AIDS) in the $i^{th}$ and $j^{th}$ province.

Under random conditions, the index hypothesis of an approximately normal distribution is usually standardised by the $Z$ score. The value of Moran’s $I$ ranges between approximately -1 and 1. The magnitude of $I$ stands for the strength of the spatial autocorrelation, where $I=1$ denotes a strong positive spatial autocorrelation between study regions and $I=-1$ a strong negative spatial autocorrelation.

Local spatial autocorrelation

The global Moran’s $I$ statistic only discloses whether a distribution pattern is clustered or random without giving detailed cluster information about site or features, such as high-high or low-low clusters, high-low or low-high outliers (Moraga and Montes, 2011). The local spatial autocorrelation makes up the disadvantage of the global Moran’s $I$ statistic. It is used to examine whether there are regional patterns involving specific areas (Barrell and Grant, 2013). The local Moran’s $I$ index was used to analyse cluster features and the local Getis-Ord $G^*$ statistics (Getis and Ord, 1992) to test the focal provinces of the HIVAIDS epidemic. The local Moran’s $I$ was calculated as follows:

$$I_i = \frac{n}{\sum_{j=1}^{n} w_{ij}(x_j - \bar{x})(x_j - \bar{x})}{\sum_{j=1}^{n} (x_j - \bar{x})^2}, \quad i \neq j$$

The value of the local Moran’s $I$ ranges from $+1$ (indicating high-high or low-low clusters) through $0$ (=random pattern) to $-1$ (indicating high-low or low-high outliers). The local $G$ index was calculated as...
The meaning of the parameters in both equations are same as those described for the general Moran's $I$ formula. If $G'_i$ is positive, the local clusters are identified as high-value correlations, i.e. statistically significant hotspots; if $G'_i$ is negative, the local clusters can be identified as low-value correlations, i.e. statistically significant coldspots.

ArcGIS software version 10.2 (ESRI, Redlands, CA, USA; http://www.esri.com/software/arcgis/arcgis-for-desktop), was used to map the province-level HIV/AIDS epidemic in China. The data were subjected to spatial autocorrelation analysis and hotspot analysis. The level of significance was 0.05.

**Results**

The thematic map of the HIV/AIDS epidemic in China in 2013, stratified by province and classified into five levels with respect to incidence is shown in Figure 1. Guangxi and Yunnan had the highest rates and were thus classified as being at the fifth level, while the eight provinces with the lowest rates, including Inner Mongolia, Gansu and Ningxia among others, were classified as being at the first level.

Based on the cumulative number of reported cases of HIV/AIDS, Guangxi and Yunnan showed the largest figures and were therefore placed at the fifth level. The eight provinces with the smallest numbers of AIDS cases were positioned at the first level and included Inner Mongolia, Gansu, Ningxia and so on. In general, the regions with the most severe HIV/AIDS epidemics were concentrated in southwestern China, while regions with a low prevalence were located in north-eastern China, north-central regions, and the western plateaus including Qinghai and Tibet.

Table 1 shows the indexes for the global spatial autocorrelation. There was a statistically significant positive spatial autocorrelation

**Table 1. Global Moran's I statistics for the human immunodeficiency virus/acquired immune deficiency syndrome epidemic in China.**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Moran's $I$</th>
<th>$E (I)^\circ$</th>
<th>Var $(I)^#$</th>
<th>$Z^§$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS incidence rate</td>
<td>0.203</td>
<td>-0.033</td>
<td>0.002</td>
<td>4.746</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cumulative number of HIV/AIDS cases</td>
<td>0.075</td>
<td>-0.033</td>
<td>0.003</td>
<td>2.140</td>
<td>0.032</td>
</tr>
</tbody>
</table>

AIDS, acquired immune deficiency syndrome; HIV, human immunodeficiency virus. $^\circ$ Expected index value; $^\#$ theoretical variance; $^§$ standard normal deviation ($Z = \frac{I - E (I)}{\text{Var} (I)^\#}$). $P<0.05$ is considered significant.
between the incidence rate of AIDS and the cumulative number of HIV/AIDS cases. These findings indicate that the distribution of the HIV/AIDS epidemic in China was clustered distribution throughout the entire study region.

Based on the incidence rate of AIDS, high-high regions included Guangxi, Yunnan and Sichuan, which not only demonstrated a high incidence rate of AIDS but also were surrounded by provinces with high incidence rates of the disease. Based on the cumulative number of reported cases of HIV/AIDS, Guangxi and Yunnan were located in high-high regions. Henan was located in a high-low region, which presented a large number of cases, but was surrounded by provinces with few cumulative numbers of reported cases. In general, high-high regions were located in south-western China (Figure 2).

Figure 3 shows the hotspots, six of which were identified based on the incidence of AIDS, including the provinces of Guangxi, Sichuan, Yunnan, Guizhou, Hainan and Hunan, which represented the higher transmission of HIV and were located in the South of the country. Eight coldspots, including Beijing and the provinces of Tianjin, Hebei,
Shanxi, Shandong, Jiangsu, Inner Mongolia and Liaoning in northeastern China were identified. Based on the cumulative number of reported cases of HIV/AIDS, Chongqing and the Guangxi, Sichuan, Yunnan, Guizhou, Hunan and Hainan provinces were hotspots, while Inner Mongolia was a coldspot.

Discussion

In the present study, we took advantage of advancements in the GIS to explore geographic distribution patterns and hotspots of the HIV/AIDS epidemic in China. The main results indicated that there was a clustered spatial distribution pattern of the HIV/AIDS epidemic; hotspots were located in the southwestern part of the country and coldspots in the Northeast. Based on the spatial autocorrelation analysis, we found that the HIV/AIDS epidemic exhibited geographical variation and a clustered distribution pattern, indicating the presence of hotspots. Figure 1 shows that the most severe HIV epidemic areas were located in the Southwest, which indicates that the specific interventions targeting these areas should be strengthened and regional health planning and resource allocation adopted.

The results of the local spatial autocorrelation analysis revealed that three provinces were autocorrelated based on the incidence rate of AIDS and two were autocorrelated due to the cumulative number of reported cases of HIV/AIDS. Henan was found to be located in a high-low region, with more HIV/AIDS cases there but less HIV/AIDS cases in its neighbouring provinces. Owing to its poorer economic status, Henan accepts paid blood/plasma donation since the 1990s (Yan et al., 2013), which could have an impact on the number of cases there. Indeed, an increasing number of HIV/AIDS cases have been reported due to blood transfusion, which is the primary transmission route of HIV in Henan (Wu et al., 2006). In other provinces, risky sexual behaviour and intravenous drug use are the main routes of transmission (Zhang et al., 2011). Another reason for more HIV/AIDS cases in Henan is that the province has a big population, ranking third in this respect in China in 2010 (Hu et al., 2015). It is also worth mentioning that the incidence rate of AIDS in Henan has remained relatively stable in recent years. A convincing explanation for this phenomenon results from the great measures implemented by the government to minimise the illegal blood trade.

The Getis-Ord statistics further support the findings obtained for the local spatial autocorrelation analysis. Hotspot regions were identified in six provinces based on the AIDS incidence rate and in seven provinces based on the cumulative number of reported cases of HIV/AIDS. There are two probable explanations for the severe HIV epidemic in Guangxi and Yunnan. One explanation concerns geographical factors (Shan et al., 2013). Guangxi and Yunnan are located in the heroin trafficking route, which begins at the Golden Triangle and continues to the south-western or western provinces of the country, potentially expanding the transmission of HIV/AIDS. In addition, frontier trade and travel contribute to the growing number of HIV/AIDS cases (Li et al., 2014a). Another explanation is the presence of a large mobile population, which is a population that is associated with a higher risk of HIV transmission due to poor knowledge of the disease and thus lower usage rates of condoms and/or other high-risk behaviour (Zhou et al., 2014). Sichuan Province is adjacent to Yunnan province, where there is active drug trade. Due to its less developed economy, there is a large outflow of rural residents to places like Yunnan, who may assume drug-related habits (Li et al., 2015). In addition, some of the ethnic minorities in Sichuan have tolerant attitudes regarding sexual behaviour (Liu et al., 2013). These high-risk factors have most probably played role in the spread of AIDS. Some other risk factors may be present in other hotspots, including risky sexual behaviour (homosexual and bisexual contacts), disorderly management of blood and blood products, unsafe medical service, low levels of AIDS-related knowledge and poor economic status (Zhuang et al., 2012).

Compared with the status of hotspots, coldspots were confined to eight provinces based on AIDS incidence rates and to one province based on the cumulative number of reported cases of HIV/AIDS. There are several likely explanations for these observations: highly developed economy and culture, high-speed flow of information high level of public cognition and knowledge of AIDS and more widespread understanding of the importance of self-protection (Teng and Shao, 2011). It is also worth noting that there was an increased proportion of men having sex with men (MSM) among the annually reported HIV cases in these regions (Liu et al., 2015). The concealed features of MSM constitute a dangerous bridge for HIV transmission from high-risk groups to the general population. Comparatively, the proportion of blood transmission, mother to child transmission, and intravenous drug abuse transmission is decreasing and stabilising at low levels in these provinces, thus maintaining a small epidemic of HIV/AIDS (He et al., 2011). The spatial analysis found a spatial autocorrelation of HIV/AIDS at the provincial level, which was consistent with the spatial distribution of HIV/AIDS epidemic in Yunnan and Chongqing at the county level. Meanwhile, the hotspots observed in the south-western part of China provide rational evidence for making for a regional policy concerning AIDS prevention and control.

This analysis has a few limitations. First, our analysis is based on the cumulative numbers of reported cases of HIV/AIDS and may not be a true representation of the actual spread. However, it shows strength in reflecting the disease burden and the effect of long-term intervention as having a direct influence on the allocation of health resources. Second, due to limitations associated with the surveillance system data sources, and the unclear high-risk population base, we could not obtain detailed cumulative numbers of HIV/AIDS cases and incidence rates classified by the transmission route. In future studies, it will be attempted to estimate the high-risk population base and the incidence rate by transmission route to better explore the spatial distribution of HIV/AIDS with the goal of identifying early warning of potential new epidemics.

Conclusions

The present study analysed the spatial distribution and hotspots of the HIV/AIDS epidemic, thereby providing a scientific basis for the development of regional policies for AIDS prevention and control of the spread of HIV/AIDS epidemics.

References


